Wireless Communications and Mobile Computing

Recent Advances in Physical Layer Technologies for 5G-Enabled Internet of Things 2022

Lead Guest Editor: Xingwang Li
Guest Editors: Han Wang and Khaled Rabie

Recent Advances in Physical Layer Technologies for 5G-Enabled Internet of Things 2022

Wireless Communications and Mobile Computing

Recent Advances in Physical Layer Technologies for 5G-Enabled Internet of Things 2022

Lead Guest Editor: Xingwang Li
Guest Editors: Han Wang and Khaled Rabie

Contents

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1971902, 12 pages
https://doi.org/10.1155/2022/1971902

Research Article

Implementation of an IoT-Based Solar-Powered Smart Lawn Mower

Tayyab Tahir,[1] Adnan Khalid,[1] Jehangir Arshad,[2] Aun Haider,[1] Iftikhar Rasheed,[3] Ateeq Ur Rehman,[4] and Seada Hussen[5]

[1]*Department of Electrical Engineering, University of Management and Technology, Lahore, Sialkot Campus, Sialkot, Pakistan*
[2]*Department of Electrical & Computer Engineering, COMSATS University Islamabad, Lahore Campus, Lahore 54000, Pakistan*
[3]*Department of Information & Communication Engineering, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan*
[4]*Department of Electrical Engineering, Government College University, Lahore 54000, Pakistan*
[5]*School of Electrical and Computer Engineering, Haramaya Institute of Technology, P.O. Box 138, Diredawa, Ethiopia*

Correspondence should be addressed to Ateeq Ur Rehman; ateeq.rehman@gcu.edu.pk
and Seada Hussen; seada.hussen@aastu.edu.et

Received 16 May 2022; Revised 5 July 2022; Accepted 22 September 2022; Published 6 October 2022

Academic Editor: Khaled Maaiuf Rabie

Rapid growth in technology has created opportunities to design and develop high-end applications and tools. Conventional mowers in practice are mostly fuel-powered and require personnel assistance for operation. This work develops a smart lawn mower powered by a solar photovoltaic (PV) panel and controlled by an Internet of Things- (IoT-) based technique. The designed lawn mower comprises one brushless direct current (BLDC) motor, four gear motors, sensors, an Arduino-based charge controller, and a Raspberry Pi-powered renewable energy source making it a sustainable device. A lawn mower is operated and controlled through an Android application. Raspberry Pi is used as an edge computing device for transmitting data through the Internet and for communication with Android applications. Arduino UNO is used for energy management and motor control operation. The main novelty of this research is IOT-based motion control feature which provides the user the provision to operate the mower remotely. Results of the designed model depict an average of 89.5% electrical efficiency of the system based on varying weather conditions. Application of the designed model is golf clubs, playgrounds, and lawns eliminating operator costs, saving energy, reducing noise pollution, and achieving environmental sustainability goals.

1. Introduction

With the advancements in technology, conventional ways of performing different tasks have been replaced with improved methods. Human work is being replaced by machines fully or partially, and automation is getting involved in commercial and domestic processes. Like in different industries to increase production and to improve accuracy, the use of robotics is becoming very common day by day. Similarly, daily life domestic tasks are also getting automated with the help of recent technologies. Along with automation, solar energy is another area that is becoming popular over time. It has become an important source of power generation because it is pollution-free. Sun is the main source of solar energy, and the energy through the sun is about 12 times the total world's energy demand. PV cells are being used to convert solar energy into electrical energy, and this process has an efficiency of about 18% to 23% [1]. Automation and solar photovoltaic energy are the two modern technologies that have been used in this research to improve conventional lawn mowers.

1.1. Background. A few decades back, a lot of human effort was required for grass-cutting tasks. Over time, a large amount of human effort has been replaced with a single person and a machine. These machines were operated manually by humans. Later on, these machines were upgraded with engines and sharp blades as shown in Figure 1 [2]. These engines use conventional energy sources which are not environment friendly and result in

Figure 1: Conventional lawn mower [2].

global warming. Global warming is one of the major challenges that humanity is facing, and it needs to be addressed seriously. The engines used for mowers are getting replaced with electric motors which require electrical power for their operation, and the electrical power used for motors still comes from conventional energy sources. The same is the case with conventional lawnmowers. Such mowers also need a long cord to power the electric motors of grass cutters that is very dangerous as it operates on high voltages. It also limits the operational range of these mowers because of the limited cord length. Increasing the chord length also increases operational costs. Despite the addition of motors, lawn mowers still require a significant amount of human effort to operate. So, this human effort can further be reduced by using the latest technology as proposed in this work.

1.2. Literature Review. Solar-powered smart lawnmower developed in [3] is an advanced version of a grass-cutting machine. It gets its power from a photovoltaic module mounted on the top and uses smart control methods for the movement of the lawnmower. The electrical energy produced by the solar panel is stored in a 12 V battery by using a charge controller. The complete operation of this lawnmower is controlled by computing devices used in the design. Motors are controlled by the driver circuit. Mowing is done using a high-speed DC motor which provides the torque required to drive the blades directly coupled with the shaft of the motor. The IoT-based control of this lawnmower is based on Hypertext Transfer Protocol (HTTP) requests and a python code written to control the lawnmower. The unique Internet Protocol (IP) address of the lawnmower along with the Transmission Control Protocol (TCP) request is used to control the lawnmower. Furthermore, an Android application is designed to call HTTP addresses when any key is pressed. All the keys have different addresses assigned to them.

In [4], a solution for the solar-powered smart lawnmower is proposed, where a PIC 16F877 microcontroller for robotic operation is used. While for solar charging, they used a direct online (DOL) connection method in which the solar panel is directly connected to the battery for charging. A gyroscope sensor is used to control the directions of the system. The microcontroller reads distance in meters and moves to the left side when the meter is complete. In [5], an artificial intelligence- (AI-) based lawn mowing model is proposed. In this paper, a smart lawn mowing method is proposed using computer vision (CV) and ultrasonic sensors. A Raspberry Pi 3 is used as the main controller which runs the python code for AI. An ultrasonic sensor is used in this work for distance measurement from an obstacle. The Pi camera is used for video streaming and based on convolution neural network (CNN), the obstacle detection model is applied to the video stream. Raspberry Pi detects the obstacles and sends signals to move the motor driver according to the AI algorithm. In [6], three motors are used to make the design cost-efficient. Two motors are directly coupled with rear wheels while the front wheel is freely moveable. For solar charging, a PWM-based solar charge controller is used while Arduino is used as the main controller of the project. In this paper, the author performs a stress-strain analysis on solid works for a lawnmower. The overall system efficiency is 93.37%. Likewise, some other lawn mowers are proposed in [7, 8]. In these papers, Arduino UNO microcontroller interfaced with Bluetooth module is used for main control of lawn mower. The Bluetooth module is connected to an Android application. Users send navigation commands from the Android application. The proposed design is simulated on Proteus, and it is verified through hardware implementation [7]. While in [8], additional PIR sensors along with ultrasonic sensors are used for pathfinding and obstacle detection.

A model of solar-powered lawn mower controlling RF communication is proposed in [9]. An HT12E encoder IC is used to convert the received parallel data to serial data. The microcontroller is programmed according to the time interval of HT12E output. Each functionality is differentiated by the high duration of each channel. HT12E starts working at a low signal at the TE pin. This low signal on the TE pin starts the communication of 4 data bits. Likewise, [10] proposed a cost-efficient autonomous smart lawnmower. It is a two-wheeled robot that can also be controlled manually using the mobile app. ESP8266 is used as a microcontroller for this robot. For remote control operation, a camera is installed on ESP8266 for live video streaming on the mobile screen. L298N motor driver is also used to drive the wheel motors, whereas the speed of motors is controlled by PWM. Similarly, a vision-based solar-powered robotic mower is presented in [11]. Unlike other robotic mowers, this design requires no wires to maintain its operation within the lawn. The design also detects and avoids objects and humans on the lawn. Another solar-powered lawn mower is developed in [12]. It is comprised of a direct current (D.C) motor, a rechargeable battery, a solar panel, a stainless-steel blade, and a control switch. Performance evaluation of the developed machine is carried out with different types of grasses and the design gives good results. Moreover, in [13], another solar-powered lawn mower is implemented. A design is proposed for the lawn mower that is powered through solar energy as well as electrical energy. This reduced both environment and noise pollution.

Most recently, a fully autonomous solar-powered lawn mower is presented in [14]. In this design, 12 V batteries are used to power all the electrical components. Arduino

microcontroller is used to control the motors used for the movement of the mower and the grass cutter motor. Based on information coming from the sensors, controller makes the corresponding decisions. The 12-volt battery used in this design is continuously charged by a solar panel mounted on the top of this mower. The presented design is simple and cheaper than other solutions. Another solar-powered lawn mower is implemented on the principle of mowing in [15]. The components used in this work are a direct current (DC) motor, battery, solar panel, steel blade, and speed controller. The required torque necessary for the steel blade was achieved through the DC motor. The speed controller is used to control the speed of the DC motor. The battery involved in this work is recharged through the charging circuit containing a solar panel and charge controller. The performance of the developed mower is evaluated with various thicknesses and shapes of the cutting blade. It was found that the cutting efficiency of the mower ranges from 70.50% to 84.10%. Moreover, in [16], fabrication and working of a smart solar grass cutter are presented. An efficient and environment-friendly solar-powered lawn mower is designed for this work. A microcontroller is used to control the different lawn mowing movements and actions. Two DC motors are used for the movement of the grass cutter, whereas one DC blade motor is used for grass cutting. It is designed to be remotely controlled using an Arduino UNO microcontroller. A Smart Solar Tracker is also implemented via Bluetooth and a smartphone. It can operate for more than two hours when the battery is fully charged. A comparison of the abovementioned design is presented in Table 1.

There are some other lawn mowers available that lack the feature of being solar-powered; yet, their working and design are very significant. These are nonsolar but electrically powered and autonomously control lawn mowers, like a robotic system is developed in [17]. This work deals with the fabrication and design of the prototype. For the computer-aided design and simulation of an automated lawn mower, the author used Auto-desk Inventor 2018. The electrical system containing the connections between the microcontroller and the rest of the electrical system is designed to cut grass efficiently with minimum or zero human effort. The designed model of lawn mower contains GPS, cameras, infrared, and ultrasonic sensors for obstacle detection. Arduino microcontroller is used to control the two geared motors to control the direction of the lawn mower. The controlling is done by readings from the sensors connected to the microcontroller. The mower makes a 180-degree turn in case of any obstacle is detected in the path, and it cuts the grass with high efficiency. Similarly, another automated lawn mower is developed in [18]. This paper deals with the design of a smart lawn mower able to distinguish between obstacles using image processing. A user interfaces mobile phone application is designed so that users can choose the slow mode, inching mode, and obstacle avoidance mode. Rather than using wheels and tires, crawler tracks are used for better surface gripping. In this lawn mower, the built-in python toolkit "Tkinter" is used for the actions to be taken for the lawn mower using a human-machine interface. The Digital Signal Processing (DSP) chip is used as the core controller of the project while Raspberry Pi is used for image recognition and human-machine interface. Another fully automated lawn mower that can be used with minimum human effort is presented in [19]. The lawn mower is designed to cut the grass within a predefined ground area without any user intervention. It can also be used in any other location using radio control (RC). In this model, Raspberry Pi 3 is used as the main controller communicating with the RC transmitter and Navio2-HAT for GPS location. Another battery power lawn mower is proposed in [20]. The primary focus of this project is to get higher electrical efficiency and reduce the cost by the use of Li-ion batteries and brushless direct current (BLDC) motors. Simple electrical circuits using operational amplifiers and voltage sensors are used for the protection of motors and batteries.

With the advancement in technology, new approaches are being employed for lawn mowing. This context [21] highlights some of the problems in conventional lawn mowing techniques and models. As a solution, an electrically powered lawn mower vehicle is designed for this work. The mower is powered by Li-ion batteries, and a solar platform is installed for the charging of batteries. The mechanical structure is flexible to change the height of the blades during grass cutting. Image processing techniques are also used for obstacle detection and path planning. Another solar-powered lawn mower is developed in [22] with an ARM controller along with ultrasonic and color sensors. A color sensor is used for grass detection while an ultrasonic sensor is used for obstacle detection. The system is tested on different ground surfaces, and it is concluded that the model can be used efficiently on flat surfaces.

The sensor data is crucial for robotic navigational tasks like path planning and collision or obstacle avoidance. AI algorithms may have the ability to adapt obstacle detection and pathfinding with much higher efficiency than microcontrollers. In this context, a method for obstacle detection using digital image processing is proposed by [23]. The paper deals with object classification that usually exists on football grounds. This work deals with image filtering, segmentation, and edge detection using image processing. RoboRealm tool is used for object detection and recognition, whereas [24] proposed a model for obstacle detection in the path using machine learning. The model can estimate the distance, scale information, and distinguish pedestrians from other barriers using a camera. The machine learning model is divided into two branches: obstacle estimation and pedestrian detection. These two branches are used for infers scales and pedestrian categorical detection, respectively. The model is tested, and results are collected during a real-time experiment with a maximum distance error of 5% and 94% pedestrian detection accuracy.

Nowadays, Android phones have become an important part of our daily life. Rather than using radio control methods for robotics, engineers are working on smartphone applications to control robots directly from Android applications. There are some lawn mowers presented recently that are working with Android applications. The implementation of Android applications for mower operation results in low-cost lawn mowers. In this regard, Android-controlled lawn mowers are proposed in [25, 26]. Like other

Table 1: Literature review comparison.

Ref #	Controller	Algorithm/technique	Deficiency
[4]	PIC	Obstacle detection	No proper method for the movement of the mower
[8]	Arduino mini	Bluetooth module, PIR motion sensor	Low range communication, insecure channel
[9]	8-bit microcontroller	RF communication	No monitoring system, no sensor for obstacle
[15]	ATMEGA32 microcontroller	Self-control	No predefined path, low solar efficiency

models, an Arduino microcontroller is used as the main controller for these mowers while a Bluetooth module is used to communicate with the Android application. The Android applications for these projects are designed in MIT app inventor. The fusion of robotics with the IoT has given rise to a new idea known as IoT-aided robotics technology or the Internet of Robotic Things (IoRT). The IoRT technology has also resulted in the development of effective mobile robot applications. Such mobile robot applications can be used for services, household tasks, and healthcare. Like in [27], an IoRT lawn mower working with IBM and cloud services is presented. IBM cloud is used as a means of communication between Android applications and microcontrollers. Furthermore, an IoT-based robot lawn mower is developed [28]. In this work, the assembly of the mower is based on PVC pipes, while an additional water sprinkler feature is also added to the robotic mower. Users can operate sprinklers and control the mower through Bluetooth connectivity with an Android phone. Industrial IoT applications are carried out using a safe infrastructure to communicate with sensors and surveillance cameras. Thenceforth, the Industrial Internet of Things (IIOT) can be used with machine learning-based algorithm for monitoring and event detection [29].

There are some nonprogramming-based online tools available for the development of the Android application. MIT app inventor is one of these open-source online tools for designing and developing custom Android applications [30]. HTTP requests are one of the widely used methods for Android communication, while a microcontroller connected to the local area network is used as a server and receives HTTP requests over the same network. To control any GPIO pin of the NodeMCU, a unique request along with IP address and port number is accessed through a web browser while connected to the same network [31]. Based on such technology, a mobile robot is proposed in [32] that can be controlled using Wi-Fi. The primary aim of this robot is to expand the control radius of the robot. Wi-Fi has a wide range as compared with other communication mediums like Bluetooth and Infrared. ESP8266 development module is used in it as a microcontroller. A Telnet SSH server is created on LAN which made it possible to be connected with external clients. The PC control software sends the speed parameters of the robot motors (in the range of 0 to 100), and it is connected to the control unit of the mobile robot (server) as a client via TCP/IP connection. The speed of the rear motors can be changed by changing the PWM pulse width. The application managing TCP/IP-based connection and control of the mobile robot is built up of two classes. In the first class, there is the description of TCP/IP connection, where the responsibility of the TCP client class is to establish, maintain, and close new connections on the IP address and port. In the second class, there is the description of performing control of the mobile robot by cursor keys.

1.3. Objectives. Considering all the issues mentioned in the background and limitations in the previous work mentioned in the literature review, a solar-powered smart lawnmower is proposed as the best solution. As the world is focused on using renewable energy resources to overcome greenhouse gases. This work will be very useful for reducing fuel consumption. Moreover, it significantly reduces further human efforts required for the process of grass cutting. The designed lawnmower is applicable on large ground areas like golf clubs, cricket grounds, or parks where a large area is needed to be mowed. Conventionally, a large number of human operators or ground keepers need to be hired for lawn mowing, but this lawn mower is very useful in such areas as a single user can cut grass without operating it manually. Moreover, solar energy allows operating this mower in large areas where power cords cannot reach power grid-powered mowers. The novelty of this research involves two aspects. Firstly, the lawn mower is powered with PV that powers all the electrical equipment of lawn mower, and it supplies electrical power to charge the battery to be used during night time. Secondly, the mower is controlled using IoT-architecture which made it possible to reduce the human labor.

The section division of this paper is as follows. Section 2 outlines the design and implementation, whereas Section 3 contains Results and Discussion. Finally, Section 4 contains the conclusion of this research work.

2. Design and Implementation

The work is divided into two parts. In the first part, detail about hardware is mentioned, whereas the second part is composed of software architecture developed for the proposed design. Figure 2 is the block diagram of the proposed design.

2.1. Master Controller. Raspberry Pi is used as a main controller for the smart lawnmower. The General-Purpose Input Output (GPIO) pins generate a PWM signal to control the speed of motors. H-bridge motor driver circuit is designed to control wheel motors, as there are four-wheel motors each motor is controlled by a single H-bridge circuit. The duty cycle of generated PWM signal is based on the

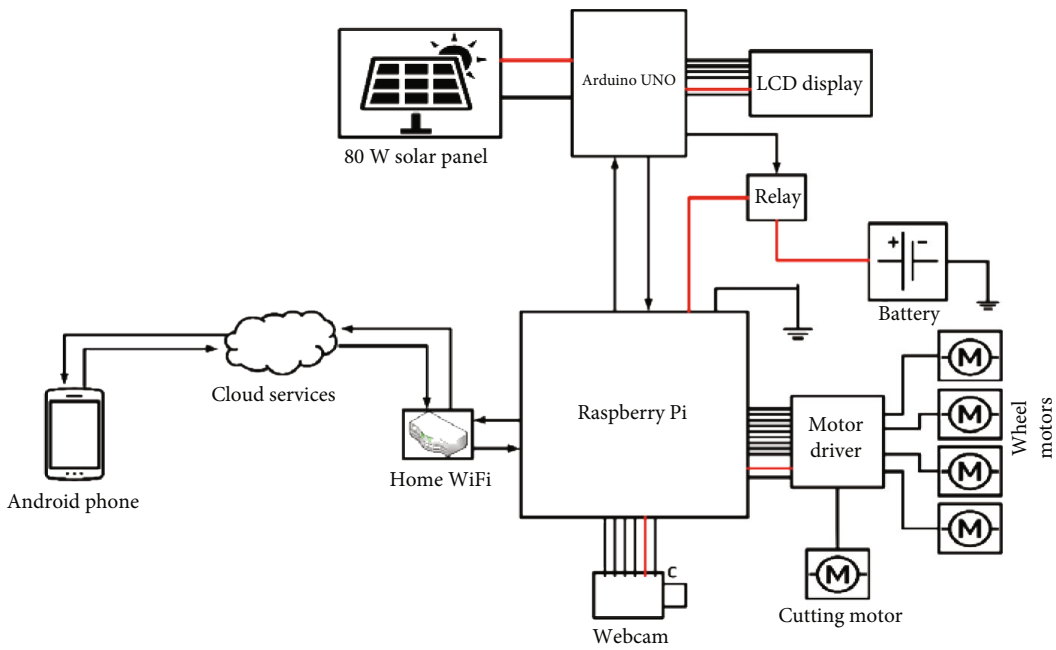


Figure 2: Block diagram of the Smart Lawn Mower.

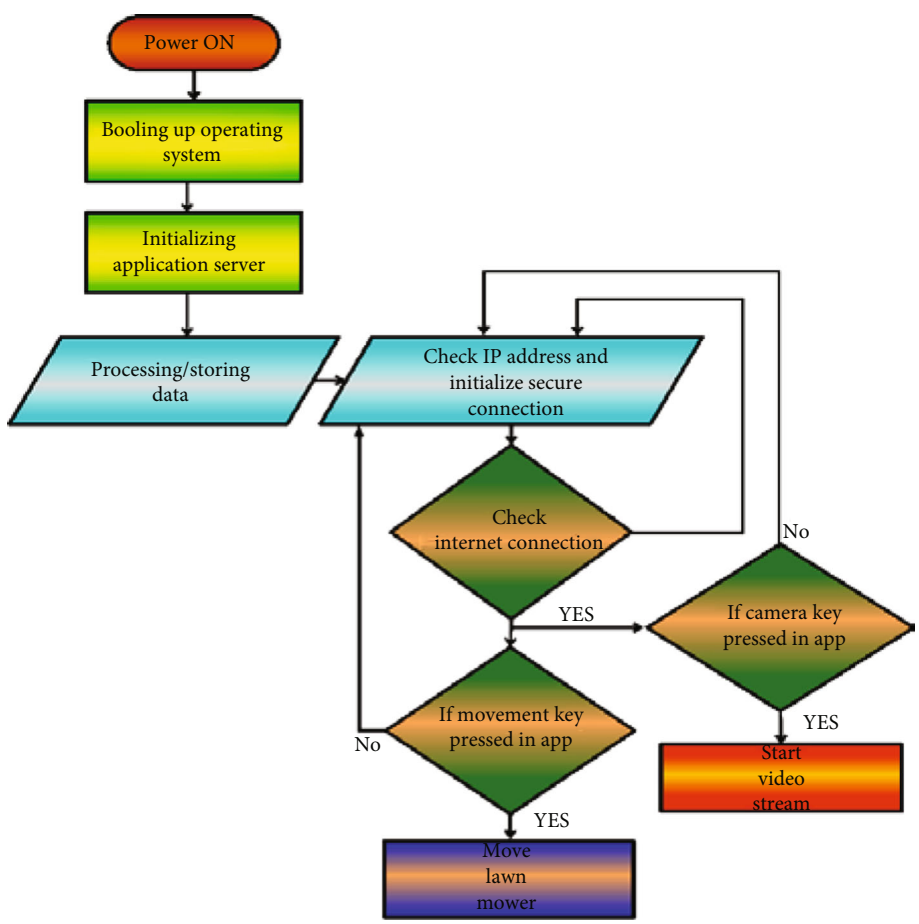


Figure 3: Local area server communication flowchart.

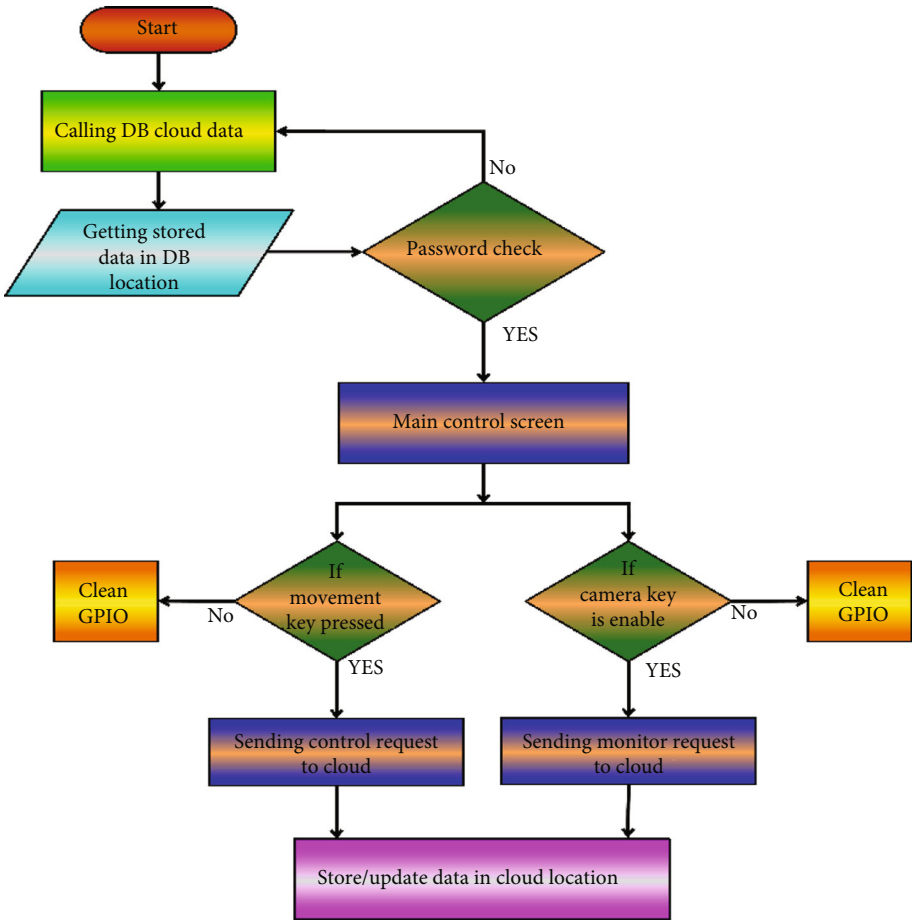


Figure 4: Android control application flowchart.

throttle value received by the operator. Raspberry Pi can generate a PWM signal ranging from 10 Hz to 8KHz. In this lawn mower, a 1 kHz frequency is set for the PWM signal. A boost converter or voltage converter is needed to change the 3.3 V signal to a 5 V signal generated from Raspberry Pi. This level shifting is important because the H-bridge motor driver circuit is based on MOSFETs. N-channel MOSFET triggers at 4.5 V gate voltage, whereas Raspberry Pi can only generate a 3.3 V low current signal. Raspberry Pi is programmed to control the movement of the mower. For this purpose, a local area network (LAN) server is created, communicating between Raspberry Pi and the Android application. A video camera is installed in front of the mower, and LAN server transmits that video to the Android application. Following Figure 3 shows the flowchart of the LAN server of Raspberry Pi.

2.2. Control Application. An Android application is designed to control a lawn mower. To connect with the mower, the user first login the application using a secure password. This password completely depends on the user and can be changed afterward. The Android application communicates to the mower by entering the IP address of Raspberry. The IP address of Raspberry Pi is set to a static IP address so that it will not change when connects to the network. When any movement key is pressed on the Android application, it sends a unique HTTP request to the server. The server receives and responds to the request by generating a PWM signal for the respected motor and also sends a response message to the application. Figure 4 shows the flowchart of the Android application.

2.3. Solar Charge Controller and Slave Device. In this lawn mower, Arduino is used as a slave device. It reads analog voltages from the voltage sensors and transmits them to the master controller (Raspberry Pi) upon request from the master. As there is no Analog to Digital converters (ADC) in Raspberry Pi, so Arduino is used for the solar charge controller. It senses the solar and battery voltages and generates PWM signals to control solar voltages. The maximum power point tracking (MPPT) algorithm has been implemented in the slave microcontroller to get maximum efficiency from the solar panel. Arduino reads the voltage and load current of the solar panel and generates a PWM signal. The generated PWM signal is provided to the buck-boost converter which controls the output of the solar panel. A 16×4 LCD is attached to the back of the mower, which displays the health of the battery and the parameters of solar. The

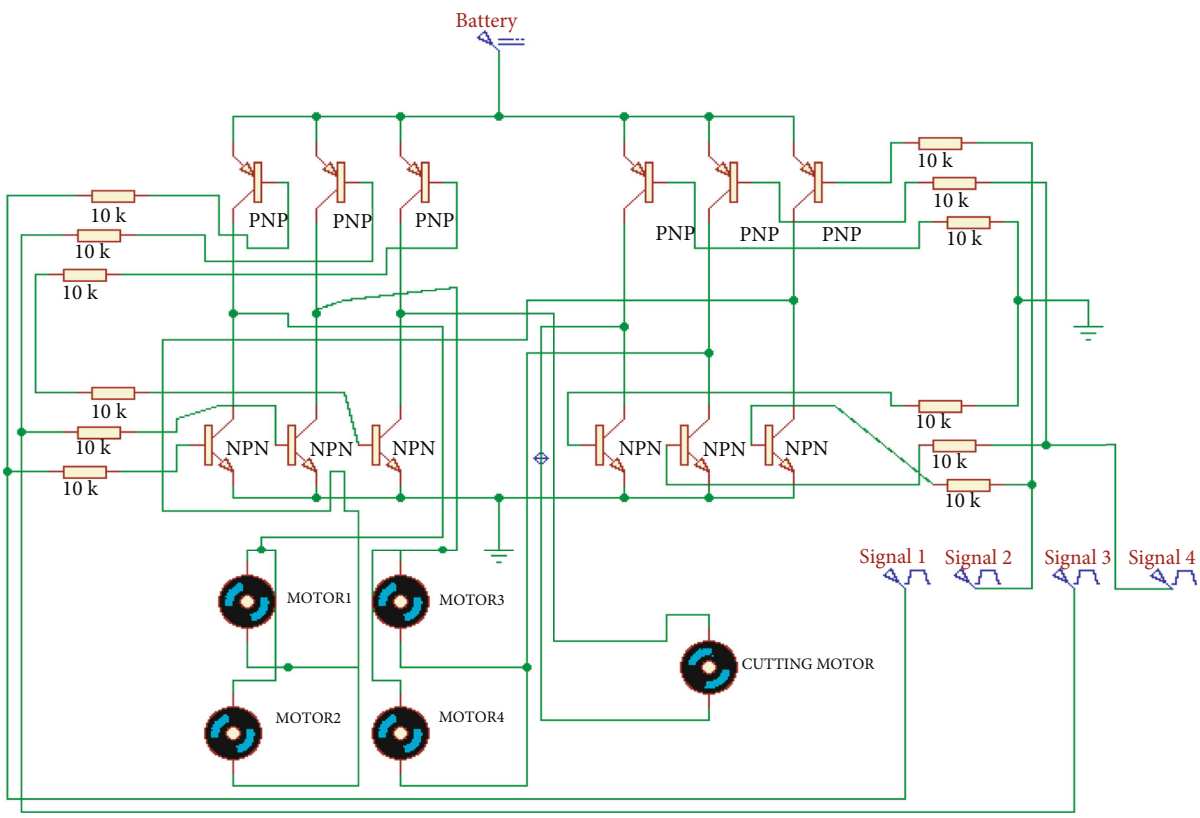


Figure 5: H-bridge motor driver circuit PROTEUS simulation.

working of the designed MPPT charge controller is based on fill factor (FF) curve. The voltage sensor of Arduino gives real-time voltage on connected load, whereas the current sensor gives load current. A V-I curve based on these voltage and current values is estimated to find the fill factor. The FF curve is estimated by recording maximum values, obtained values of voltage, and current sensors for 2000 microseconds in Arduino. Further, the duty cycle is calculated by selecting the peak point of that Fill Factor.

2.4. Motor Driver Circuit. The motor driver of this lawn mower is based on an H-bridge circuit designed by N-channel MOSFETs and BJTs, whereas in the simulation, just BJTs are used to make the circuit as simple as possible. In practical N-channel, MOSFETs are connected to the PNP sides of each H-bridge driver to quickly turn off PNP. There are four driving motors and one high-speed cutting motor. Each motor is controlled by a single H-bridge circuit. The speeds of driving motors can be controlled by changing the throttle level from the Android application, hence changing the duty cycle of the PWM signal generated by Raspberry Pi. The speed of the grass-cutting motor is fixed, and the user can only turn the on or off motor through the Android application. Figure 5 shows the simulation of the H-bridge motor driver circuit.

2.5. Circuit Diagram. Figure 6 shows the circuit diagram of the lawn mower, and the main circuit is divided into three portions. The level shifter circuit is based on NPN transistors. This circuit receives a low voltage (3.3 V) signal from the Raspberry Pi and generates 5 V signal output without changing its frequency or duty cycle. It is important to shift the voltage level because N-channel MOSFETs used in the H-bridge circuit only trigger at a 4.5 V gate signal. In this lawn mower, three different voltage levels are needed: 12 V for motors, 5 V for relays and Arduino, and 3.3 V for Raspberry Pi and camera. In the charge controller circuit, a multipurpose buck-boost converter is used which converts 12 V to these three levels. A multipurpose buck-boost converter also works for battery charging. It receives high voltages from the solar panel and modifies to optimize the power based on the PWM signal of Arduino. The charge controller circuit performs the following functions: control solar voltages, optimize power, and battery management, whereas Arduino continuously checks battery voltages and automatically cuts off load if battery voltages fall below the threshold point. A 16×4 LCD is also connected to display the charge controller parameters. The motor driver circuit contains two H-bridge drivers and one cutting motor switch circuit. The H-bridge circuit allows the user to change the direction and speeds of the driving motor. There is no need to change the direction or speed of the cutting motor; so, just a switching circuit is added to operate the cutting motor. Three NPN transistors are connected in parallel to increase the power rating of the circuit. An 8-megapixel camera is attached to the Raspberry Pi to transmit a video feed to the application.

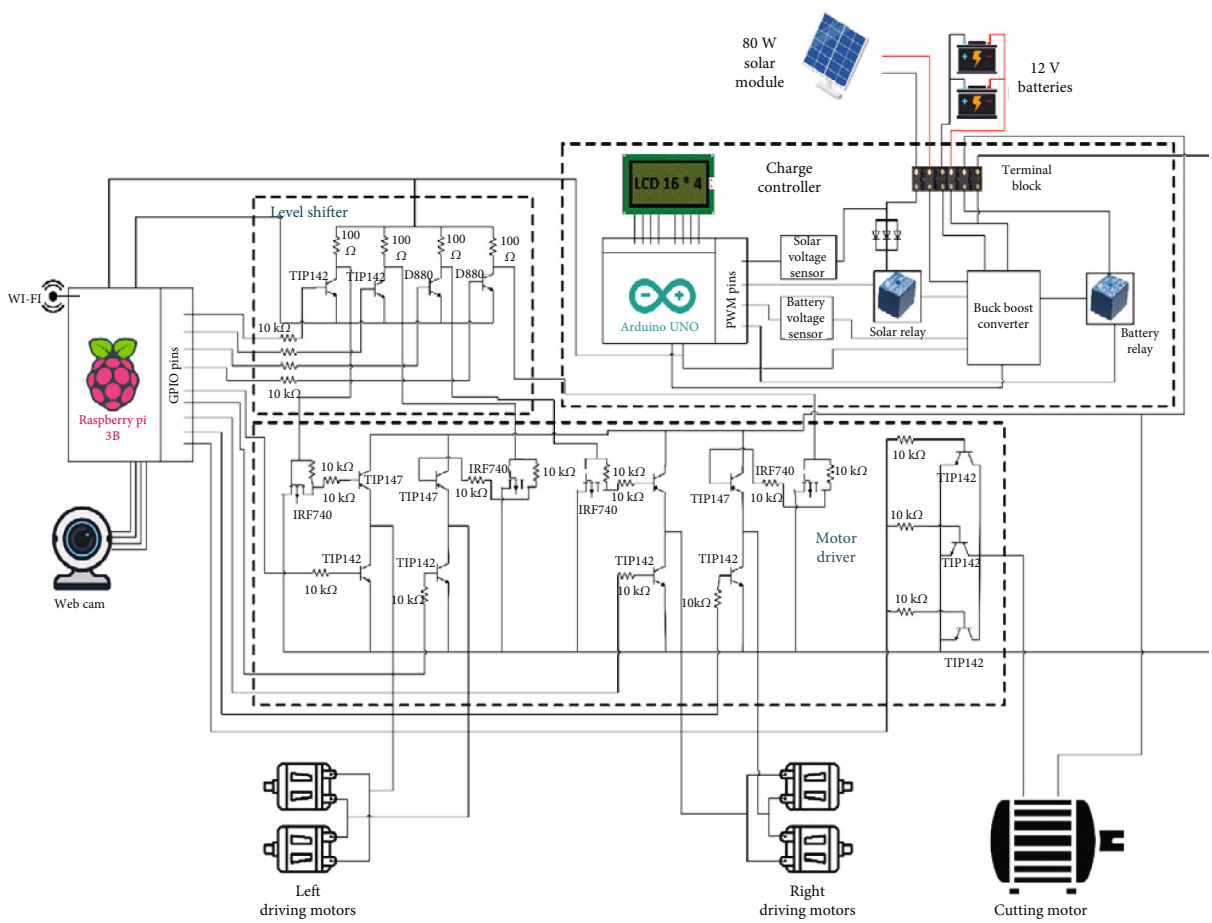


Figure 6: Complete circuit diagram of smart lawn mower.

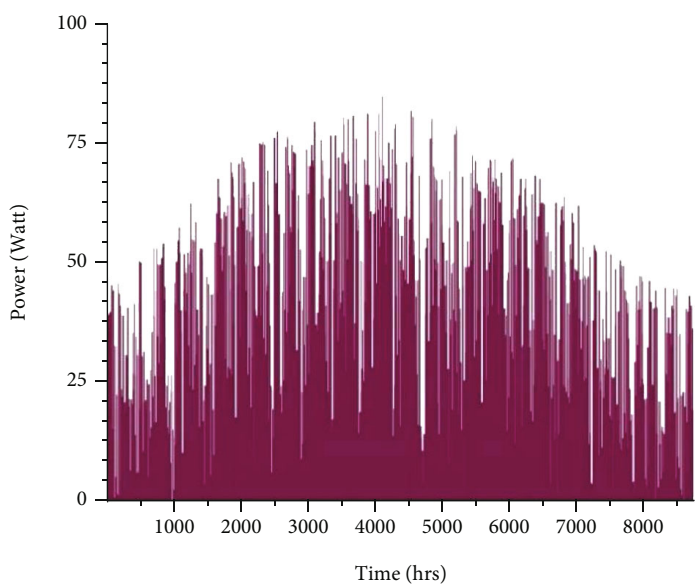


Figure 7: Power graph of the solar panel running for 1 year.

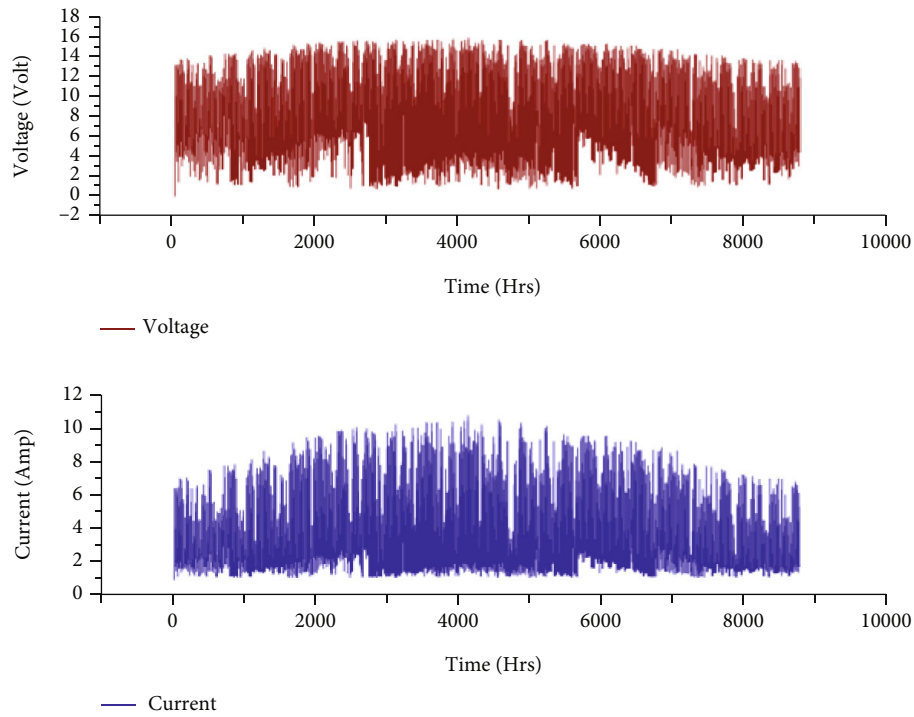


Figure 8: V-I graph of the solar panel running for 1 year.

Table 2: Values obtained during clear weather conditions.

Solar voltage (V)	Current (A)	Efficiency (%)
19	2.57	61.0
19	2.8	66.5
20	3.32	83.0
20	3.38	84.5
21	3.74	98.2

Table 3: Values obtained during clear weather conditions.

Solar voltage (V)	Current (A)	Efficiency (%)
16	2.54	50.8
17	2.61	55.5
17	2.73	58.0
18	3.02	68.0
18	3.11	70.0

The vertical tilt angle of the camera can be changed by the use of an Android application. A servo motor is attached with Raspberry Pi to control the tilt angle.

3. Results and Discussion

This research work comprises of an 80 W solar panel mounted on top of the mower to power motors and all electrical load. The overall load of the mower is calculated based on its maximum power demand to calculate the required size of the solar panel. The efficiency of a solar panel is maximum when mounted on fix surface at a specific angle, whereas it is not possible in this case. A 12 V, 7 Ah battery is added to compensate for the low efficiency and shadowing effects of solar panels. The charge controller charges the battery when generated power is more than the load demand and feeds power from the battery to the load when solar power is not enough to meet the load demand. The solar model has been simulated on TRNSYS-16 by entering the weather conditions of Lahore Pakistan. The load was entered according to the calculation, and the simulation runtime was set for 1 year (8760 hrs). Figure 7 shows the generated power graph of the solar panel running for 8760 hours. The center higher lines show the summer period while the left and right lines are for the winter season, whereas Figure 8 shows the V-I graph of the solar panel.

The solar charge controller is designed using the principle of MPPT where maximum power points are estimated using the FF curve through Arduino. The designed charge controller has the properties of a battery management system. Arduino continuously reads the voltages of the battery and solar panel, to protect the battery from overcharge and overdischarge. It automatically stops charging when battery voltages reach 14 V. Overdischarge protection gets into the process when battery voltages fall below 10 V. The efficiency of the charge controller is calculated by using equation (1).

$$\eta = \frac{\text{Solar voltage} \times \text{Current}}{\text{Maximum power of panel}} \times 100. \quad (1)$$

The solar charge controller is tested under different weather conditions, and the results obtained are shown in Tables 2 and 3.

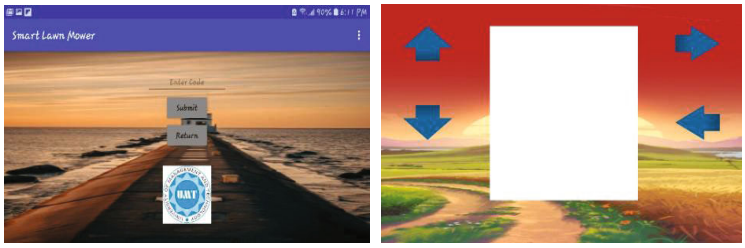


Figure 9: Android control application user interface.

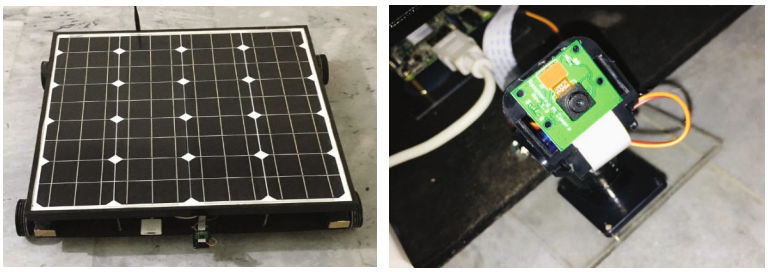


Figure 10: Hardware model of smart lawn mower.

The maximum efficiency of the charge controller during clear weather is about 98% at maximum solar voltages, while maximum efficiency during cloudy weather conditions is about 70% at a maximum 18 V.

Inside the smart lawn mower, there is a Raspberry Pi acting as the main controller of the lawn mower, which communicates with the Android application. Android application as shown in Figure 9 is built on the MIT app inventor and programmed to send an HTTP request to control.

Each button in the Android application is programmed to send a unique HTTP request to the server when such a button is pressed. Raspberry Pi receives this HTTP request and generates PWM signal to control the corresponding motor of the lawn mower. As a confirmation, Raspberry Pi sends an HTTP response message to the Android application. In case of no response, a message is received from the server, and the Android application will show an error notifier for 2 seconds on the screen. This application consists of two windows: one window is for passwords, and the other is for main control. The main window is only accessed when a user enters the correct password. When the user enters an incorrect password, a notification of an invalid password appears on the screen. Figure 10 shows the final model of the proposed IOT-based solar-powered smart lawn mower.

4. Conclusion

In this research work, a solar-powered smart lawn mower using Raspberry Pi and controlled by an Android application is developed. It has special features of solar charging and power management through an Arduino microcontroller. The designed mower can reduce human effort and power consumption using smart controllers and by deploying solar photovoltaic energy. A low voltage MPPT charge controller with a three-stage battery management system is designed for higher power efficiency. The mowing and video streaming is achieved by a local area network (LAN) connection. An Android application is used for remote monitoring and control while connecting to the same network connection. The proposed design of a smart lawn mower is successfully implemented and tested, and all the goals are also achieved successfully. The average electric power efficiency is 89.5% which is much better than conventional and electric-powered mowers. An optimal design is implemented using an IOT infrastructure, and an efficient solution is provided for grass cutting. Applications of the mower are large grounds used for different sports activities, lawns, and gardens. This concept can be extended to perform complex features like object detection through computer vision and machine learning for fully autonomous operation.

Abbreviations

ADC: Analog to digital converter
AI: Artificial intelligence
BLDC: Brushless direct current
CNN: Convolutional neural network
CV: Computer vision
DSP: Digital Signal Processing
HTTP: Hypertext Transfer Protocol
IOT: Internet of Things
IoRT: Internet of Robotic Things
IP: Internet Protocol
LAN: Local area network
MPPT: Maximum power point tracking
PIR: Passive infrared
PWM: Pulse width modulation
RF: Radio frequency
SSH: Secure Socket Shell
TCP: Transmission Control Protocol.

Data Availability

No data were used in this research.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Tayyab Tahir, Adnan Khalid, Jehangir Arshad, Aun Haider, Iftikhar Rasheed, Ateeq Ur Rehman, and Seada Hussen contributed to actualization, validation, methodology, formal analysis, investigation, software, and initial draft. All authors read and approved the final version. Tayyab Tahir and Jehangir Arshad are co-first authors.

References

[1] R. Rajesh and M. C. Mabel, "A comprehensive review of photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 231–248, 2015.

[2] 9 steps to winterize a lawn mower | lawn mower storage guide 2022, 2022, January 2022, https://www.popularmechanics.com/home/lawn-garden/how-to/a24355/how-to-winterize-your-lawn-mower/.

[3] P. P. Dutta, A. Baruah, A. Konwar, and V. Kumar, "A technical review of lawn mower technology," *ADBU Journal of Engineering Technology*, vol. 179, no. 1, pp. 179–182, 2016.

[4] P. S. M. Patil, B. Prajakta, K. Snehal, and P. Dhanashri, "Smart solar grass cutter with lawn coverage," *International research journal of engineering and technology*, vol. 5, no. 3, pp. 3476–3479, 2018.

[5] S. V. Tanaji, C. S. Chandrakant, P. S. Shashikant, and G. O. Raju, "Automated mower robo," *International Research Journal of Engineering and Technology*, vol. 5, no. 1, pp. 343–346, 2018.

[6] F. B. Ismail, N. F. O. Al-Muhsen, F. A. Fuzi, and A. Zukipli, "Design and development of smart solar grass cutter," *International Journal Of Engineering and Advanced Technology*, vol. 9, no. 2, pp. 4137–4141, 2019.

[7] A. Balaji, E. R. Shanthini, M. J. Rooshan, and V. Sivaram, "Robust Design of Smart Lawn Cutter for gardening companion," *IOP Conference Series: Materials Science and Engineering*, vol. 1012, no. 1, article 012005, 2021.

[8] B. Ibrahim, V. S. Brahmaiah, and P. Sharma, "Design of smart autonomous remote monitored solar powered lawnmower robot," *Materials Today: Proceedings*, vol. 28, pp. 2338–2344, 2020.

[9] M. A. Shapamohan, K. V. Tale, K. A. Shende, and U. S. Raut, "IRJET-remotely control lawn mower using solar system," *International Research Journal of Engineering and Technology*, vol. 6, no. 1, p. 2395, 2019.

[10] K. N. Baluprithviraj, R. Harini, M. M. Janarthanan, and C. Jasodhasree, "Design and development of smart lawn mower," in *2nd International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1215–1219, Trichy, India, 2021.

[11] S. Jain, A. Khalore, and S. Patil, "Self-efficient and sustainable solar powered robotic lawn mower," *International Journal of Trend in Research and Development*, vol. 2, no. 6, pp. 2394–9333, 2015.

[12] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq, and D. Hussain, "A trustworthy SIoT aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, p. 918, 2020.

[13] J. Arshad, M. Aziz, A. A. al-Huqail et al., "Implementation of a LoRaWAN based smart agriculture decision support system for optimum crop yield," *Sustainability*, vol. 14, no. 2, p. 827, 2022.

[14] A. Alkhater, A. Almelehi, Y. Alqubaisi, and N. Gupta, *Solar Powered Lawn Mower*, University of Bridgeport, Bridgeport, 2019.

[15] B. Soyoye, "Development and performance evaluation of a solar powered lawn mower," *Turkish Journal of Agricultural Engineering Research*, vol. 2, no. 2, pp. 348–362, 2021.

[16] M. Jabbar, M. Waqas, A. Noman, A. A. Muneer, and A. Mazhar, "Solar powered grass cutter for domestic utilization," *Engineering Proceedings*, vol. 12, no. 1, p. 105, 2022.

[17] J. Arshad, M. A. B. Siddique, Z. Zulfiqar et al., "A novel remote user authentication scheme by using private blockchain-based secure access control for agriculture monitoring," in *2020 International Conference on Engineering & Emerging Technologies (ICEET)*, Lahore, Pakistan, 2020.

[18] J. Liao, S. H. Ching, Z. Y. Chen, B. W. W. Zhuang, and Y. J. Chen, "Designing and manufacturing of automatic robotic lawn mower," *Processes*, vol. 9, no. 2, p. 358, 2021.

[19] A. Roa, A. Diego, X. Collazo, M. Goenaga, A. A. Espinoza, and K. Vazquez, "Development and evaluation of a remote controlled electric lawn mower," in *IEEE Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 90–94, Las Vegas, NV, USA, 2019.

[20] P. Vorel and J. Martiš, "Battery powered lawn mower," *ECS Transactions*, vol. 105, no. 1, pp. 567–574, 2021.

[21] R. M. Asif, J. Arshad, M. Shakir, S. M. Noman, and A. U. Rehman, "Energy efficiency augmentation in massive MIMO systems through linear precoding schemes and power consumption modeling," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8839088, 13 pages, 2020.

[22] M. S. Ghorpade, M. P. Jadhav, M. F. Mulla, and M. P. Patil, "Intelligent solar powered grass cutting robot with obstacle avoidance," *International Journal of Modern Electronics and Communication Engineering (IJMECE)*, vol. 6, no. 3, pp. 1–4, 2018.

[23] A. U. Rehman, R. M. Asif, R. Tariq, and A. Javed, "GSM based solar automatic irrigation system using moisture, temperature and humidity sensors," in *5th International Conference on Engineering Technology & Technopreneurship (ICE2T)*, Kuala Lumpur, Malaysia, 2017.

[24] J. Li, H. Liu, Z. Sun, and R. Huang, "Real time obstacle estimation based on dense stereo vision for robotic lawn mowers," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1186–1191, Dali, China, 2019.

[25] T. Rehman and I. Ullah, "Notice of violation of IEEE publication principles: clustering schemes in MANETs: performance evaluation, open challenges, and proposed solutions," *IEEE Access*, vol. 8, no. 1, pp. 25135–25158, 2020.

[26] B. Siregar, B. M. Hutagaol, and O. S. Sitompul, "Smartphone-controllable lawn mower robot," in *IEEE International*

Conference on ICT for Smart Society (ICISS), pp. 1–5, Bandung, Indonesia, 2020.

[27] M. Ryalat, M. Alsherqatli, and H. Elmoaqet, "IoT-aided smart lawnmower," *International Symposium on Computer Science and Intelligent Control*, pp. 1–8, 2019.

[28] H. Arora, J. A. Sagor, V. Panwar et al., "Design and fabrication of autonomous lawn mower with water sprinkler," *Think India Journal*, vol. 22, no. 17, p. 2101, 2019.

[29] A. Hassan, H. M. Abdullah, U. Farooq et al., "A Wirelessly Controlled Robot-based Smart Irrigation System by Exploiting Arduino," *Journal of Robotics and Control (JRC)*, vol. 2, pp. 29–34, 2021.

[30] T. Munasinghe, E. W. Patton, and O. Seneviratne, "IoT application development using MIT App inventor to collect and analyze sensor data," in *IEEE International Conference on Big Data*, pp. 6157–6159, Los Angeles, CA, USA, 2019.

[31] J. Arshad, R. Tariq, S. Saleem et al., "Intelligent greenhouse monitoring and control scheme: an arrangement of Sensors, Raspberry Pi based Embedded System and IoT platform," *Indian Journal of Science and Technology*, vol. 13, pp. 2811–2822, 2020.

[32] J. Katona and A. Kovari, "Cost-effective wi-fi controlled mobile robot," in *11th International Symposium on Applied Informatics and Related Areas (AIS 2016)*, pp. 28–31, Hungary, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8272116, 12 pages
https://doi.org/10.1155/2022/8272116

Research Article

Multi-UAV Factor Graph Colocation Algorithm

YueXia Zhang[1,2,3] **and Pengfei Jia**[1,2,3]

[1]*Key Laboratory of Information and Communication Systems, Ministry of Information Industry, Beijing Information Science and Technology University, China*
[2]*Key Laboratory of Modern Measurement & Control Technology, Ministry of Education, Beijing Information Science & Technology University, China*
[3]*Beijing Key Laboratory of High Dynamic Navigation Technology, University of Beijing Information Science & Technology, China*

Correspondence should be addressed to YueXia Zhang; zhangyuexia@bistu.edu.cn

Received 26 May 2022; Revised 22 June 2022; Accepted 7 July 2022; Published 21 August 2022

Academic Editor: Xingwang Li

To tackle the difficulty and high cost of UAV positioning in a 5G disaster environment, this study proposes a multi-UAV coordinated positioning algorithm based on a factor graph (MUAV-FGC). A factor graph model is designed for coordinated positioning of the master-slave UAVs. In the positioning process, the influence of the speed and heading errors of the UAV on the positioning accuracy was analyzed, and the expected and variance values of the variables were used as the information transmitted between the factors. The error of the cooperative positioning algorithm was derived based on the factor graph. The coordinated positioning of multiple drones and the positioning accuracy improved. The simulation results show that, compared with the traditional extended Kalman filter (EKF) algorithm, the proposed MUAV-FGC positioning algorithm reduces the mean positioning and root mean square error by approximately 18.86% and 17.54%, respectively, thus proving its effectiveness.

1. Introduction

In recent years, various types of natural disasters affecting a wide area have occurred frequently, causing huge losses to life and property. With the continuous maturity of 5G technology and the research and development of 6G technology in the future [1], the advanced technology represented by UAV is more and more applied to the field of disaster prevention, reduction, and disaster relief and emergency management [2–4], which opens the "eye of the air" for disaster prevention, reduction, and disaster relief. In the event of a severe natural disaster, the drone can quickly cross the mountain and river barriers and get to the disaster site unaffected by the information interruptions and traffic obstructions; obtain the disaster information through photography, video, or other remote sensing methods; and provide accurate information for emergency rescue and relief in a timely manner [5, 6]. Thus, drone technologies present substantial technical advantages. When performing missions, drones need to be able to accurately determine their own location to successfully perform tasks such as disaster situation detection, search and rescue, and delivery of rescue supplies [7, 8] at the correct location in scenarios including nuclear disasters, where investigation and rescue by humans is not possible. Therefore, at present, obtaining information on the precise location of the UAV has become one of the most significant core technologies to be developed and improved [9, 10].

The traditional methods of obtaining the position of the UAVs primarily include the global positioning system (GPS) [11–13] and inertial navigation and positioning system (INS) [14–16]. Although the GPS positioning system has the advantages of globalization and real time, the signal could be easily blocked by obstructions (such as high mountains and valleys and remote areas). Moreover, electromagnetic interference, which affects its position accuracy, can easily occur. If the drone cannot receive the GPS signals or if the frequency update of the GPS signal could not meet the requirements of the aircraft to update the information, the drone will lose its

position. Conversely, INS has the advantages of small size, low cost, and independent positioning without relying on external information [17]. However, when an INS moves in a large area for a long time, the error of pose estimation will accumulate over time, leading to positioning, and the error increases sharply and deviates significantly from the actual position. When drones are used for disaster relief, they are often required to work for long hours in harsh environments, such as remote mountain areas. Thus, the GPS and INS systems cannot meet the requirements of drone positioning.

At present, scholars at home and abroad have conducted a significant amount of research on the positioning of UAVs [18–20]. Padhy et al. [21] proposed a model for autonomous navigation and collision avoidance of drones in corridor environments where GPS cannot be located; detailed experiments were conducted in different corridors to prove the effectiveness of the experimental plan. Zhang et al. proposed an autonomous navigation and positioning method for UAVs based on electric field array detection [22]. Their experiments showed the positioning effect of the positioning algorithm, and the positioning error was within the range of the user error; therefore, this positioning algorithm had good usability and application value. Liu et al. [23] proposed an interactive multimodel positioning algorithm based on a robust Cuban Kalman filter and analyzed the performance of the UAV positioning and navigation algorithm. These positioning methods are only for single UAV positioning research, and there is no research on multi-UAV coordinated positioning. However, in disaster environments, multiple UAVs often work collaboratively; thus, coordinated positioning is required to improve the positioning accuracy. Therefore, the coordinated positioning of multiple drones has become a crucial issue that must be addressed.

Regarding the research on multi-UAV colocation [24–29], Sivaneri and Gross [30] proposed a UAV colocation algorithm in a global navigation satellite system (GNSS) environment, which uses a platform where the satellite signals are interfered, except for on the ground level, for navigation. However, the limitations of the ground control station technology posed a problem. Meyer et al. [31] proposed a sigma point belief propagation positioning algorithm, which uses sampling points to characterize the state and can quickly perform collaborative positioning; this method is more suitable for indoor positioning and navigation and does not consider the use of measurement information devices such as satellites. However, a large number of coordinated drones lead to a large computational load, which remains an issue yet to be solved. Vetrella et al. [32] proposed a collaborative positioning method that integrates inertial, magnetometer, available satellite pseudo-range, cooperative drone position, and monocular camera information, which effectively improves the positioning performance of the drone swarm under the GPS-constrained conditions. However, a large number of sensors were used. Wan et al. [33] designed a dynamic nonparametric belief propagation algorithm to obtain the UAV location information, but the amount of calculation was large, and only a two-dimensional space was simulated, which is not suitable for large-scale cluster UAV use. Fan et al. [34] proposed a multi-AUV colocation method based on factor graphs and product algorithms.

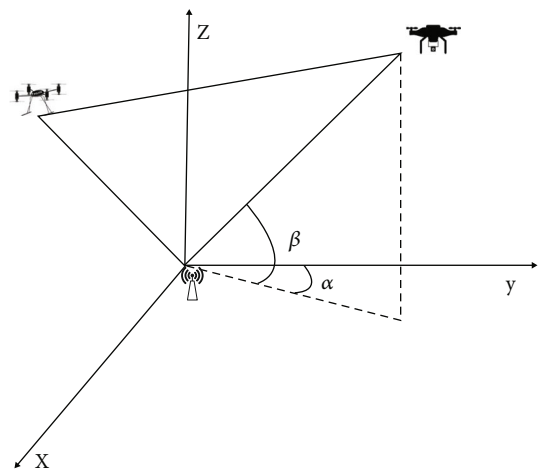


Figure 1: Multi-UAV coordinated positioning model.

Experimental results show that, its positioning accuracy is better than those of the extended Kalman filter (EKF) and unscented Kalman filter (UKF) algorithms; however, it is a centralized processing method. The aforementioned algorithms consider the difficulty of UAV positioning under conditions such as limited GNSS signals. However, problems such as the inability to eliminate the constraints of the ground control station and the requirement of a large number of sensors remain, making such methods unsuitable for use in disaster environments. The human-machine coordinated positioning does not consider the influence of the speed and heading errors of the drone on the positioning accuracy during flight. In response to the above problems, this paper proposes a multi-UAV colocation algorithm based on a factor graph (MUAV-FGC). The main contributions of the algorithm are as follows:

(1) Analyzed the problem of multi-UAV colocation error in a disaster environment, proposed a multi-UAV factor graph colocation algorithm, and designed a factor graph model of the master-slave UAV coordinated location

(2) Analyzed the influence of the UAV's speed and heading error on the positioning accuracy during the positioning process. In the factor graph model, the variable expectation and variance are used as the information transmitted between the factors and deduced the factor graph-based cooperative positioning algorithm error

(3) Finally, a simulation experiment was established and compared with the traditional UAV EKF positioning algorithm. The results show that the MUAV-FGC algorithm proposed in this study significantly improves the positioning accuracy [35]

2. Analysis of Multi-UAV Coordinated Positioning

2.1. Multi-UAV Colocation Model. The cooperative positioning model of multiple UAVs is shown in Figure 1. The

Cartesian coordinate system is established with the UAV transmitting base station B as the origin. In the figure, MUAV is the main UAV, and UAV is the UAV to be positioned. In the positioning process, it is assumed that the position of the main UAV is accurate, that is, the main UAV is an airborne 5G base station, and the position information of the main UAV is obtained by 5G data communication between the UAV and the main UAV.

The three-dimensional motion model of UAV can be expressed as

$$\begin{aligned} x &= v \cos \beta \cos \alpha, \\ y &= v \cos \beta \sin \alpha, \\ z &= v \sin \beta, \end{aligned} \tag{1}$$

where (x, y, z), v, and (α, β) represent the position coordinates of the slave drone, speed of the slave drone, and heading angles, respectively.

If the speed of the drone at time k, the measured horizontal heading angle, and the vertical heading angle are v_k, α, and β, respectively, then the position of the drone at time k is (x_k, y_k, z_k), and the position of the drone at time $k-1$ is $(x_{k-1}, y_{k-1}, z_{k-1})$, and Δt is sampling periods; thus, the following three-dimensional discrete motion model for the UAV can be obtained:

$$\begin{aligned} x_k &= x_{k-1} + v\Delta t \cos \beta \cos \alpha, \\ y_k &= y_{k-1} + v\Delta t \cos \beta \sin \alpha, \\ z_k &= z_{k-1} + v\Delta t \sin \beta. \end{aligned} \tag{2}$$

The slave and master drones communicate with each other to obtain the relative position information. The calculation equation for the relative distance information is

$$d = \sqrt{(x_k - x_k^m)^2 + (y_k - y_k^m)^2 + (z_k - z_k^m)^2}, \tag{3}$$

where (x_k^m, y_k^m, z_k^m) represents the position of the main drone at time k.

2.2. Positioning Error Analysis. According to the three-dimensional motion model of the slave drone, the speed and heading position of the slave drone measured at time a are as follows:

$$\begin{aligned} x_a^- &= x_{a-1} + v_a^- \Delta t \cos \alpha_a^- \cos \beta_a^-, \\ y_a^- &= y_{a-1} + v_a^- \Delta t \cos \alpha_a^- \sin \beta_a^-, \\ z_a^- &= z_{a-1} + v_a^- \Delta t \sin \alpha_a^-. \end{aligned} \tag{4}$$

v_a^- and α_a^-, β_a^- represent the measured value of the drone's speed and heading angles at time a; x_a^-, y_a^-, and z_a^- represent the a priori estimate of the position in the x, y, and z directions at time a, respectively; $x_{a-1}\, y_{a-1}$ and z_{a-1} represent the true value of the position in the x, y, and z directions at time $a-1$, respectively. The speed and heading angle can be further expressed as

$$\begin{aligned} v_a^- &= v_a + \tilde{v}_a, \\ \alpha_a^- &= \alpha_a + \tilde{\alpha}_a, \\ \beta_a^- &= \beta_a + \tilde{\beta}_a, \end{aligned} \tag{5}$$

where v_a represents the true speed of the slave drone at time a and $\tilde{v}_a$ represents the speed error of the slave drone at time a. α_a, β_a represents the true heading angles from the drone at time a, and$\tilde{\alpha}_a, \tilde{\beta}_a$ represents the heading errors from the drone at time a. Substituting equation (5) into equation (4), we can obtain

$$\begin{aligned} x_a^- &= x_{a-1} + \Delta t(v_a + \tilde{v}_a) \cos \left(\beta_a + \tilde{\beta}_a\right) \cos (\alpha_a + \tilde{\alpha}_a), \\ y_a^- &= y_{a-1} + \Delta t(v_a + \tilde{v}_a) \cos \left(\beta_a + \tilde{\beta}_a\right) \sin (\alpha_a + \tilde{\alpha}_a), \\ z_a^- &= z_{a-1} + \Delta t(v_a + \tilde{v}_a) \sin \left(\beta_a + \tilde{\beta}_a\right). \end{aligned} \tag{6}$$

On expanding equation (6), we obtain

$$\begin{aligned} x_a^- = x_{a-1} + \Delta t(v_a + \tilde{v}_a)\Big(&\cos \beta_a \cos \tilde{\beta}_a \cos \alpha_a \cos \tilde{\alpha}_a \\ &- \cos \beta_a \cos \tilde{\beta}_a \sin \alpha_a \sin \tilde{\alpha}_a \\ &- \sin \beta_a \sin \tilde{\beta}_a \cos \alpha_a \cos \tilde{\alpha}_a \\ &+ \sin \beta_a \sin \tilde{\beta}_a \sin \alpha_a \sin \tilde{\alpha}_a\Big), \\ y_a^- = y_{a-1} + \Delta t(v_a + \tilde{v}_a)\Big(&\cos \beta_a \cos \tilde{\beta}_a \sin \alpha_a \cos \tilde{\alpha}_a \\ &+ \cos \beta_a \cos \tilde{\beta}_a \cos \alpha_a \sin \tilde{\alpha}_a \\ &- \sin \beta_a \sin \tilde{\beta}_a \sin \alpha_a \cos \tilde{\alpha}_a \\ &- \sin \beta_a \sin \tilde{\beta}_a \cos \alpha_a \sin \tilde{\alpha}_a\Big), \\ z_a^- = z_{a-1} + \Delta t(v_a + \tilde{v}_a)\Big(&\sin \beta_a \cos \tilde{\beta}_a + \cos \beta_a \sin \tilde{\beta}_a\Big). \end{aligned} \tag{7}$$

Assuming that the speed and heading errors are infinitesimal, then

$$\begin{aligned} \sin \tilde{\alpha}_a &= \tilde{\alpha}_a, \\ \cos \tilde{\alpha}_a &= 1, \\ \tilde{v}_a \tilde{\alpha}_a &= 0, \\ \sin \tilde{\beta}_a &= \tilde{\beta}, \\ \cos \tilde{\beta}_a &= 1, \\ \tilde{v}_a \tilde{\beta}_a &= 0. \end{aligned} \tag{8}$$

Substituting equation (8) into equation (7), we can obtain

$$\begin{aligned} x_a^- &= x_{a-1} + \Delta t\Big[v_a\Big(\cos\beta_a\cos\alpha_a - \tilde{\alpha}_a\cos\beta_a\sin\alpha_a \\ &\quad - \tilde{\beta}_a\sin\beta_a\cos\alpha_a + \tilde{\alpha}_a\tilde{\beta}_a\sin\beta_a\sin\alpha_a\Big) \\ &\quad + \tilde{v}_a\cos\beta_a\cos\alpha_a\Big], \\ y_a^- &= y_{a-1} + \Delta t\Big[v_a\Big(\cos\beta_a\sin\alpha_a + \tilde{\alpha}_a\cos\beta_a\cos\alpha_a \\ &\quad - \tilde{\beta}_a\sin\beta_a\sin\alpha_a - \tilde{\alpha}_a\tilde{\beta}_a\sin\beta_a\cos\alpha_a\Big) \\ &\quad + \tilde{v}_a\cos\beta_a\sin\alpha_a\Big], \\ z_a^- &= z_{a-1} + \Delta t\Big[v_a\Big(\sin\beta_a + \tilde{\beta}\cos\beta_a\Big) + \tilde{v}_a\sin\beta_a\Big]. \end{aligned} \tag{9}$$

From equation (9), it can be concluded that the coordinate position error is directly proportional to the speed error; it is also positively correlated with heading angle error. It can be seen that both the speed and heading error will have an impact on the positioning, resulting in unmanned machine positioning error increases. Therefore, based on the above analysis, this article introduces an analysis of the speed and heading angle errors in the positioning process.

3. Multi-UAV Factor Graph Colocation Algorithm

3.1. Colocation Analysis of Factor Graph. Assuming that the UAV cannot receive the location information of the main UAV at time a, after a period of time, the location information of the main UAV is received at time k. At this time, the posterior position estimation from time a to time k is difficult to calculate, which easily leads to complicated speed and heading errors. Therefore, assuming that the amount of time from time a to time k is short, then

$$\begin{aligned} \tilde{v} &= \tilde{v}_a = \cdots = \tilde{v}_k, \\ \tilde{\alpha} &= \tilde{\alpha}_a = \cdots = \tilde{\alpha}_k, \\ \tilde{\beta} &= \tilde{\beta}_a = \cdots = \tilde{\beta}_k. \end{aligned} \tag{10}$$

Accordingly, the error $\delta x_a, \delta y_a, \delta z_a$ of the coordinate position at time a can be written as

$$\begin{aligned} \delta x_a &= x_a^- - x_a = \Delta t\Big[v_a\Big(-\tilde{\alpha}_a\cos\beta_a\sin\alpha_a \\ &\quad - \tilde{\beta}_a\sin\beta_a\cos\alpha_a + \tilde{\alpha}_a\tilde{\beta}_a\sin\beta_a\sin\alpha_a\Big) \\ &\quad + \tilde{v}_a\cos\beta_a\cos\alpha_a\Big], \\ \delta y_a &= y_a^- - y_a = \Delta t\Big[v_a\Big(\tilde{\alpha}_a\cos\beta_a\cos\alpha_a \\ &\quad - \tilde{\beta}_a\sin\beta_a\sin\alpha_a - \tilde{\alpha}_a\tilde{\beta}_a\sin\beta_a\cos\alpha_a\Big) \\ &\quad + \tilde{v}_a\cos\beta_a\sin\alpha_a\Big], \\ \delta z_a &= z_a^- - z_a = \Delta t v_a\tilde{\beta}_a\cos\beta_a + \Delta t\tilde{v}_a\sin\beta_a, \end{aligned} \tag{11}$$

where x_a^-, y_a^-, z_a^- represent a priori estimates of x_a, y_a, z_a, respectively. Because no position information is received from time a to time k, an a priori estimation is used to express the position as follows:

$$\begin{aligned} x_k^- &= x_k + \delta x_{a+1} + \cdots + \delta x_k, \\ y_k^- &= y_k + \delta y_{a+1} + \cdots + \delta y_k, \\ z_k^- &= z_k + \delta z_{a+1} + \cdots + \delta z_k. \end{aligned} \tag{12}$$

Substituting equation (11) into equation (12), we obtain equation (13) as

$$\begin{aligned} x_k^- &= x_k + \tilde{v}\Big[\Delta t(\cos\beta_{a+1}\cos\alpha_{a+1}\cdots + \cos\beta_k\cos\alpha_k) \\ &\quad - \tilde{\alpha}(\Delta t v_{a+1}\cos\beta_{a+1}\sin\alpha_{a+1} + \cdots + \Delta t\cos\beta_k\sin\alpha_k) \\ &\quad - \tilde{\beta}(\Delta t\sin\beta_{a+1}\cos\alpha_{a+1} + \cdots \Delta t\sin\beta_k\cos\alpha_k) \\ &\quad + \tilde{\alpha}\tilde{\beta}(\Delta t\sin\beta_{a+1}\sin\alpha_{a+1} + \cdots \Delta t\sin\beta_k\sin\alpha_k), \\ y_k^- &= y_k + \tilde{v}\Big[\Delta t(\cos\beta_{a+1}\sin\alpha_{a+1}\cdots + \cos\beta_k\sin\alpha_k) \\ &\quad + \tilde{\alpha}(\Delta t v_{a+1}\cos\beta_{a+1}\cos\alpha_{a+1} + \cdots + \Delta t\cos\beta_k\cos\alpha_k) \\ &\quad - \tilde{\beta}(\Delta t\sin\beta_{a+1}\sin\alpha_{a+1} + \cdots \Delta t\sin\beta_k\sin\alpha_k) \\ &\quad - \tilde{\alpha}\tilde{\beta}(\Delta t\sin\beta_{a+1}\cos\alpha_{a+1} + \cdots \Delta t\sin\beta_k\cos\alpha_k), \\ z_k^- &= z_k + \tilde{v}(\Delta t\sin\beta_{a+1} + \cdots + \Delta t\sin\beta_k) \\ &\quad + \tilde{\beta}(v_{a+1}\Delta t\cos\beta_{a+1} + \cdots + v_k\Delta t\cos\beta_k). \end{aligned} \tag{13}$$

For representation, the coordinate position error $\delta x_k, \delta y_k, \delta z_k$ at time k can be written as

$$\begin{aligned} \delta x_k &= \tilde{v}\delta x_k^{\tilde{v}} + \tilde{\alpha}\delta x_k^{\tilde{\alpha}} + \tilde{\beta}\delta x_k^{\tilde{\beta}}, \\ \delta y_k &= \tilde{v}\delta y_k^{\tilde{v}} + \tilde{\alpha}\delta y_k^{\tilde{\alpha}} + \tilde{\beta}\delta y_k^{\tilde{\beta}}, \\ \delta z_k &= \tilde{v}\delta z_k^{\tilde{v}} + \tilde{\beta}\delta z_k^{\tilde{\beta}}. \end{aligned} \tag{14}$$

It can also be rewritten as

$$\begin{aligned} \begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix} &= \begin{bmatrix} \delta x_k^{\tilde{v}} & \delta x_k^{\tilde{\alpha}} & \delta x_k^{\tilde{\beta}} \\ \delta y_k^{\tilde{v}} & \delta y_k^{\tilde{\alpha}} & \delta y_k^{\tilde{\beta}} \\ \delta z_k^{\tilde{v}} & 0 & \delta z_k^{\tilde{\beta}} \end{bmatrix} \begin{bmatrix} \tilde{v} \\ \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix}, \\ \begin{bmatrix} \tilde{v} \\ \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix} &= \begin{bmatrix} \delta x_k^{\tilde{v}} & \delta x_k^{\tilde{\alpha}} & \delta x_k^{\tilde{\beta}} \\ \delta y_k^{\tilde{v}} & \delta y_k^{\tilde{\alpha}} & \delta y_k^{\tilde{\beta}} \\ \delta z_k^{\tilde{v}} & 0 & \delta z_k^{\tilde{\beta}} \end{bmatrix}^{-1} \begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix}. \end{aligned} \tag{15}$$

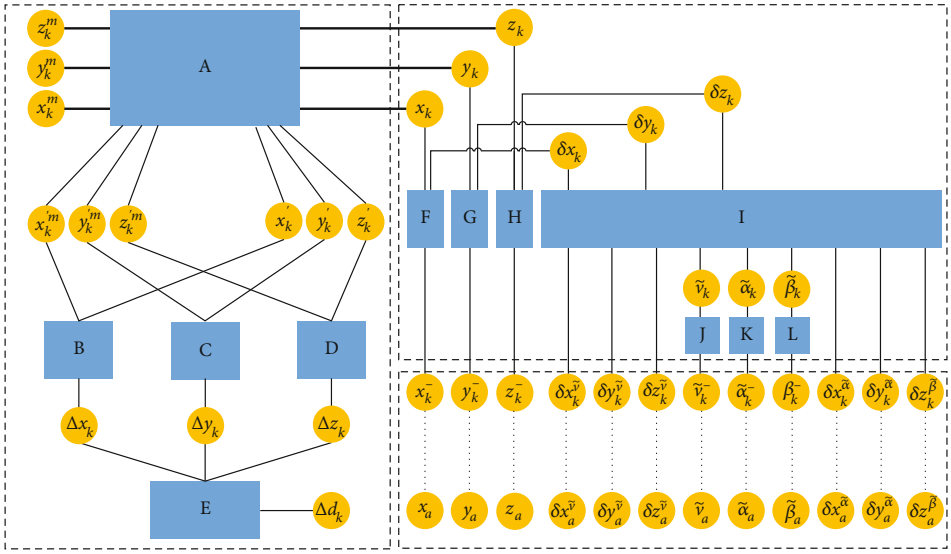


Figure 2: Multi-UAV factor graph colocation algorithm.

When the master-slave drone is colocated, the position update needs to be calculated after subtracting the corresponding estimation error from the measured value of the speed and heading:

$$\begin{aligned} x_k^- &= x_{k-1} + \left(v_k^- - \tilde{v}_k'\right)\Delta t \cos\left(\beta_k^- - \tilde{\beta}_k^-\right), \\ &\quad \cos\left(\alpha_k^- - \tilde{\alpha}_k^-\right), \\ y_k^- &= y_{k-1} + \left(v_k^- - \tilde{v}_k'\right)\Delta t \cos\left(\beta_k^- - \tilde{\beta}_k^-\right), \\ &\quad \sin\left(\alpha_k^- - \tilde{\alpha}_k^-\right), \\ z_k^- &= z_{k-1} + \left(v_k^- - \tilde{v}_k'\right)\Delta t \sin\left(\beta_k^- - \tilde{\beta}_k^-\right). \end{aligned} \tag{16}$$

According to equation (14) and the UAV's three-dimensional discrete motion and relative distance model, make the information transfer between its variables can be obtained, as demonstrated in Figure 2.The rectangular and elliptical boxes in the figure represent function and factor nodes, respectively.

3.2. Procedure of the Factor Graph Colocation Algorithm. As shown in Figure 2, the multi-UAV factor graph colocation algorithm is used to calculate the mean and variance of the transfer between the variable and function node. The specific steps are as follows:

3.2.1. Initialization. Initial state before colocation:

$$\begin{aligned} \widehat{x}_k^- &= x_0, \\ \sigma_{\widehat{x}_k^-}^2 &= \sigma_{x_0}^2, \\ \widehat{y}_k^- &= y_0, \\ \sigma_{\widehat{y}_k^-}^2 &= \sigma_{y_0}^2, \\ \widehat{z}_k^- &= z_0, \\ \sigma_{\widehat{z}_k^-}^2 &= \sigma_{z_0}^2. \end{aligned} \tag{17}$$

3.2.2. Update of A Priori Estimate $\widehat{x}_k^-, \widehat{y}_k^-, \widehat{z}_k^-$. At noninitial time, dead reckoning can be performed according to equation (2). μ_v and σ_v^2 are the expected speed and variance, and $\mu_\alpha, \sigma_\alpha^2$ and $\mu_\beta, \sigma_\beta^2$ are the expected heading and variance, respectively. Because the speed is not related to the heading, $\begin{cases} \operatorname{cov}(v, \alpha) = 0 \\ \operatorname{cov}(v, \beta) = 0 \end{cases}$, we can obtain the following formula:

$$\mu_{\widehat{x}_k^-} = \mu_{\widehat{x}_{k-1}^-} + \Delta t \mu_v \left(\frac{\cos\mu_\beta - \sigma_\beta^2 \cos\beta}{2}\right)\left(\frac{\cos\mu_\alpha - \sigma_\alpha^2 \cos\alpha}{2}\right),$$

$$\mu_{\widehat{y}_k^-} = \mu_{\widehat{y}_{k-1}^-} + \Delta t \mu_v \left(\frac{\cos\mu_\beta - \sigma_\beta^2 \cos\beta}{2}\right)\left(\frac{\sin\mu_\alpha - \sigma_\alpha^2 \sin\alpha}{2}\right),$$

$$\mu_{\widehat{z}_k^-} = \mu_{\widehat{z}_{k-1}^-} + \Delta t \mu_v \left(\frac{\sin\mu_\beta - \sigma_\beta^2 \sin\beta}{2}\right),$$

$$\begin{aligned} \sigma_{\widehat{x}_k^-}^2 &= \sigma_{\widehat{x}_{k-1}^-}^2 + \left[\left(\cos\mu_\beta \Delta t \sigma_v\right)^2 - \left(\mu_v \sin\mu_\beta \Delta t \sigma_\beta^2\right)^2\right] \\ &\quad \cdot \sin(\mu_\alpha)\sigma_\alpha^2 + \left[\left(\cos\mu_\beta \Delta t \sigma_v\right)^2 - \left(\mu_v \sin\mu_\beta \Delta t \sigma_\beta^2\right)^2\right] \\ &\quad \cdot \left[\frac{\cos(\mu_\alpha) - \sigma_\alpha^2 \cos(\mu_\alpha)}{2}\right]^2 \\ &\quad + \sin(\mu_\alpha)\sigma_\alpha^2 \mu_v^2 \left[\frac{\cos(\mu_\beta) - \sigma_\beta^2 \cos(\mu_\beta)}{2}\right]^2, \end{aligned}$$

$$\sigma^2_{\widehat{y}_k^-} = \sigma^2_{\widehat{y}_{k-1}} + \left[\left(\cos\mu_\beta \Delta t \sigma_v\right)^2 - \left(\mu_v \sin\mu_\beta \Delta t \sigma^2_\beta\right)^2\right] \cdot \cos(\mu_\alpha)\sigma^2_\alpha + \left[\left(\cos\mu_\beta \Delta t \sigma_v\right)^2 - \left(\mu_v \sin\mu_\beta \Delta t \sigma^2_\beta\right)^2\right] \cdot \left[\frac{\sin(\mu_\alpha) - \sigma^2_\alpha \sin(\mu_\alpha)}{2}\right]^2 + \cos(\mu_\alpha)\sigma^2_\alpha \mu^2_v \left[\frac{\cos\left(\mu_\beta\right) - \sigma^2_\beta \cos\left(\mu_\beta\right)}{2}\right]^2,$$

$$\sigma^2_{\widehat{z}_k^-} = \sigma^2_{\widehat{z}_{k-1}} + \left(\sin\mu_\beta \Delta t \sigma_v\right)^2 - \left(\mu_v \cos\mu_\beta \Delta t \sigma^2_\beta\right)^2. \tag{18}$$

3.2.3. First Update of $\widehat{x}_k^-, \widehat{y}_k^-, \widehat{z}_k^-$. At this time, x', z', y' after the coordinate transformation (Figure 2) have not been calculated, and the variable nodes x_k, y_k, z_k and $\widehat{x}_k^-, \widehat{y}_k^-, \widehat{z}_k^-$ are equal:

$$\begin{aligned} x_k &= \widehat{x}_k^-, \\ y_k &= \widehat{y}_k^-, \\ z_k &= \widehat{z}_k^-. \end{aligned} \tag{19}$$

3.2.4. First Update of $\widehat{x}_k^-, \widehat{y}_k^-, \widehat{z}_k^-$. To avoid the collision of the coordinate axes of the master and slave drones and to prevent the resultant large calculation errors, the coordinates of the master and slave drones are converted once in the function node to obtain x'_k, y'_k, z'_k and $x'_k m, y'_k m, z'_k m$.

$$\begin{aligned} x'_k &= x_k\left(\cos\theta_y \cos\theta_z - \sin\theta_x \sin\theta_y \sin\theta_z\right) - y_k \cos\theta_x \sin\theta_z + z_k\left(\sin\theta_y \cos\theta_z + \sin\theta_x \cos\theta_y \sin\theta_z\right), \\ y'_k &= x_k\left(\cos\theta_y \sin\theta_z + \sin\theta_x \sin\theta_y \cos\theta_z\right) + y_k \cos\theta_x \cos\theta_z + z_k\left(\sin\theta_y \sin\theta_z - \sin\theta_x \cos\theta_y \cos\theta_z\right), \\ z'_k &= x_k\left(-\cos\theta_x \sin\theta_y\right) + y_k \sin\theta_x + z_k \cos\theta_x \cos\theta_y, \\ x'_k m &= x^m_k\left(\cos\theta_y \cos\theta_z - \sin\theta_x \sin\theta_y \sin\theta_z\right) - y^m_k \cos\theta_x \sin\theta_z + z^m_k\left(\sin\theta_y \cos\theta_z + \sin\theta_x \cos\theta_y \sin\theta_z\right), \\ y'_k m &= x^m_k\left(\cos\theta_y \sin\theta_z + \sin\theta_x \sin\theta_y \cos\theta_z\right) + y^m_k \cos\theta_x \cos\theta_z + z^m_k\left(\sin\theta_y \sin\theta_z - \sin\theta_x \cos\theta_y \cos\theta_z\right), \\ z'_k m &= x^m_k\left(-\cos\theta_x \sin\theta_y\right) + y^m_k \sin\theta_x + z^m_k \cos\theta_x \cos\theta_y, \end{aligned} \tag{20}$$

where θ_x, θ_y, and θ_z are the angle of rotation along the x-, y-, and z-axes, respectively.

3.2.5. $\Delta x_k, \Delta y_k, \Delta z_k$ Update. The role of function nodes B, C, and D is to convert the relative and absolute position information. Therefore, the function probability density of these function nodes transferred to node $\Delta x_k, \Delta y_k, \Delta z_k$ is

$$\begin{aligned} &N\left(\Delta x, x'_k m - \mu_{x'_k}, \sigma^2_{x_k}\right), \\ &N\left(\Delta y, y'_k m - \mu_{y'_k}, \sigma^2_{y_k}\right), \\ &N\left(\Delta z, z'_k m - \mu_{z'_k}, \sigma^2_{z_k}\right). \end{aligned} \tag{21}$$

Owing to $d^2 = \Delta x^2_k + \Delta y^2_k + \Delta z^2_k$, the probability density function passed from the function node E to the variable node can be expressed as

$$\begin{aligned} &N\left(\Delta x_k, \pm\sqrt{d^2_k - \left(\mu_{\Delta x_k}\right)^2}, \frac{\left(\mu_{\Delta x_k}\right)^2 \sigma^2_{x_k} + d^2_k \sigma^2_{d_k}}{d^2_k - \left(\mu_{\Delta x_k}\right)^2}\right), \\ &N\left(\Delta y_k, \pm\sqrt{d^2_k - \left(\mu_{\Delta y_k}\right)^2}, \frac{\left(\mu_{\Delta y_k}\right)^2 \sigma^2_{y_k} + d^2_k \sigma^2_{d_k}}{d^2_k - \left(\mu_{\Delta y_k}\right)^2}\right), \\ &N\left(\Delta z_k, \pm\sqrt{d^2_k - \left(\mu_{\Delta z_k}\right)^2}, \frac{\left(\mu_{\Delta z_k}\right)^2 \sigma^2_{x_k} + d^2_k \sigma^2_{d_k}}{d^2_k - \left(\mu_{\Delta z_k}\right)^2}\right). \end{aligned} \tag{22}$$

3.2.6. Second Update of x'_k, y'_k, z'_k. The probability density of the function nodes B, C, and D transfer function to $(x'_k, y_k{}', z'_k)$ is

$$\begin{aligned} &N\left(x'_k, x'_k m - \mu_{\Delta x_k}, \sigma^2_{\Delta x_k}\right), \\ &N\left(y'_k, y'_k m - \mu_{\Delta y_k}, \sigma^2_{\Delta y_k}\right), \\ &N\left(z'_k, z'_k m - \mu_{\Delta z_k}, \sigma^2_{\Delta z_k}\right). \end{aligned} \tag{23}$$

3.2.7. Second Update of x_k, y_k, z_k. After calculation and finishing, the updated function probability density of (x_k, y_k, z_k) is as follows:

$$\begin{aligned} x_k &\sim N\left(x_k, \mu_{x_k}, \sigma^2_{x_k}\right), \\ y_k &\sim N\left(y_k, \mu_{y_k}, \sigma^2_{y_k}\right), \\ z_k &\sim N\left(z_k, \mu_{z_k}, \sigma^2_{z_k}\right), \\ \frac{1}{\sigma^2_{x_k}} &= \frac{1}{\sigma^2_{\widehat{x}_k^-}} + \frac{1}{\sigma^2_{x'_k{}'}}, \end{aligned}$$

$$\frac{1}{\sigma_{y_k}^2} = \frac{1}{\sigma_{\tilde{y}_k^-}^2} + \frac{1}{\sigma_{y_k''}^2},$$

$$\frac{1}{\sigma_{z_k}^2} = \frac{1}{\sigma_{\tilde{z}_k^-}^2} + \frac{1}{\sigma_{z_k''}^2},$$

$$\mu_{x_k} = \sigma_{x_k}^2 \left(\frac{\mu_{\tilde{x}_k^-}}{\sigma_{\tilde{x}_k^-}^2} + \frac{\mu_{x_k''}}{\sigma_{x_k''}^2} \right),$$

$$\mu_{y_k} = \sigma_{y_k}^2 \left(\frac{\mu_{\tilde{y}_k^-}}{\sigma_{\tilde{y}_k^-}^2} + \frac{\mu_{y_k''}}{\sigma_{y_k''}^2} \right),$$

$$\mu_{z_k} = \sigma_{x_k}^2 \left(\frac{\mu_{\tilde{z}_k^-}}{\sigma_{\tilde{z}_k^-}^2} + \frac{\mu_{z_k''}}{\sigma_{z_k''}^2} \right), \tag{24}$$

where $\mu_{x_k''}, \mu_{y_k''}, \mu_{z_k''}$ and $\sigma_{x_k''}^2, \sigma_{y_k''}^2, \sigma_{z_k''}^2$ are the information transferred to x_k, y_k, z_k after x_k', y_k', z_k' coordinate conversion. The coordinate conversion process is as follows:

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} \cos\theta_y\cos\theta_z - \sin\theta_x\sin\theta_y\sin\theta_z & \cos\theta_x\sin\theta_z & \sin\theta_y\cos\theta_z + \sin\theta_x\cos\theta_y\sin\theta_z \\ \cos\theta_y\sin\theta_z + \sin\theta_x\sin\theta_y\cos\theta_z & \cos\theta_x\cos\theta_z & \sin\theta_y\sin\theta_z - \sin\theta_x\cos\theta_y\cos\theta_z \\ -\cos\theta_x\sin\theta_y & \sin\theta_x & \cos\theta_x\cos\theta_y \end{bmatrix}^{-1} \begin{bmatrix} x_k' \\ y_k' \\ z_k' \end{bmatrix}. \tag{25}$$

3.3. Error Analysis of Factor Graph Colocation Algorithm. The error formula can be rewritten as

$$\begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix} = \begin{bmatrix} \delta x_k^{\tilde{v}} & \delta x_k^{\tilde{\alpha}} & \delta x_k^{\tilde{\beta}} \\ \delta y_k^{\tilde{v}} & \delta y_k^{\tilde{\alpha}} & \delta y_k^{\tilde{\beta}} \\ \delta z_k^{\tilde{v}} & 0 & \delta z_k^{\tilde{\beta}} \end{bmatrix} \begin{bmatrix} \tilde{v} \\ \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix},$$
$$\begin{bmatrix} \tilde{v} \\ \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix} = \begin{bmatrix} \delta x_k^{\tilde{v}} & \delta x_k^{\tilde{\alpha}} & \delta x_k^{\tilde{\beta}} \\ \delta y_k^{\tilde{v}} & \delta y_k^{\tilde{\alpha}} & \delta y_k^{\tilde{\beta}} \\ \delta z_k^{\tilde{v}} & 0 & \delta z_k^{\tilde{\beta}} \end{bmatrix}^{-1} \begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix}. \tag{26}$$

In the case of colocation, the position update needs to be calculated after subtracting the corresponding estimation error from the measured value of the speed and heading as follows:

$$\begin{aligned} x_k^- &= x_{k-1} + \left(v_k^- - \tilde{v}_k'\right)\Delta t \cos\left(\beta_k^- - \tilde{\beta}_k^-\right) \cos\left(\alpha_k^- - \tilde{\alpha}_k^-\right), \\ y_k^- &= y_{k-1} + \left(v_k^- - \tilde{v}_k'\right)\Delta t \cos\left(\beta_k^- - \tilde{\beta}_k^-\right) \sin\left(\alpha_k^- - \tilde{\alpha}_k^-\right), \\ z_k^- &= z_{k-1} + \left(v_k^- - \tilde{v}_k'\right)\Delta t \sin\left(\beta_k^- - \tilde{\beta}_k^-\right), \end{aligned} \tag{27}$$

where $\begin{cases} \tilde{v}_k' = \tilde{v}_{k-1}' + \tilde{v}_{k-1}, \tilde{v}_{k-1} = 0 \\ \tilde{\alpha}_k' = \tilde{\alpha}_{k-1}' + \tilde{\alpha}_{k-1}, \tilde{\alpha}_{k-1} = 0 \\ \tilde{\beta}_k' = \tilde{\beta}_{k-1}' + \tilde{\beta}_{k-1}, \tilde{\beta}_{k-1} = 0 \end{cases}$; therefore, the expectation and variance values can be expressed as follows:

$$\begin{aligned} \mu_{\tilde{v}_k'} &= \mu_{\tilde{v}_{k-1}'} + \mu_{\tilde{v}_{k-1}}, \\ \sigma_{\tilde{v}_k'}^2 &= \sigma_{\tilde{v}_{k-1}'}^2, \\ \mu_{\tilde{\alpha}_k'} &= \mu_{\tilde{\alpha}_{k-1}'} + \mu_{\tilde{\alpha}_{k-1}}, \\ \sigma_{\tilde{\alpha}_k'}^2 &= \sigma_{\tilde{\alpha}_{k-1}'}^2, \\ \mu_{\tilde{\beta}_k'} &= \mu_{\tilde{\beta}_{k-1}'} + \mu_{\tilde{\beta}_{k-1}}, \\ \sigma_{\tilde{\beta}_k'}^2 &= \sigma_{\tilde{\beta}_{k-1}'}^2. \end{aligned} \tag{28}$$

After receiving the measurement information of the main drone from the drone at time a and performing error estimation, $\delta x_a^{\tilde{v}} = 0$, $\delta x_a^{\tilde{v}}$ is used as the starting point of $\delta x_k^{\tilde{v}}$ for the next update. At this time, the expectation and variance of $\delta x_a^{\tilde{v}}$ are both 0. Subsequently, $\delta x_k^{\tilde{v}}$ after time a can be recursively expressed as $\delta x_k^{\tilde{v}} = \delta x_{k-1}^{\tilde{v}} + \Delta t \cos\beta_k \cos\alpha_k$; however, because the actual value of β, α cannot be determined, it is represented by $\widehat{\beta}, \widehat{\alpha}$, that is,

$$\begin{aligned} \widehat{\alpha}_k &= \alpha_k^- - \tilde{\alpha}_k', \\ \mu_{\widehat{\alpha}_k} &= \mu_{\alpha_k^-} - \mu_{\tilde{\alpha}_k'}, \\ \sigma_{\widehat{\alpha}_k}^2 &= \sigma_{\alpha_k^-}^2 + \sigma_{\tilde{\alpha}_k'}^2, \\ \widehat{\beta}_k &= \beta_k^- - \tilde{\beta}_k', \\ \mu_{\widehat{\alpha}_k} &= \mu_{\beta_k^-} - \mu_{\tilde{\beta}_k'}, \\ \sigma_{\widehat{\beta}_k}^2 &= \sigma_{\beta_k^-}^2 + \sigma_{\tilde{\beta}_k'}^2. \end{aligned} \tag{29}$$

Further, the expectation and variance of $\delta x_k^{\tilde{v}}$ can be expressed as

$$\mu_{\delta x_k^{\tilde{v}}} = \mu_{\delta x_{k-1}^{\tilde{v}}} + \Delta t \mu_{\cos \widehat{\beta}_k} \mu_{\cos \widehat{\alpha}_k},$$
$$\sigma^2_{\delta x_k^{\tilde{v}}} = \sigma^2_{\delta x_{k-1}^{\tilde{v}}} + \Delta t^2 \left(\sigma^2_{\cos \widehat{\beta}_k} \sigma^2_{\cos \alpha_k} + \sigma^2_{\cos \widehat{\beta}_k} \mu^2 + \sigma^2_{\cos \widehat{\alpha}_k} \mu^2_{\cos \widehat{\beta}_k} \right). \tag{30}$$

Similarly, the expressions of $\delta y_k^{\tilde{v}}, \delta z_k^{\tilde{v}}, \delta x_k^{\tilde{\alpha}}, \delta y_k^{\tilde{\alpha}}, \delta z_k^{\tilde{\alpha}}, \delta x_k^{\tilde{\beta}}, \delta y_k^{\tilde{\beta}}, \delta z_k^{\tilde{\beta}}$ and the corresponding expectations and variances can be obtained as follows:

$$\delta x_k^{\tilde{v}} = \delta x_{k-1}^{\tilde{v}} + \Delta t \cos \widehat{\beta}_k \cos \widehat{\alpha}_k,$$
$$\delta y_k^{\tilde{v}} = \delta y_{k-1}^{\tilde{v}} + \Delta t \cos \widehat{\beta}_k \sin \widehat{\alpha}_k,$$
$$\delta z_k^{\tilde{v}} = \delta z_{k-1}^{\tilde{v}} + \Delta t \sin \widehat{\beta}_k,$$
$$\mu_{\delta x_k^{\tilde{v}}} = \mu_{\delta x_{k-1}^{\tilde{v}}} + \Delta t \mu_{\cos \widehat{\beta}_k} \mu_{\cos \widehat{\alpha}_k},$$
$$\sigma^2_{\delta x_k^{\tilde{v}}} = \sigma^2_{\delta x_{k-1}^{\tilde{v}}} + \Delta t^2 \left(\sigma^2_{\cos \widehat{\beta}_k} \sigma^2_{\cos \widehat{\alpha}_k} + \sigma^2_{\cos \widehat{\beta}_k} \mu^2_{\cos \widehat{\alpha}_k} + \sigma^2_{\cos \widehat{\alpha}_k} \mu^2_{\cos \widehat{\beta}_k} \right),$$
$$\mu_{\delta y_k^{\tilde{v}}} = \mu_{\delta y_{k-1}^{\tilde{v}}} + \Delta t \mu_{\cos \widehat{\beta}_k} \mu_{\sin \widehat{\alpha}_k},$$
$$\sigma^2_{\delta y_k^{\tilde{v}}} = \sigma^2_{\delta y_{k-1}^{\tilde{v}}} + \Delta t^2 \left(\sigma^2_{\cos \widehat{\beta}_k} \sigma^2_{\sin \widehat{\alpha}_k} + \sigma^2_{\cos \widehat{\beta}_k} \mu^2_{\sin \widehat{\alpha}_k} + \sigma^2_{\sin \widehat{\alpha}_k} \mu^2_{\cos \widehat{\beta}_k} \right),$$
$$\mu_{\delta z_k^{\tilde{v}}} = \mu_{\delta z_{k-1}^{\tilde{v}}} + \Delta t \mu_{\sin \widehat{\beta}_k},$$
$$\sigma^2_{\delta z_k^{\tilde{v}}} = \sigma^2_{\delta z_{k-1}^{\tilde{v}}} + \Delta t^2 \sigma^2_{\sin \widehat{\beta}_k},$$
$$\delta x_k^{\tilde{\alpha}} = \delta x_{k-1}^{\tilde{\alpha}} - \widehat{v}_k \Delta t \cos \widehat{\beta} \sin \widehat{\alpha},$$
$$\delta y_k^{\tilde{\alpha}} = \delta y_{k-1}^{\tilde{\alpha}} + \widehat{v}_k \Delta t \cos \widehat{\beta} \cos \widehat{\alpha},$$
$$\delta z_k^{\tilde{\alpha}} = 0,$$
$$\mu_{\delta x_k^{\tilde{\alpha}}} = \mu_{\delta x_{k-1}^{\tilde{\alpha}}} - \Delta t \mu_{\cos \widehat{\beta}_k} \mu_{\sin \widehat{\alpha}_k},$$
$$\sigma^2_{\delta x_k^{\tilde{\alpha}}} = \sigma^2_{\delta x_{k-1}^{\tilde{\alpha}}} + \Delta t^2 \left(\sigma^2_{\cos \widehat{\beta}_k} \sigma^2_{\sin s\widehat{\alpha}_k} + \sigma^2_{\cos \widehat{\beta}_k} \mu^2_{\sin \widehat{\alpha}_k} + \sigma^2_{\sin \widehat{\alpha}_k} \mu^2_{\cos \widehat{\beta}_k} \right),$$
$$\mu_{\delta y_k^{\tilde{\alpha}}} = \mu_{\delta y_{k-1}^{\tilde{\alpha}}} + \Delta t \mu_{\cos \widehat{\beta}_k} \mu_{\cos \widehat{\alpha}_k},$$
$$\sigma^2_{\delta y_k^{\tilde{\alpha}}} = \sigma^2_{\delta y_{k-1}^{\tilde{\alpha}}} + \Delta t^2 \left(\sigma^2_{\cos \widehat{\beta}_k} \sigma^2_{\cos \widehat{\alpha}_k} + \sigma^2_{\cos \widehat{\beta}_k} \mu^2_{\cos \widehat{\alpha}_k} + \sigma^2_{\cos \widehat{\alpha}_k} \mu^2_{\cos \widehat{\beta}_k} \right),$$
$$\mu_{\delta z_k^{\tilde{\alpha}}} = 0,$$
$$\sigma^2_{\delta z_k^{\tilde{\alpha}}} = 0,$$
$$\delta x_k^{\tilde{\beta}} = 0,$$
$$\delta y_k^{\tilde{\beta}} = 0,$$
$$\delta z_k^{\tilde{\beta}} = \delta z_{k-1} + \widehat{v}_k \Delta t \cos \widehat{\beta}_k,$$
$$\mu_{\delta x_k^{\tilde{\beta}}} = 0,$$
$$\sigma^2_{\delta x_k^{\tilde{\beta}}} = 0,$$
$$\mu_{\delta y_k^{\tilde{\beta}}} = 0,$$
$$\sigma^2_{\delta y_k^{\tilde{\beta}}} = 0,$$
$$\mu_{\delta z_k^{\tilde{\beta}}} = \mu_{\delta z_{k-1}^{\tilde{\beta}}} + \Delta t \mu_{\widehat{v}_k \cos \widehat{\beta}_k},$$
$$\sigma^2_{\delta z_k^{\tilde{\alpha}}} = \sigma^2_{\delta z_{k-1}^{\tilde{\beta}}} + \Delta t^2 \sigma^2_{\widehat{v}_k \cos \widehat{\beta}_k}. \tag{31}$$

Equation (32) presents the updated formula of $\tilde{v}_k^-, \tilde{\alpha}_k^-, \tilde{\beta}_k^-$.

$$\begin{aligned} \tilde{v}_k^- &= \tilde{v}_{k-1}, \\ \tilde{\alpha}_k^- &= \tilde{\alpha}_{k-1}, \\ \tilde{\beta}_k^- &= \tilde{\beta}_{k-1}. \end{aligned} \tag{32}$$

If the slave drone does not receive the position information of the master drone at time k, then $\begin{cases} \tilde{v}_k^- = \tilde{v}_k \\ \tilde{\alpha}_k^- = \tilde{\alpha}_k \\ \tilde{\beta}_k^- = \tilde{\beta}_k \end{cases}$, and the expected variance can be expressed as follows:

$$\begin{aligned} \mu_{\delta x_k} &= \mu_{y_k^-} - \mu_{x_k}, \\ \sigma^2_{\delta x_k} &= \sigma^2_{x_k^-} + \sigma^2_{x_k}, \\ \mu_{\delta y_k} &= \mu_{y_k^-} - \mu_{y_k}, \\ \sigma^2_{\delta y_k} &= \sigma^2_{y_k^-} + \sigma^2_{y_k}, \\ \mu_{\delta z_k} &= \mu_{z_k^-} - \mu_{z_k}, \\ \sigma^2_{\delta z_k} &= \sigma^2_{z_k^-} + \sigma^2_{z_k}. \end{aligned} \tag{33}$$

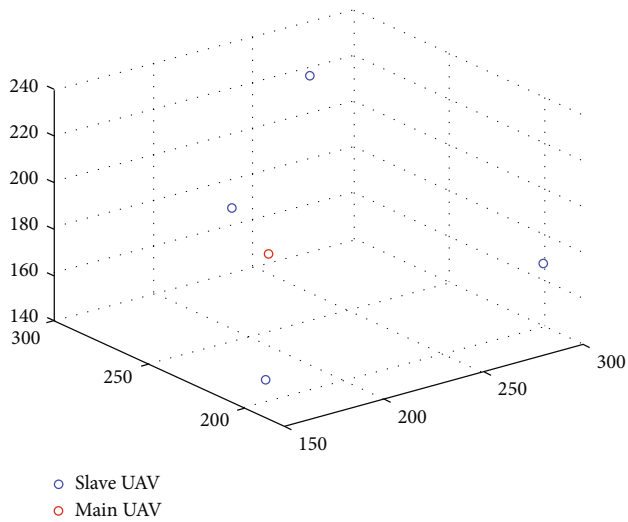


FIGURE 3: Initial position of the drone.

As shown in Figure 2, the information of the variable node $\delta x_k^{\tilde{v}}$ is obtained by the variable nodes $\delta x_{k-1}^{\tilde{v}}$, $\tilde{\alpha}_k$, and $\tilde{\beta}_k$ through the function node, which can be expressed as $N(\delta x_k^{\tilde{v}}, \mu_{\delta x_k^{\tilde{v}}}, \sigma^2_{\delta x_k^{\tilde{v}}})$. In addition, the information of the variable node $\delta y_k^{\tilde{v}}$ is obtained by the variable nodes $\delta y_{k-1}^{\tilde{v}}$, $\tilde{\alpha}_k$, and $\tilde{\beta}_k$ through the function node, which can be expressed as $N(\delta y_k^{\tilde{v}}, \mu_{\delta y_k^{\tilde{v}}}, \sigma^2_{\delta y_k^{\tilde{v}}})$. Further, for the variable node $\delta z_k^{\tilde{v}}$, its information is obtained by the variable nodes $\delta z_{k-1}^{\tilde{v}}$, $\tilde{\alpha}_k$, and $\tilde{\beta}_k$ through the $\delta y_{k-1}^{\tilde{\alpha}} \tilde{\alpha}_k N(\delta z_k^{\tilde{v}}, \mu_{\delta z_k^{\tilde{v}}}, \sigma^2_{\delta z_k^{\tilde{v}}})$ function node, which can be expressed as $N(\delta z_k^{\tilde{v}}, \mu_{\delta z_k^{\tilde{v}}}, \sigma^2_{\delta z_k^{\tilde{v}}})$.

The information of the variable node $\delta x_k^{\tilde{\alpha}}$ is obtained by $\delta x_{k-1}^{\tilde{\alpha}}$, $\tilde{\alpha}_k$, and $\tilde{v}_k$ through the function node, which can be expressed as $N(\delta x_k^{\tilde{\alpha}}, \mu_{\delta x_k^{\tilde{\alpha}}}, \sigma^2_{\delta x_k^{\tilde{\alpha}}})$; the information of the variable node $\delta y_k^{\tilde{\alpha}}$ is obtained by $\delta y_{k-1}^{\tilde{\alpha}}$, $\tilde{\alpha}_k$, and $\tilde{v}_k$ through the function node, which can be expressed as $N(\delta y_k^{\tilde{\alpha}}, \mu_{\delta y_k^{\tilde{\alpha}}}, \sigma^2_{\delta y_k^{\tilde{\alpha}}})$; and the information of the variable node $\delta z_k^{\tilde{\alpha}}$ is obtained by $\delta z_{k-1}^{\tilde{\alpha}}$, $\tilde{\alpha}_k$, and $\tilde{v}_k$ through the function node, which can be expressed as $N(\delta z_k^{\tilde{\alpha}}, \mu_{\delta z_k^{\tilde{\alpha}}}, \sigma^2_{\delta z_k^{\tilde{\alpha}}})$.

Furthermore, the information of the variable node $\delta x_k^{\tilde{\beta}}$ is obtained by $\delta x_{k-1}^{\tilde{\beta}}$, $\tilde{\beta}_k$, and $\tilde{v}_k$ through the function node, which can be expressed as $N(\delta x_k^{\tilde{\beta}}, \mu_{\delta x_k^{\tilde{\beta}}}, \sigma^2_{\delta x_k^{\tilde{\beta}}})$; the information of the variable node $\delta y_k^{\tilde{\beta}}$ is obtained by $\delta y_{k-1}^{\tilde{\beta}}$, $\tilde{\beta}_k$, and $\tilde{v}_k$ through the function node, which can be expressed as $N(\delta y_k^{\tilde{\beta}}, \mu_{\delta y_k^{\tilde{\beta}}}, \sigma^2_{\delta y_k^{\tilde{\beta}}})$; and the information of the variable node $\delta z_k^{\tilde{\beta}}$ is obtained by $\delta z_{k-1}^{\tilde{\beta}}$, $\tilde{\beta}_k$, and $\tilde{v}_k$ through the function node, which can be expressed as $N(\delta z_k^{\tilde{\beta}}, \mu_{\delta z_k^{\tilde{\beta}}}, \sigma^2_{\delta z_k^{\tilde{\beta}}})$.

The information passed from the function nodes $\tilde{v}_{k-1}$, $\tilde{\alpha}_{k-1}$, and $\tilde{\beta}_{k-1}$ to variable nodes $\tilde{v}_k$, $\tilde{\alpha}_k$, and $\tilde{\beta}_k$ can be expressed as $N(\tilde{v}_k, \mu_{\tilde{v}_k}, \sigma^2_{\tilde{v}_k})$, $N(\tilde{\alpha}_k, \mu_{\tilde{\alpha}_k}, \sigma^2_{\tilde{\alpha}_k})$, and $N(\tilde{\beta}_k, \mu_{\tilde{\beta}_k}, \sigma^2_{\tilde{\beta}_k})$, respectively.

The information transferred from the function nodes F, G, and H to the variable nodes δx_k, δy_k, and δz_k can be represented as $N(\delta x_k, \mu_{\delta x_k}, \sigma^2_{\delta x_k})$, $N(\delta y_k, \mu_{\delta y_k}, \sigma^2_{\delta y_k})$, and $N(\delta z_k, \mu_{\delta z_k}, \sigma^2_{\delta z_k})$, respectively. The information transferred from function node I to the variable node $\tilde{v}_k$ can be represented as $N(\tilde{v}_k^+, \mu_{\tilde{v}^+}, \sigma^2_{\tilde{v}^+})$, the information transferred to $\tilde{\alpha}_k$ can be represented as $N(\tilde{\alpha}_k^+, \mu_{\tilde{\alpha}^+}, \sigma^2_{\tilde{\alpha}^+})$, and the information transferred to $\tilde{\beta}_k$ can be represented as $N(\tilde{\beta}_k^+, \mu_{\tilde{\beta}^+}, \sigma^2_{\tilde{\beta}^+})$.

The a priori estimate of $\tilde{v}_k$ and the information passed from function node I to the variable node $\tilde{v}_k$ can be combined to obtain the posterior estimate of $\tilde{v}_k$ at time k:

$$\begin{aligned} \mu_{\tilde{v}_k} &= \sigma^2_{\tilde{v}_k}\left(\frac{\mu_{\tilde{v}_k^-}}{\sigma^2_{\tilde{v}_k^-}} + \frac{\mu_{\tilde{v}_k^+}}{\sigma^2_{\tilde{v}_k^+}}\right), \\ \sigma^2_{\tilde{v}_k} &= \frac{\sigma^2_{\tilde{v}_k^-} \cdot \sigma^2_{\tilde{v}_k^+}}{\sigma^2_{\tilde{v}_k^-} + \sigma^2_{\tilde{v}_k^+}}. \end{aligned} \tag{34}$$

Similarly, the posterior estimates of $\tilde{\alpha}_k$ and $\tilde{\beta}_k$ are

$$\begin{aligned} \mu_{\tilde{\alpha}_k} &= \sigma^2_{\tilde{\alpha}_k}\left(\frac{\mu_{\tilde{\alpha}_k^-}}{\sigma^2_{\tilde{\alpha}_k^-}} + \frac{\mu_{\tilde{\alpha}_k^+}}{\sigma^2_{\tilde{\alpha}_k^+}}\right), \\ \sigma^2_{\tilde{\alpha}_k} &= \frac{\sigma^2_{\tilde{\alpha}_k^-} \cdot \sigma^2_{\tilde{\alpha}_k^+}}{\sigma^2_{\tilde{\alpha}_k^-} + \sigma^2_{\tilde{\alpha}_k^+}}, \\ \mu_{\tilde{\beta}_k} &= \sigma^2_{\tilde{\beta}_k}\left(\frac{\mu_{\tilde{\beta}_k^-}}{\sigma^2_{\tilde{\beta}_k^-}} + \frac{\mu_{\tilde{\beta}_k^+}}{\sigma^2_{\tilde{\beta}_k^+}}\right), \\ \sigma^2_{\tilde{\beta}_k} &= \frac{\sigma^2_{\tilde{\beta}_k^-} \cdot \sigma^2_{\tilde{\beta}_k^+}}{\sigma^2_{\tilde{\beta}_k^-} + \sigma^2_{\tilde{\beta}_k^+}}. \end{aligned} \tag{35}$$

4. Simulation Results and Analysis

The movement of the drone was simulated in MATLAB to verify the feasibility of this algorithm. The simulation conditions were as follows: five UAVs were moving in an area of 1000 m × 600 m × 300 m. The position of the main UAV is accurately determined. The slave UAVx and main UAV communicate with each other, and the slave UAV obtains the location and observation information from the main UAV. Figure 3 illustrates the initial position of the slave and main UAV.

In the experiment, the MUAV-FGC algorithm estimates the speed error of the slave UAV, as shown in Figure 4. The abscissa in the figure represents time (s), and the ordinate is the speed error estimate (m/s). The figure depicts that, as the experiment progresses, the speed error continues to accumulate. The estimated average speed error from the UAV was 0.0644 m/s during the experiment time of 600 s. Therefore, the speed error must be compensated during the positioning process to improve positioning accuracy.

In the experiment, the MUAV-FGC algorithm estimates the dead angle error of the slave UAV, as shown in Figure 5.

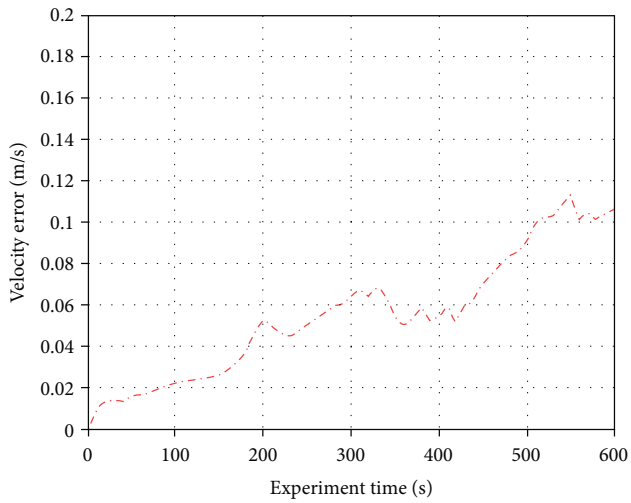


Figure 4: Speed error estimation results.

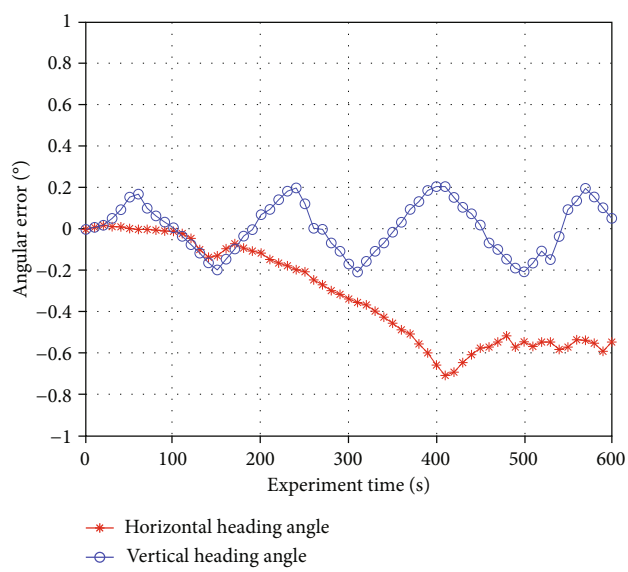


Figure 5: Dead angle error estimation results.

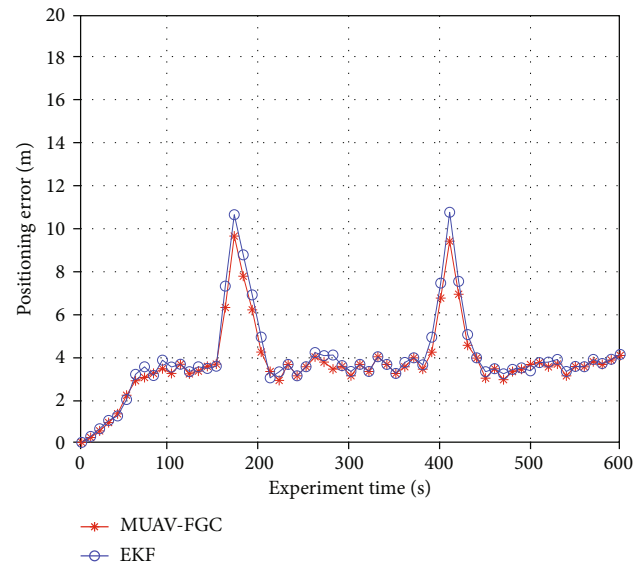


Figure 6: Positioning error curve.

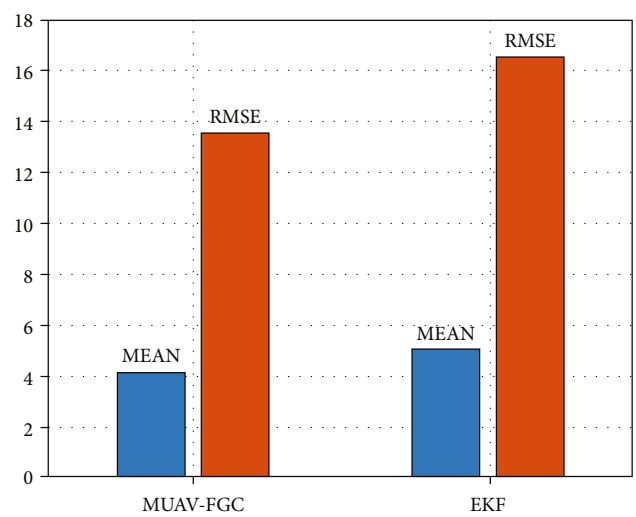


Figure 7: RMS error comparison.

In the figure, the abscissa represents the time of the experiment (s), the ordinate represents the dead-end angle error (°), the curve with "∗" is the horizontal dead-angle error curve, and the curve with "○" is the vertical dead angle curve. As shown, the vertical dead-angle error fluctuates between -0.2°and 0.2°, and the horizontal dead-angle error stabilizes at approximately -0.58°after 400 s. In addition, the error of the dead angle will affect the positioning; therefore, it is necessary to compensate for the heading angle error in the positioning process to improve the positioning accuracy.

Figure 6 presents the positioning error curves of the MUAV-FGC and the extended Kalman algorithms. In the figure, the abscissa represents the time (s), the ordinate represents the positioning error (m), the curve with "∗" is the positioning error curve of the MUAV-FGC algorithm used in this study, and the curve with "○" is the positioning error curve of the EKF algorithm. As shown in Figure 6, the MUAV-FGC algorithm curve is below the EKF algorithm curve, indicating that the average positioning error of the MUAV-FGC algorithm is smaller than that of the EKF algorithm. Thus, the positioning accuracy of the MUAV-FGC algorithm is higher than that of the EKF algorithm. This is because the EKF algorithm does not estimate and compensate for the speed and heading errors, resulting in a high positioning error, whereas the MUAV-FGC algorithm significantly reduces the positioning error.

Figure 7 demonstrates the mean and root mean square values of the positioning error of the EKF and FG algorithms in the experiment. The root mean square and mean values of the EKF algorithm are 16.47 and 5.09, whereas those of the MUAV-FGC algorithm are 13.58 and 4.13, respectively. The positioning error of the MUAV-FGC algorithm was found to be smaller than that of the EKF algorithm. Compared with the EKF algorithm, the mean positioning and

root mean square errors of the MUAV-FGC algorithm was reduced by approximately 18.86%, and 17.54%, respectively.

5. Conclusion

In this study, we addressed the problem of multi-UAV colocation in a disaster environment, proposed a multi-UAV factor graph colocation algorithm, and derived formulas for speed and heading angle. In the factor graph model, the expectation and variance of variables are used as the information transferred between factors, and the process of error estimation algorithm based on the factor graph is derived. Based on the simulation experiment results, the average positioning error and root mean square error of the MUAV-FGC algorithm are reduced by approximately 18.86% and 17.54%, respectively, along with providing a better positioning performance, compared to the EKF algorithm.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported in part by the Sub Project of National Key Research and Development Plan in 2020 (No. 2020YFC1511704), the Scientific Research Level Improvement Project to Promote the Colleges Connotation Development of Beijing Information Science and Technology University (No. 2020KYNH212 and No. 2021CGZH302), and in part by the National Natural Science Foundation of China (Grant No. 61971048).

References

[1] X. Li, Y. Zheng, W. U. Khan et al., "Physical layer security of cognitive ambient backscatter communications for green Internet-of-Things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[2] C. Xu, X. Liao, J. Tan, H. Ye, and H. Lu, "Recent research progress of unmanned aerial vehicle regulation policies and technologies in urban low altitude," *IEEE Access*, vol. 8, pp. 74175–74194, 2020.

[3] R. Yin, W. Li, Z.-Q. Wang, and X.-X. Xu, "The application of artificial intelligence technology in UAV," in *2020 5th international conference on information science, computer technology and transportation (ISCTT)*, pp. 238–241, Shenyang, China, 2020.

[4] R. J. Cooper and G. B. Basingstoke, "Image processing system[P]," Sony Corporation (Tokyo,JP):20100134516.2010-06[2022-05-26].https://www.freepatentsonline.com/y2010/0134516.html.

[5] A. V. Savkin and H. Huang, "Navigation of a network of aerial drones for monitoring a frontier of a moving environmental disaster area," *IEEE Systems Journal*, vol. 14, no. 4, pp. 4746–4749, 2020.

[6] K. Ali, H. X. Nguyen, Q. Vien, P. Shah, and M. Raza, "Deployment of drone-based small cells for public safety communication system," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2882–2891, 2020.

[7] E. T. Alotaibi, S. S. Alqefari, and A. Koubaa, "LSAR: multi-UAV collaboration for search and rescue missions," *IEEE Access*, vol. 7, pp. 55817–55832, 2019.

[8] C. Wu, B. Ju, Y. Wu et al., "UAV autonomous target search based on deep reinforcement learning in complex disaster scene," *IEEE Access*, vol. 7, pp. 117227–117245, 2019.

[9] Y. Qu and Y. Zhang, "Cooperative localization against GPS signal loss in multiple UAVs flight," *Journal of Systems Engineering and Electronics*, vol. 22, no. 1, pp. 103–112, 2011.

[10] J. Pu, S. Shi, and X. Gu, "A summary of UAV positioning technology in GPS denial environment," in *Artificial Intelligence for Communications and Networks. AICON 2020. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 356*, S. Shi, L. Ye, and Y. Zhang, Eds., Springer, Cham, 2020.

[11] Q. Xin and S. Wang, "GPS positioning method of UAV based on improved particle filter," in *2020 International conference on communications*, pp. 97–101, Information System and Computer Engineering (CISCE), 2020.

[12] J. N. Gross, Y. Gu, and M. B. Rhudy, "Robust UAV relative navigation with DGPS, INS, and peer-to-peer radio ranging," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 3, pp. 935–944, 2015.

[13] J. Zhang and H. Yuan, "Analysis of unmanned aerial vehicle navigation and height control system based on GPS," *Journal of Systems Engineering and Electronics*, vol. 21, no. 4, pp. 643–649, 2010.

[14] F. Kendoul, "Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems," *Journal of Field Robotics*, vol. 29, no. 2, pp. 315–378, 2012.

[15] E. Fresk, G. Nikolakopoulos, and T. Gustafsson, "A generalized reduced-complexity inertial navigation system for unmanned aerial vehicles," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 1, pp. 192–207, 2017.

[16] P. Marantos, Y. Koveos, and K. J. Kyriakopoulos, "UAV state estimation using adaptive complementary filters," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 4, pp. 1214–1226, 2016.

[17] G. Huang, *Visual-Inertial Navigation: A Concise Review*, Proceedings of International Conference on Robotics and Automation (ICRA), 2019.

[18] Y. Wang, G. Wei, Q. Guan, and Y. Liu, "A novel positioning system of UAV based on IMA-GPS three-layer data fusion," *IEEE Access*, vol. 8, pp. 158449–158458, 2020.

[19] S. U. Rahman, G.-H. Kim, Y.-Z. Cho, and A. Khan, "Positioning of UAVs for throughput maximization in software-defined disaster area UAV communication networks," *Journal of Communications and Networks*, vol. 20, no. 5, pp. 452–463, 2018.

[20] W. Youn, M. B. Rhudy, A. Cho, and H. Myung, "Fuzzy adaptive attitude estimation for a fixed-wing UAV with a virtual SSA sensor during a GPS outage," *IEEE Sensors Journal*, vol. 20, no. 3, pp. 1456–1472, 2020.

[21] R. Padhy, F. Xia, S. K. Choudhury, P. K. Sa, and S. Bakshi, "Monocular vision aided autonomous UAV navigation in indoor corridor environments," *IEEE Transactions on Sustainable Computing*, vol. 4, no. 1, pp. 96–108, 2019.

[22] Y. Li, W. Zhang, P. Li, Y. Ning, and C. Suo, "A method for autonomous navigation and positioning of UAV based on electric field array detection," *Sensors*, vol. 21, no. 4, p. 1146, 2021.

[23] X. Liu, W. Zhang, and Y. Yang, "Interacting multiple model UAV navigation algorithm based on a robust cubature Kalman filter," *IEEE Access*, vol. 8, pp. 81034–81044, 2020.

[24] H. Bayerlein, M. Theile, M. Caccamo, and D. Gesbert, "Multi-UAV path planning for wireless data harvesting with deep reinforcement learning," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1171–1187, 2021.

[25] D. Fan, F. Gao, B. Ai et al., "Channel estimation and self-positioning for UAV swarm," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7994–8007, 2019.

[26] W. Li, Z. Xiong, Y. Sun, and J. Xiong, "Cooperative positioning algorithm of swarm UAVs based on posterior linearization belief propagation," in *2019 IEEE 3rd information technology, networking*, pp. 1277–1282, Electronic and Automation Control Conference (ITNEC), 2019.

[27] Y. Tang, Y. Hu, J. Cui et al., "Vision-aided multi-UAV autonomous flocking in GPS-denied environment," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 1, pp. 616–626, 2019.

[28] J. Xiong, Z. Xiong, J. W. Cheong, J. Xu, Y. Yu, and A. G. Dempster, "Cooperative positioning for low-cost close formation flight based on relative estimation and belief propagation," *Aerospace Science and Technology*, vol. 106, article 106068, 2020.

[29] J. Qu, B. X. Wu, and D. Yuan, "A fault-tolerant cooperative positioning approach for multiple UAVs," *Access*, vol. 5, pp. 15630–15640, 2017.

[30] V. O. Sivaneri and J. N. Gross, "UGV-to-UAV cooperative ranging for robust navigation in GNSS-challenged environments," *Aerospace Science and Technology*, vol. 71, pp. 245–255, 2017.

[31] F. Meyer, O. Hlinka, and F. Hlawatsch, "Sigma point belief propagation," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 145–149, 2014.

[32] A. R. Vetrella, G. Fasano, and D. Accardo, "Cooperative navigation in GPS-challenging environments exploiting position broadcast and vision-based tracking," in *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, Arlington, VA, USA, June 7-10, 2016.

[33] J. Wan, L. Zhong, and F. Zhang, "Cooperative localization of multi-UAVs via dynamic nonparametric belief propagation under GPS signal loss condition," *International Journal of Distributed Sensor Networks*, vol. 10, no. 2, 562105 pages, 2014.

[34] S. Fan, Y. Zhang, C. Yu, M. Zhu, and F. Yu, "An advanced cooperative positioning algorithm based on improved factor graph and sum-product theory for multiple AUVs," *IEEE Access*, vol. 7, pp. 67006–67017, 2019.

[35] Y. Zhang and P. Jia, "Multi-UAV factor graph co-location algorithm," 21 July 2022.PREPRINT (Version 1) available at Research Square. https://www.researchsquare.com/article/rs-1265570/v1.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6785754, 9 pages
https://doi.org/10.1155/2022/6785754

Research Article

Physical Layer Security for CRNs over Beaulieu-Xie Fading Channels

Yanyang Zeng,[1] **Yang Hua**,[1] **Yiwei Fang**,[2] **and Xiaohong Wang**[3]

[1]*The College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China*
[2]*The Wuhan Maritime Communication Research Institute, Wuhan 430079, China*
[3]*The College of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China*

Correspondence should be addressed to Yanyang Zeng; zengyy@hpu.edu.cn

Received 6 May 2022; Revised 6 July 2022; Accepted 27 July 2022; Published 8 August 2022

Academic Editor: Enrico M. Vitucci

In recent years, the access of massive communication devices leads to the insufficient spectrum resources of wireless networks. One of the practical means to resolve the problem is to build cognitive radio networks (CRNs), which can realize the sharing of spectrum resources between primary and secondary users, thereby improving the utilization rate of wireless spectrum resources. To this end, the CRNs are utilized to establish the Wyner's eavesdropping model over the Beaulieu-Xie fading channels. We mainly deduce the accurate expressions of secrecy outage probability and strictly positive secrecy capacity to explore the performance of physical layer security. Moreover, the better overlap between the statistical simulation and the theoretical results indicates the correctness of the theoretical analysis equation. The interesting results are that both increasing the $P_{\max}$ and decreasing the C_{th} can improve the security performance. This work is a good reference and guidance for modeling CRNs (Internet of Things, fifth generation, cell phone networks, etc.) and security performance evaluation.

1. Introduction

With the wide application of wireless transmission, the useful information will face the risk of disclosure in the process of transmission, so it is particularly important to find measures to avoid and reduce the eavesdropping of signals. Recently, the issue has become a research hotspot in academia. Different from traditional encryption and decryption methods, physical layer security (PLS) utilizes the characteristics of the transmission channels and noise to ensure the secure transmission of information. It was first proposed in [1] and applied to many practical communication scenarios, such as Beyond-5G low-latency communication networks [2], multihop amplify-and-forward (AF) relay systems [3], and cognitive radio networks (CRNs) [4]. Furthermore, the authors of [5] proposed a scheme combining deep learning with transmit antenna selection to enhance the PLS of cellular networks. In [6], the authors outlined some low complexity PLS schemes appropriate for the Internet of Things (IoTs). A cooperative beamforming scheme was designed in [7] to enhance the PLS in CRNs. For exploring the PLS properties of $\kappa-\mu$ shadowed fading, Sun et al. in [8] deduced closed-form expressions for the lower bounds of strictly positive secrecy capacity (SPSC) and secure outage probability (SOP).

Radio waves are utilized to transmit information in wireless communication networks (WCNs), but the actual communication environment is complex and variable so that there is a difference between the received and the transmitted signal, which lead to the distorted transmission of information. To better evaluate the transmission performance of WCNs, it is necessary to build and analyze system models of different fading channels [9–16]. Zhu et al. in [9] studied the performance of a full-duplex decode-and-forward (DF) system over the Rician distribution by analyzing the exact analytical expression of the outage probability (OP). The authors in [10] proposed the $\eta-\mu$ fading model and derived precise expressions of level crossing rate (LCR) and average fading duration (AFD). Badarneh et al. in [11] obtained precise expressions for the average bit error rate (ABER) and

average channel capacity (ACC) of dual $\alpha - \mu$ distribution to explored the performance of selective combining receivers. In [12], the outage performance of relay-assisted vehicular communication under double-Rayleigh fading was investigated. The exact expressions of probability density function (PDF), LCR, and AFD of double Hoyt channel were derived in [13]. Ata in [14] analyzed the PLS of cascaded Rayleigh fading channels by solving the exact precise expressions of SOP. The authors of [15] studied the security performance of the DF relay system under the $\kappa - \mu$ shadowed fading by deducing the exact accurate expressions of SOP and SPSC. In [16], the security performance of single-input multiple-output (SIMO) system under composite distribution was studied by deriving the closed expression of ACC.

The access of a large number of wireless devices will lead to a sharp shortage of spectrum resources, and the emergence of CRNs can effectively solve the problem. The CRNs improve the utilization of wireless spectrum resources mainly by using spectrum sensing technology [17] [18]. The combination of CRNs and other technologies has been extensively studied by scholars [19–27]. Goldsmith et al. in [19] reviewed the basic capacity constraints of different wireless network design patterns based on cognitive radios. In [20], the performance of CRNs was discussed when there are multiple groups of cognitive radio relays. The authors of [21] proposed a new scheme for power allocation in CRNs with limited cross-layer interference. Zhang et al. in [22] investigated the effect of the relay selection scheme on the outage performance of CRNs by studying the exact expression of OP. Tashman and Hamouda in [23] analyzed the security performance of SIMO cognitive networks over $\kappa - \mu$ distribution by deriving the closed expression of SOP. In [24], the transmission performance of SIMO underlay CRNs over generalized K channels was studied. The authors in [25] explored the PLS of CRNs under cascaded Rayleigh channels. By analyzing the accurate expression of SOP, the authors of [26] studied the security performance when there are multiple master control units and eavesdroppers in the network. The optimal PLS scheme for users in underlay multiple-input multiple-output (MIMO) CRN was studied in [27].

Exact modeling based on wireless fading channels is important for simulating specific practical scenarios and analyzing their security performance. Different fading channels have their respective characteristics, so the channels can simulate different scenarios. For instance, the Nakagami-m distribution can vary m to control the number of clusters [28]. And the Rician distribution can simulate scenarios when line-of-sight (LOS) or deterministic components are present in the received signal [29]. Combining the advantages of the above two distributions, Beaulieu and Xie proposed a new Beaulieu-Xie (BX) fading channels [30], which can represent both specular and diffuse scattering components, that is, it can simulate both LOS and non-LOS (NLOS) scenarios [31]. Therefore, the BX fading is very suitable for modeling and researching femtocells [32] and high-speed trains [33]. The accurate formulas of the LCR and AFD for BX distribution were deduced in [34]. In [35], the closed expressions for the asymptotic upper and lower bounds of the OP and the error probability of the BX fading were deduced. The theoretical expression of the OP for two BX random variables was investigated in [36]. The authors of [37] explored precise and progressive expressions for the effective rate of MIMO networks over BX distribution. Singh and Joshi in [38] derived exact expressions for the ABER of SIMO and MIMO orthogonal frequency division multiplexing systems under BX fading.

As far as the author's knowledge, there are no relevant literatures to study the security analysis based on CRNs under BX fading channels in the published database. Therefore, this work focuses on deriving the exact closed expressions of SOP and SPSC based on CRNs over the BX fading channels. On this basis, the PLS of CRN is analyzed emphatically. Moreover, the influences of the channel parameters and the maximum interference power on the security of the considering model are both studied. Finally, the correctness of the derivations in this paper is verified by simulation analysis.

The organization of this paper is as follows: In this section, we mainly introduced the research background, current situation and contributions. In Section 2, the system model based on the underlay CRN and the accurate expression of the PDF and cumulative distribution function (CDF) of the signal-to-noise ratio (SNR) of the BX distribution were described. Section 6 obtained the precise expression of SOP. The exact expression of SPSC was derived in Section 4. Section 5 analyzed the numerical simulation results. Section 6 summarized the paper.

2. System Model and Channel Characteristics

2.1. System Model. As shown in Figure 1. The underlay CRN model includes the primary network and the secondary network, which can share frequency band resources, but the premise is that the normal communication of the primary network is not disturbed. We assume that the primary network has one transmission source (P) and N users. The secondary network consists of a signal transmitting source (S), a target receiver (D), and an eavesdropper (E), and all nodes are equipped with an antenna. The channel coefficients of $S \longrightarrow P$, $S \longrightarrow D$, and $S \longrightarrow E$ are represented as h_{SP}, h_D, and h_E, respectively. S can transmit information to P, D, and E, and the transmitted signal at S is x; then the information received at P, D, and E can be written as

$$y_{SP} = \sqrt{P_S} h_{sp} x + n_{SP}, \tag{1}$$

$$y_{SD} = \sqrt{P_S} h_D x + n_D, \tag{2}$$

$$y_{SE} = \sqrt{P_S} h_E x + n_E, \tag{3}$$

where P_S represents the transmit power of S. y_{SP}, y_{SD}, and y_{SE} are the signals received at P, D, and E, respectively. $n_i \sim CN(0, \sigma^2)$ $(i = S, P, D, E)$ are represented as additive white Gaussian noise.

2.2. Statistical Properties of the BX Distribution. In this section, the expressions of the PDF and CDF for the BX distribution are given, and to make the calculation easier and

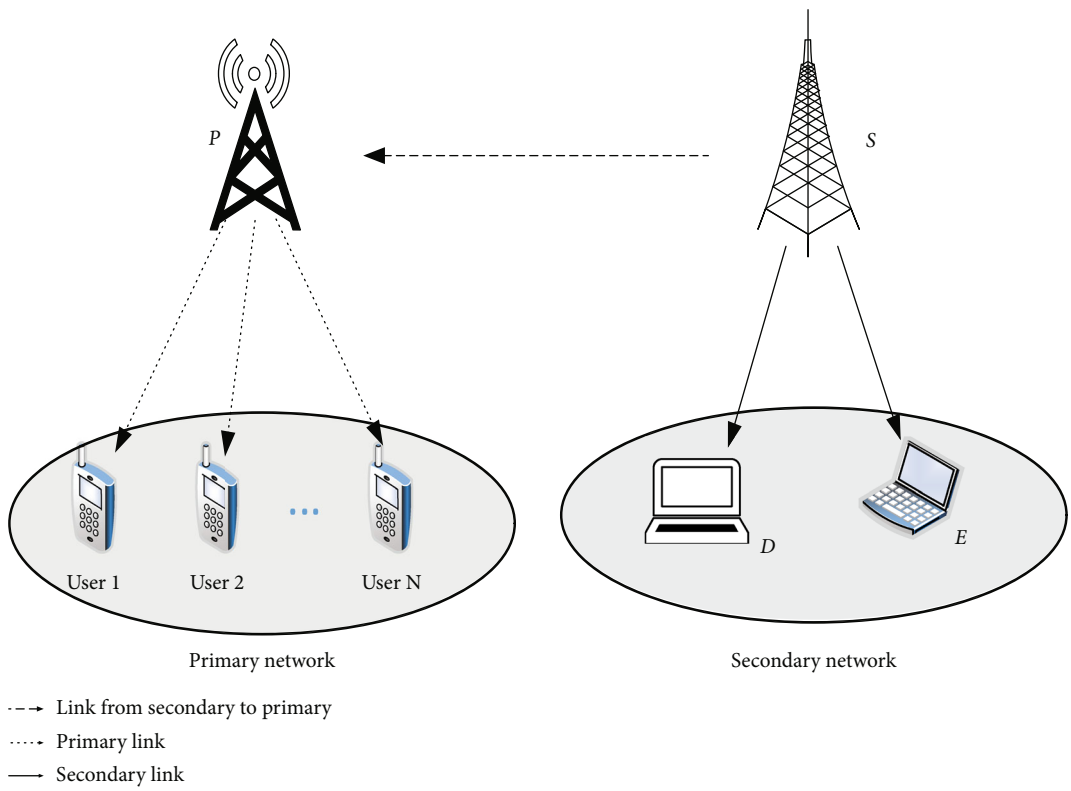


Figure 1: The system model.

more convenient, we have performed some mathematical operations on them. In [38] (eq. (3)), the PDF of the SNR for the BX channels is written as

$$f(\gamma) = \left(\frac{\sqrt{2}}{\lambda}\right)^{m-1} \left(\sqrt{\frac{m}{\Omega}}\right)^{m+1} \gamma^{m/2-1/2} e^{-\lambda^2/2} e^{-m\gamma/\Omega} I_{m-1} \cdot \left(\sqrt{\frac{2m\gamma}{\Omega}}\lambda\right), \tag{4}$$

where m and λ reflect the shape and extension, respectively, and Ω is the average SNR. $I_x(\cdot)$ is the Bessel function [39] (eq. (8.401)).

According to the formula in [39] (eq. (8.445)) of

$$I_\alpha(z) = \sum_{j=0}^{\infty} \frac{1}{j!\Gamma(\alpha+j+1)} \left(\frac{z}{2}\right)^{\alpha+2j}, \tag{5}$$

(4) can be rewritten as

$$\begin{aligned} f(\gamma) &= \left(\frac{\sqrt{2}}{\lambda}\right)^{m-1} \left(\sqrt{\frac{m}{\Omega}}\right)^{m+1} \gamma^{m/2-1/2} e^{-\lambda^2/2} e^{-m\gamma/\Omega} \cdot \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(m+k)} \left(\frac{((2/m)/\Omega)^{1/2}\gamma^{1/2}\lambda}{2}\right)^{m+2k-1} \\ &= \left(\frac{\sqrt{2}}{\lambda}\right)^{m-1} \left(\sqrt{\frac{m}{\Omega}}\right)^{m+1} e^{-(\lambda^2/2)} \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(m+k)} \cdot \left(\frac{((2/m)/\Omega)^{1/2}\lambda}{2}\right)^{m+2k-1} \gamma^{m+k-1} e^{-m\gamma/\Omega} \\ &= e^{-(\lambda^2/2)} \sum_{k=0}^{\infty} \frac{2^{-k} m^{m+k} \Omega^{-(m+k)} \lambda^{2k}}{k!\Gamma(m+k)} \gamma^{m+k-1} e^{-(m/\Omega)\gamma}. \end{aligned} \tag{6}$$

Utilizing [39] (eq. (3.351.3)), the CDF of SNR is expressed as

$$F(\gamma) = \int_0^\gamma f(x)dx = e^{-(\lambda^2/2)} \sum_{l=0}^{\infty} \frac{\left(\lambda/\sqrt{2}\right)^{2l}}{l!\Gamma(l+m)} (l+m-1)! \cdot \left(1 - e^{-(m\gamma/\Omega)} \sum_{l_1=0}^{l+m-1} \frac{(m/\Omega)^{l_1}\gamma^{l_1}}{l_1!}\right). \tag{7}$$

3. SOP Analysis

We utilize SOP as a performance evaluation metric to explore the security. SOP is defined as the probability that the value of secrecy capacity is less than the specified threshold value [40], which can be expressed as

$$SOP = \Pr\{C_S \le C_{th}\}, \tag{8}$$

where $C_S = C_D - C_E$, C_D, and C_E denote the instantaneous channel capacity of link ($S \longrightarrow D$) and the link ($S \longrightarrow E$), respectively, and C_{th} represents the specified threshold. The primary and the secondary network in the underlay CRNs must satisfy the following conditions [41]: (i) The maximum transmission power ($P_{\max}$) cannot be less than the transmission power of S. (ii) The maximum interference power (Ip) allowed by the main network cannot be less than the power from S to P. Therefore, the SOP in underlay CRNs is written as

$$\begin{aligned} SOP = \Pr\{Cs \le C_{th}\} &= \Pr\left\{C_s \le C_{th}, X \le \frac{I_p}{P_{\max}}\right\} \\ &+ \Pr\left\{C_s \le C_{th}, X \ge \frac{I_p}{P_{\max}}\right\} = \Phi_1 + \Phi_2, \end{aligned} \tag{9}$$

where X represents the instantaneous SNR of $S \longrightarrow P$. In the next subsection, we will evaluate the closed-form expressions of Φ_1 and Φ_2 in (9).

3.1. Representation of Φ_1. According to the above analysis, Φ_1 can be converted to

$$\begin{aligned} \Phi_1 &= \Pr\left\{Cs \le C_{th}, X \le \frac{Ip}{P_{\max}}\right\} \\ &= P\left\{\gamma_D \le \theta\gamma_E + \frac{\theta - 1}{\upsilon}\right\} P\left\{X \le \frac{Ip}{P_{\max}}\right\} = I_1 I_2, \end{aligned} \tag{10}$$

where $I_1 = P\{\gamma_D \le \theta\gamma_E + (\theta - 1/\upsilon)\}$, $I_2 = P\{X \le Ip/P_{\max}\}$, $\theta = e^{Cth}$, and $\upsilon = P_{\max}/\sigma^2$.

Using (6) and (7), I_1 can be obtained as

$$\begin{aligned} I_1 &= \int_0^\infty F_D\left(\theta\gamma_E + \frac{\theta - 1}{\upsilon}\right) f_E(\gamma_E) d\gamma_E \\ &= \int_0^\infty e^{-\left(\lambda_D{}^2/2\right)} \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{l!\Gamma(l + m_D)} (l + m_D - 1)! \\ &\cdot \left(1 - e^{-m_D(\theta\gamma_E + ((\theta-1)/\upsilon))/\Omega_D} \sum_{l_1=0}^{l+m_D-1} \right. \\ &\left. \cdot \frac{(m_D/\Omega_D)^{l_1} (\theta\gamma_E + ((\theta - 1)/\upsilon))^{l_1}}{l_1!}\right) \\ &\times e^{-\left(\lambda_E{}^2/2\right)} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E + k)} \\ &\cdot \gamma_E{}^{m_E+k-1} e^{-(m_E/\Omega_E)\gamma_E} d\gamma_E. \end{aligned} \tag{11}$$

Utilizing an equation in [39] (eq. (1.111)), the exact analytical solution for I_1 can be represented as

$$\begin{aligned} I_1 &= e^{-\left(\lambda_E{}^2/2\right)} e^{-\left(\lambda_D{}^2/2\right)} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E + k)} \\ &\cdot \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{l!\Gamma(l + m_D)} (l + m_D - 1)! \\ &\times \left(\int_0^\infty \gamma_E{}^{m_E+k-1} e^{-(m_E/\Omega_E)\gamma_E} d\gamma_E - e^{-((m_D((\theta-1)/\upsilon))/\Omega_D)} \right. \\ &\cdot \sum_{l_1=0}^{l+m_D-1} \frac{(m_D/\Omega_D)^{l_1}}{l_1!} \times \sum_{t_1=0}^{l_1} \binom{l_1}{t_1} \theta^{t_1} \left(\frac{\theta - 1}{\upsilon}\right)^{l_1 - t_1} \\ &\left. \cdot \int_0^\infty e^{-((m_D\theta/\Omega_D) + (m_E/\Omega_E))\gamma_E} \gamma_E{}^{m_E+k+t_1-1} d\gamma_E \right) \\ &= e^{-\left(\lambda_E{}^2/2\right)} e^{-\left(\lambda_D{}^2/2\right)} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E + k)} \\ &\cdot \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{l!\Gamma(l + m_D)} (l + m_D - 1)! \\ &\times \left(\frac{\Gamma(m_E + k)}{(m_E/\Omega_E)^{m_E+k}} - e^{-((m_D((\theta-1)/\upsilon))/\Omega_D)} \right. \\ &\cdot \sum_{l_1=0}^{l+m_D-1} \frac{(m_D/\Omega_D)^{l_1} \sum_{t_1=0}^{l_1} \binom{l_1}{t_1} \theta^{t_1} ((\theta - 1)/\upsilon)^{l_1 - t_1}}{l_1!} \\ &\left. \cdot \frac{\Gamma(m_E + k + t_1)}{((m_D\theta/\Omega_D) + (m_E/\Omega_E))^{m_E+k+t_1}} \right). \end{aligned} \tag{12}$$

Using the similar method as solving I_1, I_2 is obtained as

$$\begin{aligned} I_2 &= \Pr\left\{X \le \frac{Ip}{P_{\max}}\right\} = F_p\left(\frac{Ip}{P_{\max}}\right) \\ &= e^{-\lambda_p{}^2/2} \sum_{l=0}^{\infty} \frac{\left(\lambda_p/\sqrt{2}\right)^{2l}}{l!\Gamma(l + m_p)} (l + m_p - 1)! \\ &\cdot \left(1 - e^{-m_p(Ip/P_{\max})/\Omega_p} \sum_{l_1=0}^{l+m_p-1} \frac{(m_p/\Omega_p)^{l_1} (Ip/P_{\max})^{l_1}}{l_1!}\right). \end{aligned} \tag{13}$$

3.2. Representation of Φ_2. According to (9) and [42] (eq. (13)), the expressions for Φ_2 can be rewritten as

$$\Phi_2 = \Pr\left\{C_s \le C_{th}, X \ge \frac{I_p}{P_{\max}}\right\} = \int_{Ip/P_{\max}}^{\infty} G(x) f_s(x) dx, \tag{14}$$

where $G(x) = \int_0^\infty F_D(\theta y + ((\theta - 1)x/\varphi)) f_E(y) dy$, $\varphi = Ip/\sigma^2$.

Using (6) and (7) and an equation in [39] (eq. (3.381.4)), the exact $G(x)$ is expressed as

$$
\begin{aligned}
G(x) &= \int_0^{\infty} F_D\left(\theta y + \frac{(\theta-1)x}{\varphi}\right) f_E(y)dy \\
&= \int_0^{\infty} e^{-\lambda_D{}^2/2} \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{l!\Gamma(l+m_D)} (l+m_D-1)! \\
&\quad \times \left(1 - e^{-(m_D(((\theta-1)x)/\varphi))/\Omega_D} \sum_{l_1=0}^{l+m_D-1} \right. \\
&\quad \left. \cdot \frac{(m_D/\Omega_D)^{l_1} \sum_{t_1=0}^{l_1} \binom{l_1}{t_1} \theta^{t_1} ((\theta-1)/\varphi)^{l_1-t_1} x^{l_1-t_1}}{l_1!} y^{t_1} e^{-(\theta m_D y/\Omega_D)}\right) \\
&\quad \times e^{-(\lambda_E{}^2/2)} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E+k)} y^{m_E+k-1} e^{-(m_E/\Omega_E)y} dy \\
&= e^{-(\lambda_D{}^2/2)} \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{l!\Gamma(l+m_D)} (l+m_D-1)! e^{-(\lambda_E{}^2/2)} \\
&\quad \cdot \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E+k)} \\
&\quad \times \int_0^{\infty} \left(y^{m_E+k-1} e^{-m_E/\Omega_E y} - e^{-m_D((\theta-1)x/\varphi)/\Omega_D} \sum_{l_1=0}^{l+m_D-1} \frac{(m_D/\Omega_D)^{l_1}}{l_1!} \right. \\
&\quad \left. \cdot \times \sum_{t_1=0}^{l_1} \binom{l_1}{t_1} \theta^{t_1} \left(\frac{(\theta-1)}{\varphi}\right)^{l_1-t_1} x^{l_1-t_1} y^{m_E+k+t_1-1} e^{-(m_E/\Omega_E+\theta m_D/\Omega_D)y} \right) dy \\
&= B_1 B_2 \left(B_3 - B_4 x^{l_1-t_1} e^{-m_D(((\theta-1)x)/\varphi)/\Omega_D}\right),
\end{aligned}
\tag{15}
$$

where

$$
B_1 = e^{-(\lambda_D{}^2/2)} \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{l!\Gamma(l+m_D)} (l+m_D-1)!, \tag{16}
$$

$$
B_2 = e^{-\lambda_E{}^2/2} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E+k)}, \tag{17}
$$

$$
B_3 = \frac{\Gamma(m_E+k)}{(m_E/\Omega_E)^{(m_E+k)}}, \tag{18}
$$

$$
\begin{aligned}
B_4 = &\sum_{l_1=0}^{l+m_D-1} \frac{(m_D/\Omega_D)^{l_1} \sum_{t_1=0}^{l_1} \binom{l_1}{t_1} \theta^{t_1} ((\theta-1)/\varphi)^{l_1-t_1}}{l_1!} \\
&\cdot \frac{\Gamma(m_E+k+t_1)}{((m_E/\Omega_E)+(\theta m_D/\Omega_D))^{(m_E+k+t_1)}}.
\end{aligned}
\tag{19}
$$

Substituting (15) into (14) and using [39] (eq. (3.351.2)), Φ_2 can be represented as

$$
\begin{aligned}
\Phi_2 &= \int_{Ip/P_{\max}}^{\infty} G(x) f_s(x) dx = \int_{Ip/P_{\max}}^{\infty} B_1 B_2 \left(B_3 B_5 x^{m_E+k-1} e^{-m_E/\Omega_E x} \right. \\
&\quad \left. - B_4 B_5 x^{l_1-t_1+m_E+k-1} e^{-\left(\frac{m_D((\theta-1)/\varphi)}{\Omega_D} + \frac{m_E}{\Omega_E}\right)x} \right) dx \\
&= B_1 B_2 \left(B_3 B_5 \left(\frac{m_E}{\Omega_E}\right)^{-(m_E+k)} \Gamma\left(m_E+k, \frac{Ip}{P_{\max}} \frac{m_E}{\Omega_E}\right) \right. \\
&\quad - B_4 B_5 \left(\frac{m_D((\theta-1)/\varphi)}{\Omega_D} + \frac{m_E}{\Omega_E}\right)^{-(l_1-t_1+m_E+k)} \Gamma \\
&\quad \left. \cdot \left(l_1 - t_1 + m_E + k, \left(\frac{m_D((\theta-1)/\varphi)}{\Omega_D} + \frac{m_E}{\Omega_E}\right) \frac{Ip}{P_{\max}}\right)\right),
\end{aligned}
\tag{20}
$$

where

$$
B_5 = e^{-\lambda_E{}^2/2} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E+k)}. \tag{21}
$$

4. SPSC Analysis

For the communication network security, SPSC is also one of the important evaluation indicators. SPSC is described as the probability that the secrecy capacity is greater than zero, which can be expressed as [40]

$$
SPSC = P\{C_D - C_E > 0\} = 1 - \int_0^{\infty} F_D(\gamma_E) f_E(\gamma_E) d\gamma_E = 1 - \Psi. \tag{22}
$$

Substituting (6) and (7) into Equation (22), we can get the expression of Ψ as

$$
\begin{aligned}
\Psi &= \mathrm{P}\{\gamma_D \le \gamma_E\} = \int_0^{\infty} F_D(\gamma_E) f_E(\gamma_E) d\gamma_E \\
&= \int_0^{\infty} e^{-\lambda_D{}^2/2} \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{\Gamma(l+1)\Gamma(l+m_D)} (l+m_D-1)! \\
&\quad \cdot \left(1 - e^{-m_D \gamma_E/\Omega_D} \sum_{l_1=0}^{l+m_D-1} \frac{(m_D/\Omega_D)^{l_1} \gamma_E{}^{l_1}}{l_1!}\right) \times e^{-\lambda_E{}^2/2} \\
&\quad \cdot \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E+k)} \gamma_E{}^{m_E+k-1} e^{-m_E/\Omega_E \gamma_E} d\gamma_E \\
&= e^{-\lambda_E{}^2/2} \sum_{k=0}^{\infty} \frac{2^{-k} m_E{}^{m_E+k} \Omega_E{}^{-(m_E+k)} \lambda_E{}^{2k}}{k!\Gamma(m_E+k)} e^{-\lambda_D{}^2/2} \\
&\quad \cdot \sum_{l=0}^{\infty} \frac{\left(\lambda_D/\sqrt{2}\right)^{2l}}{\Gamma(l+1)\Gamma(l+m_D)} (l+m_D-1)! \times \left(\frac{\Gamma(m_E+k)}{(m_E/\Omega_E)^{m_E+k}} \right. \\
&\quad \left. - \sum_{l_1=0}^{l+m_D-1} \frac{(m_D/\Omega_D)^{l_1}}{l_1!} \frac{\Gamma(m_E+k+l_1)}{((m_E/\Omega_E)+(m_D/\Omega_D))^{m_E+k+l_1}} \right).
\end{aligned}
\tag{23}
$$

Combined with (22) and (23), the exact accurate analytical formula of SPSC can be obtained. In (23), with the increase of λ, λ^{2k} and $(\lambda/\sqrt{2})^{2l}$ will increase and lead to

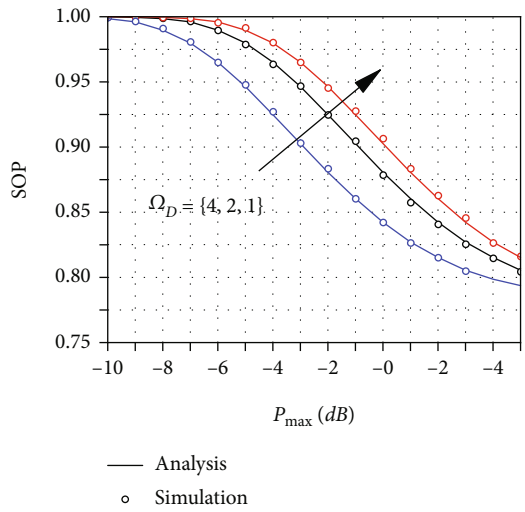


Figure 2: SOP versus $P_{\max}$ for different Ω_D.

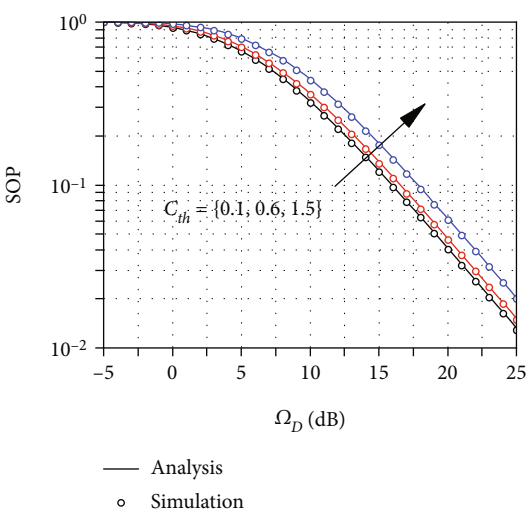


Figure 3: SOP versus Ω_D for different C_{th}.

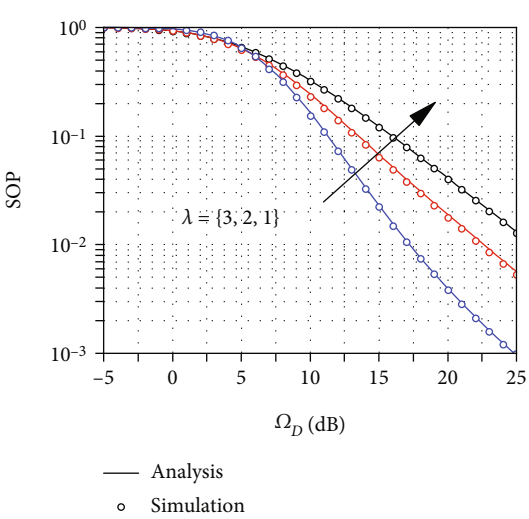


Figure 4: SOP versus Ω_D for different λ.

the improvement of SPSC; however, $e^{-\lambda^2/2}$ will increase and lead to the decline of SPSC.

5. Numerical Analysis

In this section, the simulation results of the inferences are given, and their correctness will be verified by statistical simulations. Furthermore, we analyze the performance through the simulation curves. We implement the random number generator of BX distribution by using the acceptance rejection method. It is worth noting that the SOP and SPSC contain infinite series; however, when the number of loops reaches more than 40 times, the derived formulas for SOP and SPSC will converge to a fixed value.

Figure 2 presents the variation of SOP under the conditions of different $\Omega\{\Omega_D = \Omega_E = (4, 2, 1)\}$. In all simulations, we set $\lambda_D = \lambda_E = 1$, $m_D = m_E = 1$, $Ip = 10$ dB, and $C_{th} = 0.1$ dB . The theoretical curves overlap well with the Monte Carlo simulations. Ω_D is the average SNR of the main channel, the larger Ω_D means that the better transmission quality of the signal. In this situation, the SOP should be smaller. In addition, the value of SOP decreases as $P_{\max}$ rises, which shows that increasing the value of $P_{\max}$ can enhance the confidentiality. And by increasing the Ω_D, the SOP decreases and the security is improved.

Figure 3 illustrates the change of SOP with the average SNR under different C_{th}. The parameter settings are as follows: $\lambda_D = \lambda_E = 1$, $m_D = m_E = 1$, $Ip = 10$ dB, $P_{\max} = 1$ dB, and $\Omega_E = 1$ dB. Figure 3 shows that as Ω_D rises, SOP gradually decreases, indicating that increasing the average SNR of the main channel when the SNR of the eavesdropping channel is constant can enhance security. And with the rise of the set threshold C_{th}, the security will be weakened.

Figure 4 describes the variation of SOP with the average SNR and $\lambda_D = \lambda_E = \{3, 2, 1\}$. Its parameters are set to $m_D = m_E = 1$, $Ip = 5$ dB, $P_{\max} = 1$ dB, $\Omega_E = 1$ dB, and $C_{th} = 0.1$ dB. This figure depicts the effect of the parameter λ of the BX fading on the security, Figure 4 is based on the Equations (9), (10), (12), (13), and (20). We can see that the value of λ has a large transformation on the system performance at $\Omega_D > 5$ dB. In other words, the rise of λ will greatly enhance the security under the condition of high SNR.

Figure 5 plots the change of SPSC with Ω_D and λ. The parameters are $m_D = m_E = 2$, $C_{th} = 0.1$ dB, $Ip = 0.1$ dB, $P_{\max} = 1$ dB, and $\Omega_E = 2$ dB. It can be seen that the change of λ has different effects on security under high and low SNR. When $\Omega_D < 2$ dB, the value of SPSC reduces as λ increases. That is, at low SNR, the increase of λ weakens the security. Analysis of Figures 4 and 5 shows that under high SNR, with the rise of λ, the value of SOP decreases rapidly, and the value of SPSC gradually increases. These conclusions collectively demonstrate that the increase of λ will enhance confidentiality.

Figure 6 depicts the variation of SOP with m and Ω_D. The remaining parameters are set as follows: $\lambda_D = \lambda_E = 1$, $\Omega_E = 1$ dB, $Ip = 5$ dB, $P_{\max} = 1$ dB, and $C_{th} = 0.1$ dB. From Figure 6, the value of SOP reduces rapidly as the m rises for $\Omega_D > 6$ dB, which indicates that increasing m can greatly

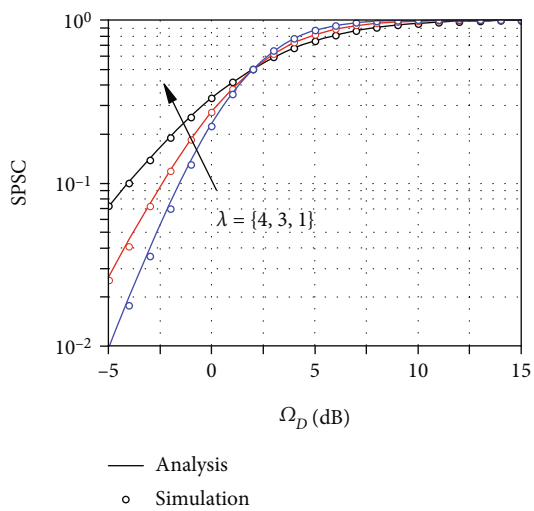


Figure 5: SPSC versus Ω_D for different λ.

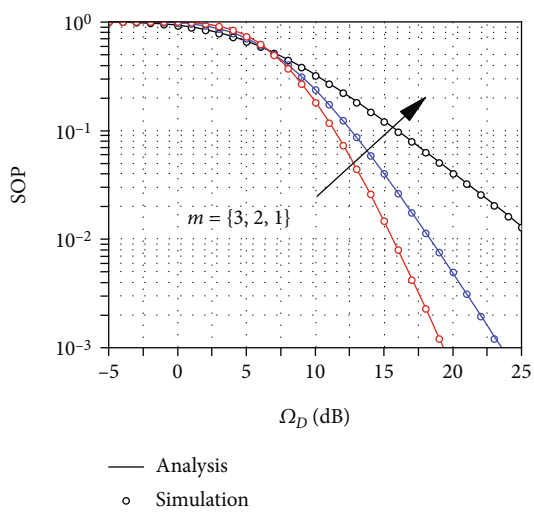


Figure 6: SOP versus Ω_D for different m.

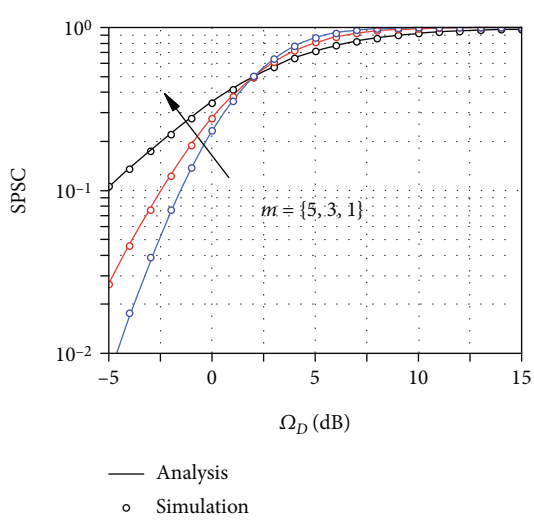


Figure 7: SPSC versus Ω_D for different m.

enhance the security, and when $\Omega_D < 6$ dB, the change of m has little effect on the confidentiality.

Figure 7 shows the variation of SPSC under different m with the average SNR. In Figure 7 we set $\lambda_D = \lambda_E = 2$, $C_{th} = 0.1$ dB, $\Omega_E = 2$ dB, $Ip = 0.1$ dB, and $P_{\max} = 1$ dB. The simulation curves are drawn based on (22) and (23). Furthermore, it can be observed that at low SNR, as m increases, the value of SPSC reduces, which proved that the confidentiality is weakened. When $\Omega_E > 3$ dB, as m adds, the security of the system is improved. This is consistent with the results obtained in Figure 6.

6. Conclusion

In this paper, we investigate the PLS of the CRNs over BX fading channels. The accurate expressions of SOP and SPSC are obtained. Then the impact of the parameters for the BX fading channels which inherits the advantages of both Nakagami-m and Rician fading channels on the security of the CRN is analyzed. Finally, we compared the theoretical derivation with Monte Carlo simulation; the coincidence of simulation curves proves the correctness of the theoretical formulas. And interestingly, rising $P_{\max}$, increasing Ω_D, decreasing C_{th}, and increasing the values of λ and m under high SNR all promote the confidentiality.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest about the publication of this paper.

Acknowledgments

This work was supported in part by the Key Scientific Research Projects of Higher Education Institutions in Henan Province under grant 20A510007 and in part by the Doctoral Fund of Henan Polytechnic University under grant B2022-13.

References

[1] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] A. K. Yerrapragada, T. Eisman, and B. Kelley, "Physical layer security for beyond 5G: ultra secure low latency communications," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 2232–2242, 2021.

[3] L. Qing, H. Guangyao, and F. Xiaomei, "Physical layer security in multi-hop AF relay network based on compressed sensing," *IEEE Communications Letters*, vol. 22, no. 9, pp. 1882–1885, 2018.

[4] Z. Lianbing, "Research on physical layer security of cognitive radio based on cooperative communication," in *Sixth International Conference on Intelligent Systems Design and*

Engineering Applications (ISDEA), vol. 2015, pp. 661–664, IEEE, Guiyang, China, 2015.

[5] L. Li, Y. Hu, H. Zhang, W. Liang, and A. Gao, "Deep learning based physical layer security of D2D underlay cellular network," *China Communications*, vol. 17, no. 2, pp. 93–106, 2020.

[6] A. Mukherjee, "Physical-Layer security in the internet of things: sensing and communication confidentiality under resource constraints," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1747–1761, 2015.

[7] F. Zhu and M. Yao, "Improving physical-layer security for CRNs using SINR-based cooperative beamforming," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1835–1841, 2016.

[8] J. Sun, X. Li, M. Huang, D. Yuan, and J. Jin, "Performance analysis of physical layer security over shadowed fading channels," *IET Communications*, vol. 12, no. 8, pp. 970–975, 2018.

[9] Y. Zhu, Y. Xin, and P. Kam, "Outage probability of Rician fading relay channels," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2648–2652, 2008.

[10] M. Yacoub, "The η-μ distribution: a general fading distribution," in *In Vehicular Technology Conference Fall 2000. IEEE VTS Fall VTC2000*, vol. 2, pp. 872–877, 52nd Vehicular Technology Conference, Boston, MA, USA, 2000.

[11] O. S. Badarneh, D. B. da Costa, M. Benjillali, and M. Alouini, "Selection combining over double α-μ fading channels," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3444–3448, 2020.

[12] Z. Li, L. Jia, F. Li, and H. Hu, "Outage performance analysis in relay-assisted inter-vehicular communications over double-Rayleigh fading channels," in *International Conference on Communications and Mobile Computing,*, vol. 2010, pp. 266–270, IEEE, Shenzhen, China, 2010.

[13] N. Hajri, N. Youssef, and M. Patzold, "A study on the statistical properties of double Hoyt fading channels," in *In 2009 6th International Symposium on Wireless Communication Systems*, pp. 201–205, IEEE, Siena, Italy, 2009.

[14] S. Ö. Ata, "Physical layer security over cascaded Rayleigh fading channels," in *In 2018 26th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, Izmir, Turkey, 2018.

[15] J. Sun, H. Bie, and X. Li, "Security performance analysis of SIMO relay systems over composite fading channels," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 1–6, 2020.

[16] J. Sun, H. Bie, X. Li, K. M. Rabie, and R. Kharel, "Average secrecy capacity of SIMO κ-μ shadowed fading channels with multiple eavesdroppers," in *IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 2020, pp. 1–6, IEEE, Seoul, Korea (South), 2020.

[17] L. Singh and N. Dutta, "Various optimization algorithm used in CRN," in *In 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, pp. 95–100, IEEE, Dubai, United Arab Emirates, 2020.

[18] D. H. Tashman and W. Hamouda, "An overview and future directions on physical-layer security for cognitive radio networks," *IEEE Network*, vol. 35, no. 3, pp. 205–211, 2021.

[19] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, 2009.

[20] S. Mishra and A. Trivedi, "Relay selection with channel allocation for cognitive radio relay channels in CRN," in *Eleventh International Conference on Wireless and Optical Communications Networks (WOCN)*, vol. 2014, pp. 1–4, IEEE, Vijayawada, India, 2014.

[21] T. Zhang, D. Zhang, J. Qiu, X. Zhang, P. Zhao, and C. Gong, "A kind of novel method of power allocation with limited cross-tier interference for CRN," *IEEE Access*, vol. 7, pp. 82571–82583, 2019.

[22] X. Zhang, Z. Yan, Y. Gao, and W. Wang, "On the study of outage performance for cognitive relay networks (CRN) with the nth best-relay selection in Rayleigh-fading channels," *IEEE Wireless Communications Letters*, vol. 2, no. 1, pp. 110–113, 2013.

[23] D. H. Tashman and W. Hamouda, "Physical-layer security on maximal ratio combining for SIMO cognitive radio networks over cascaded κ-μ fading channels," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 4, pp. 1244–1252, 2021.

[24] H. Lei, H. Zhang, I. S. Ansari, G. Pan, and K. A. Qaraqe, "Secrecy outage analysis for SIMO underlay cognitive radio networks over generalized- K fading channels," *IEEE Signal Processing Letters*, vol. 23, no. 8, pp. 1106–1110, 2016.

[25] D. H. Tashman and W. Hamouda, "Physical-layer security for cognitive radio networks over cascaded Rayleigh fading channels," in *In GLOBECOM 2020–2020 IEEE Global Communications Conference*, pp. 1–6, IEEE, Taipei, Taiwan, 2020.

[26] S. Chetry and A. Singh, "Physical layer security of outdated CSI based CRN," in *In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–5, IEEE, Bengaluru, India, 2018.

[27] N. Nandan, S. Majhi, and H. -C. Wu, "Maximizing secrecy capacity of underlay MIMO-CRN through bi-directional zero-forcing beamforming," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5327–5337, 2018.

[28] S. O. Rice, "Statistical properties of a sine wave plus random noise," *The Bell System Technical Journal*, vol. 27, no. 1, pp. 109–157, 1948.

[29] A. Glavieux, P. Y. Cochet, and A. Picart, "Orthogonal frequency division multiplexing with BFSK modulation in frequency selective Rayleigh and Rician fading channels," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1919–1928, 1994.

[30] N. C. Beaulieu and X. Jiandong, "A novel fading model for channels with multiple dominant specular components," *IEEE Wireless Communications Letters*, vol. 4, no. 1, pp. 54–57, 2015.

[31] P. S. Chauhan, S. Kumar, and S. K. Soni, "On the physical layer security over Beaulieu-Xie fading channel," *AEU–International Journal of Electronics and Communications*, vol. 113, article 152940, 2019.

[32] M. Mirahmadi, A. Al-Dweik, and A. Shami, "Interference modeling and performance evaluation of heterogeneous cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 6, pp. 2132–2144, 2014.

[33] H. S. Silva, D. B. T. Almeida, W. J. L. Queiroz, I. E. Fonseca, A. S. R. Oliveira, and F. Madeiro, "Cascaded double beaulieu-xie fading channels," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2133–2136, 2020.

[34] A. Olutayo, H. Ma, J. Cheng, and J. F. Holzman, "Level crossing rate and average fade duration for the Beaulieu-Xie fading

model," *IEEE Wireless Communications Letters*, vol. 6, no. 3, pp. 326–329, 2017.

[35] A. Olutayo, J. Cheng, and J. Holzman, "Asymptotically tight performance bounds for selection diversity over Beaulieu-Xie fading channels with arbitrary correlation," in *IEEE International Conference on Communications (ICC)*, vol. 2017, pp. 1–6, IEEE, Paris, France, 2017.

[36] H. S. Silva, D. B. T. Almeida, W. J. L. Queiroz, I. E. Fonseca, A. S. R. Oliveira, and F. Madeiro, "Outage probability of the product of two Beaulieu–Xie, η-μ, κ-μ, or α-μ random variables," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 12, pp. 2182–2186, 2020.

[37] V. Kansal and S. Singh, "Effective rate analysis of MISO over Beaulieu-Xie fading channel," *AEU-International Journal of Electronics and Communications*, vol. 138, article 153886, 2021.

[38] D. Singh and H. D. Joshi, "Generalized MGF based analysis of line-of-sight plus scatter fading model and its applications to MIMO-OFDM systems," *AEU - International Journal of Electronics and Communications*, vol. 91, pp. 110–117, 2018.

[39] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, Cambridge, MA, USA, 2007.

[40] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *IEEE International Symposium on Information Theory*, pp. 356–360, IEEE, Seattle, WA, USA, 2006.

[41] J. Li, Z. Feng, Z. Feng, and P. Zhang, "A survey of security issues in cognitive radio networks," *China Communications*, vol. 12, no. 3, pp. 132–150, 2015.

[42] H. Lei, H. Zhang, I. S. Ansari et al., "Secrecy outage performance for SIMO underlay cognitive radio systems with generalized selection combining over Nakagami-$ m $ channels," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10126–10132, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8553845, 7 pages
https://doi.org/10.1155/2022/8553845

Research Article

A Novel Modulation Technique with Low PAPR for 5G-IoT Systems

Junchao Sun, Buyu Chen, Ding Huang, Baobing Wang, Tian Xiang, and Dejin Kong

State Key Laboratory of New Textile Materials and Advanced Processing Technologies, The School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan 430200, China

Correspondence should be addressed to Dejin Kong; djkou@wtu.edu.cn

Received 13 May 2022; Accepted 10 July 2022; Published 25 July 2022

Academic Editor: Han Wang

For existing orthogonal frequency division multiplexing (OFDM) in 5G internet of things (5G-IoT) systems, one of the critical problems is the high peak to average power ratio (PAPR), which seriously degrades the energy efficiency. To this end, we propose a novel modulation technique with low PAPR for IoT systems, which preserves the advantage of low implementation complexity and ability to fight against multipath channels. The key methodology is the employment of symbol repetition in the frequency domain. Hence, by designing the appropriate phase factors on the repeated symbols, the PAPR of transmitted signals is equivalent to that of an OFDM signal with a reduced size of discrete Fourier transform (DFT). It is demonstrated that, even if there exists repetitive symbols, the proposed method can still maintain an unreduced spectral efficiency performance. Moreover, to evaluate the proposed method, the Monte-Carlo simulations are conducted for the complementary cumulative distribution function (CCDF) and the bit error rate under multipath fading channels. The simulations show that, as CCDF = 10^{-3}, the proposed method can achieve 2.5 dB gains about the PAPR compared with the original OFDM signal.

1. Introduction

In recent years, 5G-enabled internet of things (5G-IoT) have attracted much attention due to the features of low cost, low power consumption, and wide coverage, which put forward a challenge on the modulation technique. As it is well known, orthogonal frequency division multiplexing (OFDM) is one classical modulation technique and has been widely applied in many communication standards due to the low implementation complexity and ability to fight against multipath channels [1–3]. However, as one of multicarrier modulations, OFDM suffers from the problem of high peak to average power ratio (PAPR), which severely impairs the efficiency of power amplifiers and is unfavourable for the features of low power consumption and wide coverage. Consequently, OFDM may be not an optimal candidate for 5G-IoT systems.

Up to now, various approaches have been proposed to address the PAPR reduction in OFDM systems. These methods are divided into two main categories. One category requires the advance transmission of side information (SI), which costs precious time-frequency resources in control channels generally. For example, the selective mapping (SLM) method was considered in [4], and the partial transmit sequence (PTS) was proposed in [5]. In addition, high complexity is another main drawback for the selected SLM and PTS methods due to the optimal selection from multiple phase sequences. Another category of PAPR reduction methods [6–10] is SI-free approaches. In [6], the PAPR was reduced by directly clipping the transmitted signal, at the cost of degrading the performances of bit error rate (BER) and out-of-band radiation. In [7], the tone reservation method was presented to reduce the PAPR using a peak-canceling signal that generated a set of reserved tones, in which the reservation of tones reduced the spectral efficiency. In [8], an extended SLM was proposed for the PAPR reduction without SI by constructing extension matrices comprising amplitude extensions and phase rotations. In [9], a polar coding-based SI-free SLM scheme was presented by constructing coset codes of a linear polar code, and meanwhile, an adaptive trial-based blind receiver was presented for detection of the selected signal. In [10], the authors proposed a blind interleaving technique with signal space diversity, which reduces the PAPR in OFDM significantly without the need of SI.

In this paper, a novel modulation technique with low PAPR is proposed for 5G-IoT systems, in which the key methodology is the employment of symbol repetition. By designing the appropriate phase factors on the repeated symbols, the PAPR of transmitted signals is equivalent to that of an OFDM signal with a reduced size of discrete Fourier transform (DFT). With the proposed method, the peak power of the transmitted signal is low, without the signal distortion and the need of SI. Furthermore, although the proposed scheme employs the design of symbol repetition, the spectral efficiency can be maintained since the same amount of symbols is transmitted. To evaluate the proposed scheme, simulations are conducted to obtain the curves of the complementary cumulative distribution function (CCDF) and BER under multipath fading channels.

Through this paper, the remainder is organized as follows. In Section 2, the system model of OFDM is introduced. Then, the novel modulation technique is presented for 5G-IoT in Section 3, and the PAPR performance is analyzed in Section 4. Subsequently, simulation results are given in Section 5. Finally, Section 6 gives the conclusion.

2. System Model

In this section, an OFDM system is introduced briefly, in which x_m stands for the transmitted symbol of m-th subcarrier. By considering the signal superposition of M subcarriers, the transmitted signal is obtained and the transmitted signal of the OFDM system is expressed as [11, 12]

$$s(k) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} x_m e^{j2\pi mk/M}, \quad k = 0, 1, \cdots, M-1. \tag{1}$$

Subsequently, to overcome multipath channels, cyclic prefix (CP) is inserted in front of each symbol, which converts the multipath channels into a coefficient for each subcarrier. At the receiver side, after removing CP and carrying out M-point discrete Fourier transform (DFT), symbol demodulation is given as

$$y_m = H_m x_m + \eta_m, \quad m = 0, 1, \cdots, M-1, \tag{2}$$

where H_m is the channel coefficient for the m-th subcarrier and η_m is the Gaussian noise with zero mean and variance σ^2 [13].

Finally, symbol recovery is easily conducted by the single-tap equalization when it is available for the receiver, i.e.,

$$\hat{x}_m = \frac{y_m}{H_m}. \tag{3}$$

Then, we focus on the problem of high PAPR in OFDM, which has a critical influence on the power amplifier efficiency in wireless communications. The definition of PAPR in OFDM is given as

$$\mathrm{PAPR} = \frac{\max\limits_{0 \le k < M} |s(k)|^2}{\mathbb{E}\left[|s(k)|^2\right]}, \tag{4}$$

where $\mathbb{E}[\cdot]$ is the expectation operator. Note that the averaging transmitting power $\mathbb{E}[|s(k)|^2]$ is equivalent to the averaging power of x_m due to the reason that $s(k)$ is the inverse DFT (IDFT) of x_m. When the averaging power of x_m is equal to 1, the value of $\mathbb{E}[|s(k)|^2]$ will be 1. Therefore, the value of PAPR is determined by the fluctuation of the transmitted signal. As is well known, the fluctuation of the signal in OFDM becomes stronger as the subcarrier number grows, especially when the total subcarrier number is limited. As a result, OFDM systems with a larger subcarrier number usually exhibit a larger PAPR.

3. Proposed Modulation Technique with Low PAPR

In this section, a novel modulation technique with low PAPR is proposed for 5G-IoT systems. The key idea is the employment of symbol repetition in the frequency domain. SI is not required, and the low complexity is observed in the proposed method, which makes it preferred in practice. As a remark, the concept of symbol repetition has been also applied in filter-bank multicarrier for the interference cancellation [14, 15].

In the proposed scheme, the whole M subcarriers in OFDM are divided into P groups and each group consists of $R = M/P$ subcarriers. The transmitted symbols in each group are the same; i.e., symbols are repeated in each group. To maintain the same data rate, more branches can be designed to transmit symbols, and the number of branches is the same as the number of the groups.

Let $d_{m,p}$ denote the transmitted symbol, where $m \in [0, M-1]$ is the subcarrier index and $p \in [0, P-1]$ is the branch index. Different p stands for different branch. Figure 1 depicts the transmitter of different branch by changing the value of p.

The transmitted symbols in each group are the same, i.e.,

$$d_{r,p} = d_{nR+r,p}, \tag{5}$$

where $r = 0, 1, \cdots, R-1$ and $p, n \in [0, P-1]$. In other words, the symbols $d_{r,p}$ are transmitted repeatedly in the frequency domain. Note that the original data symbols x_m are divided into P groups, and the p-th group is transmitted via the p-th branch in the proposed scheme; i.e., x_{pR+r} corresponds to $d_{m,p}$. It is worthwhile to point out that p in $d_{m,p}$ is the branch index, not the time index. The symbols $d_{m,p}$ for different p can be simultaneously transmitted. During one symbol duration, total RP, i.e., M, symbols can be transmitted in the proposed scheme, exactly as the classical OFDM systems.

As one multicarrier modulation, the OFDM's orthogonality is very important, and on this basis, no interference among symbols is caused. To avoid interference among symbols, it is still important to maintain orthogonality in the proposed scheme. To maintain the orthogonality between $d_{m,p}$ for different p, transmitted symbols, $d_{m,p}$, are multiplied by the orthogonal phase sequences. For the k-th group, after

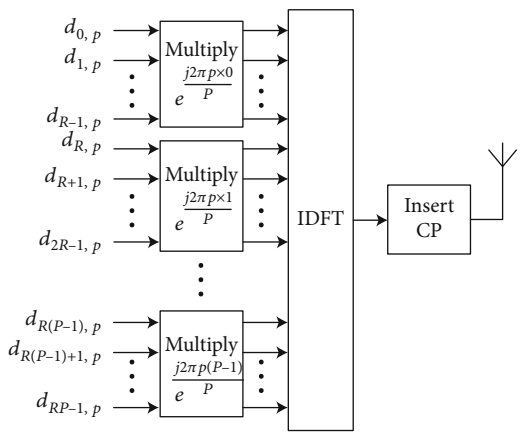


Figure 1: The transmitter of the proposed scheme.

multiplying the phase sequences, we get

$$\bar{d}_{nR+r,p} = d_{nR+r,p} e^{j2\pi np/P}, \tag{6}$$

where $p \in [0, P-1]$. Note that $nR + r$ is the subcarrier index and $nR + r \in [0, M-1]$.

Next, M-point IDFT is performed on $\bar{d}_{kR+r,p}$, which gives

$$s_p(k) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} \bar{d}_{m,p} e^{j2\pi mk/M}, \quad k = 0, 1, \cdots, M-1. \tag{7}$$

Since symbols $d_{m,p}$ for all p are simultaneously transmitted, the baseband transmitted signal can be written as

$$s(k) = \sum_{p=0}^{P-1} s_p(k) = \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} \bar{d}_{m,p} e^{j2\pi mk/M}. \tag{8}$$

Then, similar to the classical OFDM, CP is inserted and the transmitted signal gets through the multipath channel. After the CP removal and M-point DFT at the receiver, the demodulation is obtained:

$$\tilde{d}_m = \sum_{p=0}^{P-1} H_m \bar{d}_{m,p} + \eta_m. \tag{9}$$

Note that H_m stands for the channel coefficient for the m-th subcarrier and η_m satisfies the Gaussian distribution.

Subsequently, the single-tap equalization is employed, i.e.,

$$\hat{d}_m = \frac{\tilde{d}_m}{H_m} = \sum_{p=0}^{P-1} \bar{d}_{m,p} + \frac{\eta_m}{H_m}. \tag{10}$$

Based on (5), it is realized that $d_{r,p} = d_{nR+r,p}$ for $n = 0, 1, \cdots, P-1$. Accordingly, the transmitted symbol $d_{r,p}$ can be recovered as

$$\phi_{r,p} = \sum_{n=0}^{P-1} \hat{d}_{nR+r} e^{j2\pi np/P} = \sum_{n=0}^{P-1} \sum_{p_0=0}^{P-1} \bar{d}_{nR+r,p_0} e^{j2\pi np/P} + \sum_{k=0}^{P-1} \frac{\eta_{nR+r} e^{j2\pi np/P}}{H_{nR+r}}. \tag{11}$$

Following (6), equation (11) is rewritten as

$$\phi_{r,p} = \sum_{n=0}^{P-1} \sum_{p_0=0}^{P-1} d_{nR+r,p_0} e^{j2\pi np_0/P} e^{j2\pi np/P} + \xi_{r,p}, \tag{12}$$

where $\xi_{r,p} = \sum_{n=0}^{P-1} \eta_{nR+r} e^{j2\pi np/P} / H_{nR+r}$.

Due to the orthogonality of phase sequences, the following equation holds:

$$\sum_{n=0}^{P-1} e^{j2\pi np/P} e^{j2\pi np_0/P} = \begin{cases} P, & p = p_0, \\ 0, & p \neq p_0. \end{cases} \tag{13}$$

Therefore, it is obtained as

$$\phi_{r,p} = \sum_{n=0}^{P-1} \hat{d}_{nR+r} e^{j2\pi np/P} = P d_{r,p} + \xi_{r,p}. \tag{14}$$

Then, the transmitted symbol can be recovered by

$$\hat{d}_{r,p} = \frac{\phi_{r,p}}{P}. \tag{15}$$

As a remark, when the number of branches is the same as the number of the groups, the data rate of the proposed scheme is the same as that of the classical OFDM. When it is smaller, the data rate of the proposed scheme is smaller than that of the classical OFDM. When it is larger, the orthogonality disappears and the interference is caused.

As two special cases, when $P = 1$, the proposed scheme will be equivalent to the conventional OFDM systems, and when $P = M$, the transmitter of the proposed scheme will be equivalent to the single-carrier frequency domain equalization (SC-FDE) [16]. Note that, due to the signal superposition, the original OFDM signal suffers a large PAPR value, while the SC-FDE signal exhibits a very small PAPR due to the absence of the signal superposition.

4. PAPR Analysis

In this section, we give the theoretical analysis on the PAPR performance. According to (8), the transmitted signal is given as

$$s(k) = \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} \bar{d}_{m,p} e^{j2\pi mk/M}. \tag{16}$$

Compared with (1), the signal superposition of (16) comes from not only the subcarriers m but also the branches p.

By substituting (6) into (16), the following is obtained:

$$s(k) = \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} \sum_{\alpha=0}^{P-1} d_{\alpha R+r,p} e^{j2\pi\alpha p/P} e^{j2\pi(\alpha R+r)k/M}. \tag{17}$$

Due to the fact that $d_{\alpha R+r,p} = d_{r,p}$ as equation (5), (17) is rewritten as

$$\begin{aligned} s(k) &= \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} \sum_{\alpha=0}^{P-1} d_{r,p} e^{j2\pi\alpha p/P} e^{j2\pi(\alpha R+r)k/M} \\ &= \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} \sum_{\alpha=0}^{P-1} d_{r,p} e^{j2\pi rk/PR} e^{j2\pi\alpha p/P} e^{j2\pi\alpha k/P}. \end{aligned} \tag{18}$$

For the time index k in (18), it is rewritten as

$$k = \beta P + l, \tag{19}$$

where $l = 0, 1, \cdots, R-1$ and $\beta = 0, 1, \cdots, P-1$. Then, (18) is rewritten as

$$\begin{aligned} s(\beta P + l) &= \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} \sum_{\alpha=0}^{P-1} d_{r,p} e^{j2\pi r(\beta P+l)/PR} \times e^{j2\pi\alpha p/P} e^{j2\pi\alpha(\beta P+l)/P} \\ &= \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} \sum_{\alpha=0}^{P-1} d_{r,p} e^{j2\pi r(\beta P+l)/PR} \times e^{j2\pi\alpha p/P} e^{j2\pi\alpha l/P} \\ &= \frac{1}{\sqrt{M}} \sum_{p=0}^{P-1} \sum_{r=0}^{R-1} d_{r,p} e^{j2\pi r(\beta P+l)/PR} \times \sum_{\alpha=0}^{P-1} e^{j2\pi\alpha p/P} e^{j2\pi\alpha l/P}. \end{aligned} \tag{20}$$

According to (13), the following equation holds:

$$\sum_{\alpha=0}^{P-1} e^{j2\pi\alpha p/P} e^{j2\pi\alpha l/P} = \begin{cases} P, & p = l, \\ 0, & p \neq l. \end{cases} \tag{21}$$

Then, (20) can be rewritten as

$$s(\beta P + l) = \frac{P}{\sqrt{M}} \sum_{r=0}^{R-1} d_{r,l} e^{j2\pi r(\beta P+l)/PR}. \tag{22}$$

From (22), it is realized that the signal $s(\beta P + l)$ for a certain l is determined only by $d_{m,l}$. In other words, the signals of $d_{m,l}$ and $d_{m,p}$ are nonoverlapping in the time domain for $p \neq l$. For example, the transmitted signals of $d_{m,0}$ are $s(\beta P)$ with $\beta = 0, 1, \cdots, P-1$, while the transmitted signals of $d_{m,1}$ are $s(\beta P + 1)$ with $\beta = 0, 1, \cdots, P-1$. Furthermore, the amplitude of $s(\beta P + l)$ can be obtained as

$$\begin{aligned} |s(\beta P + l)| &= \frac{P}{\sqrt{M}} \left| \sum_{r=0}^{R-1} d_{r,l} e^{j2\pi r(\beta P+l)/PR} \right| \\ &= \frac{P}{\sqrt{M}} \left| \sum_{r=0}^{R-1} d_{r,l} e^{j2\pi rl/PR} e^{j2\pi r\beta/R} \right|. \end{aligned} \tag{23}$$

Defining $\tilde{d}_{r,l} = d_{r,l} e^{j2\pi rl/PR}$, (23) is rewritten as

$$|s(\beta P + l)| = \frac{P}{\sqrt{M}} \left| \sum_{r=0}^{R-1} \tilde{d}_{r,l} e^{j2\pi r\beta/R} \right|. \tag{24}$$

It is clear that the transmitted signal by the proposed scheme is equivalent to the superposition of R subcarriers, i.e., R-point IDFT. Hence, for an OFDM system with M subcarriers, its PAPR is reduced significantly by the proposed scheme, and specifically, it is equivalent to PAPR of the OFDM system with $R = M/P$ subcarriers.

As a remark, SI is not required in the proposed scheme since the phase sequences are fixed for both the transmitter and receiver, which is different from the conventional SLM method in OFDM systems. In addition, for the conventional SLM method, the optimal phase sequence is chosen by calculating and comparing multiple PAPR values, resulting in the high complexity. However, compared with the classical OFDM, the increase of complexity is mainly from the linear multiplication with phase sequences in the proposed scheme, as equation (6) at the transmitter and equation (12) at the receiver. For both equations (6) and (12), the number of complex-valued multiplications is M. Therefore, the total increased complexity is $2M$ considering both the transmitter and receiver. Note that, for the IDFT operation that can be implemented by IFFT at the transmitter, the complexity is $M \log M$. Compared to the IDFT operation, the increased complexity is acceptable in the proposed scheme due to the reason $M \ll M \log M$.

5. Simulation Results

In this section, simulations are carried out to evaluate the proposed scheme for 5G-IoT, in terms of CCDF and BER. In simulations, the following parameters in 5G NR numerology are employed: subcarrier spacing of 15 kHz, symbol duration of 66.7 μs, CP length of 5.2 μs, and symbol number per time slot of 14. To verify no intersymbol interference in the proposed scheme, no power amplifier is used in the simulations of BER. The detailed parameters of multipath channels are as follows:

(i) Number of paths: 3

(ii) Delay profile (μs): 0, 4, and 10

(iii) Power profile (dB): 0, -5, and -10

Figure 2 depicts BER of the proposed scheme under the additive white Gaussian noise (AWGN) channel and the multi-path channel, respectively. For the proposed scheme, the parameters R and M are set to 16 and 64. The simulation results demonstrate that there is no BER loss with our scheme, which means that the proposed methodology can overcome the multipath fading effectively without the intersymbol interference. It is noted that the existing clipping method suffers obvious BER loss under both the AWGN and multipath channels, which is due to the signal distortion caused by clipping.

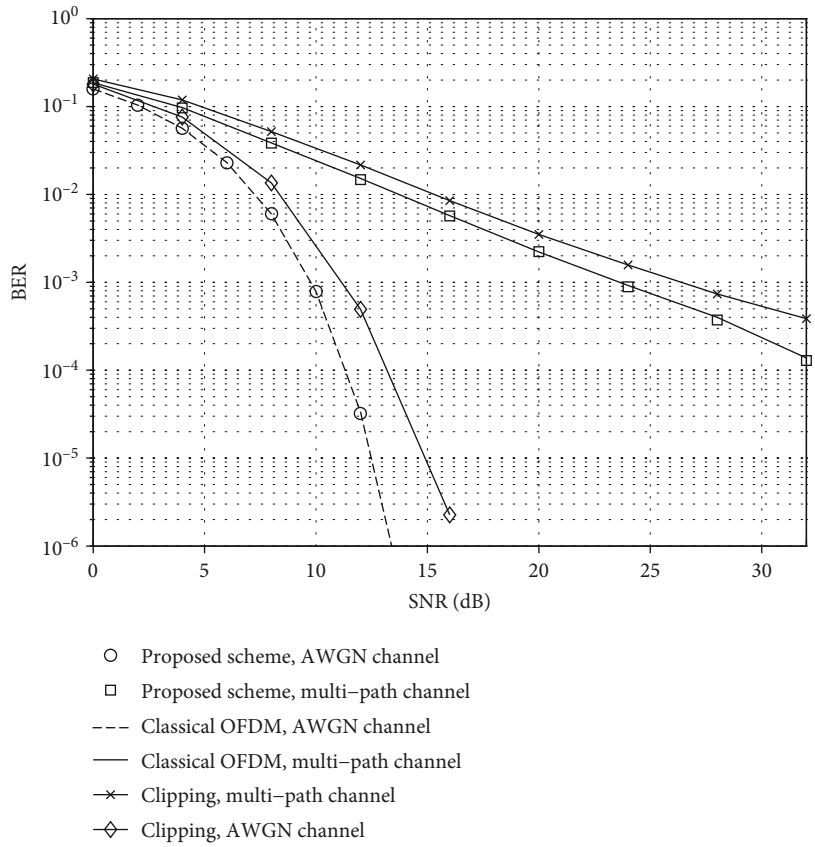


Figure 2: BER of the proposed scheme in OFDM systems for 5G-IoT, $M = 64$, and $R = 16$.

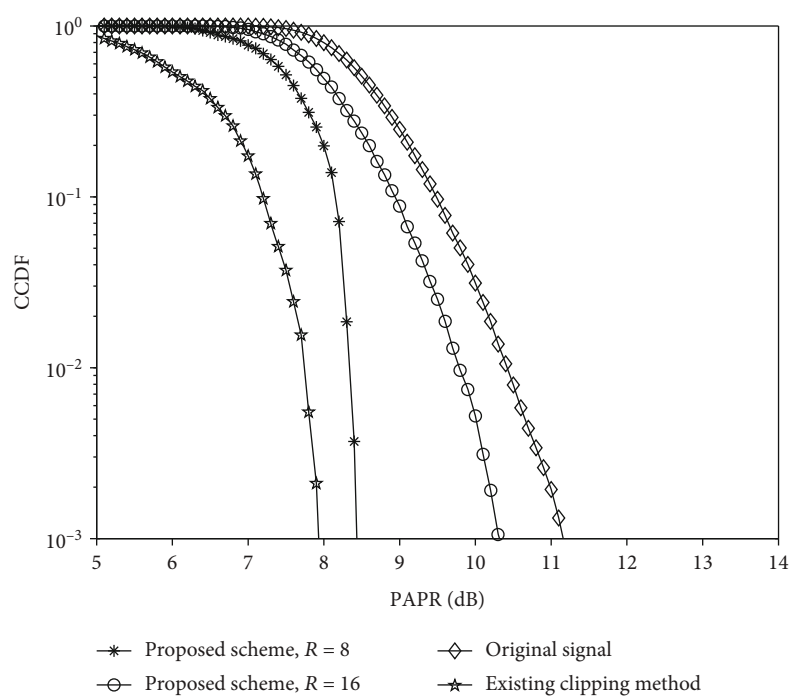


Figure 3: CCDF of the proposed scheme with 4QAM and $M = 64$.

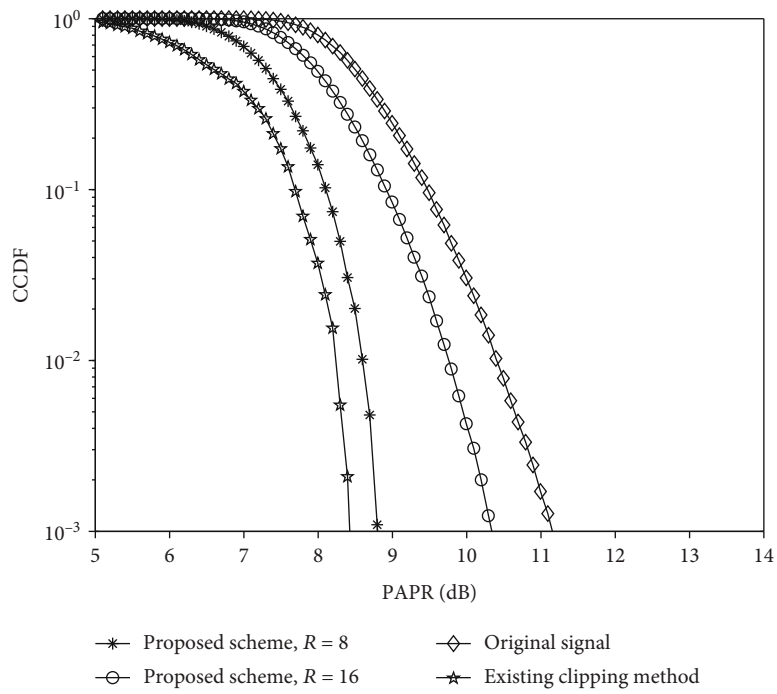


Figure 4: CCDF of the proposed scheme with 16QAM and $M = 64$.

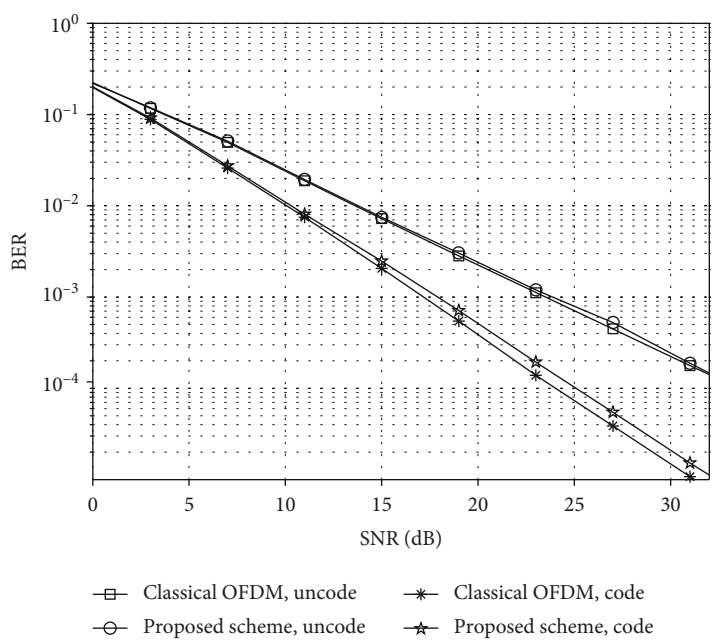


Figure 5: BER of the proposed schemes with channel codes; code rate = 0.5.

As depicted in Figures 3 and 4, the CCDFs of the proposed scheme with 4-QAM (quadrature amplitude modulation) and 16-QAM are, respectively, given under the condition $M = 64$. For comparison, we also give the CCDFs of the existing clipping method, in which no SI is required. Compared with the original signal of the OFDM system, the proposed scheme achieves lower PAPR. Specifically, the proposed scheme with $R = 16$ has about 1 dB gain at CCDF = 10^{-3}, and the proposed scheme with $R = 8$ has about 2.5 dB gain at CCDF = 10^{-3} compared with the original signal. The results are in accordance with the PAPR analysis in Section 4; i.e., the transmitted signal by the proposed scheme is equivalent to the superposition of R subcarriers, i.e., R-point IDFT. Due to the reduction of the IDFT size, PAPR of the OFDM system with M subcarriers can be reduced significantly in the proposed scheme. In addition, it can be observed from Figures 3 and 4 that the existing

clipping method achieves better PAPR performance than the proposed scheme. However, it is noteworthy that, due the signal distortion, the existing clipping method suffers obvious BER loss.

To evaluate the ability of exploiting frequency diversity, Figure 5 shows the BER comparison of uncoded and coded schemes. For comparison, the BER performances of the original OFDM are also given. In simulations, the convolution code with a code rate 0.5 is adopted. From the simulation results, compared with the uncoded scheme, an obvious gain can be achieved in the coded scheme. Therefore, the proposed scheme can also exploit frequency diversity effectively. In addition, it is observed that the coded scheme suffers a bit of BER loss at high SNR compared with the original OFDM. The reason may be that the proposed scheme tends to become a single-carrier modulation as R decreases.

6. Conclusion

In this paper, a novel modulation technique was proposed for 5G-IoT systems by employing the concept of symbol repetition in the frequency domain. It was demonstrated that the proposed scheme significantly reduced the peak power of the transmitted OFDM signal without the signal distortion and SI. Furthermore, it was noted that the number of transmitted symbols could be maintained despite of the design of symbol repetition in the proposed scheme. Numerical simulations have shown the effectiveness of the proposed method.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Junchao Sun and Buyu Chen contributed equally to this work.

Acknowledgments

This work was financially supported in part by the National Science Foundation of China with Grant number 62001333.

References

[1] H. Zhang, N. Liu, X. Chu, K. Long, A. Aghvami, and V. Leung, "Network slicing based 5G and future mobile networks: mobility, resource management, and challenges," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138–145, 2017.

[2] D. Kong, X.-G. Xia, P. Liu, and Q. Zhu, "MMSE channel estimation for two-port demodulation reference signals in new radio," *SCIENCE CHINA Information Sciences*, vol. 64, pp. 169303:1–169303:2, 2021.

[3] D. Kong, Y. Xu, G. Song, J. Li, and T. Jiang, "A CP reduction scheme based on symbol repetition for narrow-band IoT systems," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12880–12891, 2021.

[4] J.-Y. Woo, H. S. Joo, K.-H. Kim, J.-S. No, and D.-J. Shin, "PAPR analysis of class-III SLM scheme based on variance of correlation of alternative OFDM signal sequences," *IEEE Communications Letters*, vol. 19, no. 6, pp. 989–992, 2015.

[5] L. Yang, R. S. Chen, Y. M. Siu, and K. K. Soo, "PAPR reduction of an OFDM signal by use of PTS with low computational complexity," *IEEE Transactions on Broadcasting*, vol. 52, no. 1, pp. 83–86, 2006.

[6] X. Li and L. J. Cimini, "Effects of clipping and filtering on the performance of OFDM," *IEEE Communications Letters*, vol. 2, no. 5, pp. 131–133, 1998.

[7] L. Wang and C. Tellambura, "Analysis of clipping noise and tone-reservation algorithms for peak reduction in OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 3, pp. 1675–1694, 2008.

[8] W. Hu, W. Huang, Y. Ciou, and C. Li, "Reduction of PAPR without side information for SFBC MIMO-OFDM systems," *IEEE Transactions on Broadcasting*, vol. 65, no. 2, pp. 316–325, 2019.

[9] S.-C. Lim, N. Kim, and H. Park, "Polar coding-based selective mapping for PAPR reduction without redundant information transmission," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1621–1625, 2020.

[10] T. Arbi, Z. Ye, and B. Geller, "Low-complexity blind PAPR reduction for OFDM systems with rotated constellations," *IEEE Transactions on Broadcasting*, vol. 67, no. 2, pp. 491–499, 2021.

[11] S. Gopi and S. Kalyani, "An optimized SLM for PAPR reduction in non-coherent OFDM-IM," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1–971, 2020.

[12] H. Wang, P. Xiao, and X. Li, "Channel parameter estimation of mmwave MIMO system in urban traffic scene: a training channel-based method," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–9, 2022.

[13] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2020.

[14] D. Kong, X. Zheng, Y. Zhang, and T. Jiang, "Frame repetition: a solution to imaginary interference cancellation in FBMC/OQAM systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1259–1273, 2020.

[15] D. Kong, D. Qu, and T. Jiang, "Time domain channel estimation for OQAM-OFDM systems: algorithms and performance bounds," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 322–330, 2014.

[16] X. Cheng, Y. Yang, and S. Q. Li, "Joint compensation of transmitter and receiver I/Q imbalances for SC-FDE systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8483–8498, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6983655, 9 pages
https://doi.org/10.1155/2022/6983655

Research Article

Towards Security Mechanism in D2D Wireless Communication: A 5G Network Approach

Divya Gupta,[1] **Shalli Rani**,[2] **Aman Singh**,[3,4] **and Juan Luis Vidal Mazon**[5,6]

[1]*Department of Computer Science and Engineering, Chandigarh University, Mohali 140413, India*
[2]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India*
[3]*Faculty of Engineering, Universidade Internacional Do Cuanza, Estrada Nacional 250, Bairro Kaluapanda, Cuito-Bié, Angola*
[4]*Department of Engineering, Universidad Internacional Iberoamericana, Arecibo 00613, Puerto Rico, USA*
[5]*Higher Polytechnic School, Universidad Europea del Atlántico, C/Isabel Torres 21, 39011 Santander, Spain*
[6]*Department of Project Management, Universidad Internacional Iberoamericana, Campeche 24560, Mexico*

Correspondence should be addressed to Aman Singh; aman.singh@unic.co.ao

Received 4 April 2022; Revised 23 May 2022; Accepted 12 July 2022; Published 22 July 2022

Academic Editor: Han Wang

Device-to-device (D2D) communication has attracted many researchers, cellular operators, and equipment makers as mobile traffic and bandwidth demands have increased. It supports direct communication within devices with no need for any intermediate node and, therefore, offers advantage in 5G network while providing wide cell coverage range and frequency reuse. However, establishing acceptable and secure mechanism for D2D communication which ensures confidentiality, integrity, and availability is an issue encountered in this situation. Furthermore, in a resource-constrained IoT environment, these security challenges are more critical and difficult to mitigate, especially during emergence of IoT with 5G network application scenarios. To address these issues, this paper proposed a security mechanism in 5G network for D2D wireless communication dependent on lightweight modified elliptic curve cryptography (LMECC). The proposed scheme follows a proactive routing protocol to discover services, managing link setup, and for data transfer with the aim to reduce communication overhead during user authentication. The proposed approach has been compared against Diffie–Hellman (DH) and ElGamal (ELG) schemes to evaluate the protocol overhead and security enhancement at network edge. Results proved the outstanding performance of the proposed LMECC for strengthening data secrecy with approximate 13% and 22.5% lower overhead than DH and ELG schemes.

1. Introduction

1.1. Background. Device-to-device (D2D) communication eliminates the need for an intermediary node. In mobile networks, D2D communication offers several advantages [1]: Firstly, it may be used to extend the cell coverage in a cellular network while providing facilities to act as communication bridge for data transmission outside the coverage range. Secondly, by delivering data directly between devices, D2D communication helps to lower the base station's energy usage. Finally, the effectiveness of reusing the same radio frequency has been improved. The distance between devices in D2D communication is very less. This means that in a D2D communication scenario, radio frequency interference is reduced, allowing numerous data to be transmitted on the same radio frequency. Due to these benefits, the D2D communication feature is adopted by long term evolution (LTE) advanced 4G networks. [2].

However, on a mobile network, D2D communication based on proactive routing protocol has certain security

concerns [3]. Device discovery, link setup, and data transfer are the three operations that make up the proactive D2D communication mechanism [4]. There is no authentication method for confirming device identification in this process. Another node responds to a request for a setup link using an acknowledgement message. Furthermore, throughout the communication process, D2D communication does not employ any kind of encryption to provide confidentiality and message authentication for integrity. This implies the attacker can use DoS attacks, eavesdropping, and network spoofing to carry out assaults [5, 6] .

1.2. Motivation. Recently, IoT is coupled with the 5G to meet its service expectations which corresponds to massive machine type communication (mMTC) and ultrareliable low latency communication (URLLC) [7]. IoT applications, on the other hand, deal with a lot of sensitive data, yet IoT devices have restricted performance, memory, and power consumption. Because traditional security solutions cannot be applied or processed effectively, these properties of IoT make the above-mentioned security concerns crucial to get handle. We need a secure system with an appropriate authentication mechanism between devices to tackle the security problems of D2D communication. Furthermore, given the limited resources available, it must be made light.

For resource-constrained devices, lightweight cryptography may be an appropriate option. Elliptic curve cryptography (ECC) is the most common lightweight asymmetric key method. It provides 128-bit cryptographic security with a 256-bit key, which is relatively smaller than the 3072-bit key being used in public key encryption scheme RSA [8].

1.3. Contribution. The following are the key contributions of this study:

(i) This paper presents a secure D2D communication system based on lightweight modified ECC (LMECC) for a 5G IoT network

(ii) The D2D communication system has been established in three phases: service discovery, link setup, and data transfer using a proactive routing protocol

(iii) The proposed security mechanism based on LMECC has been evaluated and compared against Diffie–Hellman and ElGamal security enhancement techniques

(iv) The experiments conducted using MATLAB for overhead analysis and security enhancement proved that LMECC can better manage the 5G IoT devices

The remaining paper is structured as mentioned below: The related work done for security in D2D communication has been discussed in Section 2. Section 3 presents the proposed system model. The D2D communication process with its three operations is presented in Section 4. The existing security enhancement mechanism along with proposed security enhancement in D2D communication has been further explained in Sections 5 and 6. The Section 7 presents the results obtained after performing experimentation on proposed approach. Finally, the summary of the study is presented in the form of conclusion in Section 8.

2. Related Work

One of the most significant data transmission concerns is security. Nowadays, with the emergence of variety of smartphone apps to manage the Internet of things, mobile phone usage has increased. The extensive usage of cell phones in the industry, on the other hand, drew academics' attention to the need of protecting consumers and customers. Many models for encrypting and decrypting data outsourcing have been developed to address these difficulties. However, new dangers continue to emerge as a result of new attack tactics and hostile behavior by adversaries.

The protection of D2D communication data is critical in the face of harmful assaults [14]. Secure D2D connections between mobile devices, on the other hand, remained a problem. The work in [10] suggested a lightweight authentication technique based on ElGamal encryption. This work provides a public key infrastructure (PKI)-based authentication technique that uses a mix of ECC for key pair selection and ElGamal encryption for secret key exchange. Over public key infrastructure, another lightweight cryptography scheme considering both ECC and ElGamal has been presented in [9]. This scheme utilizes ECC for key generation and ElGamal for encryption/decryption of messages. Using smartphone sensor behavior analysis, authors in [11] suggested a D2D authentication technique. For group authentication, their authentication technique uses certificateless cryptography, and for continuous authentication, they use user behavior analysis retrieved from smartphone sensors. Further, authors in [15] designed another key exchange mechanism mainly suitable for LTE-based D2D communication which is extendable for use in 5G network. ECC-based symmetric keys are used to create their method. Similar to this, the work in [12] designed an authentication mechanism along with device detection and privacy protection with use of identity-based encryption. Another work in [13] proposed lightweight multilayer authentication scheme suitable for wireless body area networks (WBAN). To support lightweight authentication with group key design algorithm, this work again used ECC algorithm. The computation performed using Foci calculations ensures low computation cost while providing high security.

The majority of these investigations employ ECC-based cryptographic methods to enable authentication and data confidentiality/integrity. However, they have certain drawbacks, such as the inability of some of the results to give anonymity, or the fact that the studies did not discuss in detail on the data transmission stage of D2D communication (refer Table 1). Furthermore, the majority of current systems rely solely on lightweight public key techniques, rather than lightweight symmetric encryption algorithms. Our suggested system can increase the efficiency and security of D2D communication because it employs the lightweight encryption to handle all of our security concerns and all of the phases in D2D communication.

Table 1: Description of existing work.

Ref	Year	Proposed scheme	Strength	Weakness
[9]	2017	Lightweight cryptography scheme considering both ECC and ElGamal	ECC for key generation and ElGamal for encryption/ decryption of messages	High overhead
[10]	2019	Lightweight authentication technique based on ElGamal encryption	Use a mix of ECC for key pair selection and ElGamal encryption for secret key exchange	No lightweight symmetric encryption
[11]	2019	Authentication technique uses certificateless cryptography	D2D authentication technique	Asymmetric encryption with high overhead
[12]	2020	Authentication mechanism along with device detection and privacy protection with the use of identity-based encryption	Authentication, privacy protection	Weak encryption mechanism
[13]	2021	Lightweight multilayer authentication scheme	ECC algorithm with low computation cost	No lightweight symmetric encryption
Proposed work	2022	Security mechanism in 5G network for D2D wireless communication dependent on lightweight modified elliptic curve cryptography (LMECC)	Lightweight symmetric encryption with low protocol overhead	More focused on 5G D2D communication with/without various security challenges

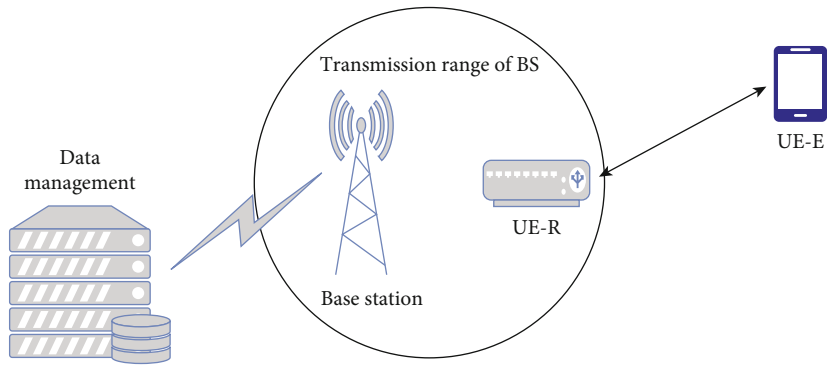


Figure 1: System model scenario.

Table 2: Abbreviation list.

Notations	Definition
DM	Data management
UE	User equipment
BS	Base station
D2D	Device to device
ECC	Elliptic curve cryptography
Non	A random value
*D*sig	Digital signature
PU_k	Public key
PR_k	Private key
key_{sec}	Secret key

3. Proposed System Model

For a 5G IoT network, this section presents a secure D2D communication system model. Figure 1 depicts the suggested secure D2D communication paradigm. 5G network components such as user equipment (UE), base station (BS), and user data management (DM) participate in D2D communication (refer Table 2). A user equipment (UE) is a physical mobile device in direct connection with other devices in proposed model. The base station (BS) connects UE to cellular networks. Within the service region, BS can work with UE-Relay (UE-R). UE-R, like other UEs, must respect the cellular network's function.

Furthermore, before the discovery process can begin, UEs must first register for proximity service discovery and D2D services. In this study, we execute procedures such as registration and authentication for all UEs. Following registration, the proximity service application on each device may begin initiating requests or monitoring the proximity services of other UEs. The BS may help in advertisement so that it is available to all D2D-enabled UEs. We further assume that in the envisioned situation, all UEs and the BS support both open proximity service and network proximity service.

4. D2D Communication Process

This section explains the complete D2D communication establishment process by providing details about service discovery, link setup, and data transfer as these are the three

operations that make up the D2D communication mechanism.

4.1. Service Discovery. The initial phase of secure D2D communication, i.e., service discovery, follows a proactive approach where BS advertises the available service information. Proactive protocol is particularly important for commercial businesses that wish to market themselves and deliver information to their clients [16]. In our approach, just one BS is taken into account for simplicity [3]. UE-R receives service advertisement information from BS and transmits it to UE-E. Any UE-E(s) that are interested should respond to this message. The communication during the device discovery phase is multicast. They do, however, exchange unicast messages after completing the D2D discovery phase. Furthermore, even though UE-E does not require special service information, a D2D connection is established between UE-R and UE-E. The PUSH mechanism is an example of this kind.

The steps involved in discovery of services have been shown in Figure 2 and are explained below:

(i) Service advertisement: Through broadcast messages, BS offers "Service advertisement" to UEs throughout its coverage

(ii) Service invitation: Being a relay device, the UE-R delivers a "Multicast D2D service invitation" to the UE-E in its close vicinity(s)

(iii) Request initialization: UE-E accepts the service invitation by sending a unicast "D2D request initialization" message to UE-R

(iv) Request for approval: UE-R requests D2D approval from BS by transmitting information about UE-E

(v) Checking operations: D2D checking operations for UE-E are represented by steps 5, 6, and 7. In step 5, "Forwards D2D request" is sent by BS to the DM. Following that, DM saves UE-E's information in its database and verifies channel capacity for it. Following this, DM responds to UE-R by sending a "D2D request admission" message via BS

4.2. Link Setup and Data Transfer. After the successful discovery of services and device for D2D communication, the next operation is to setup links and to transfer data. The complete process of link setup and data transfer has been divided in to several steps which have been illustrated in Figure 3.

(i) Step 1: UE-E sends "D2D service selection" to UE-R and, therefore, selects the exact service in which it is interested

(ii) Step 2: The request for service is passed from UE-R to the BS

(iii) The allocation of service from BS to UE-R is represented by steps 3, 4, and 5. It mandates service confirmation from DM, as well as the processing and distribution of the desired service data

(iv) Step 6: UE-R uses unicast messages to transmit information about the "Relay D2D request"

5. Existing Security Enhancement Mechanisms

This section provides knowledge on the existing authentication schemes for security enhancement based on Diffie–Hellman (DH) and ELG schemes. Both the schemes are then compared against proposed proactive service discovery protocol for security enhancement in D2D communication environment in Section 7.

5.1. Diffie–Hellman (DH) Scheme. This approach assigns responsibility to two communicating UEs to generate and use a common shared secret key for message encryption/decryption during communication process. Initially, out of all available numbers such as $1, 2, 3, \cdots, W-1$, where W is a large prime number and is known to both UEs and UE-R and UE-E generate a secret number x and y, respectively. Afterwards, a public key is computed by both UEs, UE-R computes $\mathrm{PU}_k(\mathrm{UE}-R) = h^x \bmod W$ and UE-E computes $\mathrm{PU_k}(\mathrm{UE}-E) = h^y \bmod W$. Here h is a common generator known to both UEs in advance.

Both UEs initiate the process of digital signature as listed in the following steps:

Step 1: UE-E and UE-R generate the random nonce values Non_E and Non_R, respectively.

Step 2: UE-E computes hash value of Non_E, encrypts this hash with its secret number y, and attaches to Non_E to construct digital signature of its own. Therefore, digital signature of UE-E $D_{\mathrm{sig}}(\mathrm{UE}-E)$ is represented as:

$$= [\mathrm{Encrypt}[\mathrm{hash}(\mathrm{Non}_E), y], \mathrm{Non}_E]. \quad (1)$$

Step 3: UE-R computes hash value of Non_R, encrypts this hash with its secret number x, and attaches to Non_R to construct digital signature of its own. Therefore, digital signature of UE-R $D_{\mathrm{sig}}(\mathrm{UE}-R)$ is represented as:

$$= [\mathrm{Encrypt}[\mathrm{hash}(\mathrm{Non}_R), y], \mathrm{Non}_R]. \quad (2)$$

Step 4: UE-E fetches the Non_R from D_{sig} and calculates message as:

$$\mathrm{Mes}_E = \mathrm{Encrypt}[\mathrm{hash}(\mathrm{Non}_R\mathrm{Non}_E), \mathrm{PU}_k(\mathrm{UE}-R). \quad (3)$$

Step 5: UE-R fetches the Non_E from D_{sig} and calculates message as:

$$\mathrm{Mes}_R = \mathrm{Encrypt}[\mathrm{hash}(\mathrm{Non}_E\mathrm{Non}_R), \mathrm{PU}_k(\mathrm{UE}-\mathrm{E}). \quad (4)$$

Both UEs decrypt Mes_R and Mes_E using their private key and verify the $[\mathrm{hash}(\mathrm{Non}_E\mathrm{Non}_R)]$. After successful verification, both UEs agree on formation of common secret key for encryption/decryption of rest of the messages during their communication process in network. The common

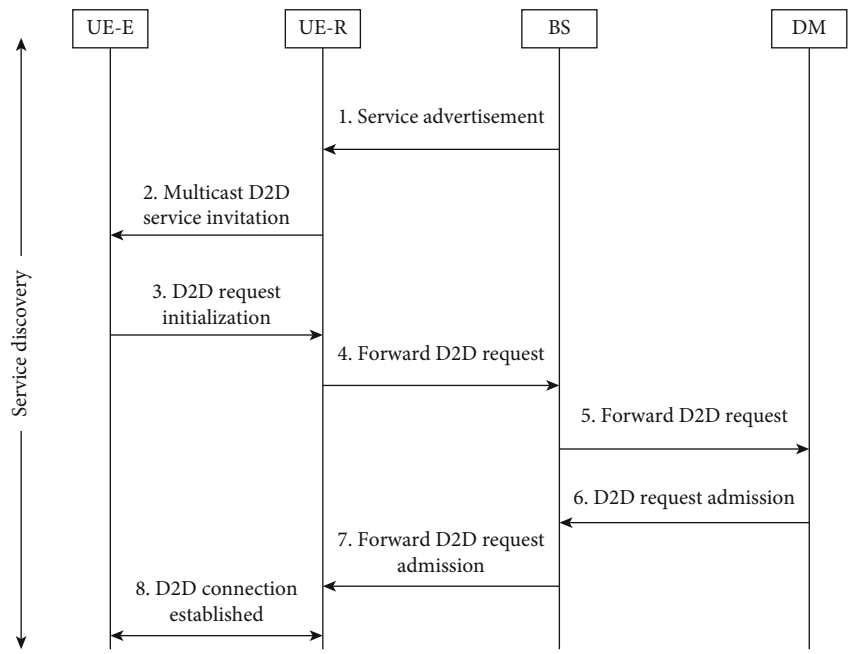


Figure 2: Service discovery process.

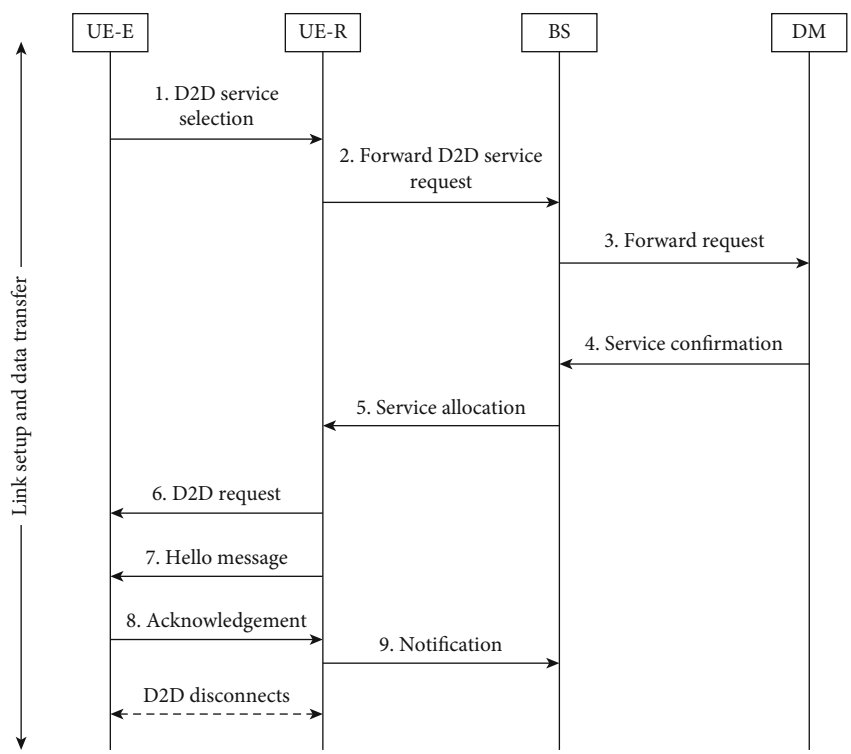


Figure 3: Link setup and data transfer process.

secret key is computed as:

$$\mathrm{Key}_{\mathrm{sec}} = (\mathrm{PU}_k(\mathrm{UE}-R))x = (\mathrm{PU}_k(\mathrm{UE}-E))y. \quad (5)$$

5.2. ELGamal (ELG) Scheme. This scheme allows exchange of secret key on an unsecured channel by users. This key is further used for message encryption [17]. Hence, security in this scheme is solely based on the difficulty lies in solving DH problem. Initially, out of all available numbers such as $1, 2, 3, \cdots, W-1$, where W is a large prime number and is known to both UEs and UE-R and UE-E generates a secret number x and y, respectively. Afterwards, a public key is

computed by both UEs, UE-R computes $\mathrm{PU}_k(\mathrm{UE}-R) = h^x \bmod W$, and UE-E computes $\mathrm{PU}_k(\mathrm{UE}-E) = h^y \bmod W$. Here, h is a common generator known to both UEs in advance.

This scheme makes use of private key and public key for encryption and decryption, respectively.

Initially, UE-E calculates the hash $\mathrm{m} = \mathrm{Hash}(\mathrm{Mes})$ to sign a message such that m is a number in the range from $0 \le m \le W-1$. Further to this, UEs start with the process of digital signature as listed in the following steps:

Step 1: UE-E selects a random number A in such a way that A is relatively prime to $W-1$, and the following conditions hold: $1 \le A \le W-1$ and $\gcd(A, W-1) = 1$

Step 2: UE-E calculates $D_1 = h^A \bmod W$

Step 3: UE-E calculates $A^{-1} \bmod (W-1)$

Step 4: UE-E computes $D_2 = A^{-1}(m - AD_1) \bmod (W-1)$

Step 4: Finally, digital signature consists of $D_{\mathrm{sig}} UE-E = (D_1, D_2)$

The UE-R verifies signature as:

Step 1: UE-R calculates $F_1 = h^m \bmod W$

Step 2: UE-R calculates $F_2 = PU_k(UE-E)^{D_1}(D_1)^{D_2} \bmod W$

The digital signature is valid if $F_1 = F_2$, then UE-R authenticate UE-E.

6. Proposed Security Mechanism in D2D Communication

The proposed mechanism for providing security in D2D communication utilizes lightweight modified elliptic curve cryptography (LMECC). The traditional elliptic curve cryptosystem (ECC) is a cutting-edge lightweight cryptosystem that uses smaller keys than other modern cryptosystems like RSA [2]. As a result, ECC can perform additive finite group operations more effectively than RSA's modular exponentiation process.

ECC follows random selection of private key. Moreover, in case the parameters picked at random are not correctly chosen, wrong calculations will lead to inaccurate plain text formation. The objective is to keep intruders out of the messages sent between UEs. We created a protocol with security enhancements for this purpose. The communication channel between UEs is open; therefore, an attacker could intercept the messages. Due to identity-oriented encryption as well as the LMECC protocol, two UEs in close proximity to each other can discover themselves, setup authentication and key agreement in this part. As shown in Figure 4, on reception of D2D invitation from UE-R, UE-E commences the security enhancement procedure.

LMECC uses asymmetric cryptography, which consists of both private and public keys. The user who is in charge of the private key is responsible for its safety. A shared key generation point PT is agreed upon by two communicating users. Let communicating users UE-R and UE-E's private keys be $\mathrm{Key}_{\mathrm{Pr}_R}$ and $\mathrm{Key}_{\mathrm{Pr}_E}$, respectively. Their public keys are then computed as $\mathrm{Key}_{\mathrm{PU}_R} = \mathrm{Key}_{\mathrm{Pr}_R}.\mathrm{PT}$ and $\mathrm{Key}_{\mathrm{PU}_E} = \mathrm{Key}_{\mathrm{Pr}_E}.\mathrm{PT}$, respectively.

The authentication procedure begins with the selection of domain parameters, followed by computation using LMECC and the Diffie–Hellman key exchange protocol. LMECC is a two-factor authentication system.

(i) Step 1: To pick the elliptic curve parameters, two users are UE-R and UE-E

(ii) Step 2: User UE-R selects PT on the selected elliptic curve and transmits it to user UE-E

(iii) Step 3: UE-R chooses the private key, $\mathrm{Key}_{\mathrm{Pr}_R}$ to be kept with him

(iv) Step 4: $\mathrm{Key}_{\mathrm{Pr}_R}$ generates the public key, which is then sent to UE-E

(v) Step 5: The private key $\mathrm{Key}_{\mathrm{Pr}_E}$ is chosen by UE-E and kept by him

(vi) Step 6: The public key after generation is forwarded to the UE-R

(vii) Step 7: UE-R computes the last verification point, $K_{\mathrm{UE}_R} = \mathrm{Key}_{\mathrm{Pr}_R}.\mathrm{Key}_{\mathrm{PU}_E}$

(viii) Step 8: User UE − E calculates the final verification point as follows: $K_{\mathrm{UE}_E} = \mathrm{Key}_{\mathrm{Pr}_E}.\mathrm{Key}_{\mathrm{PU}_R}$

(ix) Step 9: The concept of a shared secret key is implemented as:-
$K_{\mathrm{UE}_R} = \mathrm{Key}_{\mathrm{Pr}_R}.\mathrm{Key}_{\mathrm{PU}_E} = \mathrm{Key}_{\mathrm{Pr}_R}.\mathrm{Key}_{\mathrm{Pr}_E}.\mathrm{PT} = K_{\mathrm{UE}_E}$

7. Results and Discussions

The performance of the proposed LMECC is being evaluated for overhead analysis using MATLAB simulation environment. The considered scenario consists of 100 devices uniformly distributed in a $100\,\mathrm{m} \times 100\,\mathrm{m}$ cell. Inside the multicast group, including all user equipments, a portion of devices is served according to proposed approach, while those in worst channel conditions receive data via D2D connections. A bandwidth of 20 MHz with 100 RBs is available. The results retrieved from overhead calculation will decide the suitability of security scheme for D2D communication.

7.1. Overhead Analysis. The amount of overhead associated with the proposed design is measured as the count of service discovery messages needed to establish a D2D session between two users. We are assuming total Q UE-Es scattered randomly inside area A and at a distance of P from UE-R. Only q UE-E(s) desire to communicate with UE-R via D2D, and suppose $R \le Q$ establish D2D pairs, requiring proximate service (ProSe) discovery. For every D2D setup based on LMECC scheme, 18 handshakes are required. In addition, the BS sends a multicast message to all UEs on a regular basis, resulting in a total of $(2 + 18R)$ handshakes for R D2D pairs. On the other hand, Diffie–Hellman protocol for security requires 20 handshakes for every D2D establishment, giving a total of $(2 + 20R)$ handshakes for R D2D

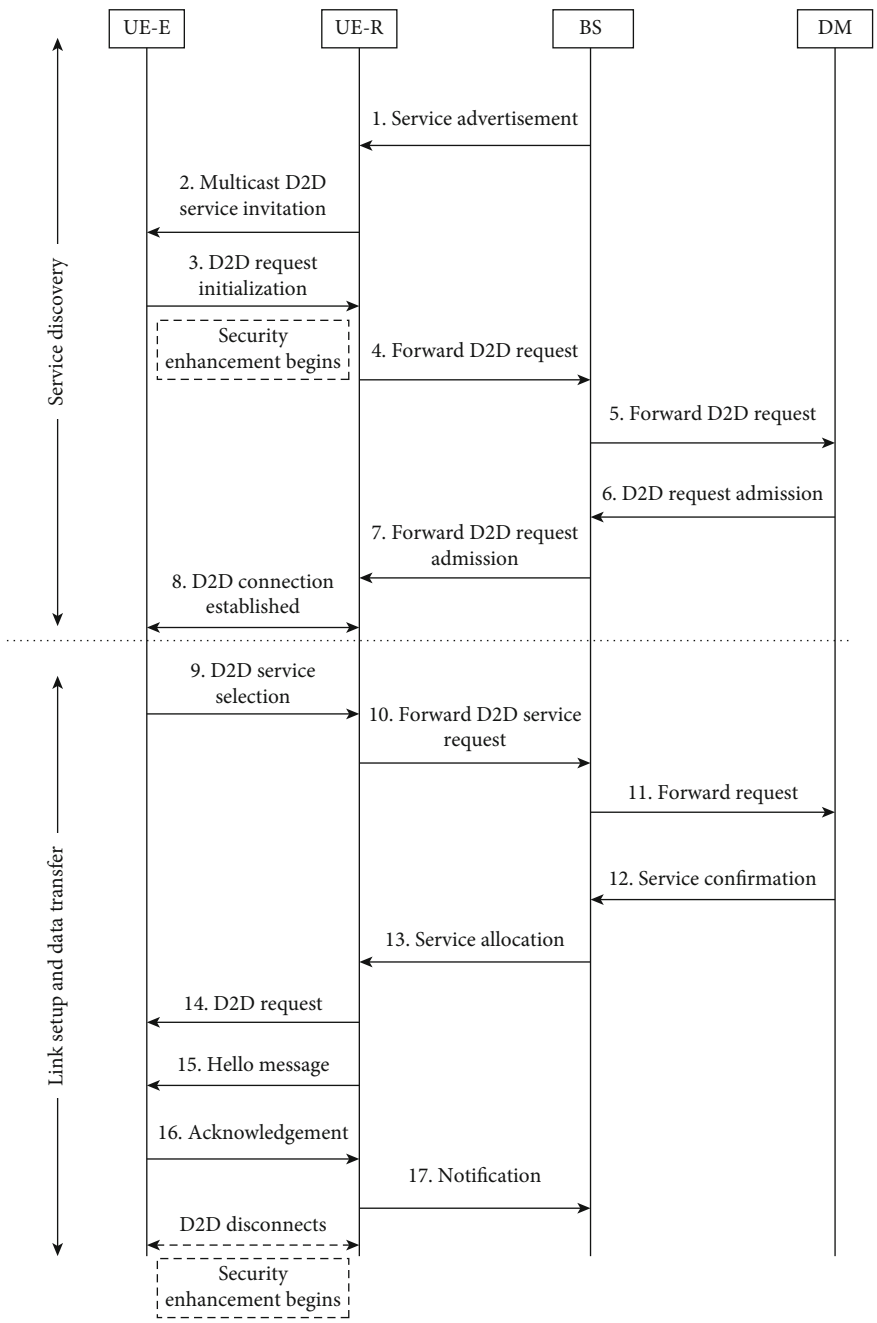


Figure 4: Proposed security in D2D communication.

Table 3: Simulation parameters.

Parameters	Values
Total UEs	20
Total timeslots	20
D2D request per timeslot	6
Number of participating UEs	15

pairs, and the ElGamal protocol requires 22 handshakes for every D2D establishment, giving a total of $(2 + 22R)$ handshakes.

For the proactive procedure, we examine two scenarios when estimating control overhead. In scenario I, the amount of D2D requests in a single timeslot is assumed to be the same across all timeslots. In scenario II, we assume that the quantity of D2D requests changes in each time slot.

Scenario I: Each time slot receives the same number of requests.

This scenario deals with the situation where each timeslot has the equal amount of D2D requests. Considering a scenario in which each timeslot's totality of device to device

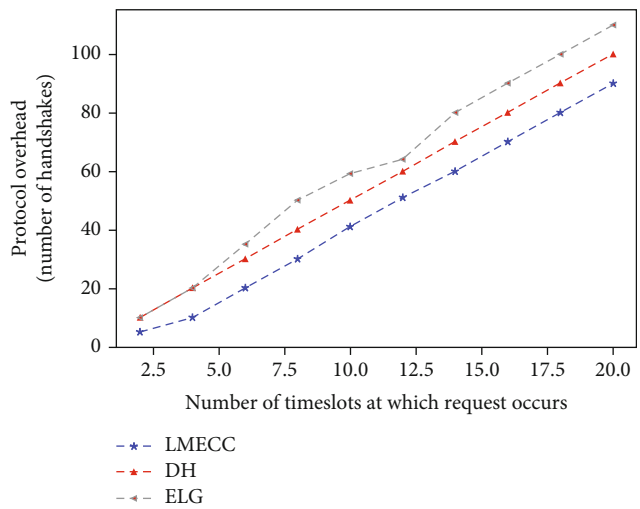


Figure 5: Protocol overhead when $R = 6$.

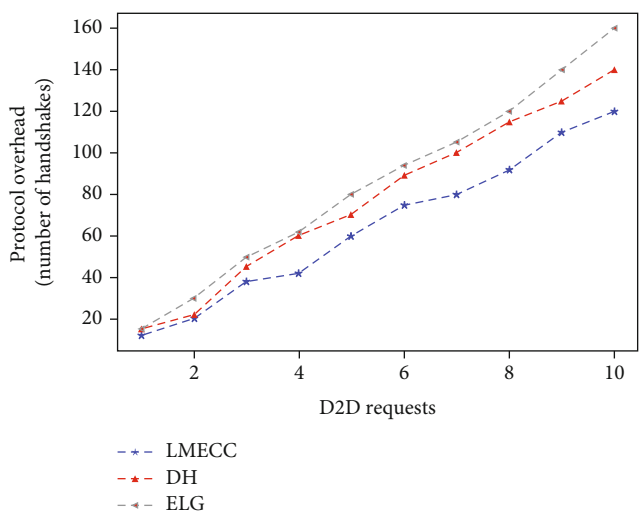


Figure 6: Protocol overhead vs D2D requests.

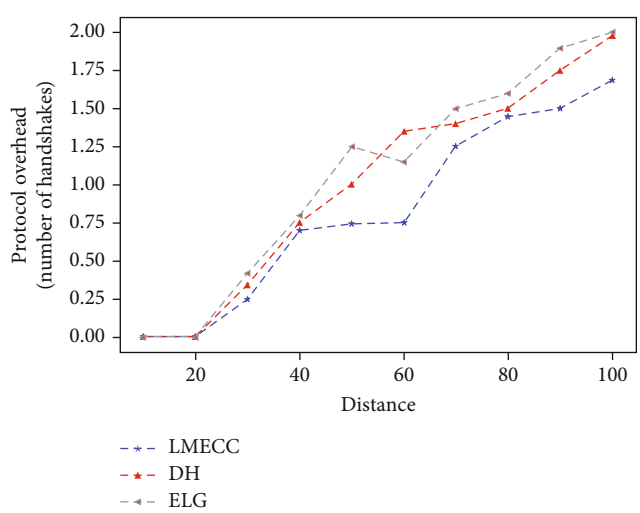


Figure 7: Protocol overhead vs distance.

requests equals to one, and another case, where each timeslot's totality of device to device requests becomes multitudinous, the number of D2D requests is considered to be six for the calculation of the second criterion. The proactive overhead is determined as follows:

$$O_{\mathrm{LMECC}} = \frac{T'(2 + (18 * R)) + \left(2 * \left(T - T'\right)\right)}{T},$$
$$O_{\mathrm{DH}} = \frac{T'(2 + (20 * R)) + \left(2 * \left(T - T'\right)\right)}{T}, \quad (6)$$
$$O_{\mathrm{ELG}} = \frac{T'(2 + (22 * R)) + \left(2 * \left(T - T'\right)\right)}{T}.$$

The parameters chosen for calculating the control overhead in scenario 1 are listed in Table 3.

As shown in Figure 5, when the number of D2D requests is six, it is evident that LMECC is the preferable option since the protocol overhead with LMECC is lower than ELG or DH cryptography, which has a higher overhead. When compared to traditional techniques like ELG or DH, elliptic curve cryptography keys are considerably small and provide equivalent security. At a given number of D2D requests to be 25, this method outperforms the DH scheme by 10.8 percent and the ELG scheme by 23.7 percent for protocol control overhead.

Scenario II: D2D requests appear at random. This scenario represents the occurrence of D2D requests at random in each time frame. As shown in Figure 6, when there is no D2D request, all three forms of proactive protocols have the same overhead. With the growing number of requests, the proactive protocol using LMECC ensures comparatively lesser overhead than using ELG or DH cryptography. Protocol overhead is decreased by 9.37 percent when using the LMECC scheme, and by 18.9 percent when using the ELG scheme, when the number of D2D requests is set to 10.

The number of UE-Es grows as the goal distance increases (refer Figure 7). There are more D2D requests when there are more UE-Es. In comparison to ELG or DH cryptography, LMECC perform better since they have less overhead. If there are multiple UE-E(s) involved in D2D communication, the LMECC is chosen; otherwise, ELG or DH cryptography is used. At a target distance of 100 m, the LMECC scheme reduces protocol overhead by 14.7 percent compared to the DH scheme and by 26.47 percent compared to the ELG scheme.

8. Conclusion

In this study, a proposal is presented to enhance the security in D2D communication networks by leveraging a proactive protocol. To accomplish this, the communication system in D2D environment has been setup in three phases such as service discovery, link setup, and data transfer. The security to the communication in D2D network has been provided through proposed lightweight modified ECC (LMECC) security enhancement scheme. The security mechanism has

been made light to meet the requirements of IoT device's limited resources availability.

In order to evaluate the performance of the proposed protocol, a simulation campaign has been conducted by using the Matlab tool. The performance of the proposed LMECC security enhancement scheme is compared to that of the DH and ELG schemes. The control overhead with the proposed LMECC security enhancement is modest, according to the results. Results proved the outstanding performance of the proposed LMECC for strengthening data secrecy with approximate 13% and 22.5% lower overhead than DH and ELG schemes. Therefore, the proposed approach can be utilized to increase the secrecy and robustness of service discovery in D2D networks in a variety of scenarios. Furthermore, the security of any communication can be enhanced by ensuring confidentiality, integrity, authentication, and availability of message transmission. The listed security parameters can be evaluated to compute the performance of the proposed security enhanced D2D communication approach. As a future work, we tend to implement the proposed approach for this variety of parameters to calculate its security measure. This could be made possible by enabling security and reliability through OFDM-SIS algorithm based on URLLC.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] M. Haus, M. Waqas, A. Y. Ding, Y. Li, S. Tarkoma, and J. Ott, "Security and privacy in device-to-device (D2D) communication: a review," *IEEE Communication Surveys and Tutorials*, vol. 19, no. 2, pp. 1054–1079, 2017.

[2] B. Seok, J. C. S. Sicato, T. Erzhena, C. Xuan, Y. Pan, and J. H. Park, "Secure D2D communication for 5G IoT network based on lightweight cryptography," *Applied Sciences*, vol. 10, no. 1, pp. 217–231, 2020.

[3] D. Fang, Y. Qian, and R. Q. Hu, "Security for 5G mobile wireless networks," *IEEE Access*, vol. 6, pp. 4850–4874, 2018.

[4] A. Hussein, S. El-Rabaie, and M. G. El-Mashed, "Proactive discovery protocol with security enhancement for D2D communication system," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 5047–5066, 2021.

[5] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, 2014.

[6] D. Samanta, A. H. Alahmadi, M. P. Karthikeyan et al., "Cipher block chaining support vector machine for secured decentralized cloud enabled intelligent IoT architecture," *IEEE Access*, vol. 9, pp. 98013–98025, 2021.

[7] S. K. Singh, M. M. Salim, J. Cha, Y. Pan, and J. H. Park, "Machine learning-based network sub-slicing framework in a sustainable 5g environment," *Sustainability*, vol. 12, no. 15, pp. 6250–6272, 2020.

[8] W. Stallings, *Cryptography and Network Security, 4/E*, Pearson Education India, 2006.

[9] Y. Javed, A. S. Khan, A. Qahar, and J. Abdullah, "EEoP: a lightweight security scheme over PKI in D2D cellular networks," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 9, no. 3-11, pp. 99–105, 2017.

[10] A. Abro, Z. Deng, and K. A. Memon, "A lightweight elliptic-Elgamal-based authentication scheme for secure device-to-device communication," *Future Internet*, vol. 11, no. 5, pp. 108–120, 2019.

[11] H. Tan, Y. Song, S. Xuan, S. Pan, and I. Chung, "Secure D2D group authentication employing smartphone sensor behavior analysis," *Symmetry*, vol. 11, no. 8, pp. 969–980, 2019.

[12] Y. Sun, J. Cao, M. Ma, H. Li, B. Niu, and F. Li, "Privacy-preserving device discovery and authentication scheme for D2D communication in 3GPP 5G Het net," in *In 2019 International Conference on Computing, Networking and Communications (ICNC), IEEE*, pp. 425–431, Honolulu, HI, USA, 2019, February.

[13] P. Das, A. Vashisth, D. Chadha, S. Ananda Kumar, A. Banerjee, and S. Shiaeles, "LIMAP: a lightweight multilayer authentication protocol for WBAN," *Wireless Personal Communications*, vol. 121, no. 4, pp. 2857–2884, 2021.

[14] S. Rani, D. Gupta, S. Garg, M. Jalilpiran, and M. S. Hossain, "Consumer electronic devices: evolution and edge security solutions," *IEEE Consumer Electronics Magazine*, vol. 11, no. 2, pp. 15–20, 2022.

[15] S. B. M. Baskaran and G. Raja, "A lightweight incognito key exchange mechanism for LTE-A assisted D2D communication," in *2017 ninth international conference on advanced computing (ICoAC)*, pp. 301–307, IEEE, 2017.

[16] S. K. Singh, Y. Pan, and J. H. Park, "OTS scheme based secure architecture for energy-efficient iot in edge infrastructure," *CMC-COMPUTERS MATERIALS CONTINUA*, vol. 66, no. 3, pp. 2905–2922, 2021.

[17] G. Zheng, B. Gong, and Y. Zhang, "Dynamic network security mechanism based on trust management in wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2021, 10 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1932708, 14 pages
https://doi.org/10.1155/2022/1932708

Research Article

Uplink Multiple Access for Reconfigurable Intelligent Surface-Aided Wireless Systems

Nhan Duc Nguyen,[1] **Chi-Bao Le**,[2] **and Munyaradzi Munochiveyi**[3]

[1]*Faculty of Mechanical-Electrical and Computer Engineering, School of Engineering and Technology, Van Lang University, 69/68 Dang Thuy Tram Street, Ward 13, Binh Thanh District, Ho Chi Minh City 70000, Vietnam*
[2]*Faculty of Electronics Technology, Industrial University of Ho Chi Minh City (IUH), Ho Chi Minh City, Vietnam*
[3]*Electrical and Electronic Engineering Department, University of Zimbabwe, Mount Pleasant, Harare, Zimbabwe*

Correspondence should be addressed to Munyaradzi Munochiveyi; mmunochiveyi@eng.uz.ac.zw

Received 1 March 2022; Revised 18 April 2022; Accepted 1 July 2022; Published 21 July 2022

Academic Editor: A.H. Alamoodi

The integration of reconfigurable intelligent surface- (RIS-) aided wireless communication and multiple access is an attractive and promising scheme for next-generation wireless networks. In this research work, separate uplink RIS-aided nonorthogonal multiple access (NOMA) and uplink relay-assisted NOMA schemes are studied, where the RIS and relay devices are deployed to enhance the coverage of an obstructed single-antenna far user by assisting it to communicate with a single-antenna base station. In each scenario, both perfect successive interference cancellation (pSIC) and imperfect successive interference cancellation (ipSIC) operations are considered in the proposed multiple access network. To characterize the system performance, the associated residual interference caused by ipSIC and relay loop self-interference is characterized using the Rayleigh fading model; subsequently, new channel statistics are derived based on the Gauss-Laguerre polynomial. Consequently, the closed-form approximate outage probability expressions are derived for each scenario in the high signal-to-noise ratio (SNR) regime. To gain further insight, the system throughput in the delay-limited transmission is also obtained for each scenario. The formulated expressions are validated via Monte-Carlo simulations. Finally, the obtained simulation results demonstrate and validate the superiority of the RIS system over the relay device under several system parameters of interest despite the limitation of ipSIC.

1. Introduction

The prospect of large low-power machine-type communications (mMTC), very low-delay communication, Tbps data rates, massive connectivity, and other extraordinary services and applications in 5G and beyond networks has created renewed research interest in new multiple access schemes [1, 2]. In 5G and beyond, multiple access technologies require a redesign due to the combination of low-energy Internet of Things (IoT) and massive connectivity [2]. Existing carrier sense multiple access (CSMA) and noncontention orthogonal multiple access (OMA) techniques found in 1G-4G cellular networks do not scale well in massive connectivity scenarios where a massive number of devices are aimed at accessing a single base station (BS) simultaneously [2]. Thus, new multiple access techniques that enhance scale and reliability need to be considered [2].

Recently, nonorthogonal multiple access (NOMA) is a promising approach to address the aforementioned challenges by serving many users within the same resource block [3–5]. Several recent research works such as [6–9] have demonstrated via numerical results the significant enhancement obtained by NOMA over OMA under certain channel conditions. However, this is not necessarily the case as the authors in [6] showed that at low signal-to-noise ratios (SNRs) in finite blocklength schemes, OMA-based users with good channel conditions outperform users relying on NOMA, albeit, at high SNRs, the reverse is true as the overall link-layer rate of NOMA outperforms OMA when the delay exponent is loose. In [7], the authors investigated

cooperative NOMA networks by considering generic $\alpha - \mu$ fading channel under the impact of residual transceiver hardware impairments (RTHIs). To be practical, imperfect channel state information (CSI) and imperfect successive interference cancellation (SIC) are taken into account. More particularly, two representative NOMA scenarios are proposed, namely, noncooperative NOMA and cooperative NOMA. The authors in [8] investigated the physical layer security (PLS) of the ambient backscatter NOMA systems to raise problems of reliability and security. They considered the realistic scenarios of channel estimation errors (CEEs) and imperfect successive interference cancellation (ipSIC) and residual hardware impairments (RHIs) and how these parameters affect system performance. In [9], the authors considered downlink NOMA over fading channels under full and partial channel state information at the transmitter (CSIT). Furthermore, the authors showed that NOMA is more effective than OMA in terms of average sum rate while maintaining user fairness in fading channels.

There are two ways of implementing NOMA which are power domain or code domain [3–5]. In this article, we focus on power-domain NOMA (PD-NOMA). In PD-NOMA, the users' receiver has additional complexity from the successive interference cancellation (SIC) technique employed to mitigate interference among users [3–5]. However, as indicated in [10], despite PD-NOMA achieving user fairness via power allocation, this improvement in the performance of users with poor channel conditions may end up impacting users with good channel conditions. The integration of reconfigurable intelligent surfaces (RISs)—also known as intelligent reflecting surfaces (IRS)—and PD-NOMA [11] has been proposed as a way to address this disadvantage of PD-NOMA. RIS devices are made up of massive passive/active low-cost reflecting elements capable of smartly configuring the phase shift of incident electromagnetic signals via a RIS controller [11]. Hence, RIS-assisted NOMA systems can provide additional channel paths which are useful in providing PD-NOMA gain in unfavorable scenarios such as when the channel strengths of the far and near user are similar or mutually orthogonal as in downlink multiple-input-single-output (MISO) networks [10]. In such scenarios, NOMA may not be beneficial or offer many advantages over OMA [12]. Another benefit of the additional channel paths provided by RIS is in blockage impacted communication such as in mmWave networks [13] where mmWave communications are vulnerable to blockages caused by vehicles, buildings, pedestrians, and trees. Further, in [14], the authors demonstrated via simulation results that the proposed downlink RIS-aided mmWave-NOMA system outperforms the traditional mmWave-NOMA without RIS in terms of achievable sum rate.

1.1. Related Works on Downlink RIS-NOMA Systems. Recently, research into RIS-aided NOMA communication systems has gained considerable attention. However, most of the research works are focused on the performance of downlink transmission in RIS-NOMA systems. In [15], the authors studied a downlink multicluster MISO RIS-NOMA system with a joint beamforming design. In this system, the authors demonstrated that using second-order cone programming- (SOCP-) based algorithm obtains a locally optimal solution for the total power minimization problem subject to base station beamforming matrices and RIS reflection amplitude and phase shift constraints. This enables the proposed RIS-NOMA system with 32 RIS elements to achieve a 2.5 dB gain over traditional MIMO systems with 64 transmit antennas. In [16], the authors demonstrated that RIS devices are a promising technology for improving NOMA systems in downlink RIS-assisted multicell NOMA networks, as RIS enables blocked users to transmit information. Using stochastic geometry, the authors were able to show the benefits of integrating RIS with NOMA by deriving closed-form and asymptotic coverage probability expressions for paired NOMA users. Simulation results proved that networks aided by RIS significantly outperform conventional networks without RIS. Additionally, the SIC order can be modified as RISs can alter the channel quality of NOMA users.

Differently, in [17], the authors considered the downlink of RIS-aided cooperative nonorthogonal multiple access (C-NOMA) systems. Here, the authors via simulation results showed that the RIS-assisted C-NOMA systems empowered by iterative penalty function-based semidefinite programming (SDP) and the successive refinement approach-based algorithm outperformed traditional C-NOMA systems without RIS. Similarly, in [18], the authors considered the downlink transmission RIS-assisted C-NOMA system with half-duplex (HD) and full-duplex (FD) relay modes under consideration. Here, the system is comprised of a BS, two NOMA users, and a RIS device. The motivation of the authors was to minimize the total transmit power of the BS and user-cooperating relay by jointly optimizing the BS power allocation coefficients, the relay user transmit power coefficient, and the RIS passive beamforming while subject to power budget, SIC, and the cellular user minimum required quality of service (QoS). The authors proposed an alternating optimization method to solve the highly coupled optimization variables. Simulation results showed that RIS introduces total transmit power gain in C-NOMA networks.

Moreover, in terms of artificial intelligence (AI) empowered downlink RIS systems, the authors in [19] investigated the benefits of deep learning (DL) and reinforcement learning (RL) approaches in empowering both RIS-NOMA and RIS-OMA multiuser downlink communication systems over fading channels. The authors also considered the time overhead required for configuring the RIS elements at the start of each fading channel. Moreover, the authors maximized the effective throughput by jointly optimizing the RIS phase shift and the access point (AP) power allocation for each channel block. Numerical results highlighted that for the proposed scheme, NOMA achieves about 42% gain over OMA; additionally, RL and DL agents achieve similar performance in Rician channels; however, in Rayleigh channels, RL is the best. However, we should refer to uplink to look at how NOMA works under the case RIS is enabled as the consideration in the next section.

1.2. Related Works on Uplink RIS-NOMA Systems. In terms of uplink RIS-NOMA systems, there have been several recent research works written on this topic. In [20], the authors considered the benefits of integrating RIS into UAV-assisted multiuser NOMA networks to enhance energy efficiency. The authors also proposed a single-leader multiple-follower Stackelberg Game to jointly design the UAV received signal strength and NOMA users' achieved energy efficiency. Simulation results demonstrated significant energy efficiency with the introduction of RIS. Differently, in [21], the authors studied the spectral and energy efficiency of RIS-NOMA-enabled simultaneous wireless information and power transfer (SWIPT) techniques in the Internet of Things (IoT) networks. The authors also utilized the single-leader multiple-follower Stackelberg Game to jointly optimize the RIS phase shifts, IoT devices' downlink allocated powers, and their harvested energy. Numerical results demonstrated the gain obtained by the introduction of RIS to the NOMA-enabled SWIPT network of IoT devices.

Furthermore, in [22], the authors considered security concerns when the design of RIS at downlink exhibits beneficial to security improvement by jointly designing the transceivers and RIS phase shifts. The authors also proposed performance analysis to confirm the role of the number of metasurface in RIS for security improvement. Numerical results demonstrated that in terms of total transmit power, the framework outdid different frameworks. Differently, the authors in [23] considered the uplink and how the system deals with downlink and uplink when we have different requirements of services at both sides. Simulation results showed the benefits of the proposed scheme in enhancing the performance at base station and mobile users.

On the other hand, in [24], the authors considered both downlink and uplink IRS-assisted NOMA and OMA networks. Here, the IRS is deployed to assist cell-edge users to communicate with a BS. The authors characterized the system performance by investigating new channel statistics of the BS-IRS-user links under Nakagami-m fading conditions. Furthermore, the authors also formulated outage probability and ergodic capacity closed-form expressions for each scenario. For further insights, the authors also obtained the diversity order and high SNR slope based on the approximations in the high-SNR regime. Simulation results showed that IRS reflecting elements and Nakagami-m fading parameters impacted the diversity order but did not impact the high-SNR slope. Numerical results showed that the IRS is superior to FD decode-and-forward relays.

1.3. Motivations and Contributions. Motivated by the above benefits of the addition of RIS in uplink NOMA networks, this article considers the impact on outage probability when RIS is introduced into uplink NOMA systems and explores the benefits of RIS under blocking, different from the works in [20–24] which lacked a detailed study on imperfect SIC (ipSIC) operation in uplink RIS-aided NOMA.

In this article, we consider the scenario of both perfect SIC (pSIC) and ipSIC uplink RIS-aided NOMA and uplink relay-assisted NOMA loop self-interference scenarios, where the RIS and relay devices are deployed to enhance the coverage of an obstructed far user by assisting it to communicate with a BS. Furthermore, to characterize the system performance of each scenario, we adopt the Rayleigh fading model for both the residual interference caused by ipSIC and relay loop self-interference. Afterward, we formulate new channel statistics of these interferences based on the Gauss-Laguerre integration. Subsequently, the closed-form outage expressions are determined for each scenario. To gain further insight, the system throughput in the delay-limited transmission is also obtained for each scenario.

Our main contributions can be written as follows:

(i) We provide an analytical study on outage probability and throughput of RIS-aided uplink multiple access transmissions under blocking conditions

(ii) To confirm the strong contribution of RIS in enhancing the performance of a distant user-facing obstruction, we compare it against traditional relaying technology facing similar obstruction conditions as seen in Figures 1 and 2

(iii) We consider the scenario of both pSIC and ipSIC operations in the NOMA system. To characterize the system performance, we adopt the Rayleigh fading model for both the residual interference caused by ipSIC and relay loop self-interference. Subsequently, new channel statistics of these interferences based on the Gauss-Laguerre integration are derived

(iv) We focus on outage probability and throughput for the above two mentioned practical situations. To further provide insights into such RIS and relay-aided systems, we derive closed-form approximations for both RIS and relay schemes at a high SNR utilizing the Gauss-Laguerre parameter $T = 40$ to yield close approximate equations. The derived equations are validated by Monte Carlo simulations

(v) Finally, we analyze and compare the outage probability and throughput expressions of the two schemes. In particular, we find that ipSIC and relay loop self-interference are the main limitations on both system outage and throughput performance. Furthermore, we observe that based on these impacts, an error floor is imposed on the system outage probability and conversely, a ceiling is placed on the system throughput. Despite these limitations, our numerical results showed that the proposed RIS-aided NOMA scheme can still outperform the relay-aided scheme in the high-SNR regime

1.4. Organization. The rest of this paper is organized as follows: Section 2 presents the system models and channel statistics of the RIS-aided and relay-aided systems in the presence of a blockage. In Section 3, we analyze the outage probability of the proposed uplink systems. Section 4 provides the closed-form outage probability approximations of both systems as well as the throughput performance. Section 5 provides numerical simulations. Section 6 concludes the paper.

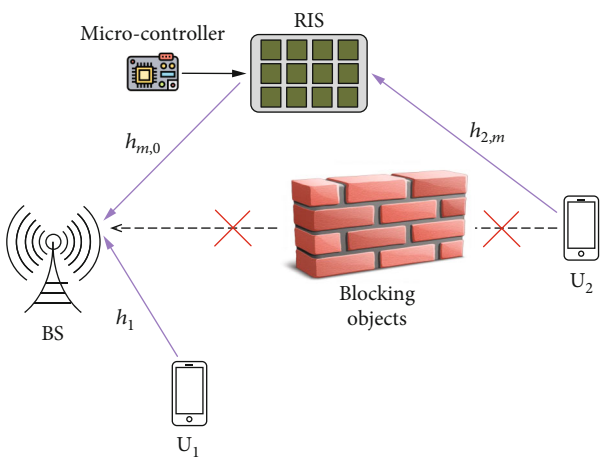


Figure 1: Uplink RIS-aided multiple access systems under blocking.

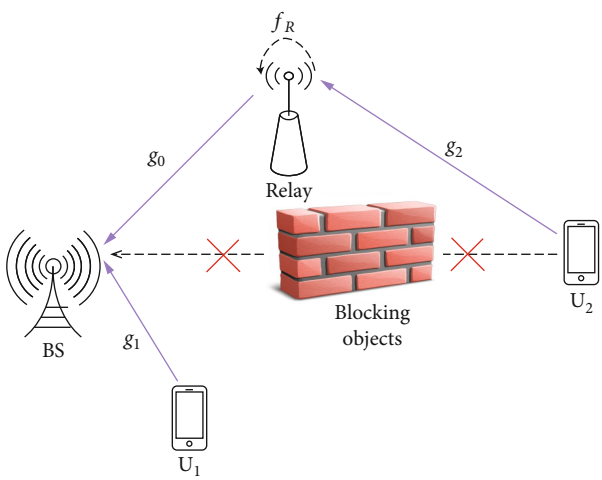


Figure 2: Uplink relay-aided multiple access systems under blocking.

2. System Models and the Channel Models

2.1. Scheme I: Uplink RIS. Figure 1 shows the uplink RIS-aided multiple access systems under blocking. The proposed system consists of a single-antenna BS, two NOMA single-antenna users—U_1 and U_2—a RIS device, and a blocking object obstructing the direct transmission between the far-user U_2 and the BS. The following transmit power constraint is imposed at BS in [24]

$$\bar{y}_S = \sqrt{P_U} h_1 \bar{x}_1 + \sqrt{P_U}\left(\sum_{m=1}^{M} h_{m,0} h_{2,m} \varepsilon_m\right)\bar{x}_2 + \bar{\omega}_S, \quad (1)$$

where P_U denotes each of the two users' transmit power; $\varepsilon_m = \varpi_m(\phi_m)e^{j\phi_m}$ is the RIS reflection coefficient produced by the mth reflector; $\varpi_m(\phi_m) = 1$ is the ideal phase shifts $(m = 1, 2, 3, \cdots, M)$; $\bar{x}_1$ and $\bar{x}_2$ denote the U_1 and U_2 transmitted signals, respectively; $\bar{\omega}_S$ denotes the AWGN at the BS with variance N_0; and $\phi_m \in [0, 2\pi)$ is the RIS phase-shift variable of the mth element $\forall m \in M$. Moreover, we have $h_1 = \bar{h}_1/\sqrt{d_1^\alpha}$, $h_{2,m} = \bar{h}_{2,m}e^{-j\vartheta_m}/\sqrt{d_2^\alpha}$, and $h_{m,0} = \bar{h}_{m,0}e^{-j\theta_m}/\sqrt{d_0^\alpha}$. According to [25], we assume that RIS knows perfectly ϑ_m and θ_m.

The BS decodes U_1 first by regarding the signal from U_2 as interference, and the corresponding signal to interference and noise ratio (SINR) is written as

$$\bar{\gamma}_{S,x_1} = \frac{P_U d_1^{-\alpha}\left|\bar{h}_1\right|^2}{N_0 + P_U d_0^{-\alpha} d_2^{-\alpha}\left|\sum_{m=1}^{M} e^{j(\phi_m - \vartheta_m - \theta_m)}\bar{h}_{m,0}\bar{h}_{2,m}\right|^2}. \quad (2)$$

As in [26], we obtain the maximal SINR using $\phi_m = \vartheta_m + \theta_m$. Then, we can rewrite (2) as

$$\bar{\gamma}_{S,x_1} = \frac{\rho d_1^{-\alpha} A^2}{1 + \rho d_0^{-\alpha} d_2^{-\alpha} B^2}, \quad (3)$$

where $\rho = P_U/N_0$ is the transmit SNR of each two users, $A\Delta = |\bar{h}_1|$, and $B\Delta = \sum_{m=1}^{M}|\bar{h}_{m,0}||\bar{h}_{2,m}|$ with $\bar{h}_{m,0} \sim CN(0, \lambda_0)$, $\bar{h}_{1,m} \sim CN(0, \lambda_1)$, and $\bar{h}_{2,m} \sim CN(0, \lambda_2)$.

After implementing SIC, U_2's signal is detected with the SNR given by

$$\bar{\gamma}_{S,x_2}^{\text{ipSIC}} = \frac{d_0^{-\alpha} d_2^{-\alpha} \rho B^2}{1 + \bar{\ell}\rho|h_I|^2}, \quad \text{with } 0 < \bar{\ell} \le 1, \quad (4a)$$

$$\bar{\gamma}_{S,x_2}^{\text{pSIC}} = \rho d_0^{-\alpha} d_2^{-\alpha} B^2, \quad \text{with } \bar{\ell} = 0, \quad (4b)$$

where $\bar{\ell} = 0$ and $\bar{\ell} \le 1$ represent the pSIC and ipSIC operations. We assume that the residual ipSIC interference is characterized as Rayleigh fading, with $h_I \sim CN(0, \lambda_I)$ being the associated complex channel coefficient with λ_I representing the level of residual interference caused by ipSIC in [27].

2.2. Scheme II: Uplink Relay. Figure 2 depicts the uplink relay-aided multiple access systems under blocking. The proposed system consists of a single-antenna BS, two NOMA single-antenna users—U_1 and U_2, a relay device, and a blocking object obstructing the direct transmission between the far-user U_2 and the BS. The following transmit power constraint is imposed at the BS.

$$\bar{y}_R = \frac{g_2}{\sqrt{d_2^\alpha}}\sqrt{P_2}\bar{x}_{U_2} + f_R\sqrt{P_R}\bar{x}_R + \bar{\omega}_R, \quad (5)$$

where P_2 and P_R are the transmission powers for U_2 and relay, respectively, and $\bar{\omega}_R$ denotes the AWGN with mean power N_0. Moreover, the loop self-interference is expressed as a Rayleigh fading channel with coefficient f_R, and λ_R is the corresponding average power. For simplicity, it could be assumed that $P_2 = P_R$ as well as $\rho = P_2/N_0 = P_R/N_0$. Hence, the received SINR for $\bar{x}_{U_2}$ at the relay is given by

$$\bar{\gamma}_R = \frac{d_2^{-\alpha}\rho|g_2|^2}{\rho|f_R|^2 + 1}. \quad (6)$$

At the BS, the received signal is written as

$$\bar{y}_S = \frac{g_1}{\sqrt{d_1^{\alpha}}}\sqrt{P_{U_1}}\bar{x}_{U_1} + \frac{g_0}{\sqrt{d_0^{\alpha}}}\sqrt{P_R}\bar{x}_R + \bar{\omega}_S, \tag{7}$$

where P_{U1} is the normalized transmission power at U_1 and $\bar{\omega}_S$ stands for AWGN with mean power N_0. Similarly, we assume $P_{U_1} = P_R$ and $\rho = P_{U_1}/N_0$.

Therefore, the received SINR at the BS is given by

$$\bar{\gamma}_{S,U_1} = \frac{d_1^{-\alpha}\rho|g_1|^2}{d_0^{-\alpha}\rho|g_0|^2 + 1}. \tag{8}$$

Next, the BS cancels the signal U_1 by utilizing SIC. Hence, the received SNR at the BS is given by

$$\bar{\gamma}_{S,R}^{\text{ipSIC}} = \frac{d_0^{-\alpha}\rho|g_0|^2}{\bar{\ell}\rho|h_I|^2 + 1}, \quad \text{with } 0 < \bar{\ell} \le 1, \tag{9a}$$

$$\bar{\gamma}_{S,R}^{\text{pSIC}} = d_0^{-\alpha}\rho|g_0|^2, \quad \text{with } \bar{\ell} = 0. \tag{9b}$$

2.3. Channel Statistics. Before computing outage probability, the probability distribution function (PDF) of channel $\bar{Z}$ in which $\bar{Z} \in \{h_1, g_0, g_1, g_2, f_R, h_I\}$, $f_{|\bar{Z}|^2}(x)$, is given by [28].

$$f_{|\bar{Z}|^2}(x) = \frac{1}{\lambda_{\bar{Z}}}e^{-(x/\lambda_{\bar{Z}})}. \tag{10}$$

In terms of the corresponding cumulative distribution function (CDF) of channel h_p, $F_{|h_{\bar{Z}}|^2}(x)$ is given as

$$F_{|h_{\bar{Z}}|^2}(x) = 1 - e^{-(x/\lambda_{\bar{Z}})}. \tag{11}$$

Based on [29], the PDF and CDF of the cascade channel gain from U_2 to RIS and then to the BS are

$$f_{B^2}(x) = \frac{2x^{(M-1)/2}}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^{M+1}}K_{M-1}\left(2\sqrt{\frac{x}{\lambda_0\lambda_2}}\right), \tag{12}$$

$$F_{B^2}(x) = 1 - \frac{2x^{M/2}}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^{M}}K_{M}\left(2\sqrt{\frac{x}{\lambda_0\lambda_2}}\right). \tag{13}$$

3. Outage Probability

3.1. Scheme I: Performance Analysis for Uplink Transmission with RIS

3.1.1. Outage Probability of U_1. The outage probability of U_1 in the uplink phase can be given as

$$\begin{aligned} P_1 &= 1 - \Pr\left(\bar{\gamma}_{S,x_1} > \varepsilon_1\right) = 1 - \Pr\left(\frac{\rho d_1^{-\alpha}A^2}{1 + \rho d_0^{-\alpha}d_2^{-\alpha}B^2} > \varepsilon_1\right) \\ &= 1 - \Pr\left(A^2 > \beta B^2 + \theta_1\right) \\ &= 1 - \int_0^{\infty} f_{B^2}(x)[1 - F_{A^2}(\beta x + \theta_1)]dx, \end{aligned} \tag{14}$$

where $\beta = (\varepsilon_1 d_0^{-\alpha}d_2^{-\alpha})/d_1^{-\alpha}$, $\theta_1 = \varepsilon_1/\rho d_1^{-\alpha}$, and $\varepsilon_i = 2^{R_i} - 1$, $i \in \{1, 2\}$, with R_i being the target rate at U_i.

Using (11) and (12) into (14), P_1 is calculated as

$$\begin{aligned} P_1 &= 1 - \int_0^{\infty} f_{B^2}(x)[1 - F_{A^2}(\beta x + \theta_1)]dx \\ &= 1 - \frac{2e^{-(\theta_1/\lambda_1)}}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^{M+1}}\int_0^{\infty} e^{-(\beta x/\lambda_1)}x^{(M-1)/2} \\ &\quad \cdot K_{M-1}\left(2\sqrt{\frac{x}{\lambda_0\lambda_2}}\right)dx. \end{aligned} \tag{15}$$

Let $t = \sqrt{x} \longrightarrow t^2 = x \longrightarrow 2tdt = dx$; then, P_1 can be rewritten as

$$P_1 = 1 - \frac{4e^{-(\theta_1/\lambda_1)}}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^{M+1}}\int_0^{\infty} e^{-(\beta/\lambda_1)t^2}t^M K_{M-1}\left(2\sqrt{\frac{1}{\lambda_0\lambda_2}}t\right)dt. \tag{16}$$

Using Equation (6.631.3), P_1 is given as

$$P_1 = 1 - \frac{e^{(\lambda_1/2\beta\lambda_0\lambda_2)-(\theta_1/\lambda_1)}}{\left(\sqrt{\beta\lambda_1^{-1}\lambda_0\lambda_2}\right)^{M}}W_{-M/2,(M-1)/2}\left(\frac{\lambda_1}{\beta\lambda_0\lambda_2}\right), \tag{17}$$

where $W_{a,b}(c)$ is a Whittaker function.

3.1.2. Outage Probability of U_2. We assume that in the first time slot, the BS successfully processes messages $\bar{x}_1$ at U_1.

If $0 < \bar{\ell} < 1$, P_2^{ipSIC} can be formulated by

$$P_2^{\text{ipSIC}} = \Pr\left(\bar{\gamma}_{S,x_2}^{\text{ipSIC}} < \varepsilon_2\right) = 1 - \Pr\left(\bar{\gamma}_{S,x_2}^{\text{ipSIC}} > \varepsilon_2\right). \tag{18}$$

Proposition 1. *The closed-form approximation of outage probability at user U_2 can be given by*

$$P_2^{ipSIC} \approx 1 - \frac{1}{\Gamma(M)}\sum_{t=1}^{T} H_t G_{0,2}^{2,0}\left(\frac{\varphi\lambda_I q_t + \theta_2}{\lambda_0\lambda_2}\middle|\overline{M, 0}\right). \tag{19}$$

Proof. The details are given in the appendix.

Similarly, if $\bar{\ell} = 0$, P_2^{pSIC} can be first transformed into

$$\begin{aligned} P_2^{\text{pSIC}} &= 1 - \Pr\left(\bar{\gamma}_{S,x_2}^{\text{pSIC}} > \varepsilon_2\right) = 1 - \Pr\left(B^2 > \frac{\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\rho}\right) \\ &= F_{B^2}\left(\frac{\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\rho}\right). \end{aligned} \tag{20}$$

□

Based on formula (13), The outage probability of U_2 with ipSIC for the uplink is presented in the following:

$$P_2^{\mathrm{pSIC}} = 1 - \frac{2}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^M}\left(\frac{\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\rho}\right)^{M/2} \times K_M\left(2\sqrt{\frac{\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\lambda_0\lambda_2\rho}}\right). \tag{21}$$

3.2. Scheme II: Performance Analysis for Uplink Transmission with Relay

3.2.1. Outage Probability of U_1. The outage probability of U_1 in the uplink phase can be given as

$$\begin{aligned} P_1 &= 1 - \Pr\left(\bar{\gamma}_R > \varepsilon_2, \bar{\gamma}_{S,U_1} > \varepsilon_1\right) \\ &= 1 - \underbrace{\Pr\left(|g_2|^2 > \frac{\varepsilon_2}{d_2^{-\alpha}}|f_R|^2 + \frac{\varepsilon_2}{d_2^{-\alpha}\rho}\right)}_{A_1} \\ &\quad \times \underbrace{\Pr\left(|g_1|^2 > \frac{d_0^{-\alpha}\varepsilon_1}{d_1^{-\alpha}}|g_0|^2 + \frac{\varepsilon_1}{d_1^{-\alpha}\rho}\right)}_{A_2}. \end{aligned} \tag{22}$$

Furthermore, substituting (10) and (11) to (22), A_1 can be calculated as

$$\begin{aligned} A_1 &= \Pr\left(|g_2|^2 > \frac{\varepsilon_2}{d_2^{-\alpha}}|f_R|^2 + \frac{\varepsilon_2}{d_2^{-\alpha}\rho}\right) \\ &= \int_0^\infty f_{|f_R|^2}(x)\left[1 - F_{|g_2|^2}\left(\frac{\varepsilon_2}{d_2^{-\alpha}}x + \frac{\varepsilon_2}{d_2^{-\alpha}\rho}\right)\right]dx \\ &= \frac{e^{-\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\rho}}{\lambda_{f_R}}\int_0^\infty e^{-\left(\left(1/\lambda_{f_R}\right)+\left(\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\right)\right)x}dx \\ &= \frac{\lambda_{g_2}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)}e^{-\left(\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\rho\right)}. \end{aligned} \tag{23}$$

Same as A_1 and after some algebraic manipulations, A_2 is calculated as

$$A_2 = \frac{\lambda_{g_1}d_1^{-\alpha}}{\left(\lambda_{g_1}d_1^{-\alpha} + \lambda_{g_0}d_0^{-\alpha}\varepsilon_1\right)}e^{-\left(\varepsilon_1/\lambda_{g_1}d_1^{-\alpha}\rho\right)}. \tag{24}$$

Substituting (23) and (24) into (22), we can obtain the outage probability of U_1 by

$$P_1 = 1 - \frac{\lambda_{g_1}\lambda_{g_2}d_1^{-\alpha}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)\left(\lambda_{g_1}d_1^{-\alpha} + \lambda_{g_0}d_0^{-\alpha}\varepsilon_1\right)} \times e^{-\left(\varepsilon_1/\lambda_{g_1}d_1^{-\alpha}\rho\right)-\left(\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\rho\right)}. \tag{25}$$

3.2.2. Outage Probability of U_2. Based on (9a), if $0 < \bar{\ell} < 1$, then the outage probability of U_2 for uplink NOMA with ipSIC is given by

$$\begin{aligned} P_2^{\mathrm{ipSIC}} &= 1 - \Pr\left(\bar{\gamma}_R > \varepsilon_2, \bar{\gamma}_{S,R}^{\mathrm{ipSIC}} > \varepsilon_1\right) \\ &= 1 - A_1 \times \underbrace{\Pr\left(\bar{\gamma}_{S,R}^{\mathrm{ipSIC}} > \varepsilon_1\right)}_{B}. \end{aligned} \tag{26}$$

And by a series of calculations, we can figure out B is given by

$$\begin{aligned} B &= \Pr\left(\bar{\gamma}_{S,R}^{\mathrm{ipSIC}} > \varepsilon_1\right) = \Pr\left(|g_0|^2 > \frac{\varepsilon_1\bar{\ell}}{d_0^{-\alpha}}|h_I|^2 + \frac{\varepsilon_1}{d_0^{-\alpha}\rho}\right) \\ &= \int_0^\infty f_{|h_I|^2}(x)\left[1 - F_{|g_0|^2}\left(\frac{\varepsilon_1\bar{\ell}}{d_0^{-\alpha}}x + \frac{\varepsilon_1}{d_0^{-\alpha}\rho}\right)\right]dx \\ &= \frac{e^{-\left(\varepsilon_1/\lambda_{g_0}d_0^{-\alpha}\rho\right)}}{\lambda_I}\int_0^\infty e^{-\left(\left(1/\lambda_I\right)+\left(\varepsilon_1\bar{\ell}/\lambda_{g_0}d_0^{-\alpha}\right)\right)x}dx \\ &= \frac{\lambda_{g_0}d_0^{-\alpha}}{\lambda_{g_0}d_0^{-\alpha} + \lambda_I\varepsilon_1\bar{\ell}}e^{-\left(\varepsilon_1/\lambda_{g_0}d_0^{-\alpha}\rho\right)}. \end{aligned} \tag{27}$$

So, we substitute (27) and (23) into (26); we can get (28).

$$P_2^{\mathrm{ipSIC}} = \frac{\lambda_{g_0}\lambda_{g_2}d_0^{-\alpha}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)\left(\lambda_{g_0}d_0^{-\alpha} + \lambda_I\varepsilon_1\bar{\ell}\right)} \times e^{-\left(\varepsilon_1/\lambda_{g_0}d_0^{-\alpha}\rho\right)-\left(\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\rho\right)}. \tag{28}$$

Finally, if $\bar{\ell} = 0$, then the outage probability of U_2 with pSIC is calculated as

$$\begin{aligned} P_2^{\mathrm{pSIC}} &= 1 - \Pr\left(\bar{\gamma}_R > \varepsilon_2, \bar{\gamma}_{S,R}^{\mathrm{pSIC}} > \varepsilon_1\right) \\ &= 1 - A_1 \times \Pr\left(|g_0|^2 > \frac{\varepsilon_1}{d_0^{-\alpha}\rho}\right) \\ &= 1 - \frac{\lambda_{g_2}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)}e^{-\left(\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\rho\right)}\int_{\varepsilon_1/d_0^{-\alpha}\rho}^\infty f_{|g_0|^2}(x)dx \\ &= 1 - \frac{\lambda_{g_2}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)}e^{-\left(\varepsilon_1/\lambda_{g_0}d_0^{-\alpha}\rho\right)-\left(\varepsilon_2/\lambda_{g_2}d_2^{-\alpha}\rho\right)}. \end{aligned} \tag{29}$$

4. The Approximation of Outage Probability

4.1. Scheme I: The Approximation for Uplink Transmission with RIS. As $\rho \longrightarrow \infty$, according to (3) and (4a), we have $\bar{\gamma}_{S,x_1}^{\infty} = d_1^{-\alpha}A^2/d_0^{-\alpha}d_2^{-\alpha}B^2$ and $\bar{\gamma}_{S,x_2}^{\infty} = d_0^{-\alpha}d_2^{-\alpha}B^2/\bar{\ell}|h_I|^2$.

First, the asymptotic expression for U_1 is calculated by

$$\begin{aligned} P_1^{\infty} &= 1 - \Pr\left(\bar{\gamma}_{S,x_1}^{\infty} > \varepsilon_1\right) = 1 - \Pr\left(A^2 > \beta B^2\right) \\ &= 1 - \int_0^{\infty} f_{B^2}(x)[1 - F_{A^2}(\beta x)]dx \\ &= 1 - \frac{2}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^{M+1}} \int_0^{\infty} e^{-(\beta/\lambda_1)x} x^{(M-1)/2} \\ &\quad \times K_{M-1}\left(2\sqrt{\frac{x}{\lambda_0\lambda_2}}\right)dx. \end{aligned} \tag{30}$$

In (30), we pay attention on $K_a(\cdot)$ which can be expressed in terms of the Meijer-G function as in [30], Equation (9.34.3).

$$K_{M-1}\left(2\sqrt{\frac{x}{\lambda_0\lambda_2}}\right) = \frac{1}{2} G_{0,2}^{2,0}\left(\frac{x}{\lambda_0\lambda_2}\middle| \frac{\bar{M}-1}{2}, \frac{1-M}{2}\right). \tag{31}$$

Then, P_1^{∞} can be reformulated as below:

$$\begin{aligned} P_1^{\infty} &= 1 - \frac{1}{\Gamma(M)\left(\sqrt{\lambda_0\lambda_2}\right)^{M+1}} \int_0^{\infty} e^{-(\beta/\lambda_1)x} x^{(M-1)/2} \\ &\quad \times G_{0,2}^{2,0}\left(\frac{x}{\lambda_0\lambda_2}\middle| \frac{\bar{M}-1}{2}, \frac{1-M}{2}\right)dx. \end{aligned} \tag{32}$$

Applying [30], Equation (7.813.1) from the table of integrals, (32) can be reformulated by

$$\begin{aligned} P_1^{\infty} &= 1 - \frac{1}{\Gamma(M)\left(\sqrt{\beta\lambda_1^{-1}\lambda_0\lambda_2}\right)^{M+1}} \\ &\quad \times G_{1,2}^{2,1}\left(\frac{\lambda_1}{\beta\lambda_0\lambda_2}\middle| \frac{(1-M)/2}{(M-1)/2, (1-M)/2}\right). \end{aligned} \tag{33}$$

Next, the asymptotic expression for U_2 with ipSIC is calculated by

$$\begin{aligned} P_2^{\infty,\text{ipSIC}} &= 1 - \Pr\left(\gamma_{S,x_2}^{\infty} > \varepsilon_2\right) = 1 - \Pr\left(B^2 > \varphi|h_I|^2\right) \\ &= 1 - \int_0^{\infty} f_{|h_I|^2}(x)[1 - F_{B^2}(\varphi x)]dx \\ &= 1 - \frac{1}{2^{M-1}\Gamma(M)\lambda_I} \int_0^{\infty} e^{-(x/\lambda_I)} \\ &\quad \cdot \left(\sqrt{\frac{4\varphi x}{\lambda_0\lambda_2}}\right)^M K_M\left(\sqrt{\frac{4\varphi x}{\lambda_0\lambda_2}}\right)dx \\ &= 1 - \frac{1}{\Gamma(M)\lambda_I} \int_0^{\infty} e^{-(x/\lambda_I)} G_{0,2}^{2,0}\left(\frac{\varphi x}{\lambda_0\lambda_2}\middle| \bar{M}, 0\right)dx. \end{aligned} \tag{34}$$

By using Equation (7.813.1) of [30], the asymptotic outage probability of U_2 with ipSIC at high SNRs can be claimed by

Table 1: Main system parameters.

Monte Carlo simulations	10^7 iterations
The targeted data rates for users	$R_1 = 0.5, R_2 = 1$ (BPCU)
The distance from U_2 to RIS	$d_2 = 20$ m
The distance from RIS to BS	$d_0 = 30$ m
The distance from U_1 to BS	$d_1 = 10$ m
The path loss exponent	$\alpha = 2$

$$P_2^{\infty,\text{ipSIC}} = 1 - \frac{1}{\Gamma(M)} G_{1,2}^{2,1}\left(\frac{\lambda_I\varphi}{\lambda_0\lambda_2}\middle| \begin{matrix} 0 \\ M, 0 \end{matrix}\right). \tag{35}$$

Finally, the asymptotic outage probability of U_2 with pSIC at high SNRs for the situations $M = 1$ and $M \geq 2$ is given by the following equations:

$$P_2^{\infty,\text{pSIC}} = -\frac{2\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\rho} \ln\left(\sqrt{\frac{\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\rho}}\right), \quad M = 1, \tag{36a}$$

$$P_2^{\infty,\text{pSIC}} = \frac{\varepsilon_2}{d_0^{-\alpha}d_2^{-\alpha}\rho(M-1)}, \quad M \geq 2. \tag{36b}$$

Proof. To make the computation easier, we use the series form of the Bessel function $K_n(x)$ to approximate the high SNR. $K_n(x)$ can be approximated when $n = 1$ and $n \geq 2$ as

$$K_1(x) \approx \frac{x}{2} \ln\left(\frac{x}{2}\right) + \frac{1}{x}, \tag{37a}$$

$$K_n(x) \approx \left[\frac{2^n(n-1)!}{x^n} - \frac{2^{n-2}(n-2)!}{x^{n-2}}\right]\frac{1}{2}. \tag{37b}$$

Equations (36a) and (36b) can be obtained by putting (37a) and (37b) into (21), respectively. The proof is finished. □

4.2. Scheme II: The Approximation for Uplink Transmission with Relay. The results can be obtained easily by using three approximate equations, namely, $\bar{\gamma}_R \approx d_2^{-\alpha}|g_2|^2/|f_R|^2$, $\bar{\gamma}_{S,U_1}^{\infty} \approx d_1^{-\alpha}|g_1|^2/d_0^{-\alpha}|g_0|^2$, and $\bar{\gamma}_{S,R}^{\text{ipSIC}} \approx d_0^{-\alpha}|g_0|^2/\bar{\ell}|h_I|^2$. Thereby at high SNRs, we can write the approximate expression for the U_1, U_2 with ipSIC, and U_2 with pSIC as follows:

$$\begin{aligned} P_1^{\infty} &= 1 - \Pr\left(\bar{\gamma}_R^{\infty} > \varepsilon_2, \bar{\gamma}_{S,U_1}^{\infty} > \varepsilon_1\right) \\ &= 1 - \frac{\lambda_{g_1}\lambda_{g_2}d_1^{-\alpha}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)\left(\lambda_{g_1}d_1^{-\alpha} + \lambda_{g_0}d_0^{-\alpha}\varepsilon_1\right)}, \end{aligned} \tag{38a}$$

$$\begin{aligned} P_2^{\infty,\text{ipSIC}} &= 1 - \Pr\left(\bar{\gamma}_R^{\infty} > \varepsilon_2, \bar{\gamma}_{S,R}^{\infty\text{ipSIC}} > \varepsilon_1\right) \\ &= \frac{\lambda_{g_0}\lambda_{g_2}d_0^{-\alpha}d_2^{-\alpha}}{\left(\lambda_{g_2}d_2^{-\alpha} + \lambda_{f_R}\varepsilon_2\right)\left(\lambda_{g_0}d_0^{-\alpha} + \lambda_I\varepsilon_1\bar{\ell}\right)}, \end{aligned} \tag{38b}$$

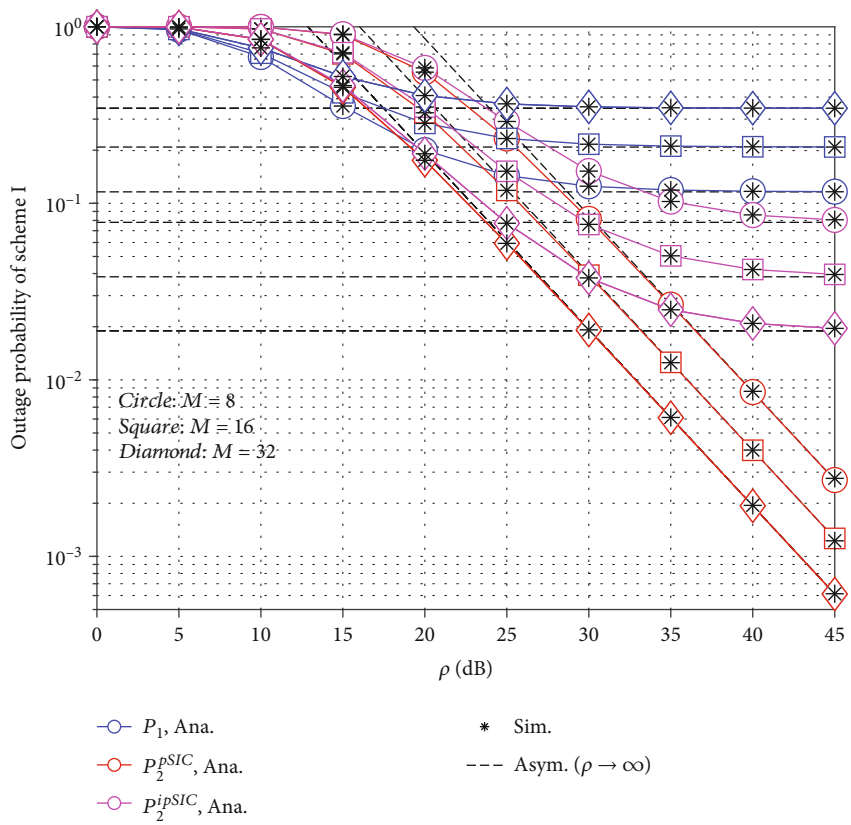


Figure 3: Outage probability of uplink RIS-aided NOMA system versus SNR in decibels with $\lambda_I = -30$ (dB).

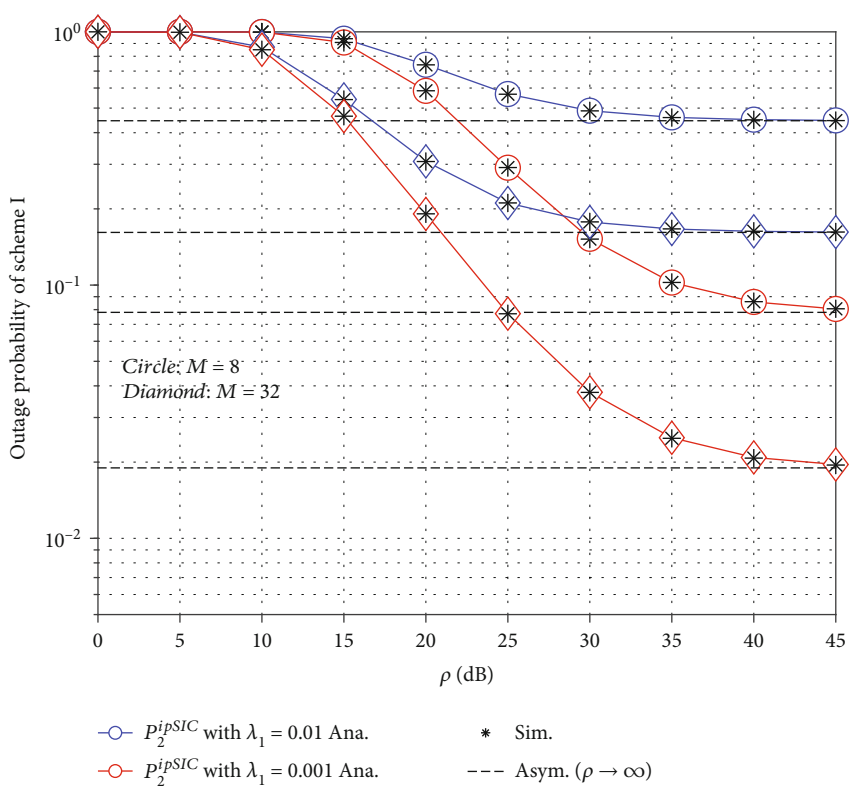


Figure 4: Outage probability of uplink RIS-aided NOMA system versus SNR in decibels varying λ_I.

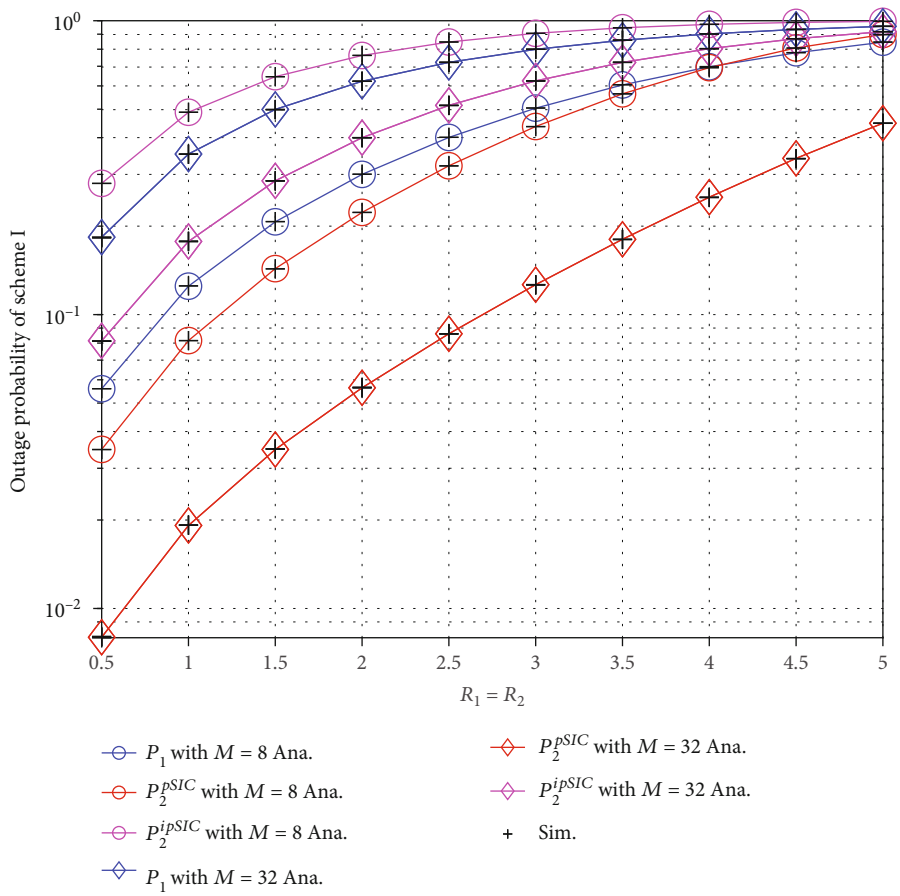


Figure 5: Outage probability of uplink RIS-aided NOMA system versus $R_1 = R_2$ with $\lambda_I = 0.001$ and $\rho = 30$ (dB).

$$P_2^{\infty,\mathrm{pSIC}} = 1 - \Pr\left(\bar{\gamma}_R > \varepsilon_2, \bar{\gamma}_{S,R}^{\mathrm{pSIC}} > \varepsilon_1\right) = 1 - \frac{\lambda_{g_2} d_2^{-\alpha}}{\left(\lambda_{g_2} d_2^{-\alpha} + \lambda_{f_R} \varepsilon_2\right)}. \tag{38c}$$

4.3. Consideration on System Throughput Performance. Although outage probability plays an important role, another metric is also necessary to evaluate system performance which is given by [31] (Equation (42))

$$\begin{aligned} \tau_I^\star &= (1 - P_1)R_1 + (1 - P_2^\star)R_2, \\ \tau_{II}^\star &= (1 - P_1)R_1 + (1 - P_2^\star)R_2, \end{aligned} \quad \star \in \{\mathrm{ipSIC}, \mathrm{pSIC}\}. \tag{39}$$

We expect to evaluate outage performance, and then, throughput can be used to provide useful guidelines for design of such RIS-aided system. More importantly, key parameters can be determined to control the quality as expected.

5. Numerical Results

In this section, we simulate the obtained outage probability and further verify with Monte Carlo simulation. The main parameters used are summarized in Table 1, where BPCU stands for a bit per channel use and we set the Gauss-Laguerre parameter as $T = 40$ to yield a close approximation; the variances of complex channel coefficients are set to be $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$, and $\lambda_I = 0.001$, respectively.

Figure 3 depicts the outage probability versus transmit SNR in decibels in uplink RIS-aided NOMA when the residual interference caused by ipSIC is $\lambda_I = -30$ dB. Paying special attention to the outage performance of RIS-assisted uplink transmission from U_2, we observe that the best outage performance is achieved when the SIC operation is perfect and the RIS phase shifts are set at $M = 32$. On the other hand, when the SIC operation is imperfect, the outage performance approaches a floor in the high-SNR region. Similarly, the uplink outage probability from U_1 experiences the same effect of approaching a floor in the high SNR region. The main reason is that both floors of the RIS-aided NOMA network are also determined by the efficiency of the SIC.

The impact of residual interference is observed in Figure 4. Particularly, the worst performance is observed for the case of $\lambda_I = 0.01$. This means that pSIC allows the system to work at acceptable outage performance levels. We can also see the performance could be better by adjusting the RIS phase shifts reasonably and the outage performance becomes suitable for continuous work of this system.

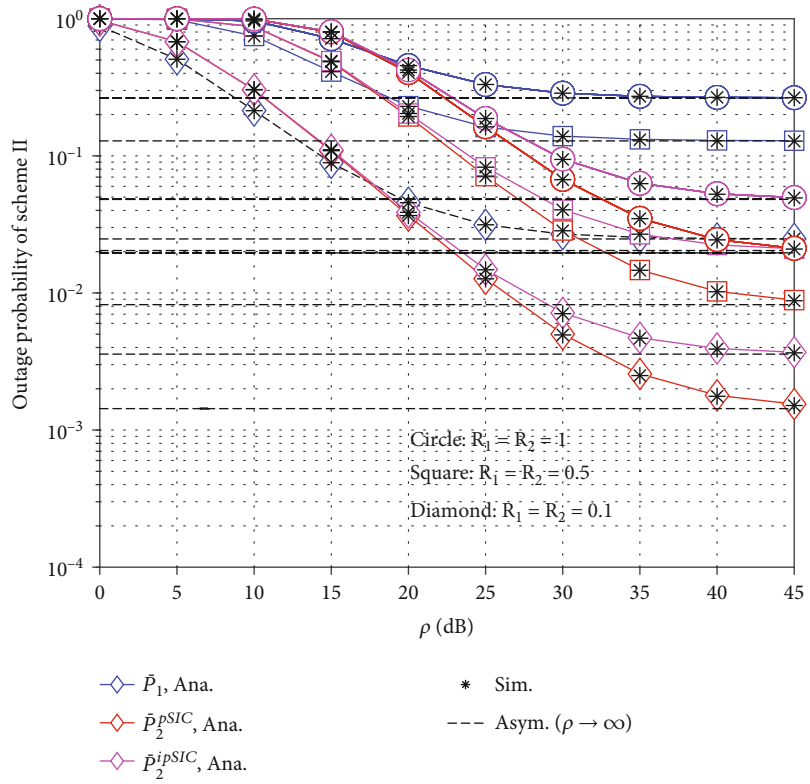


Figure 6: Outage probability of uplink relay-aided NOMA system versus SNR in decibels with $\lambda_{f_R} = 0.01$ and $\lambda_I = 0.01$.

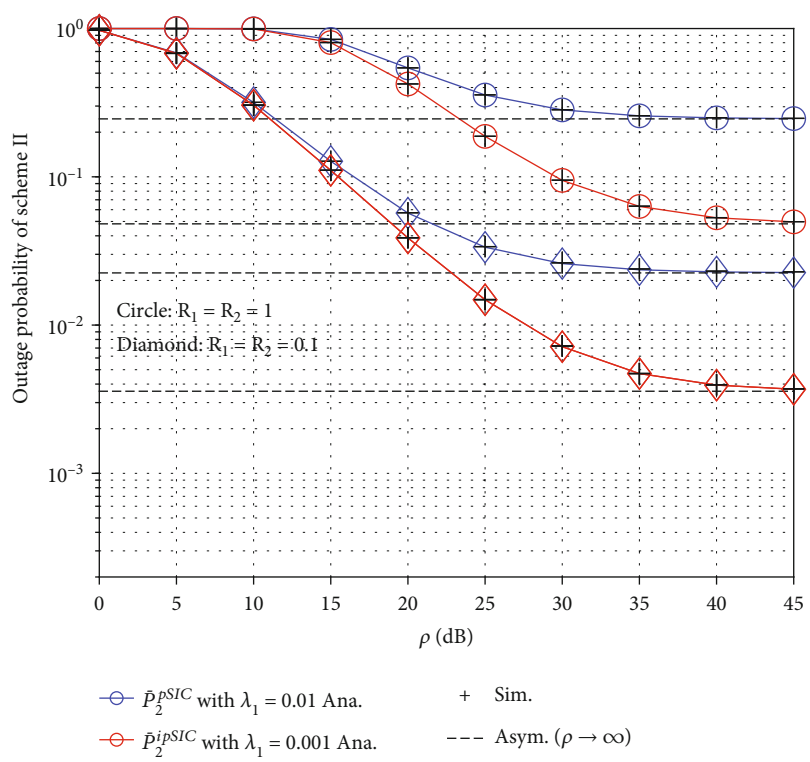


Figure 7: Outage probability of uplink relay-aided NOMA system versus SNR in decibels with $\lambda_{f_R} = 0.01$.

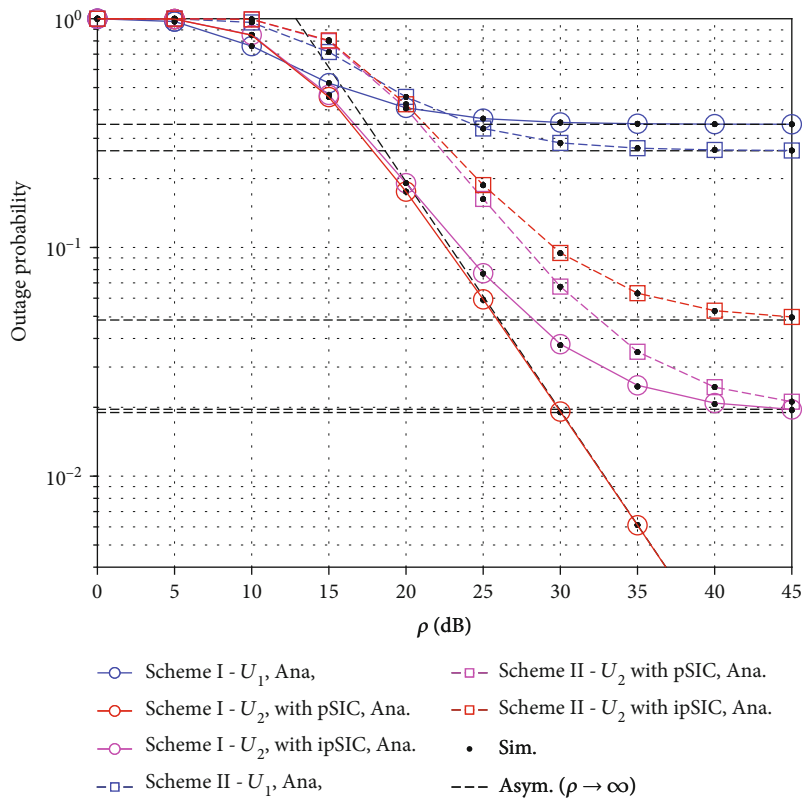


Figure 8: Comparison of outage probability between scheme I and II with $\lambda_{f_R} = 0.01$ and $M = 32$.

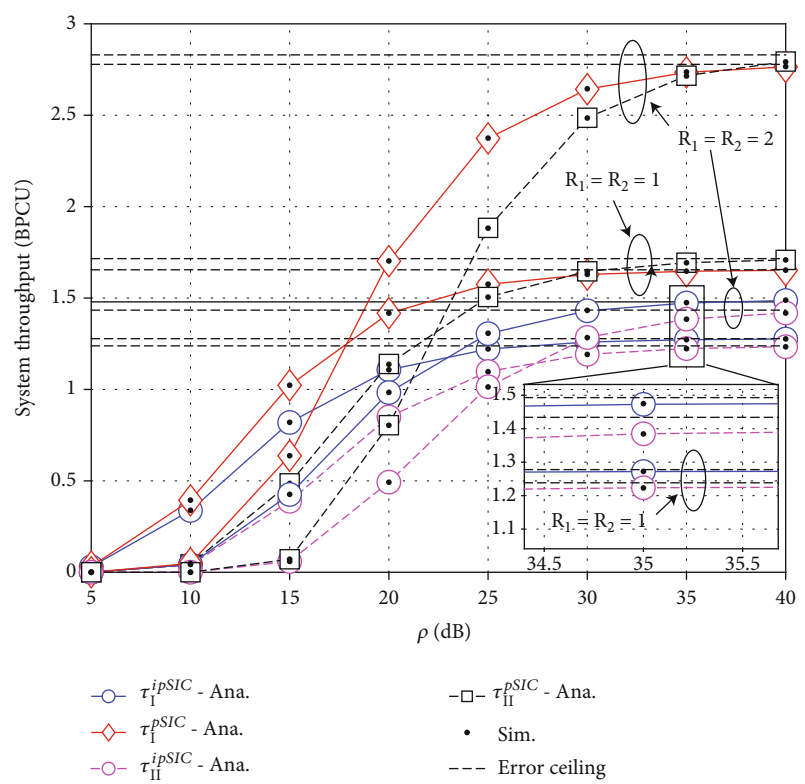


Figure 9: Comparison of system throughput between scheme I and scheme II with $M = 32$, $\lambda_{f_R} = 0.001$, and $\lambda_I = 0.0316$.

In Figure 5, we analyze the outage probability versus targeted data rates $R_1 = R_2$ with residual interference $\lambda_I = 0.001$ and transmit SNR $\rho = 30$ (dB). Here again, we can see the impact of ipSIC and RIS phase shifts with increasing throughput. The best outage performance is achieved by lower $R_1 = R_2$ rates. It is important to note that for uplink RIS-NOMA systems, both Figures 3–5 highlight the contribution of RIS phase shifts to outage probability.

In Figure 6, we can observe the outage probability versus SNR in decibels for the uplink relay-aided NOMA system with $\lambda_{f_R} = 0.01$ and $\lambda_I = 0.01$ representing the relay loop self-interference and ipSIC residual interference, respectively. All the outage probability curves approach a floor in the high-SNR region. Similar to the findings in Figure 5, the best outage performance in the proposed relay-aided NOMA system is achieved by lower $R_1 = R_2$ rates.

Similarly, as in Figure 4, we can see the impact of residual interference on outage performance in relay-aided NOMA systems in Figure 7. We can observe the best outage performance is obtained with both lower $R_1 = R_2$ rates as well as lower residual interference λ_I.

Figure 8 compares the outage probability of the two proposed schemes versus SNR in dB with $\lambda_{f_R} = 0.01$ and $M = 32$. This figure confirms the superiority of the RIS-aided system over the relay-aided system.

In Figure 9, we can see the trend of throughput between the two schemes. The curves approach an error ceiling in the high-SNR region due to the residual and loop interference.

Finally, we can observe for all Figures 3–9 that the simulated curves and asymptotic curves closely match in the high SNR region thus, confirming the validity of the derived high-SNR approximations.

6. Conclusion

We have investigated the system outage and throughput performance of uplink RIS-assisted multiple access systems by exploring the impact of pSIC and ipSIC. To gain further insights into the benefits of including RIS in uplink multiple access systems, we compared it to an uplink relay-aided multiple access system. In each scenario, the RIS and relay devices are deployed to enhance the coverage of an obstructed far user by assisting it to upload information to a BS. To provide system performance analysis, we adopted the Gauss-Laguerre integration to achieve the closed-form approximate outage expressions in the high-SNR regime. We evaluate the derived expressions to understand whether the outage and throughput can be improved. Numerical results of ipSIC scenarios indicate that the RIS-NOMA system can still perform despite the impact of ipSIC. Furthermore, the results demonstrate the superiority of RIS over the relay in the high-SNR region. However, it is important to note the existence of outage error floors and throughput ceilings in a high-SNR region. Therefore, this regard could be target outage and throughput performance once we improve other system parameters. In future work, we will investigate the impact of introducing multiple RIS devices in the system to further enhance the system uplink performance.

Appendix

Proof of Proposition 1

The outage of U_2 under ipSIC is given as

$$\begin{aligned} P_2^{\text{ipSIC}} &= 1 - \Pr\left(\gamma_{S,x_2} > \varepsilon_2\right) = 1 - \Pr\left(B^2 > \varphi|h_I|^2 + \theta_2\right) \\ &= 1 - \int_0^\infty f_{|h_I|^2}(x)[1 - F_{B^2}(\varphi x + \theta_2)]dx \\ &= 1 - \frac{1}{2^{M-1}\Gamma(M)\lambda_I}\int_0^\infty e^{-(x/\lambda_I)}\left(\sqrt{\frac{4(\varphi x + \theta_2)}{\lambda_0\lambda_2}}\right)^M \\ &\quad \times K_M\left(\sqrt{\frac{4(\varphi x + \theta_2)}{\lambda_0\lambda_2}}\right)dx, \end{aligned} \tag{A.1}$$

where $\theta_2 = \varepsilon_2/(d_0^{-\alpha}d_2^{-\alpha}\rho)$, $\varphi = \varepsilon_2\bar{\ell}/(d_0^{-\alpha}d_2^{-\alpha})$, $G_{p,q}^{m,n}[.]$ is the Meijer G-function given in [30], Equation (9.301), $H_t = (T!)^2 q_t/[L_{T+1}(q_t)]^2$, where H_t is the weight of the Gauss-Laguerre integration [32] and q_t is the q-th zero of Laguerre polynomial $L_T(q_t)$. The parameter T is to ensure a complexity-accuracy tradeoff.

With the help of the [30], Equation (9.34.3) and after some manipulation, we have

$$P_2^{\text{ipSIC}} = 1 - \frac{1}{\Gamma(M)\lambda_I}\int_0^\infty e^{-(x/\lambda_I)}G_{0,2}^{2,0}\left(\frac{\varphi x + \theta_2}{\lambda_0\lambda_2}\middle|\begin{matrix} - \\ M, 0 \end{matrix}\right)dx, \tag{A.2}$$

where $G_{p,q}^{m,n}[.]$ is the Meijer G-function given in [30], Equation (9.301). Specifically, we let $q = x/\lambda_I$ and with the help of Gauss-Laguerre integration [33], Equation (25.4.45). The closed-form approximation of the P_2^{ipSIC} can be given as

$$\begin{aligned} P_2^{\text{ipSIC}} &= 1 - \frac{1}{\Gamma(M)\lambda_I}\int_0^\infty e^{-(x/\lambda_I)}G_{0,2}^{2,0}\left(\frac{\varphi x + \theta_2}{\lambda_0\lambda_2}\middle|\begin{matrix} - \\ M, 0 \end{matrix}\right)dx \\ &= 1 - \frac{1}{\Gamma(M)}\int_0^\infty e^{-q}G_{0,2}^{2,0}\left(\frac{\varphi\lambda_I q + \theta_2}{\lambda_0\lambda_2}\middle|\begin{matrix} - \\ M, 0 \end{matrix}\right)dq \\ &\approx 1 - \frac{1}{\Gamma(M)}\sum_{t=1}^{T} H_t G_{0,2}^{2,0}\left(\frac{\varphi\lambda_I q_t + \theta_2}{\lambda_0\lambda_2}\middle|\begin{matrix} - \\ M, 0 \end{matrix}\right), \end{aligned} \tag{A.3}$$

where $H_t = (T!)^2 q_t/[L_{T+1}(q_t)]^2$, where H_t is the weight of the Gauss-Laguerre integration and q_t is the q-th zero of Laguerre polynomial $L_T(q_t)$.

The proof of Proposition 1 is completed.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work has been supported by Van Lang University, Ho Chi Minh City, Vietnam under the Project 1000.

References

[1] A. Gupta and R. K. Jha, "A survey of 5G network: architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[2] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[3] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2347–2381, 2017.

[4] M.-S. Van Nguyen, D. Dinh-Thuan, S. Al-Rubaye, S. Mumtaz, A. Al-Dulaimi, and O. Dobre, "Exploiting impacts of antenna selection and energy harvesting for massive network connectivity," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7587–7602, 2021.

[5] D.-T. Do, L. Anh-Tu, Y. Liu, and A. Jamalipour, "User grouping and energy harvesting in UAV-NOMA system with AF/DF relaying," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 11855–11868, 2021.

[6] D. T. Do, M. S. Van Nguyen, M. Voznak, A. Kwasinski, and J. N. de Souza, "Souza, "Performance analysis of clustering car-following V2X system with wireless power transfer and massive connections,"," *IEEE Internet of Things Journal*, 2021.

[7] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[8] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[9] D. T. Do, M. S. Van Nguyen, A. T. Le, K. M. Rabie, and J. Zhang, "Joint fullduplex and roadside unit selection for NOMA-enabled V2X communications: ergodic rate performance," *IEEE Access*, vol. 8, pp. 140348–140360, 2020.

[10] G. Yang, X. Xu, Y. -C. Liang, and M. D. Renzo, "Reconfigurable intelligent surface-assisted non-orthogonal multiple access," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3137–3151, 2021.

[11] M. Munochiveyi, A. C. Pogaku, D. -T. Do, A. -T. Le, M. Voznak, and N. D. Nguyen, "Reconfigurable intelligent surface aided multi-user communications: state-of-the-art techniques and open issues," *IEEE Access*, vol. 9, pp. 118584–118605, 2021.

[12] M. Fu, Y. Zhou, Y. Shi, and K. B. Letaief, "Reconfigurable intelligent surface empowered downlink non-orthogonal multiple access," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3802–3817, 2021.

[13] C. Pan, H. Ren, K. Wang et al., "Reconfigurable intelligent surfaces for 6G systems: principles, applications, and research directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14–20, 2021.

[14] Y. Xiu, J. Zhao, W. Sun et al., "Reconfigurable intelligent surfaces aided mmWave NOMA: joint power allocation, phase shifts, and hybrid beamforming optimization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 8393–8409, 2021.

[15] Y. Li, M. Jiang, Q. Zhang, and J. Qin, "Joint beamforming design in multi-cluster MISO NOMA reconfigurable intelligent surface-aided downlink communication networks," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 664–674, 2021.

[16] C. Zhang, W. Yi, Y. Liu, K. Yang, and Z. Ding, "Reconfigurable intelligent surfaces aided multi-cell NOMA networks: a stochastic geometry model," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 951–966, 2021.

[17] J. Zuo, Y. Liu, and N. Al-Dhahir, "Reconfigurable intelligent surface assisted cooperative non-orthogonal multiple access systems," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6750–6764, 2021.

[18] M. Elhattab, M. A. Arfaoui, C. Assi, and A. Ghrayeb, "Reconfigurable intelligent surface enabled full-duplex/half-duplex cooperative non-orthogonal multiple access," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3349–3364, 2021.

[19] R. Zhong, Y. Liu, X. Mu, Y. Chen, and L. Song, "AI empowered RIS-assisted NOMA networks: deep learning or reinforcement learning?," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 182–196, 2022.

[20] M. Diamanti, M. Tsampazi, E. E. Tsiropoulou, and S. Papavassiliou, "Energy efficient multi-user communications aided by reconfigurable intelligent surfaces and UAVs," in *2021 IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 371–376, Irvine, CA, USA, August 2021.

[21] M. Diamanti, E. E. Tsiropoulou, and S. Papavassiliou, "The joint power of NOMA and reconfigurable intelligent surfaces in SWIPT networks," in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 621–625, Lucca, Italy, September 2021.

[22] D. Dinh-Thuan and A.-T. Le, "Physical layer security for Internet of Things via reconfigurable intelligent surface," *Future Generation Computer Systems*, vol. 126, pp. 330–339, 2022.

[23] D. T. Do, T. T. T. Nguyen, T. N. Nguyen, X. Li, and M. Voznak, "Uplink and downlink NOMA transmission using full-duplex UAV," *IEEE Access*, vol. 8, pp. 164347–164364, 2020.

[24] Y. Cheng, K. H. Li, Y. Liu, K. C. Teh, and H. Vincent Poor, "Downlink and uplink intelligent reflecting surface aided networks: NOMA and OMA," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3988–4000, 2021.

[25] L. Yang, Y. Yang, M. O. Hasna, and M. -S. Alouini, "Coverage, probability of SNR gain, and DOR analysis of RIS-aided communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 8, pp. 1268–1272, 2020.

[26] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[27] X. Yue and Y. Liu, "Performance analysis of intelligent reflecting surface assisted NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 4, pp. 2623–2636, 2022.

[28] C.-B. Le and D.-T. Do, "Joint evaluation of imperfect SIC and fixed power allocation scheme for wireless powered D2D-NOMA networks with multiple antennas at base station," *Wireless Networks*, vol. 25, no. 8, pp. 5069–5081, 2019.

[29] C.-B. Le, D.-T. Do, X. Li, Y.-F. Huang, H.-C. Chen, and M. Voznak, "Enabling NOMA in backscatter reconfigurable intelligent surfaces-aided systems," *IEEE Access*, vol. 9, pp. 33782–33795, 2021.

[30] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Elsevier, Amsterdam, The Netherlands, 2007.

[31] D.-T. Do, C.-B. Le, and F. Afghah, "Enabling full-duplex and energy harvesting in uplink and downlink of small-cell network relying on power domain based multiple access," *IEEE Access*, vol. 8, pp. 142772–142784, 2020.

[32] X. Yue, Y. Liu, Y. Yao, X. Li, R. Liu, and A. Nallanathan, "Secure communications in a unified non-orthogonal multiple access framework," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2163–2178, 2020.

[33] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, NY, USA, 1972.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7268083, 9 pages
https://doi.org/10.1155/2022/7268083

Research Article

Multiobjective Optimization Designs in Downlink NOMA Transmission Systems

Zhe Li

School of Information Engineering, Department of Public Security, Railway Police College, Zhengzhou, China

Correspondence should be addressed to Zhe Li; lizhe@rpc.edu.cn

Received 7 April 2022; Revised 11 May 2022; Accepted 24 May 2022; Published 20 July 2022

Academic Editor: Xingwang Li

To further balance the high-capacity transmission requirement and the scarce spectrum resources status quo, we investigate the joint optimization of energy efficiency (EE) and spectral efficiency (SE) in a downlink multiuser NOMA system. While ensuring the user quality of service (QoS), a power allocation algorithm based on two-layer optimization is proposed by transforming the original multiobjective optimization problem into a univariate problem through the linear weighted sum method. Based on the optimal user power allocation coefficient solution obtained from the inner layer, the outer layer applies the bisection principle to search for the optimal system transmit power according to the uniqueness of the maximum value of pseudo-concave functions. Simulation results show that the proposed algorithm always converges within fewer iterations and can achieve the flexible tradeoff between EE and SE. This is of great significance for improving the energy utilization of actual scenarios with time-varying communication demands.

1. Introduction

The application and popularization of 5G not only accelerates the development of the Internet of Things (IoT) but also promotes the emergence of many new communication services, which brings more strict performance requirements, such as better spectrum efficiency, denser network device access, and ultra-low transmission delay [1–3]. Taking into account the nearly exhausted frequency band resources, relevant scholars and research institutions begin to seek new access technologies to deal with the above problem [4]. Therefore, how to efficiently utilize the limited energy to meet the ever-increasing communication demand becomes a major topic in the current mobile communication field.

Coincidentally, as a prominent wireless radio technique, nonorthogonal multiple access (NOMA) can transmit multiple users' information simultaneously in the same frequency by introducing the power domain dimension, which generates excellent spectral efficiency (SE) [5, 6]. Especially, the key technology determining the performance of NOMA is power multiplexing [7]. Therefore, in the background of the high energy consumption and great demand for data rates in wireless communication system, it desperately needs more efficient power allocation schemes that consider altogether energy efficiency (EE), SE, and quality of service (QoS) to maintain the superiority of NOMA.

Generally, the power allocation in NOMA is mainly discussed from two aspects, i.e., EE and SE. From EE vantage point, a fast convergence algorithm based on the Dinkelbach-Like theory was presented in [8], which contributes to obtain the optimal EE for a multiuser NOMA system. In [9], Z. Ma et al. investigated the problem about jointing user and subcarrier allocation while considering user fairness in a downlink multicarrier (MC) NOMA system. In [6], K. Cumanan et al. studied how the beamforming technology in MISO NOMA systems can better improve the EE. In addition, extensive researches have also been done from the SE aspect of NOMA. For instance, an effective resource schedule strategy is proposed to enhance the SE for a MC-NOMA system in [10]. In [11], a method for jointing optimal beamforming and power allocation all introduced to improve the throughput of the NOMA networks

in multicell. In [12], a power allocation and spectrum sharing technology for the two-slot and secondary NOMA relay network is given to maximize the overall achievable rate of cellular users.

However, the above-related works all consider the single-objective optimization (SOO) with EE or SE. In fact, the system metric often needs to be flexibly adjusted according to actual communication scenes. In some scenes with high real-time demands, the devices often need to quickly respond to the emergency requirements, such as medical care or alert information, which is SE critical [13]. But in other times, the devices can be served just to maintain its basic operation, and which requires optimal EE at this moment. Thus, it is significant to study how improve the energy utilization by weighing EE and SE in the real-time transmission scenarios. In [14], Y. Hei et al. conducted the parameter analysis and simulation verification for the EE-SE tradeoff in DCO-OFDM communication system. In [15], O. Aydin et al. studied the tradeoff about EE-SE in 5G multioperator heterogeneous networks. Additionally, the closed expressions for the relationship between EE and SE in CRN network were derived in [16], which also claimed that EE-SE are not contradictory in CRN networks and can achieve the optimal tradeoff. In NOMA system, the tradeoff about EE and SE for a single-carrier and multiuser network is investigated in [17], but it still adopted the constraint method, which makes EE as the main optimization objective, and SE as a constraint, in which the acquisition of EE's value is more experience-based and less flexible. In [18], the EE-SE tradeoff in a NOMA system with multiuser pairs was studied, and it assumed that there was no interference between user pairs, which indicates that the power allocation is only based on every two users. An energy-efficient resource allocation technique for a hybrid time division multiple access—NOMA system—was proposed in [19], the available time for transmission is divided into several subtime slots, and a subtime slot is allocated to serve a group of users (i.e., cluster). A network network nonorthogonal multiple access (N-NOMA) technique for the uplink coordinated multipoint transmission (CoMP) was applied in [20] to improve the system throughput, and the numerical results are presented to show the accuracy of the analytical results and also demonstrate the superior performance of the proposed N-NOMA scheme.

Based on the above researches, for improving the energy utilization, we consider both EE and SE as primary optimization goals simultaneously in a downlink multiuser NOMA system. By applying the linear weighted sum function method [21], the original multiobjective optimization (MOO) problem is firstly transformed into a SOO problem through a positive weighting factor. While satisfying users' QoSs, we finally propose a dual-layer power allocation algorithm based on bisection searching. The optimal closed-form expressions for power allocation coefficient of all connected users and the optimal system transmit power can be, respectively, obtained in inner and outer layer optimization. Simulation results demonstrate the validity of the proposed algorithm.

2. System Model

In this paper, a downlink NOMA transmission scenario with multiple users is considered in Figure 1, wherein one centrally located base station (BS) serves M single-antenna users simultaneously. The channel coefficient from the BS to the mth user ($1 \le \mathrm{m} \le M$) is represented by $h_m = g_m d^{-\alpha/2}$, where g_m denotes the Rayleigh fading, d is the distance between the mth user and the BS, and α is the path-loss exponent. Without loss of generality, we assume that the BS can obtain the instantaneous user channel state information (CSI), and the channel gains of users are sorted as: $0 < |h_1|^2 \le |h_2|^2 \cdots \le |h_M|^2$.

According to the protocol of NOMA, the BS sends a superimposed signal composed of all users' information to each user through the code superposition technology. At the receiving end, the successive interference cancelation (SIC) is adopted to eliminate the inter-user interference. That is, each user first decodes signal of the user whose channel gain is lower than that of its own, and subtracts it from its received signal. Afterwards, it decodes own signal by treating the signal of users with a higher channel gain than itself as interference. Specifically, the mth user first decodes the ith ($i = 1, 2, \cdots, m-1$) user's information and subtracts it from the superimposed signal it receives, and then the information of the jth ($j = m+1, m+2, \cdots, M$) user is regarded as interference.

Consequently, the achievable data rate of the mth user can be written as:

$$R_m = \log_2\left(1 + \frac{a_m P_T |h_m|^2}{P_T |h_m|^2 \sum_{i=m+1}^{M} a_i + \sigma^2}\right), \ \mathrm{m} = 1, \cdots, M, \quad (1)$$

where P_T represents the maximum available transmit power at the BS, and a_m is the ratio of the allocated power of the mth user to P_T.

Furthermore, the sum rate of the system is expressed as:

$$R = \sum_{\mathrm{m}=1}^{M} R_m. \quad (2)$$

Assuming that the system has the unit bandwidth, the SE and EE can be defined as follows [16]:

$$\eta_{SE} = R, \quad (3)$$

$$\eta_{EE} = \frac{R}{P + P_c}, \quad (4)$$

where P_c is the circuit power consumption, and P is the actual consumed transmit power by BS.

3. Problem Formation

According to previous studies on EE and SE for NOMA systems, SE always grows with the increasing of the transmit power of the BS, while EE first increases and then begins to decrease after achieving its maximum value. This means

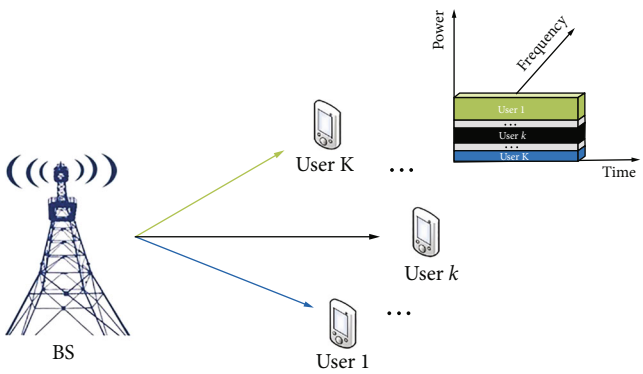


Figure 1: System model.

that EE and SE are contradictory when the BS power is large enough, and also indicates that there is a tradeoff between EE and SE.

In this section, we investigate the joint optimization of EE and SE to achieve a flexible tradeoff between the two. The primordial MOO problem can be formulated as:

$$\mathscr{P}1: \max_{P} \{\eta_{EE}(P), \eta_{SE}(P)\}, \tag{5}$$

where $0 \le P \le P_T$.

Unlike SOO problem, MOO problem can have multiple optimal solutions. For $\mathscr{P}1$, we suppose P^* is the transmit power that enables EE to achieve its maximum value. When $P^* \ge P_T$, $\mathscr{P}1$ only has one solution, i.e., the global optimal solution P_T. When $P^* < P_T$, the objective can be expressed as:

$$\begin{aligned} \mathscr{P}2: \ & \max_{P} \{\eta_{EE}(P), \eta_{SE}(P)\}, \\ & s.t. P^* \le P \le P_T. \end{aligned} \tag{6a}$$

In order to address $\mathscr{P}2$, we first normalize η_{SE} and η_{EE} to make them additive and comparable, and then, referring to the common conversion method for MOO problems, the weighted sum function is adopted. Afterwards, a new system metric is defined as:

$$W(\lambda_{EE}, \lambda_{SE}) = \omega \cdot \lambda_{EE} + (1-\omega)\lambda_{SE}, \tag{7}$$

where $\lambda_{EE} = \eta_{EE}/\eta_{EE}^{\max}$ and $\lambda_{SE} = \eta_{SE}/\eta_{SE}^{\max}$. $\eta_{EE}^{\max}$ and $\eta_{SE}^{\max}$ are, respectively, the maximum value that EE and SE can achieve within the range of a given BS power. It is evident in (7) that the flexible tradeoff between EE and SE can be realized by adjusting ω ($\omega = \{\omega | 0 \le \omega \le 1\}$), which is a preference factor.

Moreover, substituting (3) and (4) into (7), we get

$$\lambda_{EE-SE} = \left[\omega \frac{1}{\eta_{EE}^{\max}(P+P_c)} + (1-\omega)\frac{1}{\eta_{SE}^{\max}}\right] R. \tag{8}$$

As a result, $\mathscr{P}2$ can be further described as:

$$\mathscr{P}3: \max \ \lambda_{EE-SE}, \tag{9a}$$

$$s.t. R_m \ge R_m^{\min}, 1 \le m \le M, \tag{9b}$$

$$\mathrm{P} \le P_T \text{ and } P = \sum_{m=1}^{M} a_m P_T, \tag{9c}$$

where constraint (9a) denotes the minimum target rate of the mth user, which is required by its QoS and denoted as $R_m^{\min}$ for $1 \le m \le M$. (9b) shows the constraint for maximum transmit power of the BS.

To ensure problem (9) is solvable, the maximum BS transmit power must be large enough to support all users' QoS mentioned in (9a), which further introduces the minimum transmit power constraint. We define $P_{\min}$ as the total minimum transmit power that can just satisfy all users' QoS requirements, and thus, problem (9) is feasible only when $\mathrm{P} \ge P_{\min}$. Therefore, it is important to first establish the feasible range of P, and the derivation of which is discussed as follows.

Denote p_m as the allocated power of the mth user, and the problem of calculating $P_{\min}$ can be described as

$$P_{\min} \triangleq \min_{p_k, 1 \le m \le M} \sum_{m=1}^{M} p_m, \tag{10a}$$

$$s.t. p_m \ge A_m \left(\sum_{i=m+1}^{M} p_i + \frac{\sigma^2}{|h_m|^2} \right), 1 \le m \le M, \tag{10b}$$

where $A_{\mathrm{m}} = 2^{R_m^{\min}} - 1$. And (10b) originates from the minimum target rate constraints in (9b). The solution to problem (10) can be found by using the following theorem.

Theorem 1. *Denoted $\{p_m^{\min}\}_{m=1}^{M}$ as the optimal solution to the objective (10), which is given as*

$$P_m^{\min} = A_m \left(\sum_{i=m+1}^{M} P_i^{\min} + \frac{\sigma^2}{|h_m|^2} \right), 1 \le m \le M. \tag{11}$$

Proof. We prove the theorem according to the contradiction principle. Assuming $\{p_m^*\}_{m=1}^{M}$ is a set of allocated power of all users and also is the optimal solution to problem (10) with at least one constraint in (10b) inactive. More generally, we then set the kth constraint as inactive, i.e.,

$$p_k^* > A_{\mathrm{k}} \left(\sum_{i=k+1}^{K} P_i^{\min} + \frac{\sigma^2}{|h_k|^2} \right). \tag{12}$$

□

We next create $\{P_m^{**}\}_{m=1}^{M}$ as a new set by defining $p_m^{**} = p_m^*$ for $k \ne m$ and setp_k^{**} to the right-hand side (RHS) of (12). By observing the structure of the constraints in (10b), for an arbitrary m, we can find that the RHS of (10b) is a monotonically nondecreasing function of p_i for $1 \le i \le M$. In the result, the setting of p_k^{**}, whose value is less than p_k^{**}, ensures that all the constraints in

(10b) hold for the newly created set $\{P_m^{**}\}_{m=1}^M$. But then, we can get $\sum_{m=1}^M p_m^{**} < \sum_M^M M$ with the definition of $\{P_m^{**}\}_{m=1}^M$. It obviously contradicts to the assumption that $\{P_m^{**}\}_{m=1}^M$ is the optimal solution to problem (10). Therefore, we can conclude that the function (10a) is minimized when all the constraints in (10b) are active. The proof is complete.

Afterwards, when all constraints in (10b) are active, the optimal solution to objective (10), i.e.,$\{P_m^{**}\}_{m=1}^M$, can be calculated sequentially in the order $\mathrm{m} = M, M-1, \cdots, 1$, this is because that$P_{\mathrm{m}}^{\min}$ can be calculated by $\{p_{i+1}^{\min}\}_{i=m+1}^M$, and $p_M^{\min} = A_M\sigma^2/|h_m|^2$is determinable. Subsequently, we obtain$P_{\min} = \sum_{m=1}^K p_m^{\min}$ and the feasible region of the actual consumed power is $P_{\min} \le P \le P_T$.

4. Proposed Optimal Solutions

It is apparent that $\mathscr{P}3$ is nonconvex, so it is difficult to find the global optimal solution in polynomial time. Therefore, we decompose it into a dual problem, by searching the optimal power allocation in inner layer and optimal available transmit power in outer layer, the final solution of $\mathscr{P}3$ can be obtained.

4.1. Inner Layer: Optimal Power Allocation Scheme. For a given P within the feasible range $P_{\min} \le P \le P_T$, $\mathscr{P}3$ can be regarded as an optimization problem that only includes the power allocation factor a_{m}. Additionally, for any given ω, the part in square brackets on the RHS of (8) is a constant. At this time, problem $\mathscr{P}3$ is equivalent to maximize the overall sum rate R.

Moreover, substituting (1) into (2), R can recast as:

$$\begin{aligned} R = {} & \log_2\left(P_T|h_1|^2\sum_{i=1}^M a_i + \sigma^2\right) \\ & + \sum_{\mathrm{m}=1}^{M-1}\left[\log_2\left(P_T|h_{m+1}|^2\sum_{i=m+1}^M a_i + \sigma^2\right)\right. \\ & \left. - \log_2\left(P_T|h_{\mathrm{m}}|^2\sum_{i=m+1}^M a_i + \sigma^2\right)\right] - \log_2\left(\sigma^2\right). \end{aligned} \tag{13}$$

In order to simplify the notation, we then define:

$$\Gamma_m \triangleq P_T|h_{\mathrm{m}}|^2, 1 \le \mathrm{m} \le M, \tag{14}$$

$$\theta \triangleq \sum_{i=1}^M a_i = \frac{P}{P_T}, \tag{15}$$

$$\beta_m \triangleq \sum_{i=m+1}^M a_i, 1 \le m \le M-1, \tag{16}$$

$$F_m(\beta_m) \triangleq \log_2\left(\Gamma_{m+1}\beta_m + \sigma^2\right) - \log_2\left(\Gamma_m\beta_m + \sigma^2\right). \tag{17}$$

Based on the above definitions, R can be reformulated as:

$$R = \log_2\left(1 + \Gamma_1\theta/\sigma^2\right) + \sum_{\mathrm{m}=1}^{M-1} F_m(\beta_m). \tag{18}$$

Here, we further emphasize that $\theta = \{\theta|0 \le \theta \le 1\}$ is the power consumption factor, which denotes the ratio of the actually utilized power P to the maximum available BS transmit powe rP_T. Therefore, the power allocation of the inner layer is to allocate parameter $\{\mathrm{a}_{\mathrm{m}}^*(\theta)\}_{\mathrm{m}=1}^M$ by treating θ as a constant, and its solution is a function of θ. Moreover, it can be easily observed that the first term in (16) is a constant for any given θ, and then the optimization problem for maximizing R is equivalent to:

$$\begin{aligned} \mathscr{P}4: & \max_{a_m, 1\le m\le M} \sum_{\mathrm{m}=1}^{M-1} F_m(\beta_m), \\ & s.t.\ R_m \ge R_m^{\min}, \\ & \sum_{m=1}^M a_m = \theta, \end{aligned} \tag{19a}$$

where the optimal power allocation coefficient of the mth user, i.e.,$a_{\mathrm{m}}^*(\theta)$, can be obtained by [22]:

$$a_{\mathrm{m}}^*(\theta) = \begin{cases} \dfrac{A_{\mathrm{m}}}{2^{R_{\mathrm{m}}^{\min}}}\left[\theta - \displaystyle\sum_{i=1}^{m-1} a_i^*(\theta) + \frac{\sigma^2}{P|h_m|^2}\right], \mathrm{m} \ne \mathrm{M} \\ \theta - \displaystyle\sum_{i=1}^{m-1} a_i^*(\theta), \quad \mathrm{m} = \mathrm{M} \end{cases}. \tag{20}$$

4.2. Outer Layer: Optimal Available Transmit Power. In inner layer problem, we have obtained the optimal closed-form solution for the power allocation among users, in which θ is the only variable. Furthermore, in outer layer, $\mathscr{P}3$ can be transformed into a univariate optimization problem about θ.

Based on the optimal power allocation demonstrated in (20), (8) can be rewritten as:

$$\lambda_{EE-SE}'(\theta) = Q(\theta)R(\theta), \tag{21}$$

where

$$Q(\theta) = \frac{\omega}{\eta_{EE}^{\max}(\theta P_T + P_C)} + \frac{1-\omega}{\eta_{SE}^{\max}}, \tag{22}$$

$$R(\theta) = \log_2\left(1 + \Gamma_1\theta/\sigma^2\right) + \sum_{\mathrm{m}=1}^{M-1} F(\beta_m^*(\theta)), \tag{23}$$

and $\beta_{\mathrm{m}}^*(\theta) = \sum_{i=m+1}^M a_i^*(\theta)$, $P = \theta P_T$.

The problem in outer layer can be finally formulated as:

$$\mathscr{P}5: \max_{\theta} \lambda'_{EE-SE}(\theta), \tag{24a}$$

$$\text{s.t.} P_{\min}/P_T \le \theta \le 1.. \tag{24b}$$

In order to obtain the optimal solution of $\mathscr{P}5$, we first introduce the following theorem:

Theorem 2. *$\lambda'_{EE-SE}(\theta)$ is a strict pseudo-concave function with respect to θ.*

Proof. Combining (22) and (23), $\lambda'_{EE-SE}(\theta)$ can be re-expressed as:

$$\lambda'_{EE-SE}(\theta) = \frac{\omega R(\theta)}{\eta_{EE}^{\max}(\theta P_T + P_C)} + \frac{1-\omega}{\eta_{SE}^{\max}} R(\theta), \tag{25}$$

where the convexity of $R(\theta)$ can be firstly verified. It is clear that the first term of $R(\theta)$, i.e., $\log_2(1+\Gamma_1\theta/\sigma^2)$, is a logarithmic function with respect to θ, which indicates that it is strictly concave. Moreover, the convexity of the second term, $F_m(\beta_m^*(\theta))$, can be derived recursively from the convexity of $F_m(\beta_m)$. The first-order derivative of $F_m(\beta_m)$ aboutβ_m is given as:

$$\frac{dF_m(\beta_m)}{d\beta_m} = \frac{(\Gamma_{\mathrm{m}+1} - \Gamma_m)\sigma^2}{\ln 2(\Gamma_{\mathrm{m}+1}\beta_m + \sigma^2)(\Gamma_{\mathrm{m}}\beta_m + \sigma^2)} \ge 0, \tag{26}$$

and the second-order derivative can be further derived as:

$$\frac{d^2 F_m(\beta_m)}{d\beta_m{}^2} = \frac{-(\Gamma_{m+1} - \Gamma_m)D\sigma^2}{\ln 2[(\Gamma_{\mathrm{m}+1}\beta_m + \sigma^2)(\Gamma_{\mathrm{m}}\beta_m + \sigma^2)]^2} \le 0, \tag{27}$$

where $D = (\Gamma_{\mathrm{m}+1}\beta_m + \sigma^2)\Gamma_{\mathrm{m}} + (\Gamma_{\mathrm{m}}\beta_m + \sigma^2)\Gamma_{\mathrm{m}+1}$. □

Obviously, the second-order derivative of $F_m(\beta_m)$ is nonpositive, which reveals that $F_m(\beta_m)$ is concave with respect to β_m for $1 < m < M-1$. On this basis, we can further come to the conclusion that $F_{\mathrm{m}}(\beta_m^*(\theta))$ is also concave with respect to θ. This is due to the fact that $\{a_i^*(\theta)\}_{i=1}^{M}$ and $\{\beta_{\mathrm{m}}^*(\theta)\}_{m=1}^{M-1}$ are both affine mappings according to their linear expressions, which preserves the convexity of $F_{\mathrm{m}}(\beta_m^*(\theta))$ about θ. As yet, actually, we have proved that $R(\theta)$ is a strictly concave function with respect to θ, this is because it is the summation about $\log_2(\Gamma_1\theta/\sigma^2+1)$ and $F_{\mathrm{m}}(\beta_m^*(\theta))$, which are all strict concave, and the convexity of functions is preserved by addition operations.

Moreover, for arbitrary ω, $\eta_{EE}^{\max}$, and $\eta_{SE}^{\max}$, both are deterministic quantities. Additionally, $\theta P_T + P_C$ is an affine function of θ. Therefore, since the numerator and denominator of the first term on the RHS of (25) are, respectively, strict concave and affine, it can be easily concluded that the first term of $\lambda'_{EE-SE}(\theta)$ is strict pseudo-concave about θ. Clearly, the second term on the RHS of (25) is also strict concave. Finally, based on the additivity of concave functions, we can declare that $\lambda'_{EE-SE}(\theta)$ is a strict pseudo-concave function about θ. The proof is completed.

For any pseudo-concave function, there is a unique maximizer [20], which guarantees the existence and uniqueness of the global optimal solution of objective (24a), i.e., the root of the equation $d\lambda'_{EE-SE}(\theta)/d\theta = 0$, and the expression for $d\lambda'_{EE-SE}(\theta)/d\theta$ is shown as:

$$\frac{d\lambda'_{EE-SE}(\theta)}{d\theta} = \frac{-P_T\omega}{\eta_{EE}^{\max}(\theta P_T + P_c)^2} R(\theta) + \frac{dR(\theta)}{d\theta} Q(\theta), \tag{28}$$

where

$$\frac{dR(\theta)}{d\theta} = \frac{1}{\ln 2}\left(\frac{\Gamma_M(\mathrm{d}a_M^*(\theta)/d\theta)}{\Gamma_M a_M^*(\theta) + \sigma^2}\right),$$
$$\frac{da_{\mathrm{m}}^*(\theta)}{d\theta} = \begin{cases} \dfrac{A_{\mathrm{m}}}{2^{R_{\mathrm{m}}^{\min}}}\left(1 - \displaystyle\sum_{i=1}^{m-1}\frac{da_i^*(\theta)}{d\theta}\right), & m \ne M \\ 1 - \displaystyle\sum_{i=1}^{M-1}\frac{da_i^*(\theta)}{d\theta}, & m = M \end{cases}. \tag{29}$$

At this time, the outer layer optimization problem is transformed into searching the maximum value of the pseudo-concave function $\lambda'_{EE-SE}(\theta)$, which can turn to the bisection algorithm for help.

4.3. Power Allocation Scheme of Joint Optimization the EE and SE. In this section, we proposed the bisection searching-based power allocation algorithm for the joint optimization of EE and SE, which is shown in Algorithm 1, and the process of it can be described as follows, where t is used to record the number of iterations before the algorithm converges.

At the beginning, the BS calculates the minimum transmit power $P_{\min}$ and then determines whether it is in the feasible range $(P_{\min}/P_T) \le \theta \le 1$. If the condition is satisfied, we can further figure out the root of $\lambda'_{EE-SE}(\theta)$, and the value of $d\lambda'_{EE-SE}(\theta)/d\theta$ with the range $[P_{\min}/P_T, 1]$ can be divided into the following three cases:

Case 1: If $d\lambda'_{EE-SE}(\theta)/d\theta|_{\theta=1} \ge 0$, which indicates that λ'_{EE-SE} keeps increasing in feasible range, and the optimal power consumption factor is $\theta^* = 1$.

Case 2: If $\mathrm{d}\lambda'_{EE-SE}/\mathrm{d}\theta|_{\theta=P_{\min}/P_T} \le 0$, which indicates that λ'_{EE-SE} keeps decreasing in feasible range. At this time, the optimal power consumption factor is $\theta^* = P_{\min}/P_T$.

Case 3: If $\mathrm{d}\lambda'_{EE-SE}/\mathrm{d}\theta|_{\theta=P_{\min}/P_T} > 0$ and $\mathrm{d}\lambda'_{EE-SE}/\mathrm{d}\theta|_{\theta=1} < 0$, which indicates that λ'_{EE-SE} first increases and then decreases. Thus, the optimal power consumption factor is $\theta^* \in [P_{\min}/P_T, 1]$ and can be found by using the bisection searching method.

5. Simulation Results

The validity of Algorithm 1 is verified using computer simulations. We assume that the distance d from all users to

Input: $P_{\min}, P_T, t, \varepsilon$;
Initialization: $\theta_U = 1$, $\theta_L = P_{\min}/P_T$, $t = 1$, $\varepsilon = 10^{-3}$;
Given $\theta_U = 1$, the optimal power allocation coefficient
can be obtained by using (20), and then calculate Q_U:
$Q_U = d\lambda'_{EE-SE}/d\theta|_{\theta=\theta_U}$;
if $Q_u \geq 0$
  $\theta^* = \theta_U$;
else
  Given $\theta_L = P_{\min}/P_T$, the optimal power allocation
  can be obtained by using (20), and then calculate
  Q_L:
  $Q_L = (d\lambda'_{EE-SE}/d\theta)|_{\theta=\theta_L}$;
  if $Q_L \leq 0$
    $\theta^* = \theta_L$;
  else
    $\theta_M = \theta_L + \theta_U/2, Q_M = (d\lambda'_{EE-SE}/d\theta)|_{\theta=\theta_M}$;
    while $|Q_M| \geq \varepsilon$
      if $Q_L \cdot Q_M > 0$
        $\theta_L = \theta_M, Q_L = Q_M$;
      else
        $\theta_U = \theta_M, Q_U = Q_M$;
      end
      t = t + 1;
      $\theta_M = \theta_L + \theta_U/2$, $Q_M = (d\lambda'_{EE-SE}/d\theta)|_{\theta=\theta_M}$;
      end while
      $\theta^* = \theta_M$;
    end
  end
Output: $\theta^* P_T$, t.

Algorithm 1: Bisection searching-based power allocation algorithm.

Table 1: Simulation parameters.

Parameters	Characteristic
Number of users	$M = 2, 3, 4$
Noise power	σ^2 = -70dBm
Path-loss exponent	$\alpha = 3$
Maximum available transmit power of the BS	$P_T = 5 \sim 30$dBm
Circuit power consumption	$P_c = 28.30$dBm

the BS is the same and $d = 80$m. Besides, we also suppose that all users have the same QoS requirements, which means that $R_m^{\min} = R^{\min} = 1bps/\text{Hz}$ for $1 \leq m \leq M$. Finally, the settings of other parameters are shown in Table 1. Note that if the maximum available transmit power of the BS cannot support all users' minimum rate constraints, we default that the system does not work.

Figure 2 reveals the EE versus SE with variable P_T, and the variation range of P_T is $5 \sim 30$ dBm. As P_T increases, it can be seen that EE first increases with SE, but when EE reaches its optimal value, EE begins to go down while we continue to increase the P_T, and the system power utilization will decrease. This indicates that EE and SE are contradictory when the available BS power is large enough, and there is a tradeoff between the two.

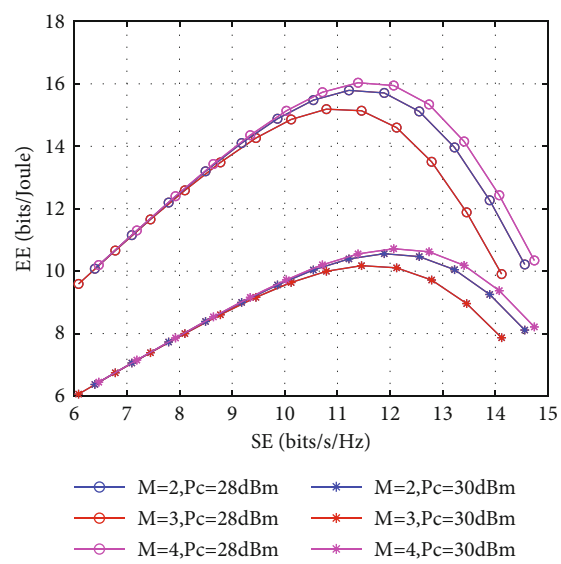


Figure 2: The EE versus SE with different M and P_c..

Figure 3 shows the EE versus P_T, in which M is set to 2, 3, 4. And "max-SE" means that when the system focuses on maximizing SE, it always applies the full power allocation strategy among users. However, from the EE's point of view, it is obviously not optimal. Because when the power is larger

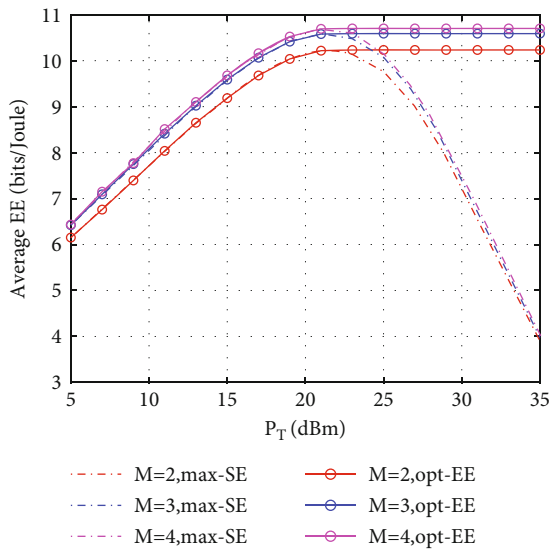


Figure 3: The EE versus P_T of the BS with different M.

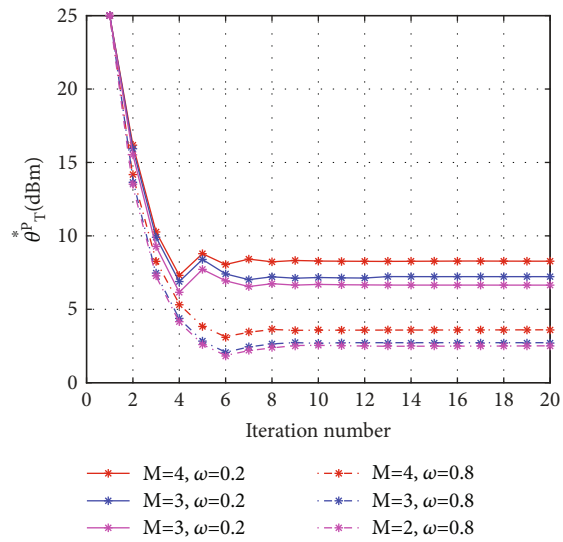


Figure 5: Optimal transmit power $\theta^* P_T$ versus t with different M.

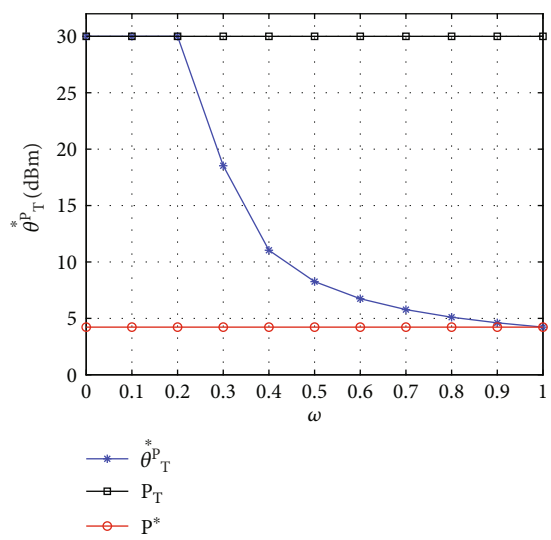


Figure 4: Optimal transmit power $\theta^* P_T$ versus ω.

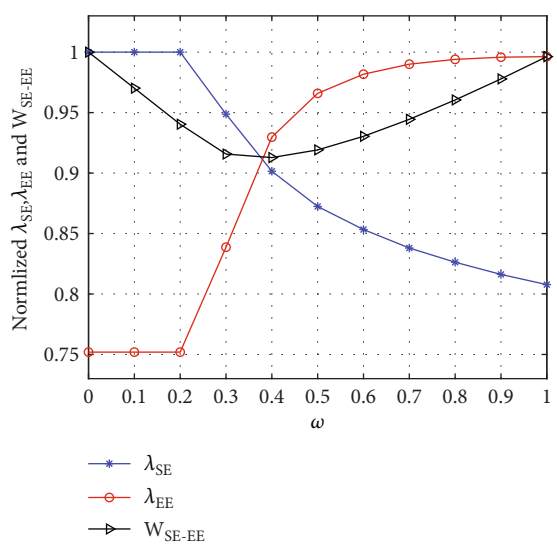


Figure 6: The λ_{EE}, λ_{SE}, and $W(\lambda_{EE}, \lambda_{SE})$ versus ω.

than P^*, EE will increase firstly and then begin to decrease with the increasing of the power. "Opt-EE" means that in the proposed joint optimization algorithm, the system treats EE as an indicator when $\omega = 1$, and the system power can always be maintained at the optimal value that maximizes EE.

Figure 4 depicts the relationship between the optimal transmit power $\theta^* P_T$ and the weighting factor ω. The maximum available transmit power of the BS is set to 30 dBm, i.e., $P_T = 30$dBm, and M is set to 3. It can be seen that as ω increases, $\theta^* P_T$ first remains unchanged at the P_T, and then gradually approaches the P^*. This is because when ω is smaller, the system metric is equivalent to SE, and BS uses as much transmit power as possible to maximize the SE. With the increasing of ω, the system metric gradually transforms to be EE. Therefore, $\theta^* P_T$ gradually approaches P^* that maximizes EE.

Figure 5 reveals $\theta^* P_T$ versus t with different M in the proposed joint optimization algorithm, where P_T is set to 25 dBm. It can be firstly observed that the actual consumed power grows with the increasing of M. This is because the BS needs more power to provide guaranteeing for the increased users' QoS, but actually, the computational complexity is not significantly enhanced. Because the proposed joint algorithm can always achieve convergence around 7 iterations. More importantly, we can find that the value of $\theta^* P_T$ with $\omega = 0.8$ is obviously smaller than that $\omega = 0.2$. This is due to that the main system metric is the SE, and at this time, the BS always tends to use all power to maximize the system capacity. Oppositely, the main system metric is EE when $\omega = 0.8$, the system always tries to keep the EE at the optimal value, and the value is always less than the power value required to maximize the SE.

Figure 6 investigates the λ_{EE} and λ_{SE}, and the joint objective function $W(\lambda_{EE}, \lambda_{SE})$varying with ω. It can be seen that EE gradually increases with ω, while the SE gradually goes down. However, when the value of ω is close to zero, both EE and SE remain unchanged, which is in consistent with the trend of $\theta^* P_T$ demonstrated in Figure 3. In particular, when $\omega = 0$, the maximum SE can be achieved, and the EE reaches the optimal value when $\omega = 1$.

6. Conclusion

In order to better maintain the system performance and reduce energy consumption, we study the joint optimization of EE and SE in a downlink NOMA system in this paper. We first define a new system performance measurement indicator and then transform the original MOO problem into a SOO problem. Under the condition that guarantees users' QoSs, a two-layer bisection searching-based power allocation algorithm is presented finally. Simulation results verify that the proposed power allocation scheme can achieve the opportune tradeoff between EE and SE by flexibly adjusting the weighting factor in the given indicator function from zero to 1. The method is universal and suitable for improving the energy utilization of actual scenarios with time-varying communication demands.

Data Availability

The data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] X. Fang, W. Feng, T. Wei, Y. Chen, N. Ge, and C. X. Wang, "5G embraces satellites for 6G ubiquitous IoT: basic models for integrated satellite terrestrial networks," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14399–14417.

[2] K. Ali, H. X. Nguyen, Q. T. Vien et al., "Review and implementation of resilient public safety networks: 5G, IoT, and emerging technologies," *IEEE Network*, vol. 35, no. 2, pp. 18–25, 2021.

[3] X. Li, Y. Zheng, W. U. Khan et al., "Physical layer security of cognitive ambient backscatter communications for green Internet-of-Things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[4] A. Zappone and E. Jorswieck, "Energy effificiency in wireless networks via fractional programming theory," *Foundations and Trends in Communication and Information Theory*, vol. 11, pp. 185–396, 2014.

[5] X. Li, M. Huang, Y. Liu, V. G. Menon, A. Paul, and Z. Ding, "I/Q imbalance aware nonlinear wireless-powered relaying of B5G networks: security and reliability analysis," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 4, pp. 2995–3008, 2021.

[6] M. Shirvanimoghaddam, M. Dohler, and S. J. Johnson, "Massive non-orthogonal multiple access for cellular IoT: potentials and limitations," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 55–61, 2017.

[7] Y. Liu, X. Li, F. R. Yu, H. Ji, H. Zhang, and V. C. M. Leung, "Grouping and cooperating among access points in user-centric ultra-dense networks with non-orthogonal multiple access," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2295–2311, 2017.

[8] H. Wang, Z. Zhang, and X. Chen, "Energy-efficient power allocation for non-orthogonal multiple access with imperfect successive interference cancellation," in *International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Nanjing, 2017.

[9] A. J. Muhammed, Z. Ma, P. D. Diamantoulakis, L. Li, and G. K. Karagiannidis, "Energy-efficient resource allocation in multi-carrier NOMA systems with fairness," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8639–8654, 2019.

[10] Z. Zhu, N. Wang, W. Hao, Z. Wang, and I. Lee, "Robust beamforming designs in secure MIMO SWIPT IoT networks with a nonlinear channel model," *IEEE Internet of Things Journal (Accepted)*, vol. 8, no. 3, pp. 1702–1715.

[11] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, "QoE-based resource allocation for multi-cell NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6160–6176, 2018.

[12] X. Sun, N. Yang, S. Yan et al., "Joint beamforming and power allocation in downlink NOMA multiuser MIMO networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5367–5381, 2018.

[13] B. Chen, Y. Chen, Y. Cao, N. Zhao, and Z. Ding, "A novel spectrum sharing scheme assisted by secondary NOMA relay," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 732–735, 2018.

[14] Y. Hei, Y. Kou, G. Shi, W. Li, and H. Gu, "Energy-spectral efficiency tradeoff in DCO-OFDM visible light communication system," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9872–9882, 2019.

[15] O. Aydin, E. A. Jorswieck, D. Aziz, and A. Zappone, "Energy-spectral efficiency tradeoffs in 5G multi-operator networks with heterogeneous constraints," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5869–5881, 2017.

[16] W. Zhang, C. Wang, D. Chen, and H. Xiong, "Energy–spectral efficiency tradeoff in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2208–2218, 2016.

[17] Q. Liu, F. Tan, T. Lv, and H. Gao, "Energy efficiency and spectral-efficiency tradeoff in downlink NOMA systems," in *IEEE International Conference on Comm. Workshops (ICC Workshops)*, pp. 247–252, Paris, 2017.

[18] S. Cetinkaya and H. Arslan, "Energy and spectral efficiency tradeoff in NOMA: multi-objective evolutionary spproaches," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Dublin, Ireland, 2020.

[19] X. Wei, H. M. Al-Obiedollah, K. Cumanan, Z. Ding, and O. A. Dobre, *Energy Efficiency Maximization for Hybrid TDMA-NOMA System with Opportunistic Time Assignment*, IEEE Transactions on Vehicular Technology, 2022.

[20] Y. Sun, Z. Ding, X. Dai, and O. A. Dobre, "On the performance of network NOMA in uplink CoMP systems: a stochastic

geometry approach," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5084–5098, 2019.

[21] O. Amin, E. Bedeer, M. H. Ahmed, and O. A. Dobre, "Energy efficiency-spectral efficiency tradeoff: a multiobjective optimization approach," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 1975–1981, 2016.

[22] Y. Zhang, H. Wang, T. Zheng, and Q. Yang, "Energy-efficient transmission design in non-orthogonal multiple access," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2852–2857, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4481416, 9 pages
https://doi.org/10.1155/2022/4481416

Research Article

Few-Shot Specific Emitter Identification Based on Variational Mode Decomposition and Meta-Learning

CunXiang Xie,[1] **LiMin Zhang**,[1] **and ZhaoGen Zhong**[2]

[1]*Department of Information Fusion, Naval Aviation University, Yantai 264001, China*
[2]*The School of Aviation Basis, Naval Aviation University, Yantai 264001, China*

Correspondence should be addressed to ZhaoGen Zhong; zhongzhaogen@163.com

Received 24 May 2022; Revised 30 June 2022; Accepted 2 July 2022; Published 14 July 2022

Academic Editor: Xingwang Li

Recently, deep learning has become the mainstream solution to solve specific emitter identification (SEI) problems. However, because large amounts of labeled signal samples cannot be obtained in noncooperative scenarios, the performance of deep learning-based data-driven methods for SEI was limited. As a result, a novel SEI method targeted on few-shot was proposed in this study. First, the received signal was preprocessed based on variational mode decomposition and the Hilbert analysis to obtain the Hilbert time-frequency spectrum. Subsequently, a classification neural network model was built and trained with a small number of Hilbert time-frequency spectrum samples through meta-learning. This model could identify specific emitters with limited training samples. The experimental results showed that this method accomplishes network training with as few as 80 training samples while obtaining a good level of generalization and effectively identifying different emitter individuals. In addition, this method exhibits a strong robustness to noise by maintaining an identification accuracy of more than 80% in channels with low signal-to-noise ratios. Finally, the proposed method demonstrated better identification performances than other existing methods with its capability to effectively solve SEI problems in the few-shot scenario.

1. Introduction

Specific emitter identification (SEI) refers to a technology that correlates a received signal to its emitting source, only utilizing certain external features of the signal [1–3]. Such external features, often referred to as radiofrequency fingerprints (RFFs) of the specific emitting source, are generated by the nonlinear characteristics of the hardware inside the emitter. They are unique, independent of the signal content, and consistent in different signals from the same emitter. However, signals emitted from different emitters exhibit distinct features, i.e., they are distinguishable RFFs, even if the emitter devices are from the same manufacturer, of the same model, or even of the same batch [4–6].

SEI is a crucial technology for cognitive radio [7, 8]. Dynamic spectrum access (DSA), an effective method of improving spectrum utilization, allows cognitive radio networks to use the current spectral resources more efficiently [9–12]. Moreover, to solve the problem of mutual interferences in DSA networks, identity authentication has usually been used to ensure the security of wireless communication systems. Through SEI, subtle features of the wireless radio signals can be determined with the help of signal processing techniques. Subsequently, together with the critical system, a hardware–software dual authentication system can be established to improve the security performance of the wireless system significantly.

In contrast, wireless radio technologies have been widely used across various industries. Mutual interferences between radio services and spectrum congestions must be prevented and avoided to meet the wireless industry's increasing demand for radio spectrums. For this purpose, legitimate radio stations must be monitored for strict compliance with the designated working parameters. Moreover, pirate radios must be tracked down. By applying SEI techniques to the captured signals in the radio monitoring equipment, it becomes possible to distinguish between the legitimate and pirate radio stations by analyzing the working parameters within the frequency bands allocated to them and extracting their hardware features. In this manner, regulating the usage of spectrum resources can

be achieved. Moreover, it enhances the radio equipment's ability to distinguish users of different nature, as well as in the detection of unknown interference signals.

In recent years, extensive research on SEI has been conducted. Padilla et al. [13] proposed a method to extract RFFs which utilizes the spectrum information of signal preambles, with which several Wi-Fi devices have been successfully identified. Yuan et al. [14] designed an algorithm that can extract 13 features of the time-frequency power spectrum of RF signals to construct a feature vector for SEI. Although redundancies have been found in the extracted features, the represented information has still been considered insufficient, limiting the identification performance. Zhang et al. [15] extracted the energy entropy, first moment, and second moment of the time-frequency power spectrum as RFFs. However, the time-frequency power spectrum has not been partitioned, implying a lack of consideration of the power distribution across the spectrum, the impact of which on the identification performance can hardly be neglected. Satija et al. [16] decomposed the received RF signals into finite modal components through variational mode decomposition (VMD). With the VMD entropy and cumulant constructed as RFFs of the emitter individuals, the SEI problem in a Rayleigh channel has been effectively solved. It is still not possible to establish an accurate mathematical model that explains the generation mechanism of RFFs. Therefore, deep learning techniques have been widely adopted to avoid expert intervention for SEI. Zhang et al. [17] have proposed a method of SEI using a convolutional neural network (CNN). In reference [18], seven ZigBee devices have been successfully identified, utilizing the error of the received baseband signal as the input of CNN. The long short-term memory (LSTM) technique has been adopted in reference [19] to learn the high-order correlation of received signals to identify multiple USRP devices effectively. In reference [20], transient signals have been processed by recurrence plots (RP) transforms, continuous wavelet transforms, and short-time Fourier transforms, respectively, before being fed to the CNN.

Consequently, 12 wireless devices have been successfully identified. Wong et al. [21] proposed a CNN model devised to estimate the gain deviation and phase deviation of the in-phase and quadrature components of the RF signals to identify the emitter individuals. He and Wang [22] proposed three signal preprocessing schemes: empirical mode decomposition (EMD), intrinsic time-scale decomposition (ITD), and VMD. Subsequently, an LSTM network has been used to extract RFFs to perform SEI.

The existing methods have promoted the development of SEI technologies; however, these methods can only achieve good identification performance when there are sufficient training samples. In practice, SEI technologies are mainly used in noncooperative scenarios, where it is difficult to obtain many signal samples with labeled information. Therefore, the traditional deep learning models are incapable of obtaining good training results, resulting in limited identification performance of the system. In this study, an SEI method in the few-shot scenario was proposed.

First, a signal preprocessing method based on VMD and the Hilbert analysis was introduced, which transformed the original RF signal into a Hilbert time-frequency spectrum to highlight its RFFs and thus enhance the distinguishability of different types of RF signals. Next, the classification neural network model was built, with the Hilbert time-frequency spectrum as its training samples, and the meta-learning algorithm was employed to conduct the training. Subsequently, the proposed network's ability to uncover the intrinsic and general characteristics of the signal data under test with only a few training samples was verified. Therefore, the proposed method has good identification performance and can effectively solve the few-shot SEI problem.

The remainder of this paper is organized as follows. In Section 2, we introduce the signal preprocessing scheme based on variational mode decomposition and the Hilbert analysis. In Section 3, we introduce the proposed few-shot SEI method based on meta-learning. In Section 4, we present and discuss the experimental results. Finally, we conclude the paper.

2. Signal Preprocessing

2.1. Variational Mode Decomposition. As a new signal decomposition technique, VMD can decompose nonlinear signals in the time domain and frequency domain to obtain finite modal components [23], which can be realized by modeling the constrained variational problem as follows:

$$\min_{v_k,\omega_k}\left\{\sum_{k=1}^{K}\left\|\frac{\partial}{\partial t}\left[\left(\delta(t)+\frac{j}{\pi t}\right)\times v_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\} \quad \text{s.t.} \sum_{k=1}^{K} v_k(t)=f(t), \tag{1}$$

where $v_k(t)$ represents the k-th modal component, ω_k represents the corresponding central frequency, $f(t)$ represents the original signal, and $\delta(t)$ represents the impulse function.

The constrained variational problem can be solved by the augmented Lagrangian multiplier algorithm, where the augmented Lagrangian function $D(\{v_k(t)\},\{\omega_k\},\lambda(t))$ can be expressed as follows:

$$D(\{v_k(t)\},\{\omega_k\},\lambda(t))=\left(\sum_{k=1}^{K}\left\|\frac{\partial}{\partial t}\left[\left(\delta(t)+\frac{j}{\pi t}\right)\times v_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right) + \left\|f(t)-\sum_{k=1}^{K}v_k(t)\right\|_2^2+\lambda(t)\left(f(t)-\sum_{k=1}^{K}v_k(t)\right), \tag{2}$$

where $\lambda(t)$ represents the Lagrange multiplier. The modal component $\{v_k(t)\}$ can be obtained by solving the saddle point according to the augmented Lagrangian function $D(\{v_k(t)\},\{\omega_k\},\lambda(t))$. The alternate direction method of multipliers (ADMM) algorithm was used to resolve the problem through the following sequence:

Step 1: update the modal component as follows:

$$V_k^{n+1}(\omega)=\frac{F(\omega)-\sum_{j\neq k}V_j^n(\omega)+\Lambda(\omega)/2}{1+2\alpha(\omega-\omega_k^n)^2} \tag{3}$$

Step 2: update the central frequency as follows:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |V_k(\omega)|^2 d\omega}{\int_0^\infty |V_k(\omega)|^2 d\omega} \tag{4}$$

Step 3: update the Lagrange multiplier as follows:

$$\Lambda(\omega) = \Lambda(\omega) + \varsigma \left[F(\omega) - \sum_{k=1}^{K} V_k(\omega) \right] \tag{5}$$

Step 4: repeat Steps 1 to 3 until the following condition is achieved:

$$\sum_{k=1}^{K} \left\| V_k^{n+1}(\omega) - V_k^n(\omega) \right\|_2^2 < \varepsilon, \tag{6}$$

where $F(\omega)$ represents the spectrum of the original signal $f(t)$, $V_k(\omega)$ represents the spectrum of the modal component $v_k(t)$, $\Lambda(\omega)$ represents the spectrum of the Lagrange multiplier $\lambda(t)$, n represents the iterative variable, and ε represents the threshold coefficient (assuming $\varepsilon = 10^{-6}$ in this study).

2.2. Hilbert Analysis. First, the Hilbert transformation is applied [24] to the decomposed modal components as follows:

$$\hat{v}_k(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{v_k(t)}{t - \tau} d\tau. \tag{7}$$

Consequently, the instantaneous amplitude $a_k(t)$ and instantaneous frequency $\omega_k(t)$ of each modal component are calculated as follows:

$$a_k(t) = \sqrt{v_k^2(t) + \hat{v}_k^2(t)}, \tag{8}$$

$$\theta_k(t) = \arctan \frac{\hat{v}_k(t)}{v_k(t)}, \tag{9}$$

$$\omega_k(t) = \frac{d\theta_k(t)}{dt}. \tag{10}$$

The Hilbert time-frequency spectrum of the signal after Hilbert transformation is as follows:

$$H(t, \omega) = \mathrm{Re} \left[\sum_{k=1}^{K} a_k(t) e^{j \int \omega_k(t) dt} \right]. \tag{11}$$

3. Few-Shot Specific Emitter Identification Method Based on Meta-Learning

3.1. Classification Neural Network Model. The classification neural network model [25] was used to process RF signals to extract RFFs comprehensively, and its structure is shown in Figure 1.

The Hilbert time-frequency spectrum of the original RF signal was used as the network input vector $\boldsymbol{x}$. First, it was fed into a CNN of multiple convolution layers (CNN I, II, and III) for feature extraction. The convolution layers contained several convolution kernels, which were used to convolute the feature map from the previous layer. The ReLU function was used to activate the convolution layers and introduce nonlinear factors, which was helpful in the extraction of complex deep features [26, 27]. Subsequently, vectorization of distributed features was realized using the extracted feature maps through the fully connected layers (Dense I and II). Finally, the output vector of the last fully connected layer Dense II was mapped by the softmax function to obtain C different output probabilities, among which the highest probability corresponded to the specific type of the identified emitting source.

The training loss function of the classification neural network model can be expressed as follows:

$$J(\boldsymbol{x}; \theta) = -\frac{1}{C} \sum_{i=1}^{C} \sum_{n=1}^{N} \boldsymbol{y}_i \times \log \left[f_\theta \left(\boldsymbol{x}_i^{(n)} \right) \right], \tag{12}$$

where $\boldsymbol{x}_i^{(n)}$ represents the n-th Hilbert time-frequency spectrum sample corresponding to RF signals of class i, $f_\theta(\cdot)$ represents the mapping function of the classification neural network model with weight parameter θ, and $\boldsymbol{y}_i$ represents the label vector of class i, which is represented by the single one-hot encoding vector.

3.2. Meta-Learning. Traditional machine learning algorithms required multiple signal samples with labels to train the classification neural network model introduced in the previous section. However, in practice, it is challenging to label multiple signal samples. Consequently, unfit training networks and unsatisfactory results were obtained. Therefore, the meta-learning [28] algorithm was proposed, which utilized only a small number of labeled signal samples for network training. Unlike traditional machine learning techniques, the basic training unit of meta-learning is a task [29], generated using the following procedure: (1) assume that the training data set D_{train} contains C classes of signals, and $C_1 (C_1 < C)$ classes of signals are randomly selected to form a training subdataset $D_{sub-train}$; (2) for each class that belongs to C_1, P training samples are randomly sampled to form a support set $S_m = \{(\boldsymbol{x}_{sm}, y_{sm})\}$, which contains a total of $C_1 \times P$ signal training samples; (3) a query set $Q_m = \{(\boldsymbol{x}_{qm}, y_{qm})\}$ is formed by sampling R training samples from the remaining data of each signal class, which contains a total of $C_1 \times R$ signal training samples; (4) finally, a task $T_m = (S_m, Q_m)$is formed by support set S_m and query set Q_m, and accordingly, the universal set of tasks $T = \{T_1, T_2, \cdots, T_m, \cdots, T_M\}$ is obtained for meta-learning.

For the task T_m, the parameter θ of the meta-learner is assigned to a temporary network, which has the same structure as the meta-learner, and its network parameter is defined as φ_m. Next, the temporary network is trained using a support set S_m, whose loss function can be expressed as follows:

$$L_{T_m}((\boldsymbol{x}_{sm}, y_{sm}); \theta) = -\frac{1}{C_1} \sum_{i=1}^{C_1} \sum_{n=1}^{P} \boldsymbol{y}_{sm,i} \times \log \left[f_\theta \left(\boldsymbol{x}_{sm,i}^{(n)} \right) \right]. \tag{13}$$

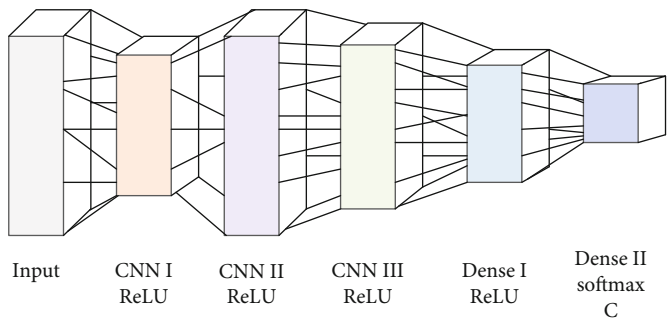


Figure 1: Structure of the classification neural network model.

Consequently, the parameter φ_m of the temporary network can be updated as follows:

$$\varphi_m = \theta - \alpha \cdot \frac{\partial L_{T_m}((\boldsymbol{x}_{sm}, y_{sm}); \theta)}{\partial \theta}, \tag{14}$$

where α represents the learning rate used to train the temporary network.

Subsequently, the temporary network is tested by query set Q_m, and the gradient loss is calculated to optimize the meta-learner's parameters as follows:

$$L_{T_m}\left(\left(\boldsymbol{x}_{qm}, y_{qm}\right); \varphi_m\right) = -\frac{1}{C_1}\sum_{i=1}^{C_1}\sum_{n=1}^{R} \boldsymbol{y}_{qm,i} \times \log\left[f_{\varphi_m}\left(\boldsymbol{x}_{qm,i}^{(n)}\right)\right], \tag{15}$$

$$\theta \longleftarrow \theta - \beta \times \nabla_{\varphi_m} \frac{\partial L_{T_m}\left(\left(\boldsymbol{x}_{qm}, y_{qm}\right); \varphi_m\right)}{\partial \varphi_m}, \tag{16}$$

where β represents the learning rate used to train the meta-learner.

It is discernible that the temporary network focuses on exploring the characteristics of a single task through data analysis and modeling, while the meta-learner specializes in understanding the common feature representation of all tasks through summarizing and analyzing the training results of the temporary network. Specifically, the common feature representation of the tasks represents the general and intrinsic characteristics implied in the training data. A good understanding of such common characteristics acquired by the meta-learner indicates its promising capability for generalization. Accordingly, the meta-learner can provide more optimal initial values to the temporary network on the new tasks, thus making the temporary network achieve training fitting through a few iterations on the condition of few samples. The training results are fed back to the meta-learner to improve its generalization performance further [30]. Therefore, the meta-learning algorithm is considered suitable for few-shot training.

3.3. Algorithm Flow. Based on the theoretical analyses presented in previous sections, the algorithm flow of specific emitter identification based on meta-learning is summarized in this section.

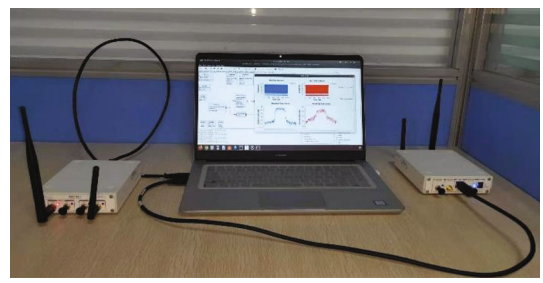


Figure 2: SDR platform.

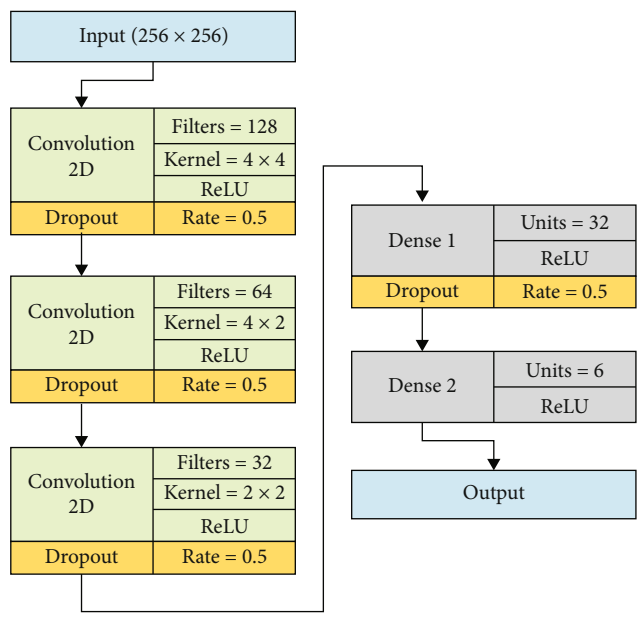


Figure 3: Structure of the classification neural network model.

Step 1: obtain the Hilbert time-frequency spectrum through preprocessing the received RF signals, which is used as the training sample to construct the training data set D_{train} that contains C classes of signals.

Step 2: acquire M meta-learning tasks $T = \{T_1, T_2, \cdots, T_m, \cdots, T_M\}$ by sampling the training data set D_{train}.

Step 3: for task T_m, train the temporary network with support set S_m according to Equation (13); and then update the temporary network parameter φ_m following Equation (14).

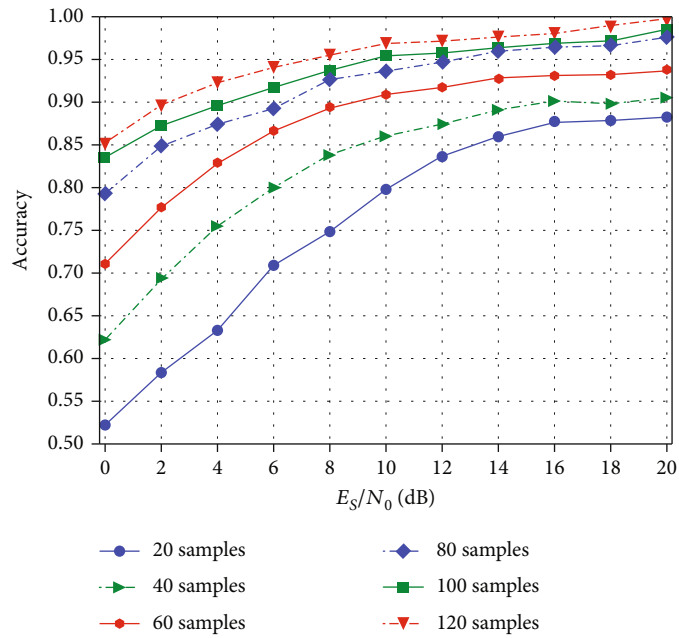


Figure 4: Identification accuracy on different numbers of training samples.

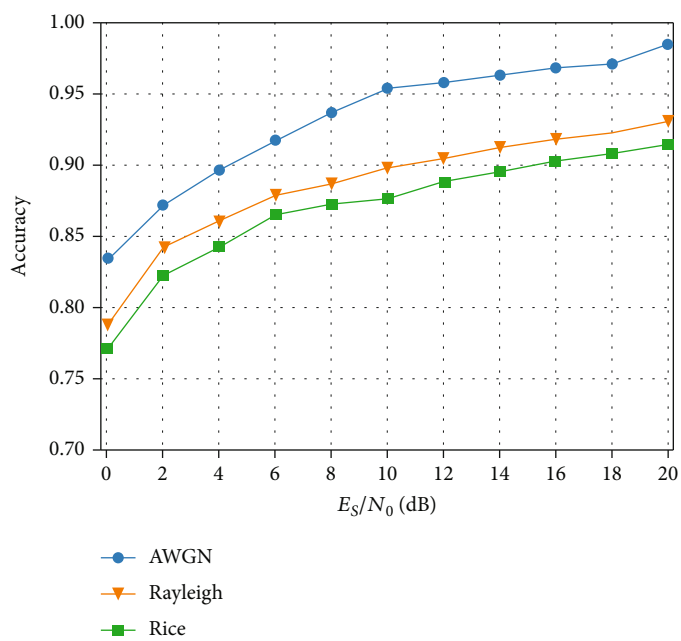


Figure 5: Identification accuracy of different propagation channels.

Step 4: test the temporary network with query set Q_m; calculate the gradient loss based on Equation (15); and then optimize the meta-learner parameters according to Equation (16).

Step 5: repeat Steps 3 and 4 until meta-learning of all tasks is completed.

4. Experiment Results and Analyses

4.1. Data Acquisition and Network Modeling. This section explains the experiments' data acquisition through a software-defined radio (SDR) platform, as shown in Figure 2.

Seven USRP devices of the same model were used to obtain signals from different emitter individuals, with six working as radio emitters and one working as the radio receiver. Each side was connected to a PC, defining either the transmitting or receiving behavior through a GNU radio. The receiver captured six different types of RF signals emitted from the defined emitters. The working frequency of the transmitted signal was set to 2.4 GHz, and the sampling frequency at the receiver end was 16 MHz. QPSK modulation was applied to all the six emitters with a bandwidth of 1.2 MHz. Each received signal was divided to record 1000 units of data, each with 256 sampling points.

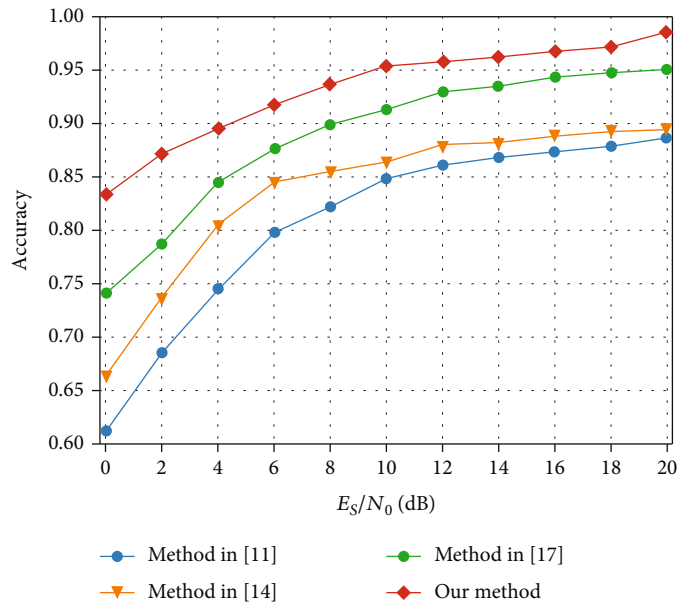


Figure 6: Identification accuracy of different methods.

Signal transmission and capture were carried out in a laboratory environment to ensure a high data quality. It is necessary to test the identification performance under different signal-to-noise (SNR) conditions by artificially adding noises to the signal to verify the algorithm's robustness to noises. The originally captured signal data was sent to MATLAB for the additional noises in the experiments, with the SNR set to 0, 2,…, 20 dB.

The neural network model—comprising the input layer, CNN layers, fully connected layers, and the softmax classification activation function—was built based on the structure presented in Figure 1. The structure of the classification neural network model is shown in Figure 3.

4.2. Impact of the Number of Training Samples on Identification Performance. This section focuses on the impact of the number of training samples on identification performance. We train CNN through meta-learning, where the number of labelled training samples is set to 20, 40, 60, 80, 100, and 120. We then test the identification accuracy under different SNRs, and the identification accuracy curve is shown in Figure 4.

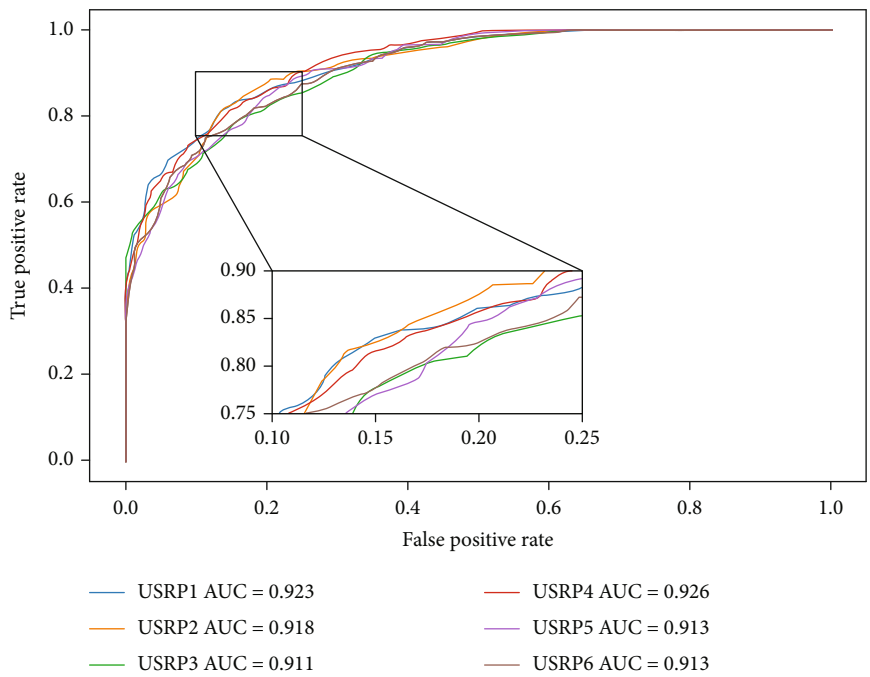
As shown in Figure 4, as the number of labelled samples increases, the overall identification accuracy also increases. However, when the number of training signal samples reaches 80, the identification performance of the system is stable and maintains a high level, indicating that the proposed method can well adapt to the few-shot specific emitter identification task. Therefore, with only a few labelled signal samples for training, the system can exhibit satisfactory identification performances.

4.3. Robustness to Noise. As RFFs are subtle and vulnerable to noise interference during wireless channel transmission, the robustness of the proposed method to noise is crucial to its identification performance. This section prepared different propagation channels with additive white Gaussian noise (AWGN), Rayleigh noise, and Rice noise added to the signal data, respectively. The identification accuracy of each channel under different SNRs was obtained and plotted in Figure 5.

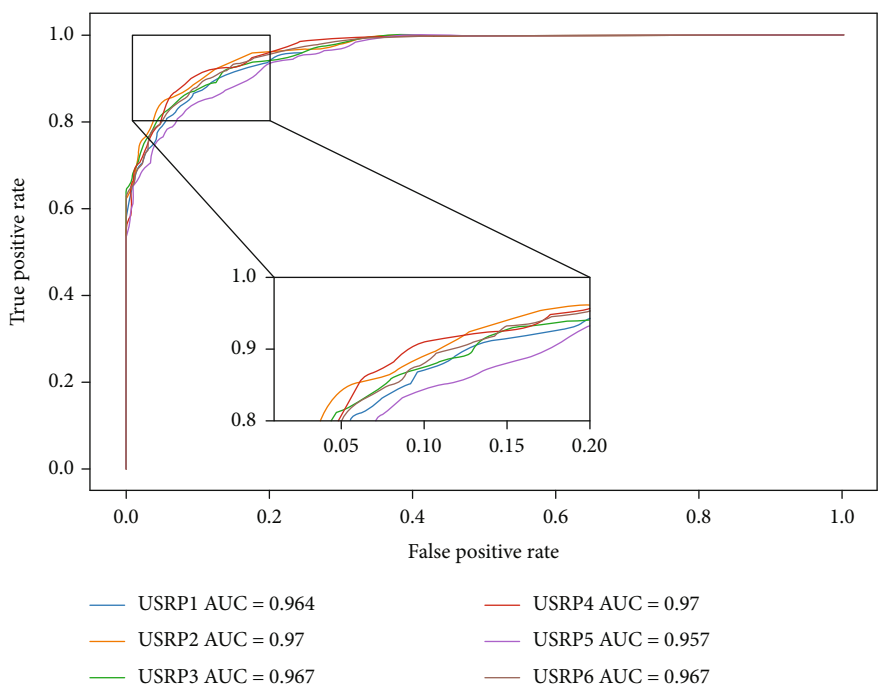
As shown in Figure 5, the best identification performance is achieved in the AWGN channel, with an identification accuracy of 90% at 4 dB and more than 95% at 10 dB. In contrast, the identification performances were poorer in both Rayleigh and Rice channels because the noises in Rayleigh and Rice channels are multiplicative noises, which have a more significant impact on RFFs extraction than the additive noises in AWGN channels. However, under the interference of multiplicative noises, the system's identification accuracy did not deteriorate further. For the Rayleigh channel, an accuracy of 85% at 4 dB and 90% at 10 dB was observed. For the Rice channel, the accuracy at 4 dB and 14 dB was 85% and 90%, respectively. The experimental results implied that the proposed method was robust to noises of different wireless transmission channels.

4.4. Other Indices to Evaluate the Identification Performance. The identification accuracy is defined as the ratio of the number of correctly classified samples to the total number of samples under test, which can only evaluate the identification performance of the system as a whole rather than the identification performance of each type of signal. In this section, the receiver operating characteristic (ROC) was used as an index to further evaluate the identification performance of the proposed method, as shown in Figure 6, with the SNR set to 6, 12, and 20 dB, respectively.

As shown in Figure 7, all the ROC curves of the proposed method are generally distributed in the upper left region of the plot—the higher the SNR, the more concentrated the ROC curve is within the region. Both the high true-positive rate and a low false-positive rate achieved for identifying each USRP device show the suggested method's promising performance in the area of SEI. In addition, the

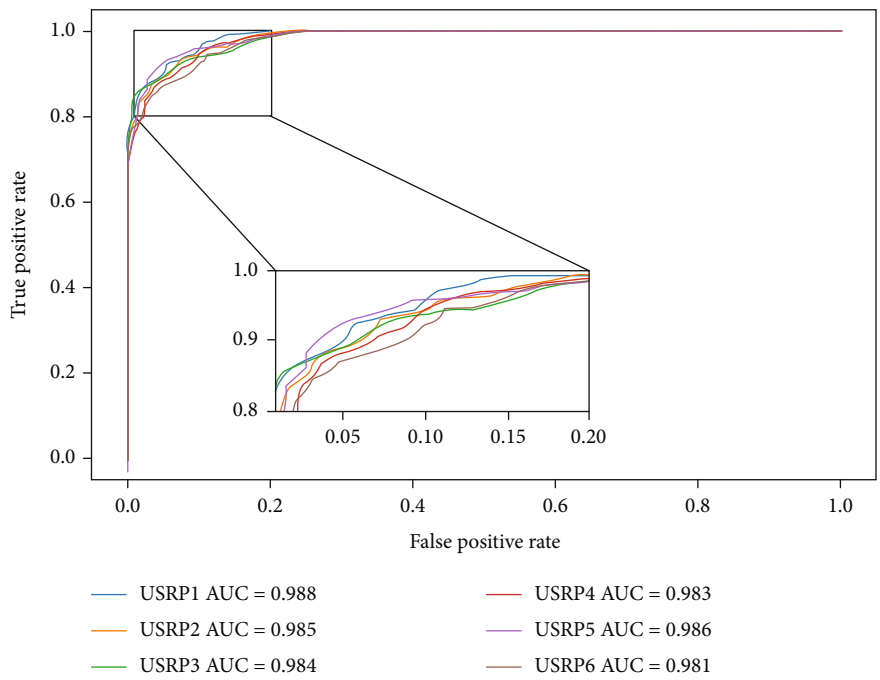


(a)



(b)

Figure 7: Continued.



(c)

Figure 7: ROC curves of the proposed method at 6 dB, 12 dB, and 20 dB, respectively. (a) SNR = 6 dB. (b) SNR = 12 dB. (c) SNR = 20 dB.

area under the curve (AUC) was calculated. In Figure 6, the AUC value of each ROC curve is greater than 0.91 at 6 dB, greater than 0.95 at 12 dB, and greater than 0.98 at 20 dB, further proving the good identification performance of the proposed method quantitatively.

4.5. Comparison with Other Identification Methods. The above experiments proved the effectiveness and efficiency of the proposed method in SEI for a few-shot scenario. This section benchmarks the study with those proposed in references [11, 14, 17] to examine its advantages. The experimental results of the methods mentioned above are shown in Figure 6.

As shown in Figure 6, the identification accuracy of this method is higher than that of other methods in each defined SNR scenario, and the overall identification accuracy is 5–15% higher than the other methods. In addition, in a low SNR scenario, the proposed method can still achieve an identification accuracy above 0.8, whereas all the other methods showed deteriorated performances. The proposed method outperforms the experiments due to its adoption of meta-learning, which can thoroughly learn the general and intrinsic characteristics of the signals. On the contrary, the methods for benchmarking rely on a large quantity of samples to train the neural networks and establish simple mapping relationships between signal data and classes. When the number of training samples is reduced, the generalization ability of these networks is degraded, and their identification performances are affected.

5. Conclusion

In this study, a method of few-shot specific emitter identification is proposed. First, the received RF signal was preprocessed based on VMD. Next, the Hilbert time-frequency spectrum was obtained through the Hilbert analysis of the modal components. Such spectrums were used as the training samples of the classification neural network to carry out meta-learning to ensure that the neural network achieves good training results with a limited number of training samples. The experimental results showed that this method could achieve excellent identification performance and strong robustness to noise with only 80 training samples. Finally, the proposed method demonstrated better identification performances than other existing methods because it can effectively solve SEI problems in the few-shot scenario.

In future work, focus will be on identifying specific emitters in open-set few-shot scenarios. In other words, the method used to detect open-set RF signals while normally identifying different types of closed-set RF signals will be investigated based on the results of this study. In addition, real-world RF signals were captured for experimental verifications to justify further and improve the versatility of the proposed method.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grant 91538201, in part by the Taishan Scholar Project of Shandong Province under grant

ts201511020, and in part by the Project supported by Chinese National Key Laboratory of Science and Technology on Information System Security under grant 6142111190404.

References

[1] L. Li, H.-B. Ji, and L. Jiang, "Quadratic time–frequency analysis and sequential recognition for specific emitter identification," *IET Signal Process*, vol. 5, no. 6, pp. 133–568, 2011.

[2] S. Rajendran, Z. Sun, F. Lin, and K. Ren, "Injecting reliable radio frequency fingerprints using metasurface for the internet of things," *IEEE Transactions on Information Forensics and Security*, vol. 32, no. 1, pp. 27–33, 2007.

[3] K. Sa, D. Lang, C. Wang, and Y. Bai, "Specific emitter identification techniques for the internet of things," *IEEE Access*, vol. 8, pp. 1644–1652, 2020.

[4] K. I. Talbot, P. R. Duley, and M. H. Hyatt, "Specific emitter identification and verification," *Technology Review Journal*, vol. 113, pp. 130–133, 2003.

[5] A. E. Spezio, "Electronic warfare systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 633–644, 2002.

[6] M. Liu and J. F. Doherty, "Nonlinearity estimation for specific emitter identification in multipath channels," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1076–1085, 2011.

[7] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[8] S. Khoshabi Nobar, K. Adli Mehr, and J. Musevi Niya, "RF-powered green cognitive radio networks: architecture and performance analysis," *IEEE Communications Letters*, vol. 20, no. 2, pp. 296–299, 2016.

[9] A. Ghosh and W. Hamouda, "Cross-layer antenna selection and channel allocation for MIMO cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3666–3674, 2011.

[10] V. Shakhov and I. Koo, "Analysis of a network stability-aware clustering protocol for cognitive radio sensor networks," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 12476-12477, 2021.

[11] O. B. Akan, O. B. Karli, and O. Ergul, "Cognitive radio sensor networks," *IEEE Network*, vol. 23, no. 4, pp. 34–40, 2009.

[12] M. Zheng, C. Wang, M. Du, L. Chen, W. Liang, and H. Yu, "A short preamble cognitive MAC protocol in cognitive radio sensor networks," *IEEE Sensors Journal*, vol. 19, no. 15, pp. 6530–6538, 2019.

[13] P. Padilla, J. L. Padilla, and J. F. Valenzuela-Valdés, "Radiofrequency identification of wireless devices based on RF fingerprinting," *Electronics Letters*, vol. 49, no. 22, pp. 1409-1410, 2013.

[14] Y. J. Yuan, Z. T. Huang, H. Wu, and X. Wang, "Specific emitter identification based on Hilbert–Huang transform-based time–frequency–energy distribution features," *IET Communications*, vol. 8, no. 13, pp. 2404–2412, 2014.

[15] J. Zhang, F. Wang, O. A. Dober, and Z. Zhong, "Specific emitter identification via Hilbert-Huang transform in single-hop and relaying scenarios," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1192–1205, 2016.

[16] U. Satija, N. Trivedi, G. Biswal, and B. Ramkumar, "Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 581–591, 2019.

[17] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074–11082, 2017.

[18] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.

[19] Q. Wu, C. Feres, D. Kuzmenko et al., "Deep learning based RF fingerprinting for device identification and wireless security," *Electronics Letters*, vol. 54, no. 24, pp. 1405–1407, 2018.

[20] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks," *Electronics Letters*, vol. 55, no. 2, pp. 90–92, 2019.

[21] L. J. Wong, W. C. Headley, and A. J. Michaels, "Specific emitter identification using convolutional neural network-based IQ imbalance estimators," *IEEE Access*, vol. 7, pp. 33544–33555, 2019.

[22] B. X. He and F. G. Wang, "Cooperative specific emitter identification via multiple distorted receivers," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3791–3806, 2020.

[23] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[24] H. Li, D. Xiao, S. Kwong, L. Yang, D. Huang, and D. Xiao, "Hilbert-Huang transform for analysis of heart rate variability in cardiac health," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 6, pp. 1557–1567, 2011.

[25] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[26] L. Han, J. Z. Sun, and W. Zhang, "Convolutional neural network for convective storm nowcasting using 3-D doppler weather radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 2, pp. 1487–1495, 2019.

[27] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[28] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," *In Proceeding of the 34th International Conference on Machine Learning*, vol. 70, pp. 1126–1135, 2017.

[29] M. Andrychowicz, M. Denil, S. Gomez et al., "Learning to learn by gradient descent by gradient descent," *Advances in Neural Information Processing Systems*, vol. 29, pp. 3981–3989, 2016.

[30] T. Y. Mu, H. Z. Wang, C. N. Wang, Z. Liang, and X. Shao, "Auto-CASH: a meta-learning embedding approach for autonomous classification algorithm selection," *Information Sciences*, vol. 591, pp. 344–364, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2419430, 12 pages
https://doi.org/10.1155/2022/2419430

Research Article

Security-Reliability Tradeoff Analysis of Untrusted Full-Duplex Relay Networks

Xingang Zhang,[1] Dechuan Chen,[2,3,4] Jin Li,[2] Zhipeng Wang,[2] and Xiaotan Li[5]

[1]*The College of Computer Science and Technology, Nanyang Normal University, Nanyang 473061, China*
[2]*The College of Physics and Electronic Engineering, Nanyang Normal University, Nanyang 473061, China*
[3]*The Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications), Ministry of Education, Nanjing 210003, China*
[4]*The Henan Engineering Research Center for Radio Frequency Front End and Antenna of Millimeter Wave Wireless Communication System, Nanyang 473061, China*
[5]*The College of Mechanical and Electrical Engineering, Sichuan Agricultural University, Yaan 625014, China*

Correspondence should be addressed to Dechuan Chen; chenchuan927@163.com

Received 24 May 2022; Revised 19 June 2022; Accepted 29 June 2022; Published 11 July 2022

Academic Editor: Xingwang Li

In this work, we investigate the physical layer security problem of wireless cooperative network, where the communication from a source to a destination is assisted by an untrusted full-duplex amplify-and-forward (AF) relay. In order to realize a positive secrecy rate, cooperative jamming is exploited at the destination. The secrecy outage probability (SOP), connection outage probability (COP), and effective secrecy throughput (EST) are, respectively, derived in closed-form expressions for the fixed gain relaying (FGR) scheme and the variable gain relaying (VGR) scheme. Subsequently, we conduct an asymptotic analysis for both the relaying schemes in the high signal-to-noise ratio (SNR) region to offer valuable insights into practical design. Theory and simulation results demonstrate that the self-interference caused by full-duplex transmission is harmful for the reliability performance, while it is beneficial to the security performance. Moreover, the EST of the considered system increases as the self-interference level decreases, which implies that the self-interference deteriorates an overall performance in terms of the security-reliability tradeoff.

1. Introduction

Due to the broadcast nature of wireless transmission, the communications between legitimate users are particularly vulnerable to be overheard by malicious users [1, 2]. Traditionally, cryptographic mechanism in the upper layers of the network based on computational complexity has been employed to guarantee secure communication in wireless networks [3]. However, as the computing power of malicious users is continuously enhancing (e.g., by adopting the quantum computing technology), cryptographic mechanism is being increasingly challenged. Unlike the cryptographic mechanism, physical layer security, which explores the time-varying properties of physical channels to achieve perfect security, has attracted considerable attention from academia and industry [4, 5].

In recent years, full-duplex relay techniques, which can transmit and receive simultaneously in the same frequency band, have been exploited to improve physical layer security performance of wireless networks. The authors in [6] proposed a newly full-duplex jamming relay scheme, which has better secrecy performance than the half duplex scheme. For full-duplex two-way relay networks, [7] exploited artificial noise to strengthen the sum secrecy rate. In [8], a power allocation strategy between data and jamming was developed to effectively improve the tradeoff between security and reliability, where both the relay and the destination operated in full-duplex mode. Subsequently, full-duplex

techniques were extended to secure multirelay networks [9] and nonorthogonal multiple access (NOMA) networks [10].

The works in [6–10] were focused on the security of full-duplex relay networks, where the eavesdroppers are external nodes in addition to legitimate nodes. However, even in the absence of external eavesdroppers, the security caused by legitimate nodes in the network may still be a concern [11–13]. For example, the relay may be untrusted, which means that they can possibly try to eavesdrop the source's confidential information. The untrusted relay is also a potential eavesdropper, even though it complies with the communication protocols to assist the source forwarding information to the destination [14–16]. In practice, untrusted relay scenario may occur in Internet of Things (IoT), government intelligence networks, and device-to-device (D2D) communications, where not all users do have the same security clearance.

Recently, a few works about untrusted full-duplex relay networks have been considered in the context of physical layer security. Based on the source jamming scheme, [17] derived the secrecy outage probability for untrusted full-duplex relay networks. The impact of outdated channel state information (CSI) on the ergodic secrecy rate of untrusted full-duplex relay networks was investigated in [18]. The authors of [19] proposed an artificial noise-aided secure transmission scheme to confuse the untrusted full-duplex relay and external eavesdroppers. In [20], an optimal beamforming scheme was design to maximize the secrecy sum rate for full-duplex multiple-input multiple-output two-way untrusted relay networks. The aforementioned contributions are mainly concerned on enhancing the secrecy rate of untrusted full-duplex relay networks without paying much attention to the communication reliability. In particular, the impact of self-interference caused by full-duplex transmission on the security-reliability tradeoff is indeed worthy of our attention.

Motivated by the above considerations, we explore the physical layer security of untrusted full-duplex relay networks, where a source operates in half duplex mode and an untrusted relay and a destination operate in a full-duplex mode. In addition, the source equipped with multiple antennas uses a maximal ratio transmission (MRT) scheme to increase reliability. Depending on the availability of the CSI, the relay amplifies and forwards the source's signal with a fixed-gain or variable-gain factor. Different from the source-based jamming scheme in [17], we employ destination-assisted jamming for secrecy improvement. Compared with [18–20], we consider the outage performance of the untrusted full-duplex relay network, where the secrecy outage probability (SOP) is used to evaluate the security performance and the connection outage probability (COP) is used to characterize the reliability performance. The main contributions of this paper are summarized as follows:

(i) We first present exact expressions for the SOP and COP of the fixed gain relaying (FGR) scheme and the variable gain relaying (VGR) scheme and then derive closed-form expressions for the EST of both the relaying schemes, which provide an efficient means to evaluate the reliability and security performance comprehensively

(ii) To gain further insights, we conduct an asymptotic analysis in the high signal-to-noise ratio (SNR) region for the SOP, COP, and EST. Based on the asymptotic analysis, we find that there exists a EST floor for both the relaying schemes in the high SNR region and the floor of the VGR scheme is higher than that of the FGR scheme

(iii) The results demonstrate the intuition that the self-interference caused by full-duplex transmission is harmful for the reliability performance, while it is beneficial to the security performance. In addition, the EST of the considered system increases as the self-interference level decreases, which implies that the self-interference deteriorates an overall performance in terms of the security-reliability tradeoff

The remainder of this paper is organized as follows. Section 2 gives the system model and describes the key metric in evaluating the performance of physical layer security, including the SOP, COP, and EST. In Section 3, we present the closed form expressions of the SOP, COP, and EST for the VGR scheme and the FGR scheme and analyze the corresponding asymptotic behavior in the high SNR region. Numerical results and conclusions are, respectively, provided in Sections 4 and 5.

2. System Model

We consider an untrusted full-duplex relay network shown in Figure 1, where a source S communicates with a destination D via an untrusted relay R. The source has N_s antennas and operates in the half-duplex mode, while the destination and the relay are equipped with dual antennas and operate in the full-duplex mode. It means that the destination and the relay can receive and send data concurrently. The system considered in this work is applicable to numerous practical full-duplex relay scenarios, for instance, the wireless sensor networks and IoT. Moreover, the direct link between the source and the destination is unavailable, due to severe shadow fading or significant obstacle [21]. Thus, source-destination communication can be performed only through the relay. In addition, all channels are modeled as quasistatic block fading channels, following the Rayleigh distribution. As in [13, 15], we assume that the channel between the relay and the destination is reciprocal.

Because the relay is untrusted, the source wishes to keep the information secret from the relay while simultaneously making use of it to forward the information. To maintain the confidentiality of the source information, the destination sends a jamming signal to the relay when the source transmits information to the relay. Thus, the received signal at the relay can be expressed as

$$y_r[i] = h_{sr} w_{sr} x_s[i] + h_{dr} x_d[i] + h_{rr} x_r[i] + n_r[i], \tag{1}$$

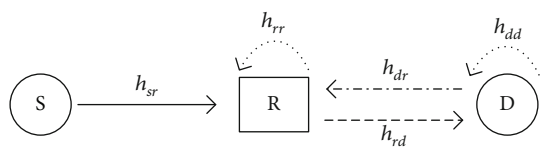


Figure 1: System model, where S, R, and D denote the source, untrusted relay, and destination, respectively.

where h_{sr} is the $1 \times N_s$ channel vector between the source and the relay and its entries follow independent and identically distributed (i.i.d.) complex Gaussian distribution with zero mean and variance $\bar{\gamma}_{SR}$. $w_{sr} = h_{sr}^{\dagger}/\|h_{sr}\|$ is the transmit precoding vector at the source. h_{dr} is the channel coefficient between the relay and the destination with parameter $\bar{\gamma}_{RD}$. h_{rr} is the loopback interference channel coefficient between transmitting and receiving antennas at the relay. $x_s[i]$ is the confidential signal transferred from the source at time slot i and satisfies $E\{|x_s[i]|^2\} = \beta P$, where $E\{\cdot\}$ represents the expectation operation. $x_d[i]$ is the jamming signal transferred from the destination at time slot i and satisfies $E\{|x_d[i]|^2\} = (1-\beta)P$. Here, $\beta \in (0,1)$ is the power allocation factor between the source and the destination. $x_r[i]$ is the loopback interference signal transferred from the relay at time slot i and satisfies $E\{|x_r[i]|^2\} = P$. $n_r[i]$ is the additive white Gaussian noises (AWGN) at the relay with zero mean and variance N_0. Since $x_r[i]$ is known to the relay, it can apply multiple stage interference cancellation methods to alleviate the loopback interference. Thus, the received signal at the relay after compensation can be expressed as

$$\hat{y}_r[i] = h_{sr} w_{sr} x_s[i] + h_{dr} x_d[i] + I_r[i] + n_r[i], \tag{2}$$

where $I_r[i]$ can be modeled as a Gaussian random variable with mean zero and variance $l_r^2 P$ [18, 22]. l_r corresponds to the self-interference cancellation capability at the relay.

Then, the relay amplifies the signal $\hat{y}_r[i]$ with factor G and forwards it to the destination. The two most common factors employed at the relay are so-called fixed gain factor and variable gain factor. In the FGR scheme, the factor is constant based on the statistical CSI, resulting in an output signal with variable power. In the VGR scheme, the factor is variable depended on the instantaneous CSI, resulting in an output signal with fixed power. The factors of both FGR and VGR schemes can be given by

$$G = \begin{cases} E\left(\sqrt{\dfrac{P}{\beta P\|h_{sr}\|^2 + (1-\beta)P|h_{dr}|^2 + l_r^2 P + N_0}}\right), & \text{with FGR,} \\ \sqrt{\dfrac{P}{\beta P\|h_{sr}\|^2 + (1-\beta)P|h_{dr}|^2 + l_r^2 P + N_0}}, & \text{with VGR.} \end{cases} \tag{3}$$

Hence, the received signal at the destination is given by

$$y_d[i] = h_{rd} x_r[i] + h_{dd} x_d[i] + n_d[i] = h_{rd} G \hat{y}_r[i] + h_{dd} x_d[i] + n_d[i], \tag{4}$$

where h_{dd} is the loopback interference channel coefficient between transmitting and receiving antennas at the destination and $n_d[i]$ is the AWGN at the destination with zero mean and variance N_0. Analogously, the received signal at the destination after self-interference cancellation can be written as

$$\hat{y}_d[i] = h_{rd} G h_{sr} w_{sr} x_s[i] + h_{rd} G I_r[i] + h_{rd} G n_r[i] + I_d[i] + n_d[i], \tag{5}$$

where $I_d[i]$ can be modeled as a Gaussian random variable with mean zero and variance $l_d^2(1-\beta)P$. l_d corresponds to the self-interference cancellation capability at the destination.

According to (2), the achievable eavesdropping rate with the FGR scheme and the VGR scheme can be obtained as

$$C_R^{\mathrm{FGR}} = C_R^{\mathrm{VGR}} = \log_2\left(1 + \frac{\beta\lambda\|h_{sr}\|^2}{(1-\beta)\lambda|h_{rd}|^2 + l_r^2\lambda + 1}\right), \tag{6}$$

where $\lambda = P/N_0$ denotes the transmit SNR. According to (3) and (5), the achievable data rate with the FGR scheme and the VGR scheme can be, respectively, obtained as

$$C_D^{\mathrm{FGR}} = \log_2\left(1 + \frac{\beta\lambda^2\|h_{sr}\|^2|h_{rd}|^2}{(\lambda + \lambda^2 l_r^2)|h_{rd}|^2 + C}\right), \tag{7}$$

$$C_D^{\mathrm{VGR}} = \log_2\left(1 + \frac{\beta\lambda^2\|h_{sr}\|^2|h_{rd}|^2}{C_1\|h_{sr}\|^2 + C_2|h_{rd}|^2 + C_3}\right), \tag{8}$$

where $C = (1 + l_r^2\lambda + \beta\lambda N_s\bar{\gamma}_{SR} + (1-\beta)\lambda\bar{\gamma}_{RD})(1 + l_d^2\lambda(1-\beta))$, $C_1 = \beta\lambda(1 + l_d^2\lambda(1-\beta))$, $C_2 = \lambda + \lambda^2 l_r^2 + \lambda(1-\beta)(1 + l_d^2\lambda(1-\beta))$, and $C_3 = (1 + l_r^2\lambda)(1 + l_d^2\lambda(1-\beta))$.

As discussed in [23], when the achievable eavesdropping rate is larger than the positive rate difference between the codeword transmission rate R_0 and the confidential information rate R_s, there will be a secrecy outage event. When the achievable data rate is less than R_0, there will be a connection outage event. Thus, we can formulate the SOP and COP of our system as

$$P_{so} = \Pr(C_R > R_0 - R_s), \tag{9}$$

$$P_{co} = \Pr(C_D < R_0). \tag{10}$$

The SOP and COP give practical insights into security and reliability performance of the transmission system, respectively. To establish the direct relationship between security and reliability, we adopt effective secrecy throughput (EST)

to measure the overall efficiency of our system, which is given by

$$\varsigma = R_s \cdot \Pr(C_R < R_0 - R_s, C_D > R_0). \tag{11}$$

3. Secrecy Performance Analysis

In this section, we endeavor to analyze both the reliability and security performance comprehensively for the untrusted full-duplex relay network. Specifically, we first derive closed form expressions of the SOP, COP, and EST for the FGR scheme and the VGR scheme. Then, the asymptotic analysis of the SOP, COP, and EST is provided to reveal additional insights on the secrecy performance.

3.1. Fixed Gain Relaying

3.1.1. Secrecy Outage Probability. Combining (6) and (9), the SOP of the untrusted full-duplex relay network with the FGR scheme can be given by

$$\begin{aligned} P_{so}^{\mathrm{FGR}} &= \Pr\left(\frac{\beta\lambda\|h_{sr}\|^2}{(1-\beta)\lambda|h_{rd}|^2 + l_r^2\lambda + 1} > t_1\right) \\ &= \Pr\left(\|h_{sr}\|^2 > \frac{(1-\beta)t_1|h_{rd}|^2}{\beta} + \frac{l_r^2 t_1\lambda + t_1}{\beta\lambda}\right), \end{aligned} \tag{12}$$

where $t_1 = 2^{R_0 - R_s} - 1$. Noting that $X = \|h_{sr}\|^2$ is a central chi-square distribution random variable with $2N_s$ degrees of freedom, its cumulative distribution function (CDF) is given by

$$F_X(x) = 1 - e^{-x/\bar{\gamma}_{SR}} \sum_{m=0}^{N_s-1} \frac{x^m}{m!\bar{\gamma}_{SR}^m}. \tag{13}$$

On the other hand, $Y = |h_{rd}|^2$ is an exponential variable with probability density function (PDF) given by

$$f_Y(y) = \frac{1}{\bar{\gamma}_{RD}} e^{-y/\bar{\gamma}_{RD}}. \tag{14}$$

Based on (13) and (14), we can rewrite (12) as

$$\begin{aligned} P_{so}^{\mathrm{FGR}} &= \int_0^\infty \left(1 - F_X\left(\frac{(1-\beta)t_1 y}{\beta} + \frac{l_r^2 t_1\lambda + t_1}{\beta\lambda}\right)\right) f_Y(y)dy \\ &= \sum_{m=0}^{N_s-1} \frac{e^{-t_1 l_r^2\lambda + t_1/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m\beta^m\lambda^m} \int_0^\infty e^{-(((1-\beta)t_1/\beta\bar{\gamma}_{SR}) + (1/\bar{\gamma}_{RD}))y} \\ &\quad \times \left(t_1 + t_1 l_r^2\lambda + (1-\beta)t_1\lambda y\right)^m dy. \end{aligned} \tag{15}$$

Then, using $(t_1 + t_1 l_r^2\lambda + (1-\beta)t_1\lambda y)^m = \sum_{n=0}^m \binom{m}{n} (t_1 + t_1 l_r^2\lambda)^{m-n}(1-\beta)^n t_1^n \lambda^n y^n$, we can calculate P_{so}^{FGR} as

$$\begin{aligned} P_{so}^{\mathrm{FGR}} &= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \frac{\left(t_1 l_r^2\lambda + t_1\right)^{m-n}(1-\beta)^n t_1^n\lambda^n}{(m-n)!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m\beta^m\lambda^m} \\ &\quad \times \left(\frac{(1-\beta)t_1}{\beta\bar{\gamma}_{SR}} + \frac{1}{\bar{\gamma}_{RD}}\right)^{-n-1} e^{-t_1 l_r^2\lambda + t_1/\beta\lambda\bar{\gamma}_{SR}}. \end{aligned} \tag{16}$$

From (16), we find that the SOP increases as the power allocation factor β increases from 0 to 1. This is because that increasing β results in decreasing transmission power of the jamming signal. Moreover, the SOP decreases as l_r increases, which means that the self-interference caused by full-duplex transmission is beneficial to the security performance.

3.1.2. Connection Outage Probability. Combining (7) and (10), the COP of the untrusted full-duplex relay network with the FGR scheme can be given by

$$P_{co}^{\mathrm{FGR}} = \Pr\left(\frac{\beta\lambda^2\|h_{sr}\|^2|h_{rd}|^2}{(\lambda + \lambda^2 l_r^2)|h_{rd}|^2 + C} < t_2\right) = \Pr\left(\|h_{sr}\|^2 < \frac{t_2(1+\lambda l_r^2)}{\beta\lambda} + \frac{t_2 C}{\beta\lambda^2|h_{rd}|^2}\right), \tag{17}$$

where $t_2 = 2^{R_0} - 1$. Based on (13) and (14), we can rewrite (17) as

$$\begin{aligned} P_{co}^{\mathrm{FGR}} &= \int_0^\infty F_X\left(\frac{t_2(1+\lambda l_r^2)}{\beta\lambda} + \frac{t_2 C}{\beta\lambda^2 y}\right) f_Y(y)dy \\ &= 1 - \sum_{m=0}^{N_s-1} \frac{e^{-t_2\lambda l_r^2 + t_2/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m} \int_0^\infty \left(\frac{t_2 + t_2\lambda l_r^2}{\beta\lambda} + \frac{t_2 C}{\beta\lambda^2 y}\right)^m \\ &\quad \times e^{-\left(t_2 C/\beta\lambda^2\bar{\gamma}_{SR}y\right) - \left(y/\bar{\gamma}_{RD}\right)} dy. \end{aligned} \tag{18}$$

Then, with the help of the binomial theorem and ([24], 3.471.9), the exact COP of the untrusted full-duplex relay network with the FGR scheme can be derived as

$$\begin{aligned} P_{co}^{\mathrm{FGR}} &= 1 - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \binom{m}{n} \frac{2t_2^m C^n\left(1+\lambda l_r^2\right)^{m-n}}{m!\beta^m\lambda^{m+n}\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m} e^{-t_2\lambda l_r^2 + t_2/\beta\lambda\bar{\gamma}_{SR}} \\ &\quad \times \left(\frac{t_2 C\bar{\gamma}_{RD}}{\beta\lambda^2\bar{\gamma}_{SR}}\right)^{1(-n)/2} K_{1-n}\left(\sqrt{\frac{4t_2 C}{\beta\lambda^2\bar{\gamma}_{SR}\bar{\gamma}_{RD}}}\right), \end{aligned} \tag{19}$$

where $K_{1-n}(\cdot)$ is the Bessel functions of imaginary argument. From (19), we find that the COP decreases as the power allocation factor β increases from 0 to 1. The main reason is that when β increases, the achievable data rate is larger which makes reliability better. In addition, the COP increases as l_r or l_d increases, which means that the self-interference caused by full-duplex transmission is harmful for the reliability performance.

3.1.3. Effective Secrecy Throughput. Combining (6), (7), and (11), the EST of the untrusted full-duplex relay network with the FGR scheme can be given by

$$\varsigma^{\mathrm{FGR}} = R_s \Pr\left(\|h_{sr}\|^2 < \frac{(1-\beta)t_1|h_{rd}|^2}{\beta} + \frac{l_r^2 t_1 \lambda + t_1}{\beta\lambda}, \|h_{sr}\|^2 > \frac{t_2(1+\lambda l_r^2)}{\beta\lambda} + \frac{t_2 C}{\beta\lambda^2|h_{rd}|^2} \right). \tag{20}$$

Since $((1-\beta)t_1|h_{rd}|^2/\beta) + (l_r^2 t_1\lambda + t_1/\beta\lambda)$ and $(t_2(1+\lambda l_r^2)/\beta\lambda) + (t_2 C/\beta\lambda^2|h_{rd}|^2)$ have only one point of intersection when $|h_{rd}|^2 > 0$, and the intersection is u_1. Thus, expression (20) can be expressed as

$$\begin{aligned}
\varsigma^{\mathrm{FGR}} &= R_s \int_{u_1}^{\infty} F_X\left(\frac{(1-\beta)t_1 y}{\beta} + \frac{l_r^2 t_1 \lambda + t_1}{\beta\lambda} \right) f_Y(y)dy \\
&\quad \cdot - R_s \int_{u_1}^{\infty} F_X\left(\frac{t_2(1+\lambda l_r^2)}{\beta\lambda} + \frac{t_2 C}{\beta\lambda^2 y} \right) f_Y(y)dy \\
&= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \binom{m}{n} \frac{R_s t_2^m C^n (1+\lambda l_r^2)^{m-n} e^{-t_2+t_2\lambda l_r^2/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{SR}^m \beta^m \bar{\gamma}_{RD} \lambda^{m+n}} \\
&\quad \times \int_{u_1}^{\infty} \frac{1}{y^n} e^{-\left(\left(t_2 C/\beta\lambda^2\bar{\gamma}_{SR} y\right) + \left(y/\bar{\gamma}_{RD}\right)\right)} dy - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \binom{m}{n} \\
&\quad \times \frac{R_s t_1^m (1-\beta)^n e^{-t_1\left(1+\lambda l_r^2\right)/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{SR}^m \beta^m \bar{\gamma}_{RD}} \left(\frac{1+\lambda l_r^2}{\lambda} \right)^{m-n} \\
&\quad \times \int_{u_1}^{\infty} y^n e^{-(((1-\beta)t_1/\beta\bar{\gamma}_{SR}) + (1/\bar{\gamma}_{RD}))y} dy,
\end{aligned} \tag{21}$$

where $u_1 = (-b_1 + \sqrt{b_1^2 - 4a_1c_1})/2a_1$, $a_1 = t_1\lambda^2(1-\beta)$, $c_1 = -t_2 C$, and $b_1 = \lambda(t_1 - t_2)(l_r^2\lambda + 1)$. Then, using [24, 1.211.1], [24, 3.351.2], and [24, 3.381.3], the exact EST of the untrusted full-duplex relay network with the FGR scheme can be derived as

$$\begin{aligned}
\varsigma^{\mathrm{FGR}} &= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\sum_{k=0}^{\infty} \binom{m}{n} \frac{(-1)^k R_s C^{n+k} (1+\lambda l_r^2)^{m-n}}{m!k!\bar{\gamma}_{RD}^{n+k} \lambda^{m+n+2k}} \\
&\quad \times \left(\frac{t_2}{\beta\bar{\gamma}_{SR}} \right)^{m+k} e^{-t_2+t_2\lambda l_r^2/\beta\lambda\bar{\gamma}_{SR}} \Gamma\left(1-n-k, \frac{u_1}{\bar{\gamma}_{RD}}\right) \\
&\quad \cdot - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \binom{m}{n} \frac{R_s t_1^m (1-\beta)^n e^{-t_1\left(1+\lambda l_r^2\right)/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{SR}^m \beta^m \bar{\gamma}_{RD}} \\
&\quad \times \left(\frac{1+\lambda l_r^2}{\lambda} \right)^{m-n} \left(\frac{(1-\beta)t_1}{\beta\bar{\gamma}_{SR}} + \frac{1}{\bar{\gamma}_{RD}} \right)^{-n-1} \\
&\quad \times \Gamma\left(n+1, \frac{(1-\beta)t_1 u_1}{\beta\bar{\gamma}_{SR}} + \frac{u_1}{\bar{\gamma}_{RD}}\right).
\end{aligned} \tag{22}$$

From (22), we find that the EST of the considered system is small, when the confidential information rate R_s is small. At the other extreme, when the confidential information rate R_s is too large, the secrecy performance of the considered system is poor which also results in small EST. Hence, there exists an optimal R_s to maximize the EST of the considered system. Furthermore, the EST increases as the self-interference level decreases, which means that the self-interference deteriorates an overall performance in terms of the security-reliability tradeoff.

3.1.4. Asymptotic Behavior. To extract additional insights on the FGR scheme, we now investigate the high SNR asymptotic behavior of the SOP, COP, and EST of the considered system. According (16), we have

$$\begin{aligned}
\lim_{\lambda\longrightarrow\infty} P_{so}^{\mathrm{FGR}} &= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \frac{\left(t_1 l_r^2\right)^{m-n} (1-\beta)^n t_1^n e^{-t_1 l_r^2/\beta\bar{\gamma}_{SR}}}{(m-n)!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m \beta^m} \\
&\quad \times \left(\frac{t_1 - t_1\beta}{\beta\bar{\gamma}_{SR}} + \frac{1}{\bar{\gamma}_{RD}} \right)^{-n-1}.
\end{aligned} \tag{23}$$

From (19), we have

$$\begin{aligned}
\lim_{\lambda\longrightarrow\infty} P_{co}^{FGR} &= 1 - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \binom{m}{n} \frac{2t_2^m \bar{C}^n l_r^{2m-2n} e^{-t_2 l_r^2/\beta\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m \beta^m} \\
&\quad \times \left(\frac{t_2\bar{C}\bar{\gamma}_{RD}}{\beta\bar{\gamma}_{SR}} \right)^{1-n/2} K_{1-n}\left(\sqrt{\frac{4t_2\bar{C}}{\beta\bar{\gamma}_{SR}\bar{\gamma}_{RD}}} \right),
\end{aligned} \tag{24}$$

where $\bar{C} = l_d^2(1-\beta)(l_r^2 + \beta N_s\bar{\gamma}_{SR} + \bar{\gamma}_{RD}(1-\beta))$. According to (22), we have

$$\begin{aligned}
\lim_{\lambda\longrightarrow\infty} \varsigma^{\mathrm{FGR}} &= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\sum_{k=0}^{\infty} \binom{m}{n} \frac{(-1)^k R_s t_2^{m+k} l_r^{2m-2n} \bar{C}^{n+k}}{m!k!\bar{\gamma}_{SR}^{m+k}\bar{\gamma}_{RD}^{n+k}\beta^{m+k}} \\
&\quad \times e^{-t_2 l_r^2/\beta\bar{\gamma}_{SR}} \Gamma\left(1-n-k, \frac{u_2}{\bar{\gamma}_{RD}}\right) - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m} \binom{m}{n} \\
&\quad \times \frac{R_s t_1^m (1-\beta)^n l_r^{2m-2n}}{m!\bar{\gamma}_{SR}^m \beta^m \bar{\gamma}_{RD}} e^{-t_1 l_r^2/\beta\bar{\gamma}_{SR}} \left(\frac{(1-\beta)t_1}{\beta\bar{\gamma}_{SR}} + \frac{1}{\bar{\gamma}_{RD}} \right)^{-n-1} \\
&\quad \Gamma\left(n+1, \frac{(1-\beta)t_1 u_2}{\beta\bar{\gamma}_{SR}} + \frac{u_2}{\bar{\gamma}_{RD}}\right),
\end{aligned} \tag{25}$$

where $u_2 = (-l_r^2(t_1 - t_2) + \sqrt{l_r^4(t_1-t_2)^2 + 4t_1t_2\bar{C}(1-\beta)})/2t_1(1-\beta)$. It is interesting to note that when transmit power is sufficiently large, the security and reliability of the untrusted full-duplex relay network are still depend on the power allocation factor. Thus, the power allocation factor enables a tradeoff between reliability and security for the untrusted full-duplex relay network. From the numerical results, we find that there exists an optimal power allocation factor maximizing the EST.

3.2. Variable Gain Relaying

3.2.1. Secrecy Outage Probability. The gain factor employed at the relay determines the quality of signal reception at the destination and has no influence on the eavesdropping ability at the relay. Therefore, the SOP of the VGR scheme is the same as the FGR scheme and it is omitted here.

3.2.2. Connection Outage Probability. Combining (8) and (10), the COP of the untrusted full-duplex relay network

with the VGR scheme can be given by

$$P_{co}^{\mathrm{VGR}} = \Pr\left(\frac{\beta\lambda^2\|h_{sr}\|^2|h_{rd}|^2}{C_1\|h_{sr}\|^2 + C_2|h_{rd}|^2 + C_3} < t_2\right) = \Pr\left(\left(\beta\lambda^2|h_{rd}|^2 - t_2C_1\right)\|h_{sr}\|^2 < t_2C_2|h_{rd}|^2 + t_2C_3\right). \tag{26}$$

Based on (13) and (14), we can rewrite (26) as

$$\begin{aligned} P_{co}^{\mathrm{VGR}} &= \int_0^{C_5} f_Y(y)dy + \int_{C_5}^{\infty} F_X\left(\frac{t_2C_2y + t_2C_3}{\beta\lambda^2y - t_2C_1}\right)f_Y(y)dy \\ &= 1 - \sum_{m=0}^{N_s-1}\frac{e^{-t_2C_2/\beta\lambda^2\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m}\left(\frac{t_2C_2}{\beta\lambda^2}\right)^m\int_{C_5}^{\infty}\left(1 + \frac{C_4}{y - C_5}\right)^m \\ &\quad\times e^{-\left(\left(t_2C_2C_4/\beta\lambda^2\bar{\gamma}_{SR}(y-C_5)\right)+(y/\bar{\gamma}_{RD})\right)}dy, \end{aligned} \tag{27}$$

where $C_4 = (C_3/C_2) + (t_2C_1/\beta\lambda^2)$, and $C_5 = t_2C_1/\beta\lambda^2$. Then, let $\tilde{y} = y - C_5$, and use [24, 1.211.1] and [24, 3.471.9], the exact COP of the untrusted full-duplex relay network with the VGR scheme can be derived as

$$\begin{aligned} P_{co}^{VGR} &= 1 - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\binom{m}{n}\frac{2C_4^ne^{-\left(t_2C_2/\beta\lambda^2\bar{\gamma}_{SR}\right)-(C_5/\bar{\gamma}_{RD})}}{m!\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m}\left(\frac{t_2C_2}{\beta\lambda^2}\right)^m \\ &\quad\times\left(\frac{t_2C_2C_4\bar{\gamma}_{RD}}{\beta\lambda^2\bar{\gamma}_{SR}}\right)^{(1-n)/2}K_{1-n}\left(\sqrt{\frac{4t_2C_2C_4}{\beta\lambda^2\bar{\gamma}_{SR}\bar{\gamma}_{RD}}}\right). \end{aligned} \tag{28}$$

When $N_s \longrightarrow \infty$, we have $\|h_{sr}\|^2 \approx^{(a)} N_s\bar{\gamma}_{SR}$, where (a) follows the law of large numbers. Thus, from (26), we clarify that increasing the number of antennas at the source unboundedly is not helpful because the channel between the relay and the destination will finally become the bottleneck and dominate the achievable data rate.

3.2.3. Effective Secrecy Throughput. Combining (6), (8), and (11), the EST of the untrusted full-duplex relay network with the VGR scheme can be given by

$$\begin{aligned} \varsigma^{\mathrm{VGR}} = R_s\Pr\Bigg(&\|h_{sr}\|^2 < \frac{(1-\beta)t_1|h_{rd}|^2}{\beta} \\ &+ \frac{l_r^2t_1\lambda + t_1}{\beta\lambda}, \|h_{sr}\|^2 > \frac{t_2C_2|h_{rd}|^2 + t_2C_3}{\beta\lambda^2|h_{rd}|^2 - t_2C_1}, |h_{rd}|^2 > \frac{t_2C_1}{\beta\lambda^2}\Bigg). \end{aligned} \tag{29}$$

Since $((1-\beta)t_1|h_{rd}|^2/\beta) + (l_r^2t_1\lambda + t_1/\beta\lambda)$ and $(t_2C_2|h_{rd}|^2 + t_2C_3)/(\beta\lambda^2|h_{rd}|^2 - t_2C_1)$ have only one point of intersection when $|h_{rd}|^2 > 0$, and the intersection is u_3. Thus, the expression (29) can be expressed as

$$\begin{aligned} \varsigma^{\mathrm{VGR}} &= R_s\int_{u_4}^{\infty}F_X\left(\frac{(1-\beta)t_1y}{\beta} + \frac{l_r^2t_1\lambda + t_1}{\beta\lambda}\right)f_Y(y)dy \\ &\quad - R_s\int_{u_4}^{\infty}F_X\left(\frac{t_2C_2y + t_2C_3}{\beta\lambda^2y - t_2C_1}\right)f_Y(y)dy \\ &= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\binom{m}{n}\frac{R_sC_4^ne^{-t_2C_2/\beta\lambda^2\bar{\gamma}_{SR}-C_5/\bar{\gamma}_{RD}}}{m!\bar{\gamma}_{SR}^m\bar{\gamma}_{RD}}\left(\frac{t_2C_2}{\beta\lambda^2}\right)^m \\ &\quad\times\int_{u_4-C_5}^{\infty}\frac{1}{y^n}e^{-\left(t_2C_2C_4/\beta\lambda^2\bar{\gamma}_{SR}y\right)-(y/\bar{\gamma}_{RD})}dy - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\binom{m}{n} \\ &\quad\times\frac{R_st_1^m(1-\beta)^ne^{-t_1\left(1+\lambda l_r^2\right)/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{SR}^m\beta^m\bar{\gamma}_{RD}}\left(\frac{1+\lambda l_r^2}{\lambda}\right)^{m-n} \\ &\quad\times\int_{u_4}^{\infty}y^ne^{-(((1-\beta)t_1/\beta\bar{\gamma}_{SR})+(1/\bar{\gamma}_{RD}))y}dy, \end{aligned} \tag{30}$$

where $u_4 = \max(u_3, t_2C_1/\beta\lambda^2)$, $u_3 = (-b_2 + \sqrt{b_2^2 - 4a_2c_2})/2a_2$, $a_2 = t_1\beta\lambda^3(1-\beta)$, $b_2 = t_1\beta\lambda^2(1 + l_r^2\lambda) - t_1t_2\lambda C_1(1-\beta) - \beta\lambda t_2C_2$, and $c_2 = -\beta\lambda t_2C_3 - t_1t_2C_1(1 + l_r^2\lambda)$. Then, using [24, 1.211.1], [24, 3.351.2] and [24, 3.381.3], the exact EST of the untrusted full-duplex relay network with the VGR scheme can be derived as

$$\begin{aligned} \varsigma^{\mathrm{VGR}} &= \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\sum_{k=0}^{\infty}\binom{m}{n}\frac{(-1)^kR_sC_4^{n+k}e^{-\left(t_2C_2/\beta\lambda^2\bar{\gamma}_{SR}\right)-(C_5/\bar{\gamma}_{RD})}}{m!k!\bar{\gamma}_{SR}^{m+k}\bar{\gamma}_{RD}^{n+k}} \\ &\quad\times\left(\frac{t_2C_2}{\beta\lambda^2}\right)^{m+k}\Gamma\left(1-n-k, \frac{u_4 - C_5}{\bar{\gamma}_{RD}}\right) \\ &\quad\cdot - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\binom{m}{n}\frac{R_st_1^m(1-\beta)^ne^{-t_1\left(1+\lambda l_r^2\right)/\beta\lambda\bar{\gamma}_{SR}}}{m!\bar{\gamma}_{SR}^m\beta^m\bar{\gamma}_{RD}} \\ &\quad\times\left(\frac{1+\lambda l_r^2}{\lambda}\right)^{m-n}\left(\frac{(1-\beta)t_1}{\beta\bar{\gamma}_{SR}} + \frac{1}{\bar{\gamma}_{RD}}\right)^{-n-1} \\ &\quad\times\Gamma\left(n+1, \frac{(1-\beta)t_1u_4}{\beta\bar{\gamma}_{SR}} + \frac{u_4}{\bar{\gamma}_{RD}}\right). \end{aligned} \tag{31}$$

The derived exact results in (31) provide an efficient way to holistically examine the security and reliability performance with arbitrary system parameters, i.e., the number of antennas at the source and the self-interference cancellation capability at the relay and destination. Additionally, it is worth noting that the EST of the considered system will tend toward zero when $N_s \longrightarrow \infty$. An intuitive explanation to this phenomenon is that, more antennas at the source would bring us better reliability performance, but at the same time, the grade of the security is severely degraded. Thus, the designers have to carefully take into account the number of antennas at the source.

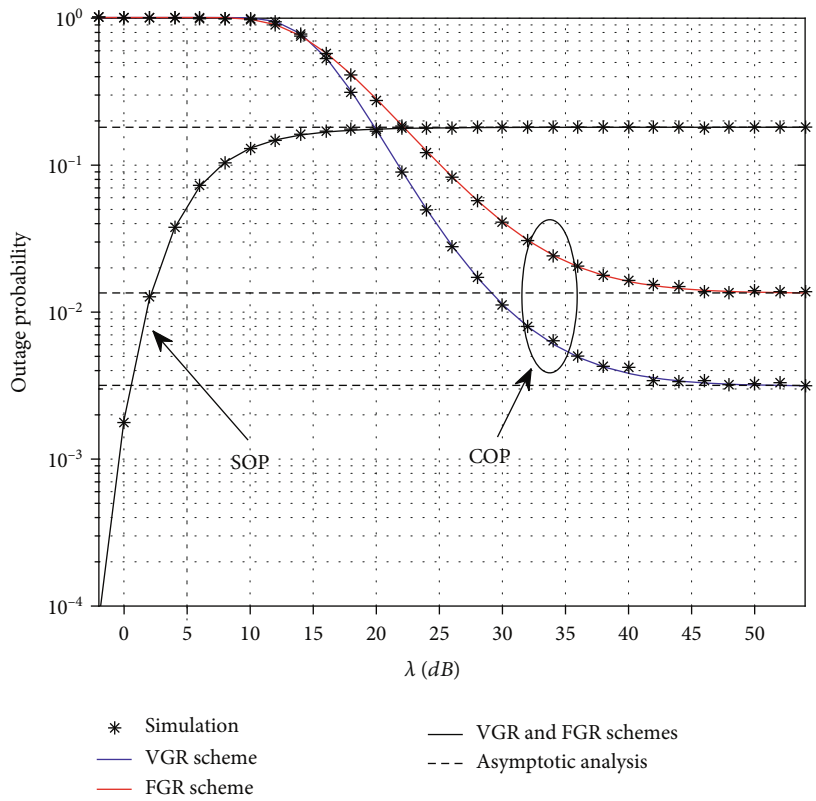


Figure 2: The outage probability of the FGR scheme and the VGR scheme versus the transmit SNR.

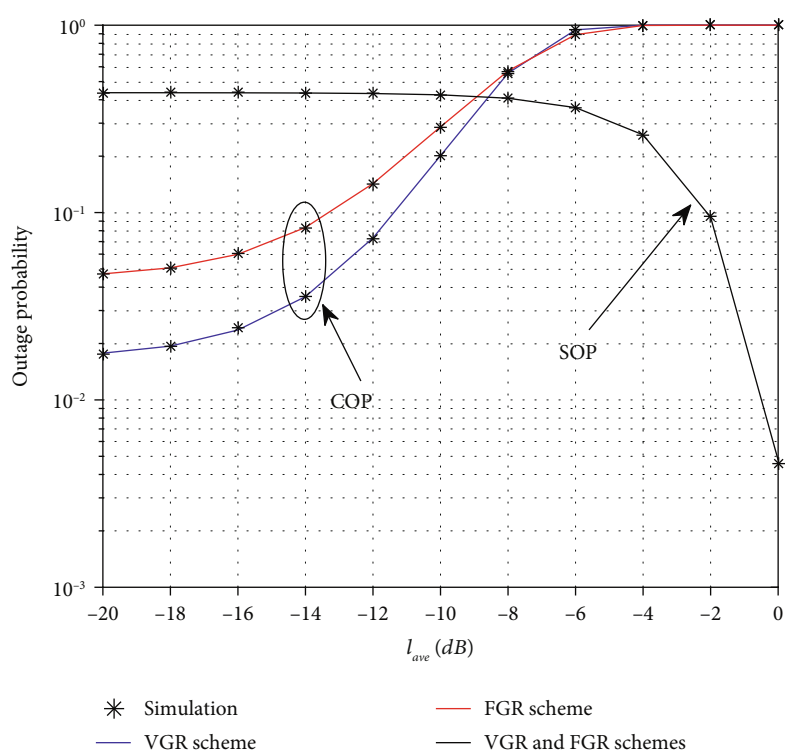


Figure 3: The outage probability of the FGR scheme and the VGR scheme versus the self-interference cancellation capability.

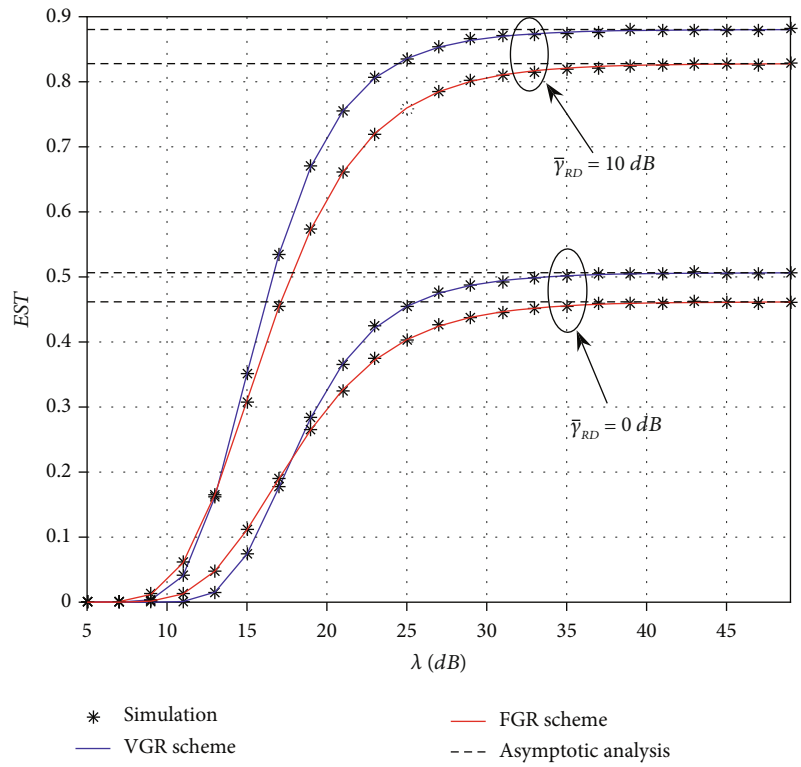


Figure 4: The EST of the FGR scheme and the VGR scheme versus the transmit SNR.

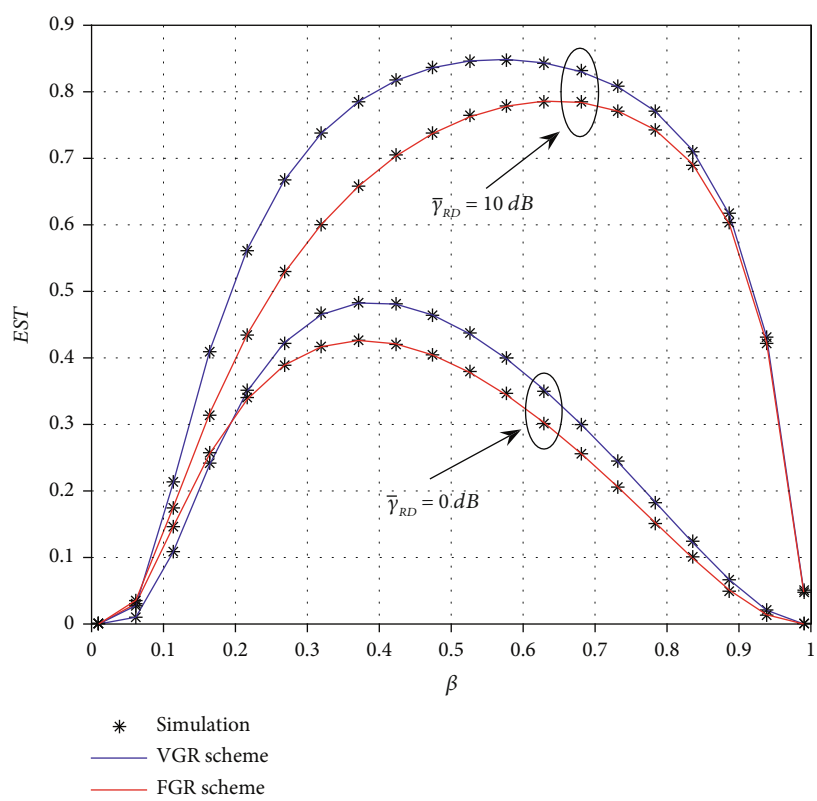


Figure 5: The EST of the FGR scheme and the VGR scheme versus the power allocation factor.

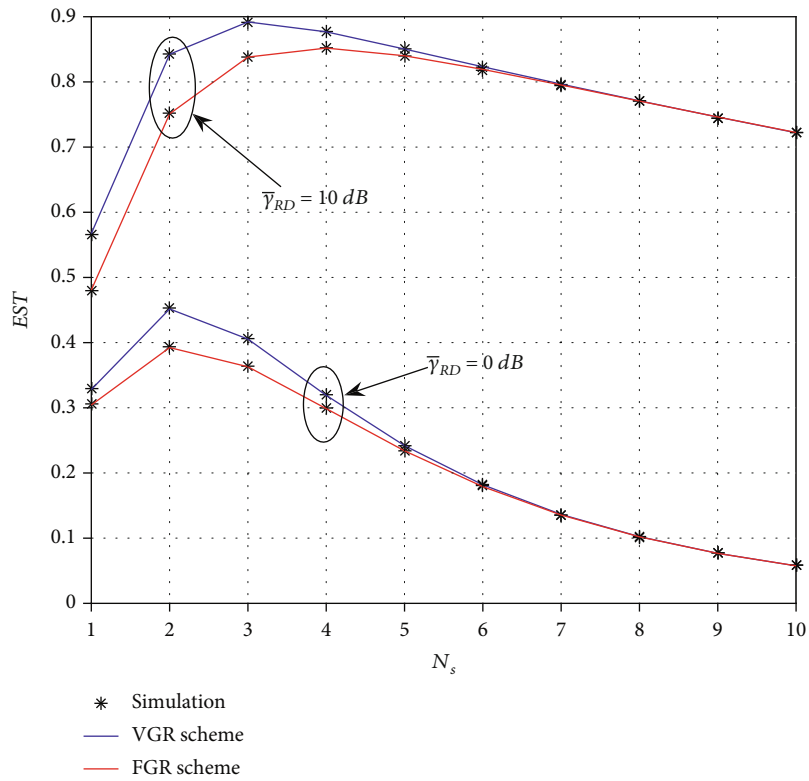


Figure 6: The EST of the FGR scheme and the VGR scheme versus the number of antennas at the source.

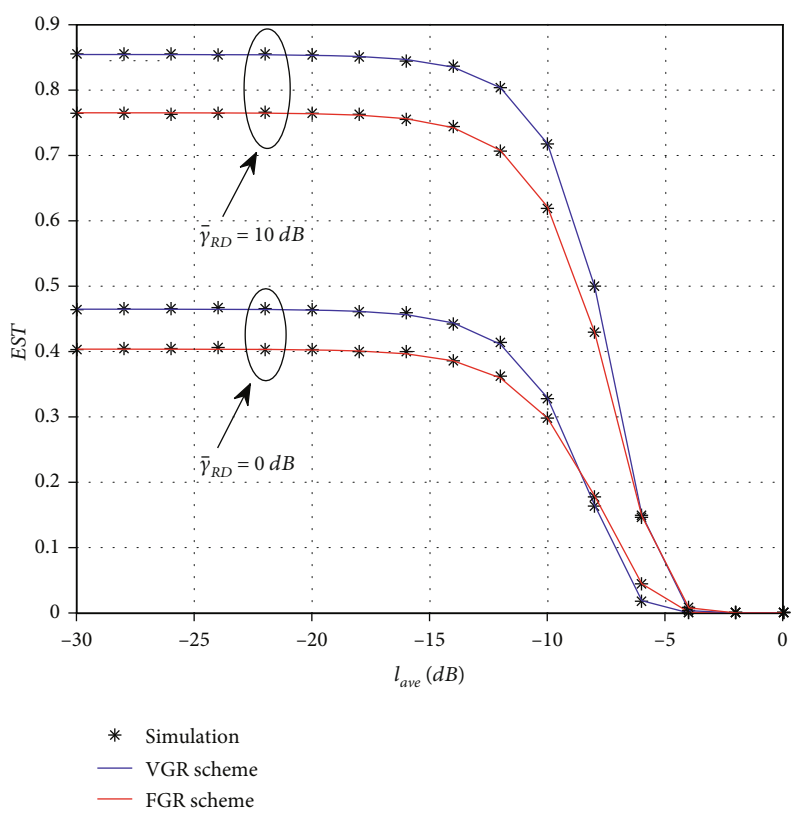


Figure 7: The EST of the FGR scheme and the VGR scheme versus the self-interference cancellation capability.

3.2.4. Asymptotic Behavior. To get more insight, we further study the asymptotic behavior of the COP and EST for the VGR scheme under high SNR case. According to (28), we have

$$\lim_{\lambda\longrightarrow\infty} P_{co}^{\mathrm{VGR}} = 1 - \sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\binom{m}{n}\frac{2t_2^m C_7^m C_9^n e^{-(t_2C_7/\beta\bar{\gamma}_{SR})-(C_{10}/\bar{\gamma}_{RD})}}{m!\beta^m\bar{\gamma}_{RD}\bar{\gamma}_{SR}^m} \times\left(\frac{t_2C_7C_9\bar{\gamma}_{RD}}{\beta\bar{\gamma}_{SR}}\right)^{(1-n)/2} K_{1-n}\left(\sqrt{\frac{4t_2C_7C_9}{\beta\bar{\gamma}_{SR}\bar{\gamma}_{RD}}}\right), \tag{32}$$

where $C_6 = \beta l_d^2(1-\beta)$, $C_7 = l_r^2 + l_d^2(1-\beta)^2$, $C_8 = l_r^2 l_d^2(1-\beta)$, $C_9 = (C_8/C_7) + (t_2C_6/\beta)$, and $C_{10} = t_2C_6/\beta$. From (31), we have

$$\begin{aligned}\lim_{\lambda\longrightarrow\infty}\varsigma^{\mathrm{VGR}} = &\sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\sum_{k=0}^{\infty}\binom{m}{n}\frac{(-1)^k R_s C_9^{n+k} e^{-(t_2C_7/\beta\bar{\gamma}_{SR})-(C_{10}/\bar{\gamma}_{RD})}}{m!k!\bar{\gamma}_{SR}^{m+k}\bar{\gamma}_{RD}^{n+k}}\\ &\times\left(\frac{t_2C_7}{\beta}\right)^{m+k}\Gamma\left(1-n-k,\frac{u_6-C_{10}}{\bar{\gamma}_{RD}}\right)\\ &\cdot -\sum_{m=0}^{N_s-1}\sum_{n=0}^{m}\binom{m}{n}\frac{R_s t_1^m(1-\beta)^n l_r^{2m-2n}}{m!\bar{\gamma}_{SR}^m\beta^m\bar{\gamma}_{RD}}\\ &\times e^{-t_1 l_r^2/\beta\bar{\gamma}_{SR}}\left(\frac{(1-\beta)t_1}{\beta\bar{\gamma}_{SR}}+\frac{1}{\bar{\gamma}_{RD}}\right)^{-n-1}\\ &\times\Gamma\left(n+1,\frac{(1-\beta)t_1u_6}{\beta\bar{\gamma}_{SR}}+\frac{u_6}{\bar{\gamma}_{RD}}\right),\end{aligned} \tag{33}$$

where $u_6 = \max(u_5, t_2C_6/\beta)$, $u_5 = (-b_3 + \sqrt{b_3^2 - 4a_3c_3})/2a_3$, $a_3 = t_1\beta(1-\beta)$, $b_3 = t_1\beta l_r^2 - \beta t_2C_7 - t_1t_2C_6(1-\beta)$, and $c_3 = -\beta t_2C_8 - t_2t_1C_6l_r^2$.

From the asymptotic behaviour of the SOP, COP, and EST, we find that the continued boosting of the transmit power will not always improve the security and reliability performance of the considered system. This is due to the fact that the self-interference caused by full-duplex transmission becomes a bottleneck of improving the security and reliability performance at the high SNR region.

4. Numerical Results

In this section, we present simulation and numerical results to illustrate tendencies of the performance of the untrusted full-duplex relay network. Unless otherwise stated, we set the codeword transmission rate, $R_0 = 2$ bits/sec/Hz, and the confidential information rate $R_s = 1$ bits/sec/Hz. Also, we set $l_r = l_d = l_{\mathrm{ave}}$. It is evident from Figures 2–7 that the derived closed form expressions of SOP, COP, and EST agree well with their Monte Carlo simulation points, which indicates the accuracy of our theoretical analysis stated above.

Figure 2 shows the SOP and COP versus the transmit SNR of the FGR scheme as well as the VGR scheme, where $N_s = 2$, $\beta = 0.5$, $l_{\mathrm{ave}} = -15dB$, $\bar{\gamma}_{SR} = 0$ dB, and $\bar{\gamma}_{RD} = 5$ dB. We first observe that the SOP of the FGR and VGR schemes increases as the transmit SNR increases, whereas the COP of the FGR and VGR schemes decrease as the transmit SNR increases. This is due to the fact that an increasing transmit SNR will lead to a higher achievable eavesdropping rate and achievable data rate. Second, we observe that the SOP and COP of the FGR and VGR schemes converge to a constant when the transmit SNR is sufficiently large. Thus, we can conclude that the transmit SNR is not the larger the better. Moreover, we observe that the COP of the VGR scheme outperforms the FGR scheme at the high SNR region.

Figure 3 shows the SOP and COP versus the self-interference cancellation capability of the FGR scheme as well as the VGR scheme, where $N_s = 2$, $\beta = 0.5$, $\lambda = 30$, $\bar{\gamma}_{SR} = 0$ dB, and $\bar{\gamma}_{RD} = 0$ dB. We observe that the SOP of the FGR and VGR schemes decreases as the self-interference cancellation capability decreases, whereas the COP of the FGR and VGR schemes increase as the self-interference cancellation capability decreases. The reason is that as the self-interference cancellation capability decreases, the loopback interference at the destination and the relay will be stronger. Thus, the self-interference cancellation capability enables a tradeoff between security and reliability of the untrusted full-duplex relay network.

Figure 4 depicts the EST as a function of the transmit SNR, where $N_s = 2$, $\beta = 0.5$, $l_{\mathrm{ave}} = -10$ dB, and $\bar{\gamma}_{SR} = 0$ dB. We first observe that the EST of the FGR scheme and the VGR scheme converge to a constant when the transmit SNR tends to infinity. This phenomenon is due to the fact that the EST is constrained by the self-interference caused by full-duplex transmission in the high SNR regime. Second, we observe that the EST of the VGR scheme is better than the FGR scheme in the high SNR region. Moreover, we observe that for a given transmit SNR, the EST of both the FGR scheme and the VGR scheme with $\bar{\gamma}_{RD} = 10$ dB outperform the EST with $\bar{\gamma}_{RD} = 0$ dB. This is because that the improved quality of the second hop channel is beneficial for secure transmission.

Figure 5 examines the impact of power allocation factor on the EST performance, where $N_s = 2$, $\lambda = 20$ dB, $l_{\mathrm{ave}} = -15$ dB, and $\bar{\gamma}_{SR} = 0$ dB. We observe that when the power allocation factor is either extremely small or large, the EST of the FGR scheme and the VGR scheme approach zero. This is because when the power allocation factor is small, it is hard to establish a reliable link from the source to the destination. On the other hand, when the power allocation factor is large, the power of the jamming signal at the destination is poor and secure transmission is impossible. Thus, there exists an optimal power allocation factor maximizing the EST. Moreover, the optimal power allocation factor gradually increases as $\bar{\gamma}_{RD}$ increases. It indicates that the destination does not need more power to transmit jamming when the channel between the relay and the destination is better.

Figure 6 illustrates the EST of the FGR scheme and the VGR scheme as a function of the number of antennas at the source, with $\lambda = 20$ dB, $\beta = 0.5$, $l_{\mathrm{ave}} = -15$ dB, and $\bar{\gamma}_{SR} = 0$ dB. We observe that the EST of the FGR scheme and the VGR

scheme first increase to a peak value and then decrease as a function of the number of antennas at the source. This phenomenon is due to the fact that increasing the number of antennas at the source leads to the enhancement of the reliability performance, but the degradation of the security performance. Second, we observe that the performance gain of the VGR scheme is much more pronounced for small N_s, and gradually diminishes when N_s becomes large.

Figure 7 plots the EST of the FGR scheme and the VGR scheme as a function of the self-interference cancellation capability, with $N_s = 2$, $\beta = 0.5$, $\lambda = 20$, and $\bar{\gamma}_{SR} = 0\,\text{dB}$. It can be seen from the figure that the EST of the FGR and VGR schemes decrease as the self-interference cancellation capability decreases. From Figure 3, we know that the self-interference cancellation capability has a positive effect on the secure transmission but incurs a negative effect on the reliable transmission. Thus, we can conclude that the self-interference caused by full-duplex transmission does more harm than good for the untrusted full-duplex relay network. When the the level of self-interference is strong, the EST is close to zero, which suggests that both reliable and secure transmission can not be achieved.

5. Conclusion

In this paper, we investigated the security-reliability tradeoff of an untrusted full-duplex relay network over the Rayleigh fading channels. Based on the availability of the CSI, we considered two types of relaying scheme, namely, the FGR scheme and the VGR scheme. The closed form expressions of the SOP, COP, and EST were, respectively, derived, which quantitatively reveal the relationship between the system performances and the number of antennas at the source, power allocation factor, and self-interference cancellation capability, as well as other various parameters. In addition, simple asymptotic results for the high SNR region, i.e., $\lambda \longrightarrow \infty$, were also provided to offer valuable insights into practical design. Analytical and numerical results demonstrated that the self-interference caused by full-duplex transmission is harmful for the reliability performance, while it is beneficial to the security performance. Moreover, the EST of the considered system increases as the self-interference level decreases, which implies that the self-interference deteriorates an overall performance in terms of the security-reliability tradeoff.

Data Availability

The simulation data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the Doctoral Research Start-up Funding of Nanyang Normal University under Grant 2022ZX017, in part by the Cultivating Fund Project for the National Natural Science Foundation of China of Nanyang Normal University under Grant 2022PY024, in part by the open research fund of Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications), Ministry of Education under Grant JZNY202107, in part by the Key Scientific Research Projects of Colleges and Universities in Henan Province of China under Grant 21A520033, and in part by the Key Scientific and Technological Research Projects in Henan Province under Grant 222102320369.

References

[1] W. Yang, X. Lu, S. Yan, F. Shu, and Z. Li, "Age of information for short-packet covert communication," *IEEE Wireless Communications Letters*, vol. 10, no. 9, pp. 1890–1894, 2021.

[2] X. Lu, W. Yang, S. Yan, Z. Li, and D. W. K. Ng, "Covertness and timeliness of data collection in UAV-aided wireless-powered IoT," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12573–12587, 2021.

[3] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[4] M. Li, X. Tao, N. Li, H. Wu, and J. Xu, "Secrecy energy efficiency maximization in UAV-enabled wireless sensor networks without eavesdropper's CSI," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3346–3358, 2022.

[5] P. Angueira, I. Val, J. Montalban et al., "A survey of physical layer techniques for secure wireless communications in industry," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 810–838, 2022.

[6] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Physical layer network security in the full-duplex relay system," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 574–583, 2015.

[7] Q. Li, W. K. Ma, and D. Han, "Sum secrecy rate maximization for full-duplex two-way relay networks using Alamouti-based rank-two beamforming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 8, pp. 1359–1374, 2016.

[8] C. Zhang, F. Jia, J. Ge, and F. Gong, "Security-reliability tradeoff for secure relaying systems with full-duplex radio," *IEEE Access*, vol. 6, pp. 60862–60868, 2018.

[9] Z. Cao, X. Ji, J. Wang et al., "Security-reliability trade-off analysis of AN-aided relay selection for full-duplex relay networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2362–2377, 2021.

[10] B. Chen, Y. Chen, Y. Chen et al., "Secure primary transmission assisted by a secondary full-duplex NOMA relay," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 7214–7219, 2019.

[11] X. He and A. Yener, "Cooperation with an untrusted relay: a secrecy perspective," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3807–3827, 2010.

[12] C. Jeong, I. M. Kim, and D. I. Kim, "Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 310–325, 2012.

[13] J. Huang, A. Mukherjee, and A. Swindlehurst, "Secure communication via an untrusted non-regenerative relay in fading channels," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2536–2550, 2013.

[14] J. Mo, M. Tao, Y. Liu, and R. Wang, "Secure beamforming for MIMO two-way communications with an untrusted relay," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2185–2199, 2014.

[15] L. Sun, P. Ren, Q. Du, Y. Wang, and Z. Gao, "Security-aware relaying scheme for cooperative networks with untrusted relay nodes," *IEEE Communications Letters*, vol. 19, no. 3, pp. 463–466, 2015.

[16] H. Xu, L. Sun, P. Ren, and Q. Du, "Securing two-way cooperative systems with an untrusted relay: a constellation-rotation aided approach," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2270–2273, 2015.

[17] S. Atapattu, N. Ross, Y. Jing, and M. Premaratne, "Source-based jamming for physical-layer security on untrusted full-duplex relay," *IEEE Communications Letters*, vol. 23, no. 5, pp. 842–846, 2019.

[18] J. Lim, T. Kim, and I. Bang, "Impact of outdated CSI on the secure communication in untrusted in-band full-duplex relay networks," *IEEE Access*, vol. 10, pp. 19825–19835, 2022.

[19] A. Mabrouk, A. E. Shafie, K. Tourki, N. Al-Dhahir, and N. Hamdi, "Securing untrusted full-duplex relay channels in the presence of multiple external cluster-based eavesdroppers," *IEEE Systems Journal*, vol. 14, no. 1, pp. 665–668, 2020.

[20] S. Zhao, J. Liu, Y. Shen, X. Jiang, and N. Shiratori, "Secure beamforming for full-duplex MIMO two-way untrusted relay systems," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3775–3790, 2020.

[21] C. Yang, W. Wang, S. Zhao, and M. Peng, "Performance of decode-and-forward opportunistic cooperation with channel estimation errors," in *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1967–1971, Istanbul, Turkey, 2010.

[22] Z. Liu, Y. Ye, G. Lu, and R. Q. Hu, "System outage performance of SWIPT enabled full-duplex two-way relaying with residual hardware impairments and self-interference," *IEEE Systems Journal*, pp. 1–12, 2022.

[23] J. Hu, W. Yang, N. Yang, X. Zhou, and Y. Cai, "On–off-based secure transmission design with outdated channel state information," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6075–6088, 2016.

[24] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, Inc., 7th edition, 2007.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6723403, 10 pages
https://doi.org/10.1155/2022/6723403

Research Article

UAV-Aided Multiuser Mobile Edge Computing Networks with Energy Harvesting

Changyu Wang,[1] Weili Yu,[1] Fusheng Zhu,[2] Jiangtao Ou,[3] Chengyuan Fan,[3] Jianghong Ou,[4] and Dahua Fan[4]

[1]*Aviation University Air Force, Changchun, Jilin, China*
[2]*Guangdong New Generation Communication and Network Innovative Institute (GDCNi), Guangzhou, China*
[3]*AI Sensing Technology, Foshan, Guangdong, China*
[4]*Starway Communication, Guangzhou, China*

Correspondence should be addressed to Fusheng Zhu; fushengzhu.gdcni@hotmail.com

Received 2 April 2022; Revised 16 April 2022; Accepted 21 April 2022; Published 21 June 2022

Academic Editor: Xingwang Li

This article studies a mobile edge computing (MEC) with one edge node (EN), where multiple unmanned aerial vehicles (UAVs) act as users which have some heavy tasks. As the users generally have limitations in both calculating and power supply, the EN can help calculate the tasks and meanwhile supply the power to the users through energy harvesting. We optimize the system by proposing a joint strategy to unpacking and energy harvesting. Specifically, a deep reinforcement learning (DRL) algorithm is implemented to provide a solution to the unpacking, while several analytical solutions are given to the power allocation of energy harvesting among multiple users. In particular, criterion I is the equivalent power allocation, criterion II is designed through equal data rate, while criterion III is based on the equivalent transmission delay. We finally give some results to verify the joint strategy for the UAV-aided multiuser MEC system with energy harvesting.

1. Introduction

In recent years, wireless communication has been put into many efforts from the researchers of both academy and industry [1, 2], which inspires a lot of practical applications, such as internet of things and video monitoring [3]. Among these applications, a key feature is that massive calculating is involved due to the massive number of accessing nodes [4]. To suppress the massive calculating, cloud computing has been proposed which assisted the task calculating through wireless transmission [5, 6]. A major limitation is that the latency and power consumption (PoC) become prohibitively high in a poor channel state, which limits the development and application of cloud computing severely.

To resolve the above disadvantages of cloud computing, mobile edge computing (MEC) has been proposed to install the calculating resources at the edge node (ENs) of the network [7–9]. In this way, the users can unpack the tasks to the nearby EN through wireless transmission, which leads to a decreased delay and PoC compared to the cloud computing. A key design in the MEC system is the unpacking scale [10, 11], which gives the number of scale of tasks to be calculated at the EN. The fundamental principle of unpacking is to jointly utilize the communication and calculating resources, through achieving a fine trade-off between the calculating and wireless transmission. Moreover, some advanced wireless techniques have been proposed to decrease the delay and PoC in the calculating and transmission [12, 13].

Another new technique to assist the calculating and communication in IoT networks is the deployment of unmanned aerial vehicles (UAVs), which are easy to be used and provide flexible ability. Moreover, the price of UAV is becoming cheaper and cheaper, which inspires a lot of applications in practice [14, 15]. For the MEC system, the UAVs can rescue the data calculating with higher priority through some intelligent path routing and scheduling, which exploits

the incremental system resources due to the usage of UAVs. The integration of UAVs into MEC systems has attracted much attention from the researchers of academy and industry, which becomes the motivation of this article.

Motivated by the above literature review, this article studies a MEC system with one EN, where multiple unmanned aerial vehicles (UAVs) act as users which have some heavy tasks. As the users generally have limitations in both calculating and power supply, the EN can help calculate the tasks and meanwhile supply the power to the users through energy harvesting. We optimize the system by proposing a joint strategy to unpacking and energy harvesting. Specifically, a deep reinforcement learning (DRL) algorithm is implemented to provide a solution to the unpacking, while several analytical solutions are given to the power allocation of energy harvesting among multiple users. In particular, criterion I is the equivalent power allocation, criterion II is designed through equal data rate, while criterion III is based on the equivalent transmission delay. We finally give some results to verify the joint strategy for the UAV-aided multiuser MEC system with energy harvesting.

2. System Model

In this paper, we consider an unloading system model in Figure 1 which has an edge node (EN) (note that the notation of "CAP" is used in some literature, while the notation "EN" is used in other literature. Both stand for the same meaning, and these two notations can be used interchangeably) surrounded by N unmanned aerial vehicles (UAVs). Specifically, the EN has an energy transmitter and a server which can provide calculating. The EN is capable of providing charging services to the UAVs, and each UAV is equipped with a limited battery capable of wireless charging. Each UAV has a calculating task l_n. Due to the UAVs' limited calculating power, each UAV unloads the calculating task to the EN in order to reduce the calculating time. The EN ensures that the UAV is always supplied with electricity, so the UAVs in this system unload tasks without considering power consumption. We will introduce the local calculating model and unloading calculating model in the next parts.

2.1. Local Calculating Model. The local calculating delay of the UAV_n is [16]

$$D_{\text{local}}^n = \frac{l_n(1-\beta_{\text{EN}}^n)c}{f_n}, \tag{1}$$

where l_n is the size of the task. β_{EN}^n is the unloading ratio from UAV_n to the EN. Moreover, c is the CPU cycles for executing one bit, and f_n is the local calculating ability. Because all UAVs calculate their tasks in parallel, we use the maximum value of local calculating as the local delay of the whole system. So, the local calculating delay of the whole system is

$$D_{\text{local}} = \max\left\{D_{\text{local}}^1, D_{\text{local}}^2, \cdots, D_{\text{local}}^N\right\}. \tag{2}$$

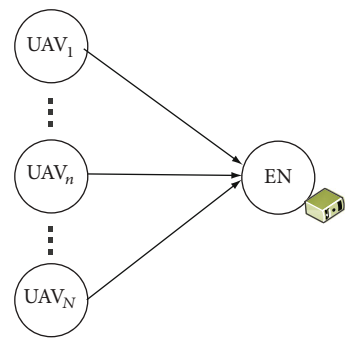


Figure 1: System model of multiuser MEC system with energy harvesting.

2.2. Unloading Calculating Model. In this paper, UAV_n will be charged by EN, and the charging process from EN to UAV_n is

$$H_n = \eta P_{\text{charge}}^n |h_{EN}^n|^2 \alpha_n \Gamma, \tag{3}$$

where notations η denotes the charging factor, P_{charge}^n is the charged power of EN, α_n is the charging time, and Γ denotes the span of each time slot.

From (3), the transmission power at the UAV_n is

$$P_{\text{tran}}^n = \frac{H_n}{(1-\alpha_n)\Gamma}. \tag{4}$$

The transmission rate between the UAV_n and the EN is

$$r_{\text{tran}}^n = W_{\text{total}} \log_2\left(1 + \frac{P_{\text{tran}}^n |h_{\text{EN}}^n|^2}{\sigma_{\text{EN}}^2}\right), \tag{5}$$

where W_{total} is the total bandwidth of the system. $h_{EN}^n \sim \mathscr{C}\mathcal{N}(0, \delta_{\text{EN}})$ is the channel parameter from UAV_n to the EN. σ_{EN}^2 is the variance of the additive white Gaussian noise at the EN. The transmission delay of the UAV_n is

$$D_{\text{tran}}^n = \frac{l_n \beta_{EN}^n}{r_{\text{tran}}^n}. \tag{6}$$

The calculating delay at the UAV_n is

$$D_{\text{com}}^n = \frac{l_n \beta_{\text{EN}}^n c}{f_{\text{EN}}}, \tag{7}$$

where f_{EN} is the calculating ability at the EN. Further, the transmission delay of all UAVs is

$$D_{\text{tran}} = \max\left\{D_{\text{tran}}^1, D_{tran}^2, \cdots, D_{\text{tran}}^N\right\}. \tag{8}$$

The calculating delay of all UAVs is

$$D_{\text{com}} = \max\left\{D_{\text{com}}^1, D_{\text{com}}^2, \cdots, D_{\text{com}}^N\right\}. \tag{9}$$

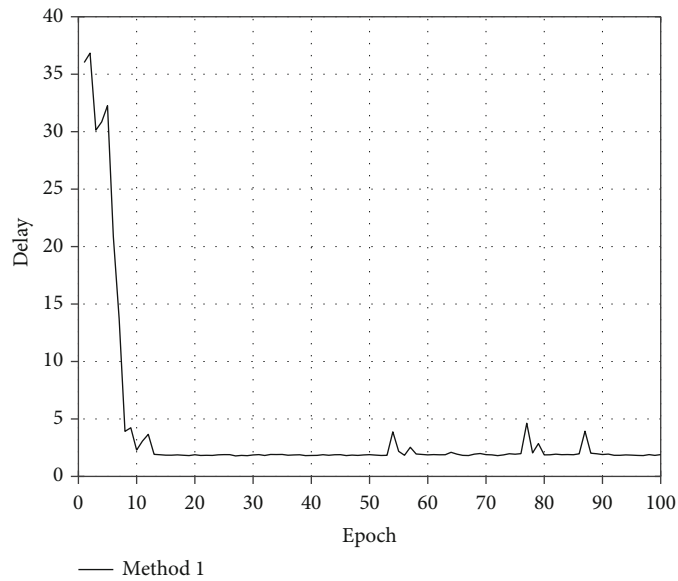


Figure 2: The convergence of method 1.

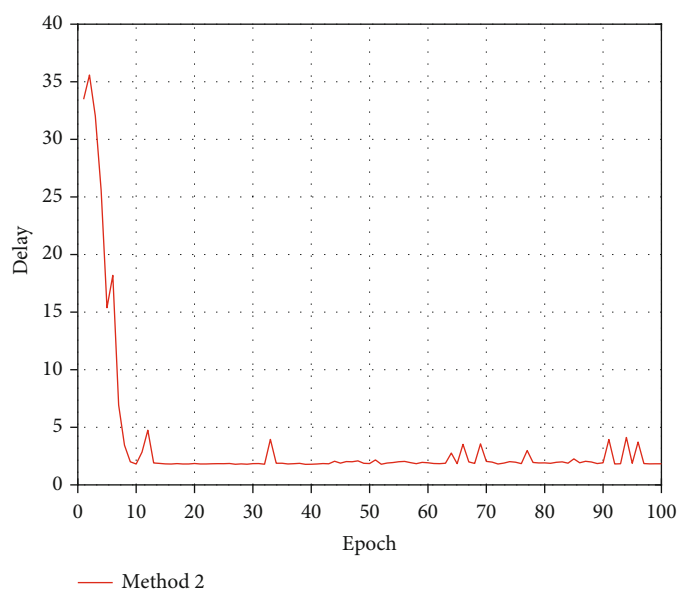


Figure 3: The convergence of method 2.

From (8) and (9), the unloading calculating of the whole system is

$$D_{\text{total}} = \max\{D_{\text{tran}}, D_{\text{com}}\}. \tag{10}$$

Therefore, the system target in this considered MEC network is

$$\begin{aligned} &\min_{\{\beta_n, P^n_{\text{charge}}\}} \Phi = D_{\text{total}} \\ &\text{s.t.} \quad \text{C}_1 : \beta^n_{\text{EN}} \in [0, 1] \\ &\text{C}_2 : \sum_{n=1}^{N} P^n_{\text{charge}} = P^{\text{total}}_{\text{charge}}, \end{aligned} \tag{11}$$

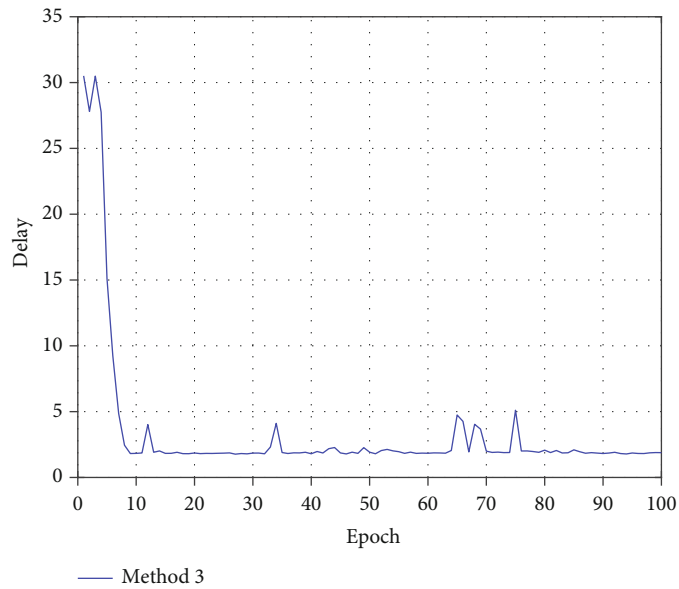


Figure 4: The convergence of method 3.

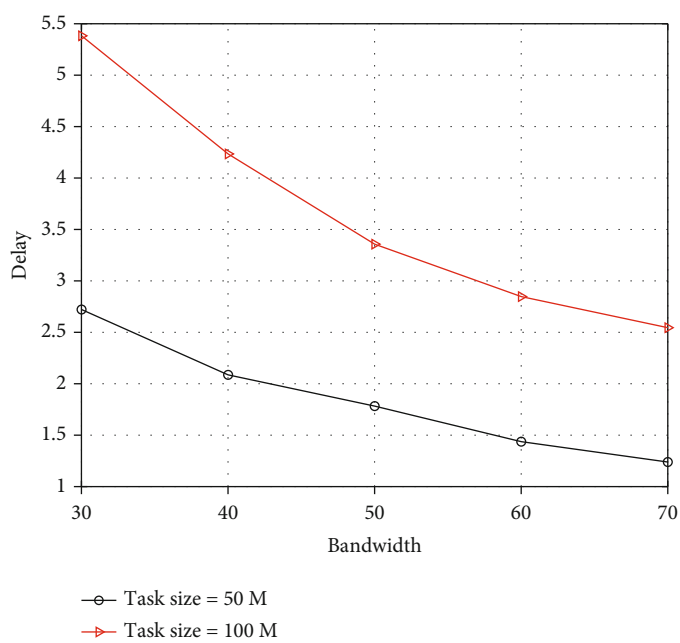


Figure 5: System delay of method 1 when the value of W_{total} ranges from 30 to 70.

where $P_{\text{charge}}^{\text{total}}$ is the total charged power of EN. In the next section, we will describe how we optimize the system target in detail.

3. System Optimization

In this section, we demonstrate our optimization scheme for the considered system target. Specifically, we first utilize deep Q-network (DQN) algorithm to obtain the task unloading strategy, and then, we proposed three methods to allocate the charged power for UAVs in the considered system. The details of our optimization scheme are expressed as follows.

3.1. Scheme on the Task Unloading. Due to the complexity of wireless link in the system, it is hard to dynamically unload the task of UAVs by traditional method. Therefore, we exploit DQN algorithm to obtain the task unloading strategy. Different from the Q-learning algorithm, DQN has an experience pool and two neural networks that include the evaluation network and the target network, to interact with the training environment and break the training data correlation.

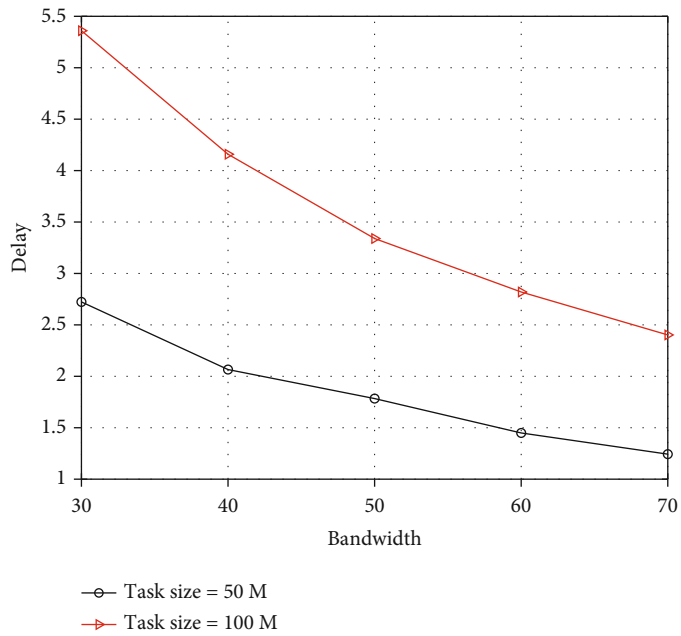


FIGURE 6: System delay of method 2 when the value of W_{total} ranges from 30 to 70.

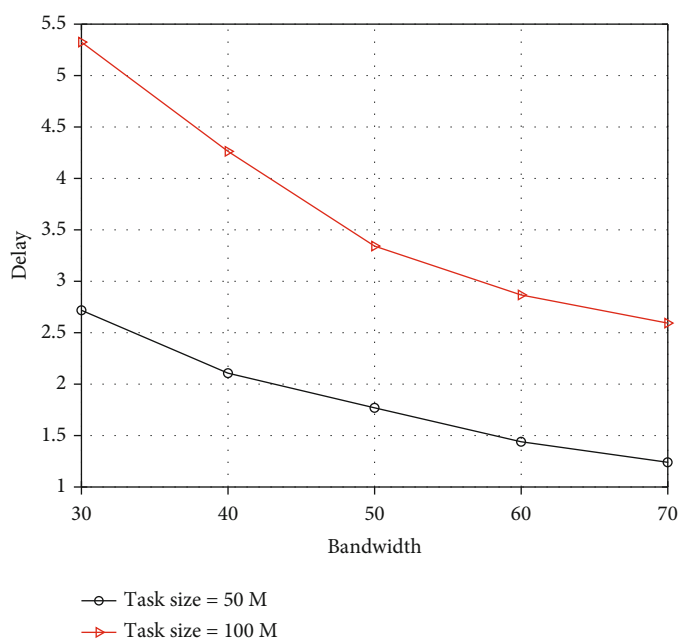


FIGURE 7: System delay of method 3 when the value of W_{total} ranges from 30 to 70.

Moreover, we use the Markov decision process (MDP) to model the consider task unloading issue. In particular, MDP generally consists of the state set $S = [s_1, s_2, \cdots, s_N]$, the action set $A = [a_1, a_2, \cdots, a_{2N}]$, and reward function $R = [0, -1, 1]$. The training process can be represented as follows: the DQN agent first initializes the system state set S, and then, it selects an action command under the current state. After the DQN agent executes the selected action command, the system state set will be updated. Further, the DQN agent will obtain a feedback according to the reward function R. Then, the DQN agent will put the previous state, the updated state, the selected action under the previous state, and the according feedback into the experience pool. When the DQN agent finishes the above process, it will obtain a state-action value $Q(S, A; \omega)$ that ω represents the network matrix of evaluation network. Then, the evaluation network will be trained by the loss function, which is

$$X = (Y - Q(S, A; \omega))^2, \tag{12}$$

where Y denotes the function of target network, which is

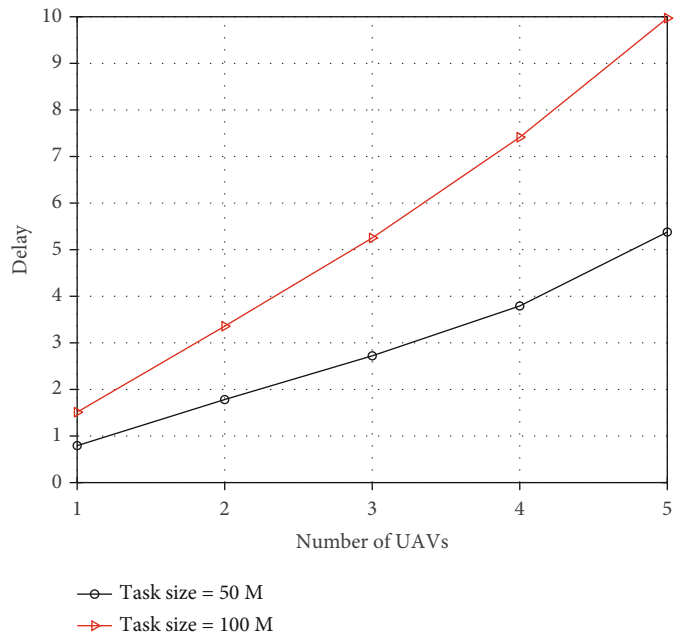


Figure 8: System delay versus the number of UAVs: method 1.

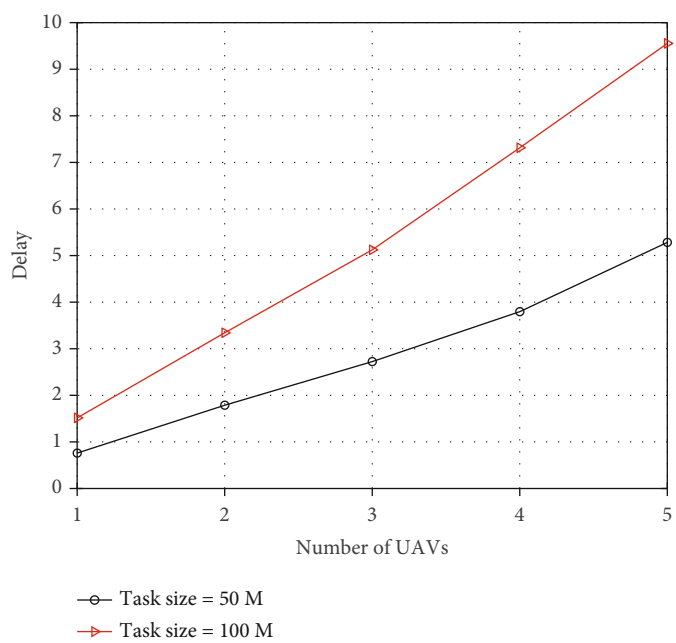


Figure 9: System delay versus the number of UAVs: method 2.

$$Y = R + \varphi \arg \max_{A^-} Q(S^-, A^-; \omega^*), \tag{13}$$

where φ represents a discount element and ω^* denotes the network matrix of the target network. It is notable that the structures of evaluation network and the target network are the same. However, different from the target network, the evaluation network will be trained in every round, and its training process can be denoted as

$$\omega^- = \omega - \xi \frac{\partial X}{\partial \omega}, \tag{14}$$

where ξ is the learning rate of the evaluation network.

3.2. Methods on the Charged Power Allocation. In this part, we will describe three methods for allocating the charged power from EN to UAV_n. Specifically, we exploit equal-

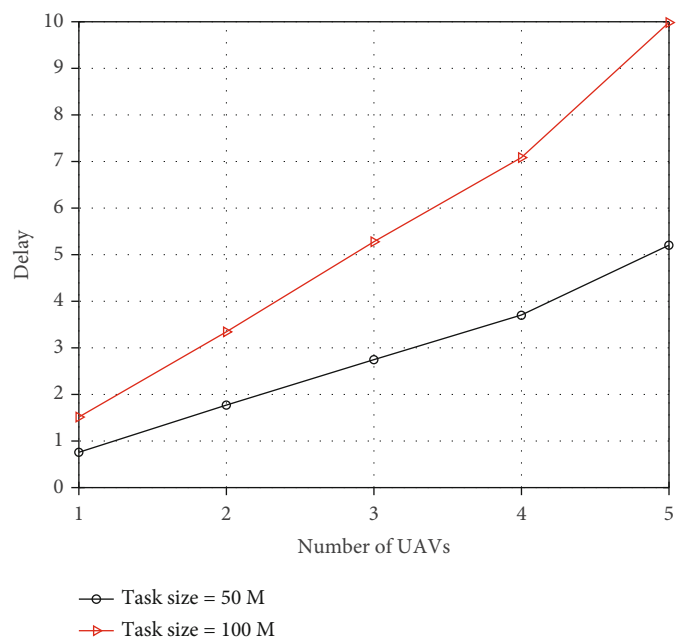


Figure 10: System delay versus the number of UAVs: method 3.

charge-power allocation method, equal-transmit-rate allocation method, and equal-charge-energy allocation.

(1) Equal-charge-power allocation method

Firstly, we allocate the charge power to UAV_n in a traditional way, so that each UAV_n can obtain the same charge power. Moreover, we define this method as equal-charge-power allocation method or method 1, and it can be denoted as

$$P_{\text{charge}}^{n} = \frac{P_{\text{charge}}^{\text{total}}}{N}, \tag{15}$$

where notation P_{charge}^{n} denotes the allocated charge power of UAV_n.

(2) Equal-transmission-rate allocation method

Secondly, we allocate the charge power to UAV_n by a method that ensure each UAV_n can obtain the same transmission rate according to (5). Moreover, we define this method as equal-transmission-rate allocation method or method 2. This method can be represented as

$$r_{\text{trans}}^{1} = r_{\text{trans}}^{2} = \cdots = r_{\text{trans}}^{N}. \tag{16}$$

From (16) and (5), we can obtain

$$W_{\text{total}} \log_2\left(1 + \frac{P_{\text{tran}}^{1}\left|h_{\text{EN}}^{1}\right|^2}{\sigma_{\text{EN}}^2}\right) = W_{\text{total}} \log_2\left(1 + \frac{P_{\text{tran}}^{2}\left|h_{\text{EN}}^{2}\right|^2}{\sigma_{EN}^2}\right) = \cdots = W_{\text{total}} \log_2\left(1 + \frac{P_{\text{tran}}^{N}\left|h_{\text{EN}}^{N}\right|^2}{\sigma_{\text{EN}}^2}\right). \tag{17}$$

By removing the common item of W_{total}, we can have

$$\log_2\left(1 + \frac{P_{\text{tran}}^{1}\left|h_{\text{EN}}^{1}\right|^2}{\sigma_{\text{EN}}^2}\right) = \log_2\left(1 + \frac{P_{\text{tran}}^{2}\left|h_{\text{EN}}^{2}\right|^2}{\sigma_{\text{EN}}^2}\right) = \cdots = \log_2\left(1 + \frac{P_{\text{tran}}^{N}\left|h_{\text{EN}}^{N}\right|^2}{\sigma_{\text{EN}}^2}\right). \tag{18}$$

From (4), we can obtain

$$\log_2\left(1 + \frac{H_1/(1-\alpha_1)\Gamma\left|h_{\text{EN}}^{1}\right|^2}{\sigma_{\text{EN}}^2}\right) = \log_2\left(1 + \frac{H_2/(1-\alpha_2)\Gamma\left|h_{\text{EN}}^{2}\right|^2}{\sigma_{\text{EN}}^2}\right) = \cdots = \log_2\left(1 + \frac{H_N/(1-\alpha_N)\Gamma\left|h_{\text{EN}}^{N}\right|^2}{\sigma_{\text{EN}}^2}\right), \tag{19}$$

$$1 + \frac{H_1/(1-\alpha_1)\Gamma\left|h_{\text{EN}}^{1}\right|^2}{\sigma_{\text{EN}}^2} = 1 + \frac{H_2/(1-\alpha_2)\Gamma\left|h_{\text{EN}}^{2}\right|^2}{\sigma_{\text{EN}}^2} = \cdots = 1 + \frac{H_N/(1-\alpha_N)\Gamma\left|h_{\text{EN}}^{N}\right|^2}{\sigma_{\text{EN}}^2}, \tag{20}$$

$$\frac{H_1/(1-\alpha_1)\Gamma\left|h_{\text{EN}}^{1}\right|^2}{\sigma_{\text{EN}}^2} = \frac{H_2/(1-\alpha_2)\Gamma\left|h_{\text{EN}}^{2}\right|^2}{\sigma_{EN}^2} = \cdots = \frac{H_N/(1-\alpha_N)\Gamma\left|h_{\text{EN}}^{N}\right|^2}{\sigma_{\text{EN}}^2}, \tag{21}$$

$$\frac{H_1}{(1-\alpha_1)\Gamma}\left|h_{\mathrm{EN}}^1\right|^2=\frac{H_2}{(1-\alpha_2)\Gamma}\left|h_{EN}^2\right|^2=\cdots=\frac{H_N}{(1-\alpha_N)\Gamma}\left|h_{\mathrm{EN}}^N\right|^2. \tag{22}$$

Moreover, from (3) and (22), we can obtain

$$\begin{aligned}\frac{\eta P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1\Gamma}{(1-\alpha_1)\Gamma}\left|h_{\mathrm{EN}}^1\right|^2&=\frac{\eta P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2\Gamma}{(1-\alpha_2)\Gamma}\left|h_{\mathrm{EN}}^2\right|^2\\&=\cdots=\frac{\eta P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N\Gamma}{(1-\alpha_N)\Gamma}\left|h_{\mathrm{EN}}^N\right|^2.\end{aligned} \tag{23}$$

After removing the comment term of Γ, we can have

$$\begin{aligned}\frac{\eta P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1}{(1-\alpha_1)}\left|h_{\mathrm{EN}}^1\right|^2&=\frac{\eta P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2}{(1-\alpha_2)}\left|h_{\mathrm{EN}}^2\right|^2\\&=\cdots=\frac{\eta P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N}{(1-\alpha_N)}\left|h_{\mathrm{EN}}^N\right|^2.\end{aligned} \tag{24}$$

Then, by further removing the comment term of η, we can have

$$\begin{aligned}\frac{P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1}{(1-\alpha_1)}\left|h_{\mathrm{EN}}^1\right|^2&=\frac{P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2}{(1-\alpha_2)}\left|h_{\mathrm{EN}}^2\right|^2\\&=\cdots=\frac{P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N}{(1-\alpha_N)}\left|h_{\mathrm{EN}}^N\right|^2,\end{aligned} \tag{25}$$

$$\frac{P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1}{(1-\alpha_1)}=\frac{P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2}{(1-\alpha_2)}=\cdots=\frac{P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N}{(1-\alpha_N)}. \tag{26}$$

For simplicity, we assume the charging time of each UAV is the same, which can be written as

$$\begin{aligned}\alpha_1=\alpha_2=\cdots=\alpha_N,\\(1-\alpha_1)=(1-\alpha_2)=\cdots=(1-\alpha_N).\end{aligned} \tag{27}$$

Therefore, from (26), we can obtain

$$\frac{P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2}{(1-\alpha_1)}=\frac{P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2}{(1-\alpha_2)}=\cdots=\frac{P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2}{(1-\alpha_N)}. \tag{28}$$

By removing the common item of $(1-\alpha_n)$ for $n\in[1,N]$, we can have

$$P_{\mathrm{charge}}^1\left|h_{EN}^1\right|^2=P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2=\cdots=P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2. \tag{29}$$

From this equation, we have

$$\begin{aligned}P_{\mathrm{charge}}^1:P_{\mathrm{charge}}^2:\cdots:P_{\mathrm{charge}}^N=\left|h_{\mathrm{EN}}^1\right|^2:\left|h_{\mathrm{EN}}^2\right|^2:\cdots\\:\left|h_{\mathrm{EN}}^N\right|^2.\end{aligned} \tag{30}$$

Then, we can further obtain

$$\frac{P_{\mathrm{charge}}^n}{P_{\mathrm{charge}}^1+P_{\mathrm{charge}}^2+\cdots+P_{\mathrm{charge}}^N}=\frac{\left|h_{\mathrm{EN}}^n\right|^2}{\left|h_{\mathrm{EN}}^1\right|^2+\left|h_{\mathrm{EN}}^2\right|^2+\cdots+\left|h_{\mathrm{EN}}^N\right|^2}. \tag{31}$$

By using the relationship of $P_{\mathrm{charge}}^1+P_{\mathrm{charge}}^2+\cdots+P_{\mathrm{charge}}^N=P_{\mathrm{charge}}^{\mathrm{total}}$, we can have

$$\frac{P_{\mathrm{charge}}^n}{P_{\mathrm{charge}}^{\mathrm{total}}}=\frac{\left|h_{EN}^n\right|^2}{\left|h_{\mathrm{EN}}^1\right|^2+\left|h_{\mathrm{EN}}^2\right|^2+\cdots+\left|h_{\mathrm{EN}}^N\right|^2}. \tag{32}$$

From this equation, we can have the power charge allocation result of method 2 as

$$P_{\mathrm{charge}}^n=\frac{\left|h_{\mathrm{EN}}^n\right|^2}{\left|h_{\mathrm{EN}}^1\right|^2+\left|h_{\mathrm{EN}}^2\right|^2+\cdots+\left|h_{\mathrm{EN}}^N\right|^2}P_{\mathrm{charge}}^{\mathrm{total}}. \tag{33}$$

(3) Equal-charge-energy allocation method

Thirdly, we allocate the charge power to UAV_n by a method that ensure each UAV_n can be charged same energy according to (3). Moreover, we define this method as equal-charge-energy allocation method or method 3, which can be represented as

$$H_1=H_2=\cdots=H_N. \tag{34}$$

From (3), we can obtain

$$\eta P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1\Gamma=\eta P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2\Gamma=\cdots=\eta P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N\Gamma. \tag{35}$$

By removing the common item of η, we can have

$$P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1\Gamma=P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2\Gamma=\cdots=P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N\Gamma. \tag{36}$$

Then, by removing the common item of Γ, we can have

$$P_{\mathrm{charge}}^1\left|h_{\mathrm{EN}}^1\right|^2\alpha_1=P_{\mathrm{charge}}^2\left|h_{\mathrm{EN}}^2\right|^2\alpha_2=\cdots=P_{\mathrm{charge}}^N\left|h_{\mathrm{EN}}^N\right|^2\alpha_N. \tag{37}$$

Since we assume that the charging time of each UAV is

the same, we can further obtain

$$P^1_{\text{charge}}\left|h^1_{\text{EN}}\right|^2 = P^2_{\text{charge}}\left|h^2_{\text{EN}}\right|^2 = \cdots = P^N_{\text{charge}}\left|h^N_{\text{EN}}\right|^2,$$
$$P^1_{\text{charge}} : P^2_{\text{charge}} : \cdots : P^N_{\text{charge}} = \left|h^1_{\text{EN}}\right|^2 : \left|h^2_{\text{EN}}\right|^2 : \cdots : \left|h^N_{\text{EN}}\right|^2, \tag{38}$$

Then, we can further obtain

$$\frac{P^n_{\text{charge}}}{P^1_{\text{charge}} + P^2_{\text{charge}} + \cdots + P^N_{\text{charge}}} = \frac{\left|h^n_{\text{EN}}\right|^2}{\left|h^1_{\text{EN}}\right|^2 + \left|h^2_{\text{EN}}\right|^2 + \cdots + \left|h^N_{\text{EN}}\right|^2}. \tag{39}$$

By using the relationship of $P^1_{\text{charge}} + P^2_{\text{charge}} + \cdots + P^N_{\text{charge}} = P^{\text{total}}_{\text{charge}}$, we can have

$$\frac{P^n_{\text{charge}}}{P^{\text{total}}_{\text{charge}}} = \frac{\left|h^n_{\text{EN}}\right|^2}{\left|h^1_{\text{EN}}\right|^2 + \left|h^2_{\text{EN}}\right|^2 + \cdots + \left|h^N_{\text{EN}}\right|^2}. \tag{40}$$

From this equation, we can have the power charge result of method 3 as

$$P^n_{\text{charge}} = \frac{\left|h^n_{\text{EN}}\right|^2}{\left|h^1_{\text{EN}}\right|^2 + \left|h^2_{\text{EN}}\right|^2 + \cdots + \left|h^N_{EN}\right|^2} P^{\text{total}}_{\text{charge}}. \tag{41}$$

In the next section, we will perform some simulations to demonstrate the effectiveness of our proposed scheme on task offloading and charged power allocation.

4. Simulation

In this section, we perform some simulations to demonstrate our proposed scheme on task offloading and charged power allocation. Specifically, the channel in the considered MEC network adopts the Gaussian channel, and the average channel gain of the wireless link from UAVs to EN is set to 1. The variance of AWGN at the EN is set 0.1. Moreover, the number of UAVs is set to 2, and the task size of UAVs is set to 50 MB. We set the calculating ability of UAVs to 1.3×10^2 cycle/s, while the calculating ability of EN is set to 1×10^7 cycle/s. The total wireless bandwidth of EN is set to 50 MHz, and the total charged power of EN is set to 20 W, while the charging time of UAV is set 0.5.

Figure 2 shows the convergence of the proposed strategy with method 1. We can find that the system delay declines rapidly and converges after 15 epochs. For example, the system delay of method 1 decreases from 35 to less than 5. Similarly, Figures 3 and 4 show the convergence of the proposed strategy with methods 2 and 3, respectively. We can find that the system delay converges after 15 epochs and the value of delay eventually stabilised below five. These results demonstrate that the proposed DRL optimization strategy can effectively reduce the system delay and find the minimum value of the system delay.

Figure 5 shows the performance of the proposed strategy with method 1, where the value of W_{total} ranges from 30 to 70. When the task size of each UAV is 100M or 50M, the system delay decreases as W_{total} increases. This is because the increase in total bandwidth speeds up the transmission from the UAV to the EN and reduces system delay effectively. For example, the system delay at $W_{\text{total}} = 70$ is lower than the delay at $W_{\text{total}} = 30$. Similarly, Figures 6 and 7 show the performance of the proposed strategy with methods 2 and 3 when W_{total} ranges from 30 to 70, respectively. We can find that system delay decreases when the total bandwidth is increasing. These results demonstrate the effectiveness of proposed optimization strategy.

Figure 8 shows the performance of the proposed strategy with method 1, where the number of UAV ranges from 1 to 5. When the task size of each UAV is 100M or 50M, system delay increases as the number of UAVs increases. This i because the increase in the number of UAVs increases system burden and calculating delay. For example, the system delay when $n = 2$ is lower than the delay when $n = 5$. Similarly, Figures 9 and 10 show the performance of the proposed strategy with methods 2 and 3 when the number of UAVs ranges from 1 to 5, respectively. We can find that system delay increases when the number of UAVs is increasing. These results demonstrate that the proposed strategy can find the lowest system delay when the number of UAV ranges from 1 to 5.

5. Conclusions

This article studied a MEC system with one EN, where multiple unmanned aerial vehicles (UAVs) acted as users which had some heavy tasks. As the users generally had limitations in both calculating and power supply, the EN could help calculate the tasks and meanwhile supply the power to the users through energy harvesting. We optimized the system by proposing a joint strategy to unpacking and energy harvesting. Specifically, a deep reinforcement learning algorithm was implemented to provide a solution to the unpacking, while several analytical solutions were given to the power allocation of energy harvesting among multiple users. In particular, criterion I was the equivalent power allocation, criterion II was designed through equal data rate, while criterion III was based on the equivalent transmission delay. We finally gave some results to verify the joint strategy for the UAV-aided multiuser MEC system with energy harvesting.

Data Availability

The data can be obtained through email to the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key-Area Research and Development Program of Guangdong Province (No. 2018B010124001).

References

[1] J. Xia, F. Zhou, X. Lai et al., "Cache aided decode-and-forward relaying networks: from the spatial view," *Wireless Communications and Mobile Computing*, vol. 2018, 9 pages, 2018.

[2] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.

[3] X. Hu, C. Zhong, Y. Zhang, X. Chen, and Z. Zhang, "Location information aided multiple intelligent reflecting surface systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7948–7962, 2020.

[4] H. Yan, L. Hu, X. Xiang, Z. Liu, and X. Yuan, "PPCL: Privacy-preserving collaborative learning for mitigating indirect information leakage," *Information Sciences*, vol. 548, pp. 423–437, 2021.

[5] Z. Su, F. Biennier, Z. Lv, Y. Peng, H. Song, and J. Miao, "Toward architectural and protocol-level foundation for end-to-end trustworthiness in cloud/fog computing," *IEEE Transactions on Big Data*, vol. 8, no. 1, pp. 35–47, 2022.

[6] M. T. Islam, S. Karunasekera, and R. Buyya, "Performance and cost-efficient spark job scheduling based on deep reinforcement learning in cloud computing environments," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 7, pp. 1695–1710, 2022.

[7] X. Lai, L. Fan, X. Lei, Y. Deng, G. K. Karagiannidis, and A. Nallanathan, "Secure mobile edge computing networks in the presence of multiple eavesdroppers," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 500–513, 2022.

[8] J. Zhao, X. Sun, Q. Li, and X. Ma, "Edge caching and computation management for real-time internet of vehicles: an online and distributed approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2183–2197, 2021.

[9] L. Chen, R. Zhao, K. He, Z. Zhao, and L. Fan, "Intelligent ubiquitous computing for future UAV-enabled MEC network systems," *Cluster Computing*, vol. 2021, no. 1, pp. 1–11, 2021.

[10] F. Zhou and R. Q. Hu, "Computation efficiency maximization in wireless-powered mobile edge computing networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3170–3184, 2020.

[11] F. Wang, H. Xing, and J. Xu, "Real-time resource allocation for wireless powered multiuser mobile edge computing with energy and task causality," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 7140–7155, 2020.

[12] W. Zhou, D. Deng, J. Xia, and Z. Shao, "The precoder design with covariance feedback for simultaneous information and energy transmission systems," *Wireless Communications and Mobile Computing*, vol. 2018, 17 pages, 2018.

[13] Q. Tao, J. Wang, and C. Zhong, "Performance analysis of intelligent reflecting surface aided communication systems," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2464–2468, 2020.

[14] S. Arzykulov, A. Celik, G. Nauryzbayev, and A. M. Eltawil, "UAV-assisted cooperative & cognitive NOMA: deployment, clustering, and resource allocation," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 263–281, 2022.

[15] R. Akbar, S. Prager, A. R. Silva, M. Moghaddam, and D. Entekhabi, "Wireless sensor network informed UAV path planning for soil moisture mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[16] J. Zhao, Q. Li, Y. Gong, and K. Zhang, "Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7944–7956, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9090494, 14 pages
https://doi.org/10.1155/2022/9090494

Research Article

Statistical Characteristics of 3D MIMO Channel Model for Vehicle-to-Vehicle Communications

Asad Saleem,[1] **Yan Xu**,[2] **Rehan Ali Khan**,[3] **Iftikhar Rasheed**,[3] **Zain Ul Abidin Jaffri**,[4] **and Md Abu Layek**[5]

[1]*Shenzhen Key Laboratory of Antennas and Propagation, The College of Electronics and Information Engineering, Shenzhen University, China 518060*
[2]*School of Opto-Electronic Engineering, Zaozhuang University, Zaozhuang, China 277160*
[3]*Department of Electrical Engineering, University of Science and Technology Bannu, Pakistan*
[4]*College of Physics and Electronic Information Engineering, Neijiang Normal University, Neijiang 641100, China*
[5]*Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh*

Correspondence should be addressed to Md Abu Layek; layek@cse.jnu.ac.bd

Received 20 February 2022; Revised 8 April 2022; Accepted 21 April 2022; Published 21 June 2022

Academic Editor: Khaled Maaiuf Rabie

Spatial and temporal characteristics of the propagation channel have a significant influence on multiantenna method applicability for fifth-generation- (5G-) enabled Internet of Things (IoT). In this paper, the statistical characteristics of a novel three-dimensional (3D) geometric-based stochastic model for next-generation vehicle-to-vehicle (V2V) multiple-input multiple-output (MIMO) communications under the nonisotropic scattering environment are investigated. In both line-of-sight (LoS) and non-line-of-sight (NLoS) conditions, the proposed model investigates the spatial, frequency, and temporal domain statistical distribution of multipath received signals by using the time-variant transfer function for indoor environments. The probability density function (PDF) of separation distance between the transceiver antennas, angle-of-arrival (AoA), and angle-of-departure (AoD) in the azimuth and elevation planes is derived by using closed-form expressions. For the space, time, and frequency correlation function (STF-CF), a precise analytical expression is derived based on MIMO antenna system. We further determine the effects of several model parameters on the V2V channel performance, such as tunnel width, antenna array spacing, Ricean K-factor, and moving velocity. The statistical characteristics of the MIMO channel model are validated by simulation results, confirming the flexibility and effectiveness of our proposed model in the tunnel scenario.

1. Introduction

Recently, the wireless communication technologies like multiple input multiple output (MIMO), channel coding, cooperative communication, and internet of vehicles (IoV) have played a crucial role towards the development of future communication networks [1–10]. MIMO channel, which uses hundreds or even more antennas, has received a lot of attention. It has been demonstrated that MIMO technique is capable to improve the energy and spectrum efficiency, especially in rich scattering environment. The rapid advancement of the 5G-enabled Internet-of-Things (IoT) communication systems for vehicle-to-vehicle (V2V) radio communications has acquired much more attentions [11–14]. Unlike traditional fixed to mobile cellular networks, the V2V channel uses low elevation for multiple antennas when both the transmitting and receiving antennas are conceived in motion. The radio propagation parameters between the transmitting (M_T) and receiving (M_R) antennas must be designed to facilitate the thorough investigation of V2V radio communications [15]. Therefore, the realistic propagation channel model is required which can provide a quick and easy algorithm to approximate the statistical characteristics of V2V radio channel [16, 17].

To determine the optimized performance of the V2V radio communications in tunnel environments, precise channel models are under consideration. The V2V channel models can be classified as stochastic models and

deterministic models (mostly based on the ray tracing (RT) method). Stochastic models are classified into two types: regular-shaped geometrical-based stochastic models (RS-GBSMs) [18–22] and non-regular-shaped geometrical-based stochastic models (NGBSMs). The latter, also named as parametric models, are established from channel measurements, whereas the former are constructed from the regular geometrical-shape of the scattering objects. In [18], it is proved that the line-of-sight (LoS) path between the transceiver antennas is likely to be interrupted by buildings and barriers. As a result, Ricean channels must be used to investigate V2V communication systems. In [19], a precise visual scattering model is proposed for the multibounced V2V propagation channels in crowded urban communication environment; however, they did not address the influence of moving car velocity on the V2V channel propagation characteristics. In addition, [20, 21] introduce the 3D RS-GBSMs for microcell and macrocell communication systems, respectively.

However, the majority of the RSGBSMs models are based on narrowband channel assumptions, where the rays have almost the same propagation delay [22], which is not an accurate portrayal of V2V wireless communication system. In [23], the narrowband and wideband V2V channel characteristics based on propagation delays have been investigated through the real-time measurement campaign. The wide-sense stationary uncorrelated scattering (WSSUS) assumptions are used mostly in previous channel models; however, it is only applicable for short time intervals [24]. The WSSUS means actually first- and second-order stationarity in both time and frequency domains (wide-sense stationarity in time, uncorrelated scattering in delay) and holds in successive time and frequency interval. According to the measurement campaign in [25, 26], the stationary interval, in which the WSSUS assumption is still valid for V2V environments, is substantially smaller than the observation interval. This implies that in V2V channels, the WSSUS assumptions are not being fulfilled in most of the cases. Therefore, the stochastic channel models must need to consider the channel's nonstationarity properties to compensate these shortcomings.

Similarly, in [27], the algebraic framework of the transceiver antennas in the Doppler spectrum based on moving velocity and arbitrary delays has been investigated and velocity vectors are considered to characterize the arbitrary moving direction in V2V radio communication environment. In [28], a novel 3D-GBSM model based on an ellipsoid Gaussian scattering distribution is proposed that can jointly characterize the elevation angles, azimuth angles, and distances from the transmitter (receiver) to the first (last)-bounce scatterers. The spatial consistency of the channel is supported by tracking the locations of the transmitter, receiver, and scatterers. Reference [29] proposes a statistical wideband V2V MIMO channel model based on the Unitary matrix transformation algorithm. The V2V MIMO channel model for other taps is estimated by examining the propagation delays of the first tap with a Unitary matrix transformation algorithm. In [30], a nonstationary stochastic model has been presented between the vehicle and terminal using 3D antenna arrays and random trajectories with fixed velocity consideration. In [31], the delay-based Doppler PDF for V2V radio channels is computed by using a prolate spheroidal system. In [32], the authors proposed a V2V scattering model with the assumption of randomly distributed scatterers across the semicircular shaped tunnel sidewalls. In the context of randomly moving clusters, [33] proposes a nonstationary GBSM channel model for future intelligent V2V communications. Because of the nonstationarity induced by moving clusters, the clusters' movements are described using a random walk procedure.

Also, [34] presents a regular-shaped 3D-GBSM model for V2V MIMO system for nonisotropic fading channel. However, the channel statistics of PDF, time delays, and tunnel width impact at different locations, which is critical for V2V MIMO systems, have not been investigated in the given model. Therefore, in this paper, the 3D MIMO V2V channel is modeled using the rectangular single-bounce scattering tunnel (RSBST) model, rather than the geometrical semicircular or circular tunnel propagation models. The inspiration came from the idea that both of the geometrical rectangular and semicircular tunnels are generally used in MIMO systems. As arched and rectangular tunnel environments have different propagation characteristics, therefore, we feel that our proposed model is simple and more suitable for indoor rectangular tunnel environment. To our knowledge, no investigation of radio wave propagation in rectangular tunnel from the perspective of scatterer distribution and movement of V2V MIMO systems along tunnel width has been performed.

In this paper, we investigate the new geometrical RSBST model and compute its geometrical statistics with other scattering models to examine the influence of size and cross-sectional shape of tunnels on the V2V radio channel configurations. The key contribution of this study is to analyze the effect of different RSBST model parameters on V2V channel statistics and to obtain an optimized expression for scatterer distribution. The distribution of scatterers has been derived by considering the z-axis influence on the system performance and computed the 3D STF-CF of the V2V propagation channel in LoS and NLoS propagation environments. The impacts of numerous model parameters, such as tunnel diameter, antenna array spacing, Ricean K-factor, and moving velocity, on the V2V channel statistics are investigated as well.

The remainder of this paper is arranged in the following manner. Section 2 explains the proposed RSBST system model. In Section 3, different characteristics of the 3D-GBSM channel, such as TVTF, STF-CF, and power angular spread, are derived. In Section 4, the analytical results are presented and critically analyzed. Finally, Section 5 summarizes the paper.

2. Analysis of Geometric RSBST Model

The geometrical RSBST propagation model for the V2V MIMO system inside a rectangular-shaped tunnel is described in this section. Figure 1 represents a typical propagation scenario in an L-length tunnel. As illustrated in

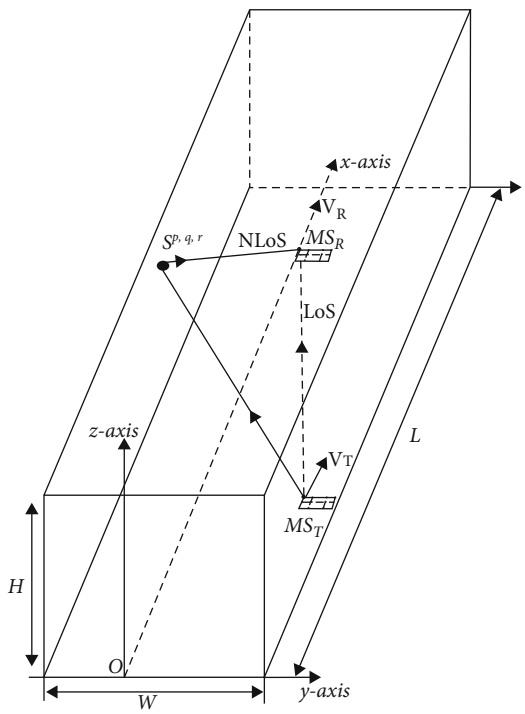


Figure 1: Propagation environments inside rectangle tunnels.

Figure 2, we assume that the scatterers are distributed randomly beside the tunnel sidewalls and ceiling. Due to cross-sectional size, shape, and geometry, the characteristics of a rectangular tunnel are substantially different from those of a circular or arch tunnel [35]. For example, in an arch tunnel, the tunnel width (W) is dependent on the tunnel radius due to semicircular geometry, whereas in a rectangular-shaped tunnel, the width (W) of the tunnel is independent from the height (H) of tunnel. As a result, when the scatterers are distributed on the tunnel ceiling and sidewalls, the z-axis affects the statistics of the V2V radio channel in a rectangular tunnel. Thus, we consider the impact of tunnels' z-axis when computing the STF-CF for next-generation V2V MIMO channel configurations in our proposed 3D RSBST model. In addition, Figure 3 shows the geometrical representation of RSBST V2V scattering model.

As illustrated in Figure 3, we employ a Cartesian coordinate approach to investigate the distribution of random scatterers S^{pqr} on the tunnel sidewalls and ceiling for $\forall m$, $1 \le p \le P$, $\forall q$, $1 \le q \le Q$, and $\forall r$, $1 \le r \le R$. On the rectangle tunnel sidewalls and ceiling, we suppose that the infinite number of scattering components (i.e., P, Q, $R \longrightarrow \infty$) is distributed randomly [36].

However, the WSSUS assumptions can be still fulfilled for V2V radio channels over the small time periods [26]. Therefore, the V2V radio channel inside the tunnel is assumed to be WSSUS in our analysis. The locations of the receiver (MS_R) and transmitter (MS_T) are defined by cartesian coordinates $(x_{R_X}, y_{R_X}, z_{R_X})$ and $(x_{T_X}, y_{T_X}, z_{T_X})$, when MS_R and MS_T are both moving in the opposite or same directions with velocity v. φ_{Rv} and φ_{Tv} determine the angle of motion for M_R and M_T antenna elements, respectively. Additionally, the MS_R and MS_T are equipped with a multiple M_R and M_T antenna elements, and the antenna element spacing is represented by δ_R and δ_T, respectively. The path length between the mth transmitting antenna A_T^m $(m = 1, 2, \cdots, M_T)$ and the scatterer S^{pq} is denoted by $d_T^{m,pq}$, whereas the distance between the nth receiving antenna A_R^n $(n = 1, 2, \cdots, N_R)$ and the scatterer S^{pq} is denoted by $d_T^{pq,n}$. Moreover, the $d_{TR}^{m,n}$ specifies the distance between the mth element of transmitting array and the nth element of receiving array in terms of LoS path. In our work, we assume single-bounce scattering (SB) approach and further analyze the scattering effects due to randomly distributed scatterers between transceiver antennas, also known as effective scattering elements. The elevation angles and orientation of the M_R (M_T) antenna array elements with respect to the xy plane can be represented by $\phi_R(\phi_T)$ and $\gamma_R(\gamma_T)$, respectively. The symbols $\beta_T^{p,q,r}$ and $\alpha_T^{p,q}$ signify the elevation angle of departure (EAoD) and azimuth angle of departure (AAoD) of the signals that are impinging on the S^{pq} scattering elements, whereas the angles $\beta_R^{p,q,r}$ and $\alpha_R^{p,q}$ represent the elevation angle of arrival (EAoA) and azimuth angle of arrival (AAoA) of the signals which are scattered from S^{pq} scattering elements, respectively. The EAoD, EAoA, AAoD, and AAoA in coordinate system (x_p, y_q) of scattering elements are given as follows:

$$\alpha_i^{p,q} = \begin{cases} f\left(x_p, y_q\right), & \text{if } y_q \ge y_i, \\ -f\left(x_p, y_q\right), & \text{if } y_q < y_i, \end{cases}$$
$$\beta_i^{p,q} = \begin{cases} g\left(x_p, y_q\right), & \text{if } z_r \ge z_i, \\ -g\left(x_p, y_q\right), & \text{if } z_r < z_i, \end{cases} \tag{1}$$

where $i = R$ or $i =$ and, the index i relates to the receiver or transmitter. The functions $f(x_p, y_q)$ and $g(x_p, y_q)$ are defined as follows:

$$f\left(x_p, y_q\right) = \cos^{-1}\left(\frac{x_p - x_i}{\sqrt{\left(x_p - x_i\right)^2 + \left(y_p - y_i\right)^2}}\right), \tag{2}$$

$$g\left(x_p, y_q\right) = \cos^{-1}\left(\frac{\sqrt{\left(x_p - x_i\right)^2 + \left(y_p - y_i\right)^2}}{\sqrt{\left(x_p - x_i\right)^2 + \left(y_p - y_i\right)^2 + \left(z_p - z_i\right)^2}}\right). \tag{3}$$

3. Statistical Characteristics of RSBST V2V Model

The reference model for the V2V MIMO channel under direct (LoS) and diffuse (NLoS) propagation conditions is presented here. The time-variant transfer function (TVTF),

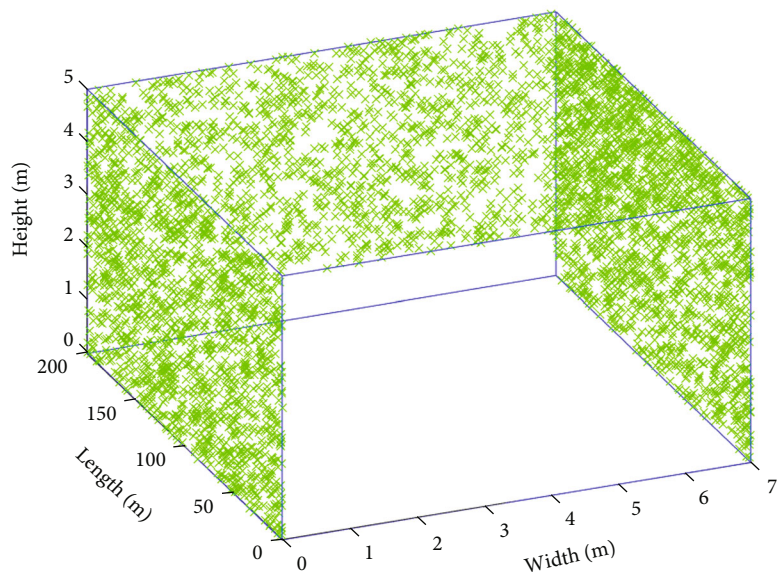


Figure 2: The distribution of scatterers along the sidewalls and ceiling of a tunnel when $W = 7$ m, $H = 5$ m, and length $L = 200$ m.

STF-CF, and power angular spread will be calculated from the geometrical RSBST scattering model and provided as a sum of LoS and diffuse components.

3.1. TVTF. The TVTF is denoted as $H_{nm}(t,f)$, which is the sum of LoS and diffuse components as

$$H_{nm}(t,f) = H_{nm}^{\mathrm{LoS}}(t,f) + H_{nm}^{\mathrm{diffuse}}(t,f), \tag{4}$$

where $H_{nm}^{\mathrm{LoS}}(t,f)$ and $H_{nm}^{\mathrm{diffuse}}(t,f)$ represent the TVTF of LoS and diffuse components, respectively. The $H_{nm}^{\mathrm{LoS}}(t,f)$ of LoS components is written as follows:

$$H_{nm}^{\mathrm{LoS}}(t,f) = \sqrt{\frac{k_{\mathrm{R}}^{\mathrm{LoS}}}{k_{\mathrm{R}}^{\mathrm{LoS}}+1}} e^{-j(2\pi/\lambda)d_{\mathrm{TR,LoS}}^{m,n}} e^{j2\pi\left[f^{o}t-\tau_{nm}^{o}f\right]}, \tag{5}$$

where

$$\begin{aligned} d_{\mathrm{TR,LoS}}^{m,n} = {} & d^{o} - (M_{\mathrm{T}} - 2m + 1)\frac{\delta_{\mathrm{T}}}{2}\cos\gamma_{\mathrm{T}}\cos\phi_{\mathrm{T}} \\ & - (M_{\mathrm{R}} - 2n + 1)\frac{\delta_{\mathrm{R}}}{2}\cos\left(\alpha_{\mathrm{T}}^{o} - \gamma_{\mathrm{R}}\right)\cos\phi_{\mathrm{R}}, \end{aligned} \tag{6}$$

$$d^{o} = \sqrt{(x_{\mathrm{R}} - x_{\mathrm{T}})^2 + (y_{\mathrm{R}} - y_{\mathrm{T}})^2 + (z_{\mathrm{R}} - z_{\mathrm{T}})^2}, \tag{7}$$

$$f^{o} = f_{\mathrm{T}}^{o} + f_{\mathrm{R}}^{o}, \tag{8}$$

$$f_{\mathrm{T}}^{o} = f_{\mathrm{T}_m}^{o}\cos\left(\beta_{\mathrm{T}}^{o}\right)\cos\left(\alpha_{\mathrm{R}}^{o} - \varphi_{v}^{\mathrm{T}}\right), \tag{9}$$

$$f_{\mathrm{R}}^{o} = f_{\mathrm{R}_m}\cos\left(\beta_{\mathrm{R}}^{o}\right)\cos\left(\alpha_{\mathrm{R}}^{o} - \varphi_{v}^{\mathrm{R}}\right). \tag{10}$$

In (8), the Doppler shifts of the LoS components generated by the movement of the M_{T} and M_{R} antennas are denoted by f_{T}^{o} and f_{R}^{o}, respectively. The angles β_{T}^{o}, α_{T}^{o}, β_{R}^{o}, and α_{R}^{o} in (9) and (10) represent the EAoD, AAoD, EAoA, and AAoA of the LoS component, respectively. The $k_{\mathrm{R}}^{\mathrm{LoS}}$ in (5) provides the Ricean K-factor, which is defined as the mean power ratio of LoS component with the diffuse components. In (5), τ_{nm}^{o} represents the propagation delays of LoS components and given by $\tau_{nm}^{o} = d_{\mathrm{TR,LoS}}^{m,n}/c_{o}$, where c_{o} is speed of light. The $H_{nm}^{\mathrm{diffuse}}(t,f)$ of diffuse components for the $A_{\mathrm{T}}^{m} - A_{\mathrm{R}}(n)$ transmission link can be determined as

$$\begin{aligned} H_{nm}^{\mathrm{diffuse}}(t,f) = {} & \lim_{P,Q,R\longrightarrow\infty}\left(\frac{1}{\sqrt{\left(k_{\mathrm{R}}^{\mathrm{diffuse}}+1\right)PQR}}\right)\sum_{p,q,r=1}^{P,Q,R} e^{j} \\ & \left[\theta^{p,q,r} + 2\pi\left(f_{\mathrm{T}}^{p,q,r} + f_{\mathrm{R}}^{p,q,r}\right)t - 2\pi\tau_{nm}^{p,q,r}f\right] \times e^{-j(2\pi/\lambda)\left(d_{\mathrm{T}}^{m,pq}+d_{\mathrm{R}}^{pq,n}\right)}, \end{aligned} \tag{11}$$

where

$$f_{\mathrm{T}}^{p,q,r} = f_{\mathrm{T}_m}\cos\left(\beta_{\mathrm{T}}^{p,q,r}\right)\cos\left(\alpha_{\mathrm{T}}^{p,q} - \varphi_{v}^{\mathrm{T}}\right), \tag{12}$$

$$f_{\mathrm{R}}^{p,q,r} = f_{\mathrm{R}_m}\cos\left(\beta_{\mathrm{R}}^{p,q,r}\right)\cos\left(\alpha_{\mathrm{R}}^{p,q} - \varphi_{v}^{\mathrm{R}}\right), \tag{13}$$

$$\begin{aligned} d_{\mathrm{T}}^{m,pq} = {} & \sqrt{(x_m - x_{\mathrm{T}})^2 + (y_m - y_{\mathrm{T}})^2 + (z_m - z_{\mathrm{T}})^2} \\ & - (M_{\mathrm{T}} - 2m + 1)\frac{\delta_{\mathrm{T}}}{2} \\ & \left[\cos\beta_{\mathrm{T}}^{p,q,r}\cos\phi_{T} \times \cos\left(\gamma_{\mathrm{T}} - \alpha_{\mathrm{T}}^{p,q}\right) + \sin\beta_{\mathrm{T}}^{p,q,r}\sin\phi_{\mathrm{T}}\right], \end{aligned} \tag{14}$$

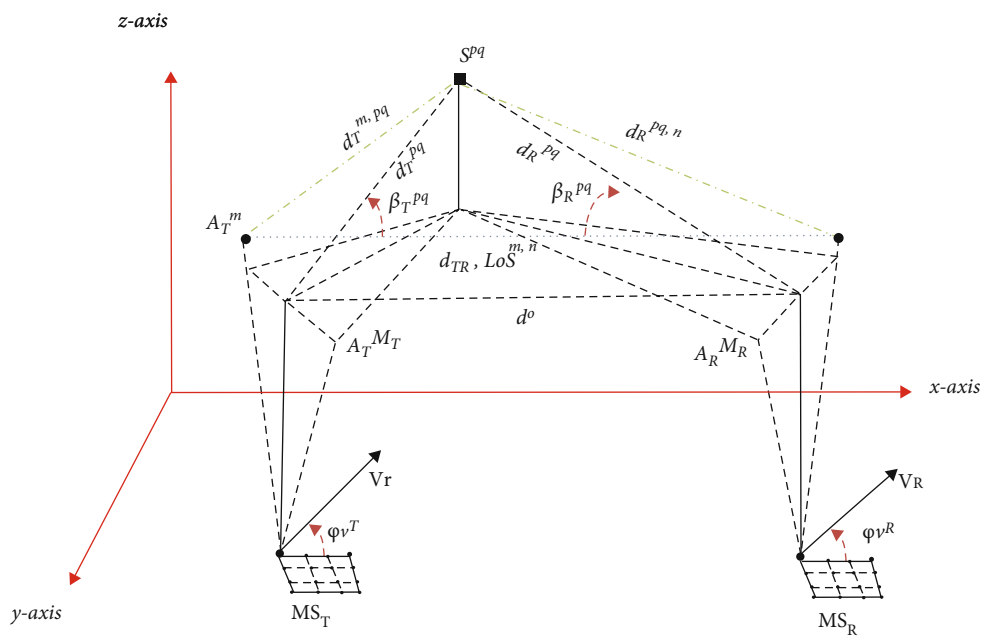


Figure 3: Geometrical RSBST scattering model in the presence of specular component (- · -) and LoS component (····) for $M_T \times M_R$ MIMO RSBST channel.

$$d_R^{pq,n} = \sqrt{(x_m - x_R)^2 + (y_m - y_R)^2 + (z_m - z_R)^2} - (M_R - 2n + 1)\frac{\delta_R}{2}\left[\cos\beta_R^{p,q,r}\cos\phi_R \times \cos\left(\gamma_R - \alpha_R^{p,q}\right) + \sin\beta_R^{p,q,r}\sin\phi_T\right], \tag{15}$$

where the symbols f_{T_m}, f_{R_m}, k_R^{diffuse}, and $\tau_{nm}^{p,q,r}$ represent the maximum Doppler frequency of MS_T and MS_R, Ricean K -factor, and the delays of diffuse components, respectively. In addition, $d_T^{m,pq}$ and $d_R^{pq,n}$ denote the distances between scatterer $S^{p,q,r}$ to mth transmitting antenna array A_T^m $(m = 1, 2, \cdots, M_T)$ and from the scatterer $S^{p,q,r}$ to the nth receiving antenna array components $A_R^n (m = 1, 2, \cdots, M_R)$, respectively.

3.2. STF-CF. The STF-CF is the correlation of $H_{nm}(t, f)$ and $H_{n'm'}(t + \tau, f + \upsilon)$ in (4), which can be calculated as

$$\begin{aligned}\rho_{nm,n'm'}(\delta_T, \delta_R, \tau, \nu) &= E[H^*{}_{nm}(t, f) \cdot H_{n'm'}(t + \tau, f + \nu)] \\ &= \rho_{nm,n'm'}^{\text{diffuse}}(\delta_T, \delta_R, \tau, \nu) + \rho_{nm,n'm'}^{\text{LoS}}(\delta_T, \delta_R, \tau, \nu),\end{aligned} \tag{16}$$

where $(\cdot)^*$ represents the conjugate value operator and $E(\cdot)$ denotes the expected value operator. $\rho_{nm,n'm'}^{\text{LoS}}(\delta_T, \delta_R, \tau, \nu)$ and $\rho_{nm,n'm'}^{\text{diffuse}}(\delta_T, \delta_R, \tau, \nu)$ represent the STF-CF of the LoS and the diffuse components, respectively. Thus, the STF-CF $\rho_{nm,n'm'}^{\text{LoS}}(\delta_T, \delta_R, \tau, \nu)$ can be calculated as follows:

$$\rho_{nm,n'm'}^{\text{LoS}}(\delta_T, \delta_R, \tau, \nu) = \frac{k_R^{\text{LoS}}}{k_R^{\text{LoS}} + 1} k_{mm'}^o k_{nn'}^o e^{j2\pi\left(f^o\tau + \tau'^{\,o}_{nm}\nu'\right)}, \tag{17}$$

where

$$k_{mm'}^o = e^{j2\pi(\delta_T/\lambda)(m-m')\cos\gamma_T\cos\varphi_T}, \tag{18}$$

$$k_{nn'}^o = e^{j2\pi(\delta_R/\lambda)(n-n')\cos(\alpha_R{}^o - \gamma_R)\cos\varphi_R}. \tag{19}$$

Moreover, $\rho_{nm,n'm'}^{\text{diffuse}}(\delta_T, \delta_R, \tau, \nu)$ can be estimated as

$$\rho_{nm,n'm'}^{\text{diffuse}}(\delta_T, \delta_R, \tau, \nu) = \lim_{P,Q,R\longrightarrow\infty}\left(\frac{1}{\sqrt{\left(k_R^{\text{diffuse}} + 1\right)PQR}}\right)\sum_{p,q,r=1}^{P,Q,R} E \left\{k_{mm'}^{p,q,r} k_{nn'}{}^{p,q,r} \times e^{j2\pi\left[f_T^{p,q,r} + \tau f_R^{p,q,r} - \tau'^{\,p,q,r}_{nm}\nu\right]}\right\}. \tag{20}$$

Note that the propagation delays $(\tau'^{\,p,q,r}_{mm})$, and the Doppler frequencies $(f_T^{p,q,r})$ are based on x_m, y_n, and z_r coordinates. Due to the infinite number of scattering elements in proposed V2V model, the random coordinates x_m, y_n, and z_r can be replaced by independent variables x, y, and z. Furthermore, the x and y are assumed to be uniformly

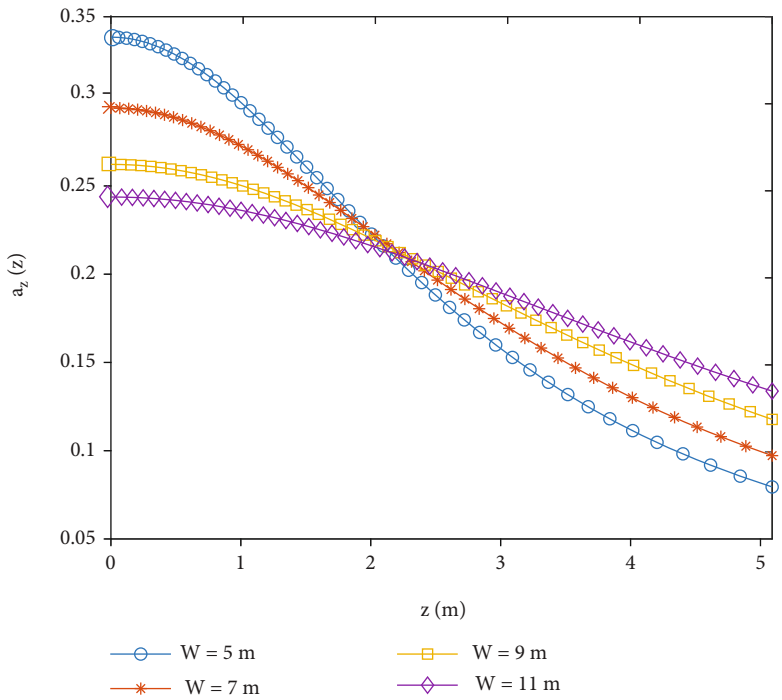


FIGURE 4: PDF $a_z(z)$ effect based on variable z and tunnel width (W).

distributed, and their PDF can be estimated as follows:

$$a_x(x) = \frac{1}{x_R - x_T}, \quad \text{if } x_R \geq x \geq x_T, \tag{21}$$

$$a_y(y) = \frac{1}{W}, \quad \text{if } \frac{W}{2} \geq x \geq -\frac{W}{2}. \tag{22}$$

The expression $z = W/2 \cdot \tan\alpha$ can be used to calculate the distribution of variable z, where z is randomly distributed in the range of $[0 \quad \tan^{-1}(2H/W)]$. By considering the transformation of given variables, the PDF of variable z can be derived as follows:

$$a_z(z) = \frac{1}{\tan^{-1}(2H/W)} \frac{2W}{W^2 + 4z^2}, \quad \text{if } H \geq z \geq 0. \tag{23}$$

The $a_z(z)$ of variable z for different widths (W) is shown in Figure 4. We can see that the $a_z(z)$ of z in (23) drops down the uniform value within the range of $z \in [0, H]$, as the width (W) of tunnel increases.

Thus, by considering PDFs of randomly distributed variables x, y, and z, the joint PDF $a_{xyz}(xyz)$ can be calculated as

$$a_{xyz}(xyz) = a_x(x) \cdot a_y(y) \cdot a_z(z),$$
$$a_{xyz}(xyz) = \frac{2}{(W^2 + 4z^2)(x_R - x_T)\tan^{-1}(2H/W)} \quad \text{if } z \in [0, H], y \in \left[\frac{W}{2}, -\frac{W}{2}\right], x \in [x_T, x_R]. \tag{24}$$

Similarly, by exploiting the joint PDF $a_{xyz}(xyz)$, the STF-CF $\rho^{\text{diffuse}}_{nm,n'm'}(\delta_T, \delta_R, \tau, \nu)$ of diffuse components in (20) can be rewritten as

$$\rho^{\text{diffuse}}_{nm,n'm'}(\delta_T, \delta_R, \tau, \nu) = A \iiint_{x_T, -W/2, 0} k_{mm'}(x, y, z) k_{nn'}(x, y, z) \times \frac{1}{W^2 + 4z^2} \times e^{j2\pi\left[f_T{}^{x,y,z} + \tau f_R{}^{x,y,z} - \tau'_{nm}{}^{p,q,r} \nu\right]} \tag{25}$$

where

$$A = \frac{2}{(k_R + 1)(x_R - x_T)\tan^{-1}(2H/W)}, \tag{26}$$

$$k_{mm'}(x, y, z) = e^{j2\pi(\delta_T/2)(m-m')\cos(\gamma_T - \alpha_T(x,y))\cos(\beta_T(x,y,z))\cos(\phi_T)} \times e^{j2\pi(\delta_T/2)(m-m')\sin(\beta_T(x,y,z))\sin(\phi_T)}, \tag{27}$$

$$k_{nn'}(x, y, z) = e^{j2\pi(\delta_R/2)(n-n')\cos(\gamma_R - \alpha_R(x,y))\cos(\beta_R(x,y,z))\cos(\phi_R)} \times e^{j2\pi(\delta_R/2)(n-n')\sin(\beta_R(x,y,z))\sin(\phi_R)}. \tag{28}$$

It should be noted that the propagation angles $\beta_T(x, y, z)$, $\beta_R(x, y, z)$, $\alpha_T(x, y, z)$, and $\alpha_R(x, y, z)$ in (27) and (28) are dependent on the x, y, and z coordinates of $S^{p,q,r}$.

3.3. Power Angular Spread. The power angular spread (PAS) can be configured by exploiting the Bartlett method [37]. For electromagnetic analysis, the Bartlett beamforming technique is used to estimate the angle of departure and angle of arrival in azimuth and elevation planes. The Bartlett

algorithm is predicated based on the steering vector $a(a_A^{p,q}, \beta_E^{p,q,r})$. The Bartlett beamforming algorithm is commonly known as the Fourier spectrum analyzing approach. The objective is to discover the weighting vector w that maximizes the power of the incoming signal. The receiving antenna array elements can receive the signals from different spatially distributed users. Furthermore, the received power includes both LoS and reflected (NLoS) path signals, all of which appear to be emanating from different directions and angles. The steering vector of the transmitted signals $x_1(t)$ can be estimated as

$$a\left(a_A^{p,q}, \beta_E^{p,q,r}\right) = \mathrm{vec}\left\{\begin{bmatrix} 1 \\ e^{jm} \\ e^{j2m} \\ \vdots \\ e^{j(w-1)m} \end{bmatrix}\left[1, e^{jn}, e^{j2n}, \cdots e^{j(l-1)n}\right]\right\}, \tag{29}$$

where $m = k_w d_y \cos a_A^{p,q} \cos \beta_E^{p,q,r}$, $n = k_w d_x \cos a_A^{p,q} \cos \beta_E^{p,q,r}$, λ is the wavelength, w is the width of antenna array, and l represents the length of antenna arrays, respectively. Also, d_y, d_x, $\beta_E^{p,q,r}$, and $a_A^{p,q}$ denote the antenna array spacing along the y-axis, antenna array spacing along the x-axis, elevation angle along the positive y-axis, and azimuth angle along the positive $x-$ axis. Furthermore, the signal vector of the arrays can be written as

$$y(t) = a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)x_1(t) + n(t)\ , \tag{30}$$

where $n(t)$ denotes the noise power. Also, if K sources have the same time slot and frequency, the received signals can be written as follows:

$$y(t) = \sum_{k=1}^{K} a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)x_k(t) + n(t). \tag{31}$$

By assuming that the received signals are coming through the azimuth direction, so the output of the array by Bartlett can be calculated as follows:

$$\max E\left\{w^H y(t) y^H(t) w\right\} = \max E\left\{x(t)|^2 \left|w^H \cdot a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)\right|^2 . \sigma^2 |w|^2\right\}, \tag{32}$$

where σ^2 represents the noise power variance and $[\cdot]^H$ specifies the transpose conjugate of the given vector. The significant solution of (29) is as follows:

$$w_B = \frac{a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)}{\sqrt{a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)^H a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)}}\ , \tag{33}$$

where $a(a_A^{p,q}, \beta_E^{p,q,r})$ denotes the normalized vector; thus, the weighting vector of the Bartlett algorithm becomes

$$w_B = a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)\ . \tag{34}$$

This means that the weighting vector is the same as the incident wave spatial characteristics. The final expression for estimating the PAS is given here only. Reference [37] contains further information about the Bartlett algorithm. The PAS of $a(a_A^{p,q}, \beta_E^{p,q,r})$ can be estimated as

$$P\left(a_A^{p,q}, \beta_E^{p,q,r}\right) = \frac{a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)^H R_{xx} a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)}{a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)^H a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)}\ , \tag{35}$$

where R_{xx} is the covariance matrix of received signals, and it may be calculated as

$$R_{xx} = \frac{1}{M_{\mathrm{R}} \times M_{\mathrm{T}}} \sum_{i=1}^{I} vec[Y(i)]\mathrm{vec}[Y(i)]^H\ , \tag{36}$$

where $Y(i) \in C^{M_{\mathrm{R}} \times M_{\mathrm{T}}}$ denotes the received signals at the ith delay bin and vec$[\cdot]$ is the linear transformation operator, which sequentially turns the matrix columns into a single column vector. If the normalized vector $a(a_A^{p,q}, \beta_E^{p,q,r})$ is used, the total PAS can be stated as

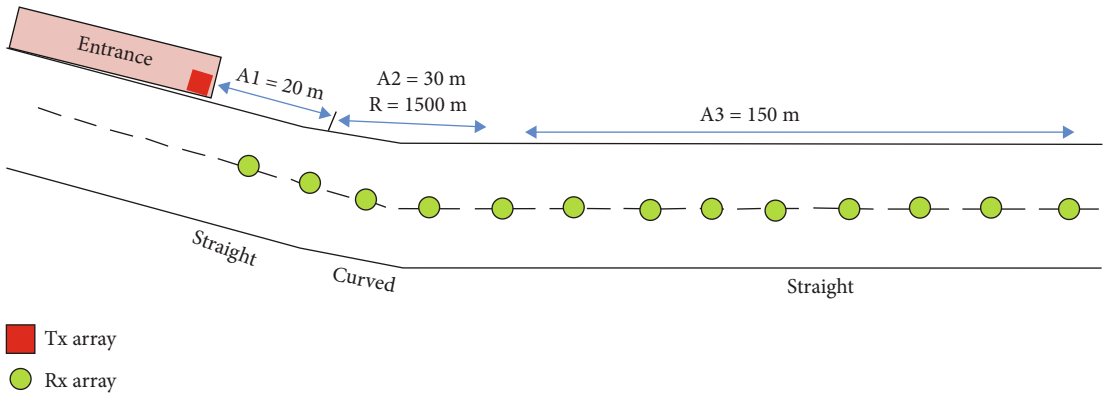
$$P\left(a_A^{p,q}, \beta_E^{p,q,r}\right) = a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)^H R_{xx} a\left(a_A^{p,q}, \beta_E^{p,q,r}\right)\ . \tag{37}$$

In V2V scenario, the geometric path length is usually time-variant due to the relative movement between the M S_{R} and MS_{T} antenna arrays. Compared to several existing 3D and 2D fixed-to-mobile (F2M) and V2V channel models, the proposed channel model can be utilized more efficiently in tunnel environment; for example, as the relative time is equal to zero, the proposed model tends to be wide-sense-stationary (WSS) channel, which is similar to the F2M channel [38]. However, the proposed model represents nonstationary V2V channels where the relative time is not equal to zero, as described in [39]. Moreover, the proposed 3D model can reflects a wide range of communication scenarios by modifying the geometrical model parameters, as indicated in Sections 2 and 3. For example, the proposed channel model can represent the 3D semiellipsoid channel when the relative time is set to zero, as demonstrated in [40].

4. Numerical Result Analysis

4.1. Simulation Scenario Description and Parameter Adjustment. The numerical findings obtained by assessing the STF-CF for the LoS and the NLoS propagation environments are presented in this section. Simulations are used to verify the accuracy of the proposed analytical results. To determine the optimized performance of the V2V MIMO channel, the STF-CF under different separating distances of transceivers, different moving speed, and different K -factor values are compared with similar dimensions of an indoor tunnel between Qihua road to University station in Shanghai, China. We have considered the 8×8 MIMO

(a)



(b)

Figure 5: Simulated environment: (a) real-time tunnel view; (b) tunnel structure.

Table 1: Parameter adjustment.

Parameters	Values
Frequency	5.6 GHz
Bandwidth	160 MHz
M_T, M_R	8, 8
δ_T, δ_R	0.5 λ
Transceiver antenna array structure	Omnidirectional
ϕ_T, ϕ_R	60°
f_{T_m}, f_{R_m}	0, [120 240 360]
ϕ_T, ϕ_R	30°
γ_T, γ_R	45°
Tunnel area	200 m (length)/7 m (width)/5 m (height)
Farthest measurement distance	200 m
Number of sampling locations	101
Ricean K-factor	3, 6

system in a rectangular-shaped tunnel with a height of $L = 200$ m, a width of $W = 7$ m, and a height of $H = 5$ m using our proposed geometrical propagation model. In our simulation results, we have considered the SNR as 20 dB. Basically, the SNR is influenced by the path loss, so we have fixed it to check the influence of multipath components effect. The real-time inner view of indoor Shanghai tunnel and geometrical representation is given in Figure 5. The tunnel is divided into two sections, the first of which is a 20 m long tunnel platform or entrance with a 6.5 m high

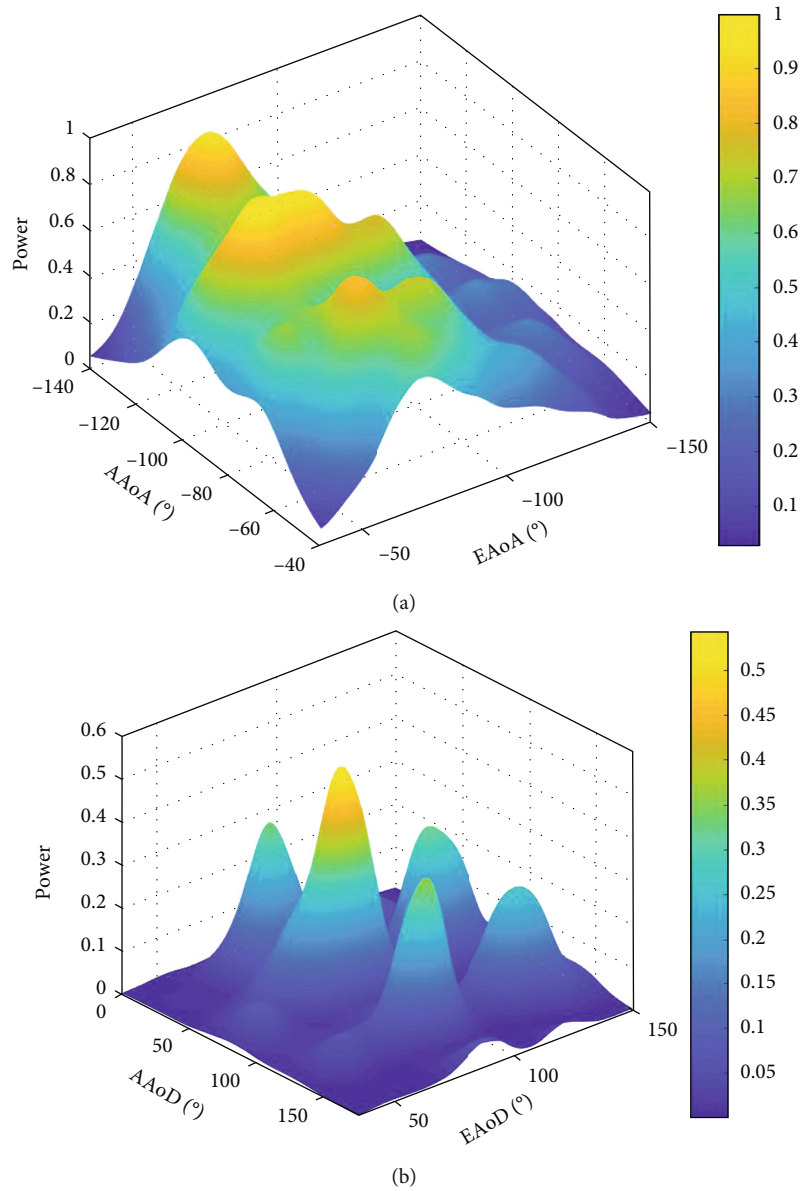


Figure 6: The azimuth and elevation polarization distribution for indoor tunnel: (a) AAAoA and EAoA; (b) AAoD and EAoD.

rectangular cross-section. A nonplatform structure with a 5.96 m high arch cross-section makes up the second section. Throughout this 200 m distance, 101 different measuring locations are taken into account, each of which is 2 m apart. The following parameters are considered to produce numerical results: $\phi_T = \phi_R = 60^\circ$, $\gamma_T = \gamma_R = 45^\circ$, $\varphi_T = \varphi_R = 30^\circ$, $(x_T = 0, y_T = 0, z_T = 3.1\,\text{m})$, $(x_R = 0, 2, \cdots, 200\,\text{m}, y_R = 3.5\,\text{m}, z_R = 2.7\,\text{m})$, and $f_{R_m}(Hz) = [120\ 240\ 360]$. The MS_T is fixed to the tunnel sidewalls at the entrance, whereas the MS_R moves along the tunnel axis. The tunnel walls are reinforced with concrete material. The propagation angles $\beta_T(x, y, z)$, $\beta_R(x, y, z)$, $\alpha_T(x, y, z)$, and $\alpha_R(x, y, z)$ are estimated according to (27) and (28). The first 50 m distance in tunnel is conceived as LoS path, where the rest of 150 m is considered as NLoS path. Table 1 also lists the additional parameters that are used to generate our simulation findings. The Monte Carlo approach [41] is used to simulate the STF-CF under various K-factors, time separations (τ), and Doppler shifts (f_{R_m}). We can use the Monte Carlo approach to multiply the iteration of the TVTF's NLoS components since the LoS components are deterministic while the NLoS components are based on a random process. The number of iterations for the Monte Carlo technique is fixed to 2×10^4 in this paper. In our simulations, we have considered the SNR as 20 dB. Basically, the SNR is influenced by the path loss, so we have fixed it to check the influence of multipath component effect. In this paper, our aim is to check the channel correlation, where the channel correlation is influenced by the SNR and channel matrix simultaneously. In our simulation results, we pay more attention to the influence of wireless channel matrix on correlation in different environments and different MS_T-MS_R distances, which requires removing

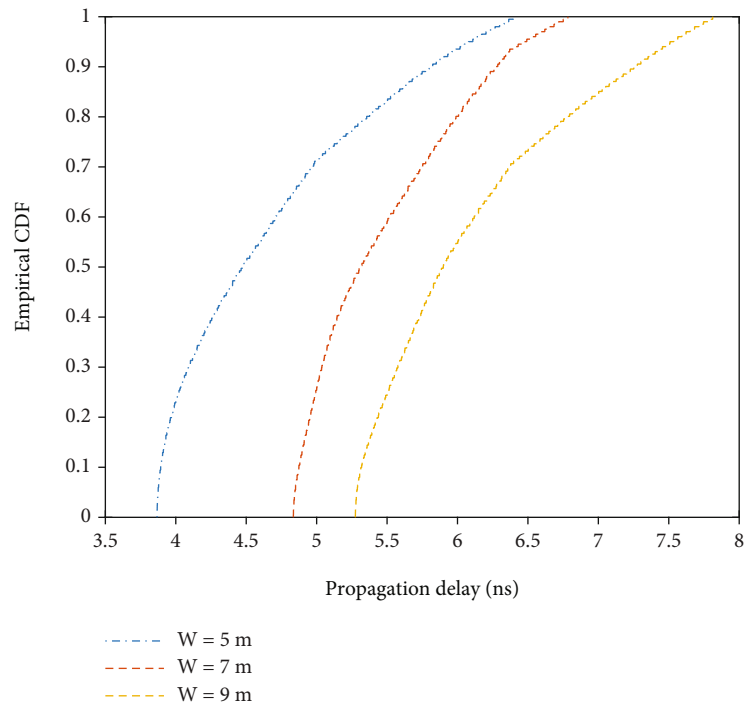


Figure 7: CDF of propagation delays with different tunnel widths (W).

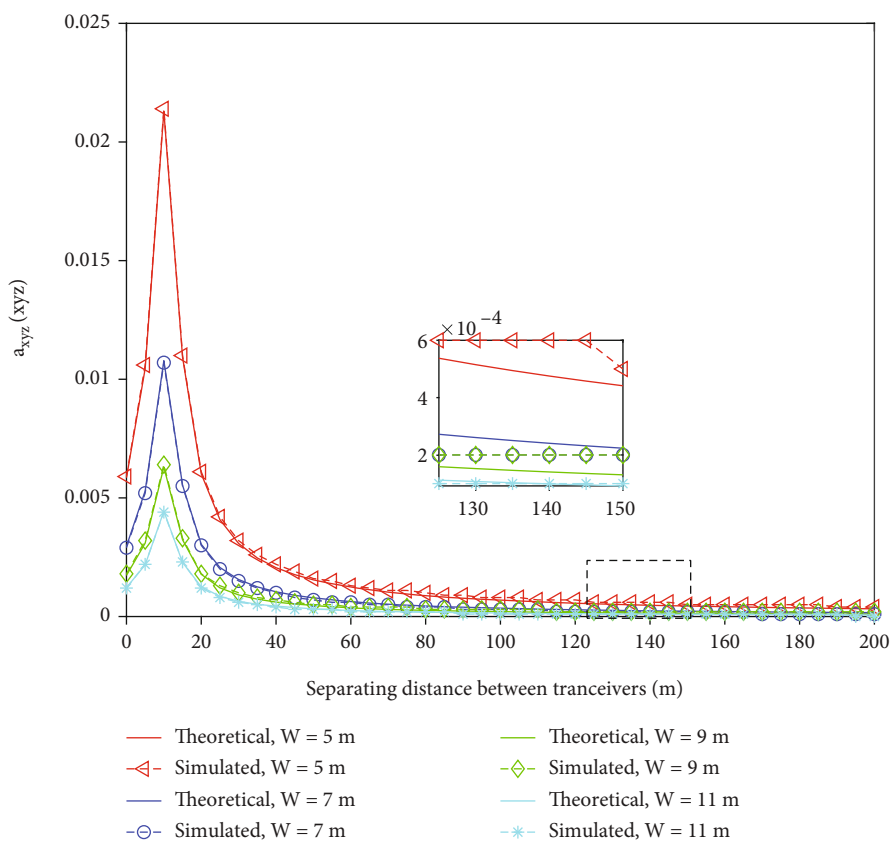


Figure 8: The joint PDF with different tunnel widths (W).

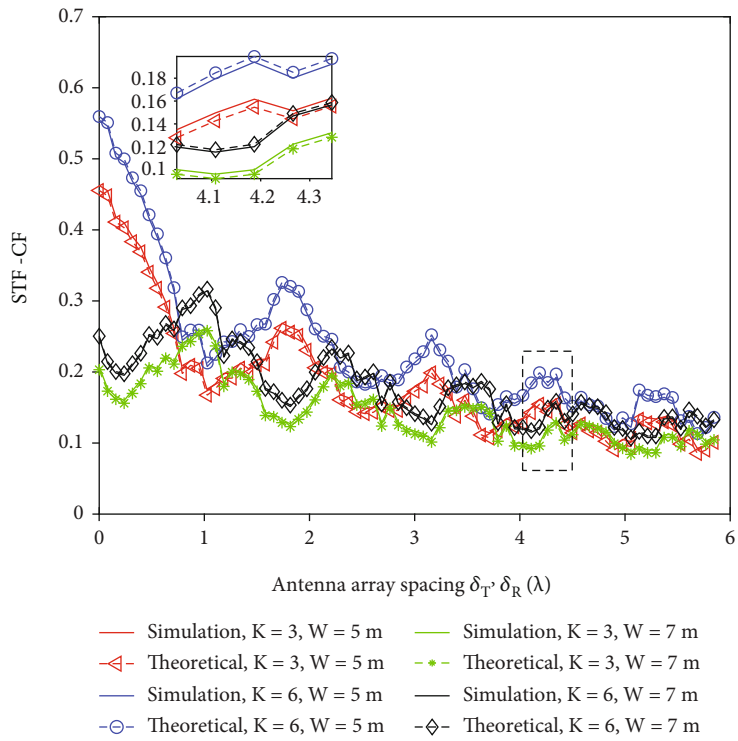


Figure 9: Magnitude of the STF-CFs of the RSBST model with different W and K-factor values.

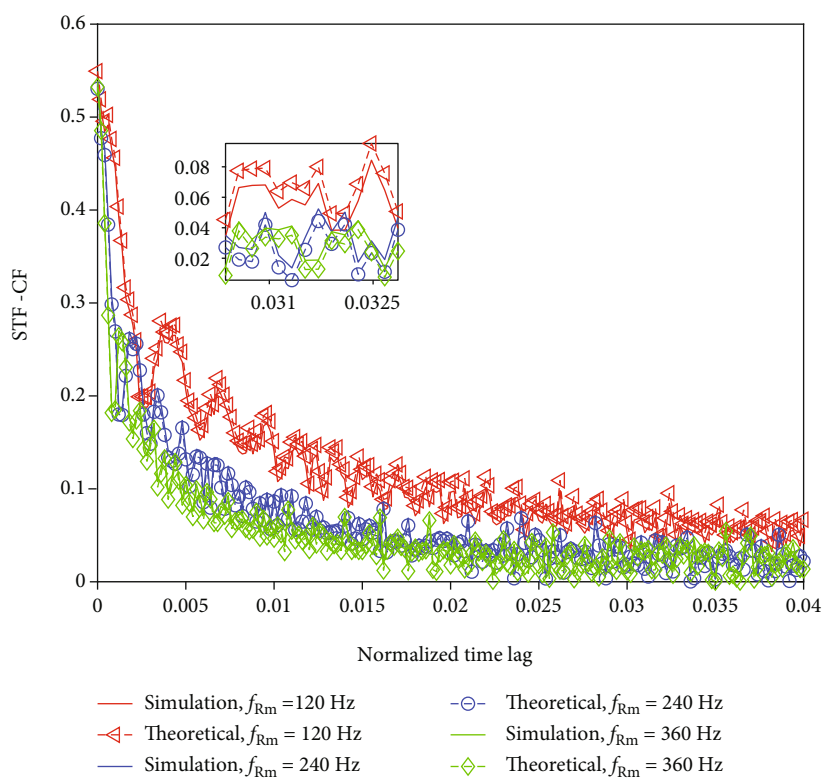


Figure 10: Magnitude of the STF-CFs of the RSBST model with different Doppler frequencies.

the influence of path loss and signal-to-noise ratio so that channel correlation can be compared in a fair way. After fixing the SNR, channel matrix can express the multipath effects in a proper way. In our simulations, the bandwidth is 160 MHz, which means the sampling interval of two adjacent multipath components can be as low as 6.25 ns (1/160 MHz). Because the coherent bandwidth of the channel is lower than the long-term-evolution for metro (LTE-M) signal bandwidth, we used wideband channel assumptions for the calibrations. In this paper, we have considered the reflections from the tunnel sidewalls and ceiling only for simulation results.

4.2. Angular Distribution and Propagation Delays. In Figure 6, we have shown power angular distribution by considering the elevation angle of departure (EAoD), azimuth angle of departure (AAoD), elevation angle of arrival (EAoA), and azimuth angle of arrival (AAoA) at 100 m separation distances between the transceiver antenna arrays ($x_R - x_T = 100$ m). Here, the number of scatterers is conceived as 2000 [42]. We can see that EAoA and AAoA both have negative magnitude which is because of $z_r < z_i$ and $y_q < y_i$, respectively. The power and angular spread is configured by using the Bartlett method [37]. From Figure 6, we can observe that the propagative angles reproduce two main beams at −58° of elevation and azimuth of −83° and −46° for EoA and AoA (see Figure 6(a)) and at 84° elevation and azimuths of 88° and 133° for EAoD and AAoD (see Figure 6(b)), respectively. Moreover, the propagation delays (τ) are shown in Figure 7 by using RSBST approach at 5.6 GHz operating frequency. The results represent the cumulative distribution function (CDF) of propagation delays in the presence of different tunnel widths ($W = 5$ m, 7 m, and 9 m). It can be seen that propagation delays are increased by increasing the tunnel width. A modal theory can explain this pretty well [43]. Multiple propagation modes are triggered at the shorter MS_T to MS_R distances. At a greater distance, particularly in the case of NLoS propagation, high-order model attenuation becomes critical, and only the low-order primary model is retained. As a result, the reflected environment's delay spread is relatively steady. The tunnel has a smaller path loss exponent in the given frequency band, but it still displays the waveguide effect. Furthermore, the signal's shorter wavelength (λ) increases the specular reflections, which boosts the waveguide effect in the tunnel for higher frequency transmissions. When the multipath components are obstructed by the tunnel sidewalls and ceiling, the waveguide effect is considerably decreased. Also, due to the enormous 1500 m radius of curvature (R), there are only a few path loss exponents in the NLoS region.

4.3. PDF and Correlation Function. In Figure 8, the theoretical vs. the simulation analysis of joint PDF of proposed RSBST model for different tunnel widths (W) is shown by using (24). Here, antenna array spacing is set as $\delta_T = \delta_R 0.5\ \lambda$. It can be shown that joint PDF grows as the separation distance between the MS_T and the MS_R reduces and decreases as the width (W) of the tunnel increases. Furthermore, the simulated results are in good agreement with the theoretical values of $a_{xyz}(xyz)$, verifying the validity of the proposed model. In a multipath channel, the initial geometric path length has a significant impact on the PDF of the transceiver separation distance. It is worth noting that the MS_R and MS_T are installed in two semiellipsoid scenarios in [44], and the PDF statistics are compatible with our proposed 3D model. Due to the symmetry of the geometric radio channel model, each joint PDF $a_{xyz}(xyz)$ curve in Figure 8 is symmetric, which is also compatible with [44]. Thus, the proposed 3D MIMO model can be used to characterize the real tunnel environment.

Figure 9 presents the STT-CF for different K-factor values, when $f_{R_m} = 360$ Hz, coherence time $\tau = 1$ ns, and separating distance between transceiver arrays is $x_R - x_T = 200$ m. As shown in Figure 9, the STF-CF decreases as the tunnel width (W) or MS_T and MS_R antenna array spacing (δ_T, δ_R) increases, and the STF-CF increases as the K-factor increases. Meanwhile, the STF-CF simulation values are in good agreement with the theoretical values, indicating that the proposed model is suitable. The LoS components have a negative impact on MIMO channel performance. Figure 10 represents the STF-CF for different normalized time lag $\tau \cdot f_{R_m}$ and Doppler shift ($f_{R_m} = [120$ Hz (23 km/h), 240 Hz (46 km/h), and 360 Hz (69 km/h)]) for MS_R when $K = 3$, and distance between MS_T and MS_R is set to $x_R - x_T = 200$ m. The STF-CF of the simulation results fits very well with the theoretical results, as illustrated in Figure 10. It is worth noting that when temporal separation increases, the STF-CF decreases. In our proposed model, we have considered the LoS and diffused components which provide different STF-CF values instead of 1 under different antenna spacing and normalized time lag, as given in [45]. When the Doppler shift f_{R_m} raises, the STF-CF reduces rapidly. The moving velocity of vehicles affects the V2V channel statistics of the MIMO system, as shown in [46], and this behavior can also be shown in Figure 10.

5. Conclusions

In this paper, we have presented a novel 3D scattering model for V2V MIMO radio communications in geometrical rectangle tunnel environment. Moreover, the statistical characteristics and analytical results of the proposed model allow one to understand the real tunnel propagation environment for 5G-enabled IoT communications, as it assumes the randomly distributed scatterers along the tunnel sidewalls and ceiling. A generic analytical expression for the STF-CF has been obtained by considering the nonisotropic scattering conditions in both LoS and NLoS propagation environments. Near- and far-field effects are taken into account in this RSBST model to characterize the spherical wavefronts, resulting in AAoA, EAoA, EAoD, and AAoD statistics and PDF and STF-CF responses on the MIMO antenna arrays. It is proved that the nonstationarity, including moving direction φ_v and propagation delays, affects the MIMO channel performance. From the analytical results, we can deduce that the STT-CF is seriously affected by propagation

delay, K-factor, movement velocity of receiving antenna array, and antenna array spacing. The proposed model results are aligned very well with previous propagation models, indicating that the proposed massive MIMO channel model can be generalized.

Data Availability

The data that support the research findings are available on request. The data are not publicly available due to the privacy of research participants.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] W. U. Khan, X. Li, A. Ihsan, M. A. Khan, V. G. Menon, and M. Ahmed, *NOMA-enabled optimization framework for next-generation small-cell IoV networks under imperfect SIC decoding (IoTV)*, EEE Transactions on Intelligent Transportation Systems, 2021.

[2] M. Asif, W. Zhou, M. Ajmal, and N. A. Khan, "A construction of high performance quasicyclic LDPC codes: a combinatoric design approach," *Wireless Communications and Mobile Computing*, vol. 2019, 10 pages, 2019.

[3] W. U. Khan, T. N. Nguyen, F. Jameel et al., *Learning-based resource allocation for backscatter-aided vehicular networks*, IEEE Transactions on Intelligent Transportation Systems, 2021.

[4] A. Saleem, F. Zhang, M. Wang, X. Yin, and G. Zheng, "Proficiency of leaky coaxial cable-based MIMO system using radiated field distribution," *International Journal of Antenna and Propagation*, vol. 2018, 13 pages, 2018.

[5] M. Asif, W. Zhou, Q. Yu, X. Li, and N. A. Khan, "A deterministic construction for jointly designed quasicyclic LDPC coded-relay cooperation," *Wireless Communications and Mobile Computing*, vol. 2019, 12 pages, 2019.

[6] W. U. Khan, M. A. Javed, T. N. Nguyen, S. Khan, and B. M. Elhalawany, *Energy-efficient resource allocation for 6G backscatterenabled NOMA IoV networks*, IEEE Transactions on Intelligent Transportation Systems, 2021.

[7] M. Asif, W. Zhou, Q. Yu, S. Adnan, M. S. Ali, and M. S. Iqbal, *Jointly designed quasi-cyclic LDPC-coded cooperation with diversity combining at receiver*, IEEE Transactions on Intelligent Transportation Systems, 2020.

[8] A. Saleem, M. Wang, G. Zheng, and X. Yin, "Spatial characteristics of wideband channels using leaky coaxial cables in tunnel scenario," *International Journal of Antenna and Propagation*, vol. 2019, 2019.

[9] M. Asif, W. U. Khan, H. Afzal et al., "Reduced-complexity LDPC decoding for next-generation IOT networks," *Wireless Communications and Mobile Computing*, vol. 2021, 10 pages, 2021.

[10] W. U. Khan, F. Jameel, N. Kumar, R. Jäntti, and M. Guizani, "Backscatter-enabled efficient V2X communication with non-orthogonal multiple access," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1724–1735, 2021.

[11] M. N. Peter and M. P. Rani, "V2V communication and authentication: the internet of things vehicles(IoTV)," *Wireless Personal Communications*, vol. 120, no. 1, pp. 231–247, 2021.

[12] C. Zhao, M. Dong, K. Ota, J. Li, and J. Wu, "Edge-MapReduce-based intelligent information-centric IoV: cognitive route planning," *IEEE Access*, vol. 7, pp. 50549–50560, 2019.

[13] S. Al-Sarawi, M. Anbar, K. Alieyan, and M. Alzubaidi, "Internet of Things (IoT) communication protocols: Review," in *2017 8th International Conference on Information Technology (ICIT)*, pp. 685–690, Amman, Jordan, 2017.

[14] N. Faisal, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Non-orthogonal HARQ for URLLC: design and analysis," *IEEE Internet of Things Journal*, vol. 8, no. 24, pp. 17596–17610, 2021.

[15] C. Zhang, K. Ota, J. Jia, and M. Dong, "Breaking the blockage for big data transmission: gigabit road communication in autonomous vehicles," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 152–157, 2018.

[16] L. Bai, Z. Huang, Y. Li, and X. Cheng, "A 3D cluster-based channel model for 5G and beyond vehicle-to-vehicle massive MIMO channels," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 9, pp. 8401–8414, 2021.

[17] A. Saleem, M. Wang, and G. Zheng, "Scattering of electric field from leaky coaxial cable in confined area," *International Journal of Antenna and Propagation*, vol. 2017, 9 pages, 2017.

[18] A. S. Akki and F. Haber, "A statistical model for mobile-to-mobile land communication channel," *IEEE transactions on vehicular technology*, vol. 35, no. 1, pp. 2–7, 1986.

[19] H. Jiang, Z. Zhang, J. Dang, and L. Wu, "Analysis of geometric multibounced virtual scattering channel model for dense urban street environments," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 1903–1912, 2017.

[20] J. Bian, C. X. Wang, J. Huang et al., "A 3D wideband non-stationary multi-mobility model for vehicle-to-vehicle MIMO channels," *IEEE Access*, vol. 13, no. 7, pp. 32562–32577, 2019.

[21] H. Jiang, M. Mukherjee, J. Zhou, and J. Lloret, "Channel modeling and characteristics for 6G wireless communications," *IEEE Network*, vol. 35, no. 1, pp. 296–303, 2021.

[22] H. Jiang, Z. Zhang, L. Wu, J. Dang, and G. Gui, "A 3-D non-stationary wideband geometry-based channel model for MIMO vehicle-to-vehicle communications in tunnel environments," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, p. 6257, 2019.

[23] I. Sen and D. W. Matolak, "Vehicle–vehicle channel models for the 5-GHz band," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 2, pp. 235–245, 2008.

[24] A. Paier, T. Zemen, L. Bernado et al., "Non-WSSUS vehicular channel characterization in highway and urban scenarios at 5.2 GHz using the local scattering function," *international itg workshop on smart antennas*, pp. 9–15, 2008.

[25] C. X. Wang, A. Ghazal, B. Ai, Y. Liu, and P. Fan, "Channel measurements and models for high-speed train communication systems: a survey," *IEEE communications surveys & tutorials*, vol. 18, no. 2, pp. 974–987, 2016.

[26] J. Chebil, H. Zormati, and J. B. Taher, "Geometry-based channel modelling for vehicle-to-vehicle communication: a review," *International Journal of Antennas and Propagation, vol.*, vol. 2021, pp. 1–10, 2021.

[27] M. Walter and A. Dammann, "Relative positioning and velocity estimation using V2V delay and Doppler information," *Proceedings of the 31st International Technical Meeting of the Satellite Division of The Institute of Navigation*, 2018.

[28] B. Ji, C. X. Wang, X. Gao, X. You, and M. Zhang, "A general 3D non-stationary wireless channel model for 5G and beyond," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, 2021.

[29] J. Hao, B. Xiong, Z. Zhang et al., "Novel statistical wideband MIMO V2V channel modeling using unitary matrix transformation algorithm," *PIEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4947–4961, 2021.

[30] A. Naeem, B. Hua, Q. Zhu, K. Mao, and J. Bao, "A novel GBSM for non-stationary V2V channels allowing 3D velocity variations," *Sensors*, vol. 21, no. 9, 2021.

[31] M. Walter, D. Shutin, and U. W. Fiebig, "Prolate spheroidal coordinates for modeling mobile-to-mobile channels," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 155–158, 2015.

[32] N. Avazov and P. Matthias, "A novel wideband MIMO car-to-car channel model based on a geometrical semi-circular tunnel scattering model," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 1070–1082, 2015.

[33] X. Baiping, Z. Zhang, J. Zhang, H. Jiang, J. Dang, and L. Wu, "Novel multi-mobility V2X channel model in the presence of randomly moving clusters," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3180–3195, 2021.

[34] J. Huang, W. Cheng-Xiang, C. Hengtai, S. Jian, and G. Xiqi, "Multi-frequency multi-scenario millimeter wave MIMO channel measurements and modeling for B5G wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2010–2025, 2020.

[35] A. Hrovat, G. Kandus, and T. Javornik, "Impact of tunnel geometry and its dimensions on path loss at UHF frequency band," *Proceedings of the International Conference on Circuits, Systems, Communications & Computers (SysCon')*, pp. 253–258, 2011.

[36] J. M. Molina-Garcia-Pardo, M. Lienard, and P. Degauque, *Wireless communications in tunnel*, In Tech, Rijeka, Croatia, 2012.

[37] M. A. Sadoon, R. Abd-Alhameed, and N. J. McEwan, "The impact of the covariance matrix sampling on the angle of arrival estimation accuracy," *Inventions*, vol. 4, no. 3, pp. 43–63, 2019.

[38] P. Petrus, J. H. Reed, and T. S. Rappaport, "Geometrical-based statistical macrocell channel model for mobile environments," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 495–502, 2002.

[39] N. Avazov and M. Pätzold, "A geometric street scattering shannel model for car-to-car communication systems," in *International Conference on Advanced Technologies for Communications*, pp. 224–230, Da Nang, Vietnam, 2011.

[40] J. Chen, S. Wu, S. Liu, C. Wang, and W. Wang, "On the 3-D MIMO channel model based on regular-shaped geometry-based stochastic model," in *2015 International Symposium on Antennas and Propagation (ISAP*, Hobart, TAS, Australia, 2015.

[41] G. Li, L. Klingbeil, F. Zimmermann, S. Huang, and H. Kuhlmann, "An integrated positioning and attitude determination system for immersed tunnel elements: a simulation study," *Sensors*, vol. 20, no. 24, p. 7296, 2020.

[42] R. Zhicheng, F. Zhang, G. Zheng, A. Saleem, and K. Guan, "A 3D non-stationary channel model with moving mobile station in rectangular tunnel," *International Journal of Antennas and Propagation*, vol. 2019, Article ID 6750153, 12 pages, 2019.

[43] D. G. Dudley, M. Lienard, S. F. Mahmoud, and P. Degauque, "Wireless propagation in tunnels," *IEEE Antennas and Propagation Magazine*, vol. 49, no. 2, pp. 11–26, 2007.

[44] M. Riaz, N. M. Khan, and S. J. Nawaz, "A generalized 3-D scattering channel model for spatiotemporal statistics in mobile-to-mobile communication environment," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4399–4410, 2015.

[45] G. Bakhshi, R. Saadat, and K. Shahtalebi, "Modeling and simulation of MIMO mobile-to-mobile wireless fading channels," *International Journal of Antennas and Propagation*, vol. 2012, 13 pages, 2012.

[46] H. Jiang, Z. Zhang, L. Wu, and J. Dang, "Three-dimensional geometry-based UAV-MIMO channel modeling for A2G communication environments," *IEEE Communications Letters*, vol. 22, no. 7, pp. 1438–1441, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9661411, 7 pages
https://doi.org/10.1155/2022/9661411

Research Article

A Robust Graph Theoretic Solution of Routing in Intelligent Networks

Muhammad Aasim Qureshi,[1] Mohd Fadzil Hassan,[2] Muhammad Khurram Ehsan,[3] Muhammad Owais Khan,[4] Md YeaKub Ali,[5] and Shafiullah Khan[6]

[1]*Department of Computer Sciences, Bahria University, Lahore Campus, Lahore, Pakistan*
[2]*Computer and Information Science Department, University Teknologi, Petronas, Malaysia*
[3]*Faculty of Engineering Sciences, Bahria University, Lahore Campus, Lahore, Pakistan*
[4]*School of Systems and Technology, University of Management and Technology, Lahore, Pakistan*
[5]*Department of Electronics and Telecommunication Engineering, Rajshahi University of Engineering & Technology (RUET), Rajshahi, Bangladesh*
[6]*Department of Electronics, Islamia College University, Peshawar, Pakistan*

Correspondence should be addressed to Md YeaKub Ali; yeakub@ete.ruet.ac.bd

Received 12 April 2022; Revised 20 May 2022; Accepted 3 June 2022; Published 20 June 2022

Academic Editor: Khaled Maaiuf Rabie

Implementation of robust routing is very critical in network communication. Connecting devices like routers maintain databases for the whole network topology in the routing table. Each router needs to keep these tables updated with the best possible routes so that an efficient communication can always take place in nondelay tolerant intelligent networks that include military and tactical systems, vehicular communication networks, underwater acoustic networks, and intelligent sensor networks. The fast construction of shortest-path tree (SPT) is important to devise an efficient routing in a nondelay tolerant networks. That is why a simple and efficient algorithm is the need of the time. A robust routing solution SPT with $O(V+E)$ time complexity is proposed to supersede the existing landmark.

1. Introduction

Network routing is a backbone of intelligent communication networks specifically nondelay tolerant networks that include military and tactical systems, vehicular communication networks (VCNs), underwater acoustic networks (UANs), and intelligent sensor networks (ISNs) to send the communication data between vertices in a certain network. For the purpose, every router maintain routing tables. These tables manage routes of various network destinations. This network routing process is generally accomplished based on these routing tables and routes defined in them. This makes the formation, written in a memory of these routers, critically important for successful and efficient routing. There exist many graph theoretic solutions of this routing problem. In a network, the vertex of a graph represents a router, and edge/link which connect these vertices represents physical links between the routers.

The needs for broadband Internet applications have increased rapidly in today's Internet. Which makes high speed routing inevitable and has become the key at open shortest-path first (OSPF). As a result of topological changes that take place due to due to an unavoidable circumstances at the OSPF, these algorithms update the routing table to handle the new topological changes. For example, if some link fails in a network, then there is no way but to recalculate these paths [1]. In such scenarios, these shortest paths are regenerated by running the algorithms again [2].

We can formally define shortest-path problem as follows:

Let $G = (V, E)$, where $V = \{v_1, v_2, v_3, v_4, \cdots, v_n\}$ and $E = \{e_1, e_2, e_3, e_4, \cdots, e_m\}$.

G is an undirected nonnegative weighted connected graph (in this scenario, it will be representing the routing network). G has a vertex "s" designated as source and another vertex, "t,", designated as destination, such that $s, t \in V$. We need to find simple path(s) from s to t with minimum incurred cost.

[3] presented a hybrid model with local and global modes. A fixed graph model is used to present the local mode connections. In the paper, they denoted graph by a unit-disk graph UDG(V): for any $u, v \in V$, {u,v} ∈ UDG(V) if and only if distance between u and v is less than 1. For the local mode, in each round, node u can send a message of $O(\log n)$ bits to node v. For the global mode, they used a variant of the node-capacitated clique (NCC) model presented in [4] that describes salient features of overlay networks. In [5], the authors explain the SPSN approach including route object modeling, capacity-oriented search, and possible route organization procedures. Prototype assessments presented significant support that SPSN is better than the legacy CGR in more than an order of magnitude in contact plans comprising thousands of contacts.

The shortest-path tree (SPT) problem is one of the most well-known problems in graph theory [6]. Since 1959, most of the advances in SPT for general graphs, directed and undirected, are based on Dijkstra's algorithm. There exist lots of applications of SPT problems like computer systems, transportation networks, and vehicle routing [7]. Time complexity of the basic algorithm of Dijkstra [8] is calculated as $O(V^2 + E)$ (where V is number of vertices and E is number of edges) if linear search is applied to find the minimum [9]. [10, 11] introduced a new data structure, heap, to find the minimum. This new data structure improved the time complexity to $O(E \log V)$. Fibonacci heap improved this bound further [12]. Authors of [12] used comparison model for the optimal implementation of Dijkstra's algorithm as Dijkstra's algorithm traverses vertices in ascending order.

In [13], with the use of optimal structure of the data, heap data structure provides the complexity of $O(E \log V)$ with implementation of Dijkstra and Fibonacci heap gives the complexity of $O(E + V \log V)$ with the implementation of Dijkstra. For the stochastic decision-making problem in [14], the stochastic dynamic programming solution is required. The generalized elementary shortest-path problem in [15] shows that two-phase heuristic graphs by including negative cycles can improve the average gap of 0.3% if compared with the best known solutions if number of vertices N are set to 100. The solution in [16] suggests that these algorithms can be implemented in all kind of networks with complexity of Dijkstra $O(V^2 + E)$, Bellman $O(V)^3$, and Floyd $O(VE)$. The paper [17] suggests a solution with $O(kV \log(kV) + kE)$ complexity for a fast delivery problem. The complexity of Bellman is $O(V)^3$, and the complexity of Dijkstra is $O(V)^2$ if the network is fully connected or near to the fully connected [18].

Invent of fusion trees and its application in Dijkstra's algorithm make randomized bound of time complexity, $O(E(\log V)1/2)$. Later, the invention of atomic heaps improved this time bound to $O(E + n \log(v/\log(\log V)))$ bound and $O(E + V(\log V + \varepsilon)1/2)$ [12]. Afterwards, in [6, 19, 20], priority queues improved to $O(E \log(\log(V))$ and $O(E + V \log V^{1+\epsilon})^{1/2}$. These worst-case time bounds are randomized supposing linear space are prerequisite. The researcher of [21] further improved it to $O(E + V(\log V \log \log V)^{1/2})$ which was improved by the authors of [16], next year, using randomized bound to $O(E + V \log V^{1+\epsilon})^{1/3}$.

[2] presented priority queue, which improved single source shortest-path cost to $O(E + V(\log C)^{1/2}$ where C is the cost of the most heavy edge. This bound was further improved by [22] to $O(E + V(3 \log C \log\log C)^{1/3})$ expected time, and [23] presented a further improvement to $O(E + V(\log C)^{1/4+\epsilon})$. The authors of [24] claimed that the algorithm they have presented will outclass Dijkstra algorithm.

In [25], the authors presented an exact adjacency-based primal algorithm to resolve the SPT problem with high resources. Their algorithm explored solution space iteratively using path adjacency-based partition. They claimed optimal convergence with good path in lesser cost. Authors of [26] introduced a multiobjective (time and cost) shortest-path tree problem for connected directed graphs. Every edge was having two cost values. They proved that uncertainty MSPP reduced the total travel time and cost.

[27] used a dual evolutionary algorithm while analyzing the problem of clustered shortest-path tree. Their first algorithm (N-LSE) targeted the minimization of search space with fewer vertices. The second algorithm named M-LSE using N-LSE creates smaller and multiple intercluster edges. [28] solved the dynamic SP problem to solve time-dependent TSP. In logistic planning where traffic data is rich enough, time-dependent path optimization techniques give considerable benefits. Dynamic shortest-path solution contains no length k cycles and resolves the complexity problem when compared to acyclic graph-based methods.

In [29], the author reviewed many popular algorithms for solving the SPT problem in weighted graphs. He proposed two variants of Dijkstra's algorithm. He presented asymptotic comparison different variants of Dijkstra's algorithm on expanded graphs. He presented $O(E' + V' \lg V')$ time algorithm where $E' = E + C^-_{\max} C^+_{\max}$ and $V' = V(C^-_{\max} + C^-_{\max})$. They used color-in and color-out scheme.

2. Concept

This work presents a novel algorithm of finding shortest-path routes in intelligent networks. This algorithm is based on bidirectional search where first search is starting from source vertex, and its target is destination vertex, while second search starts from the destination vertex, and its target is source vertex. It is analogous to two robots looking for each other from two distinct points. First robot, R_A, starts its search from its current position, and let us call it P_A, and second robot, R_B, starts its search from its current position, and let us call it P_B. R_A explores all possible paths, in

BFS fashion, looking for P_B or R_B. In the same way, R_B will be exploring all possible paths, simultaneously, looking for either P_A or R_A. Following are the three possible situations that can rise, in this bidirectional search.

(i) R_A will find P_B

(ii) R_B will find P_A

(iii) R_A and R_B meets at position, P_M

Situation i is the simplest one. Let Path_A be the path that R_A has explored and has vertices $a_1, a_2, a_3, \cdots, a_x$, where a_x is P_B.

$$\text{Required path} = \text{Path}_A = \sum_{i=1}^{x-1} e_{a_i,a_{i+1}}. \tag{1}$$

In situation ii, R_B has found P_A. It means that path is found but in reverse order. Let R_B has explored the vertices in sequence $b_1, b_2, b_3, \cdots, b_y$, where b_y is P_A.

$$\text{Path}_B = \sum_{i=1}^{y-1} e_{b_i,b_{i+1}}. \tag{2}$$

The reverse of this will generate the required path. See the following equation:

$$\text{Required path} = \sum_{i=y}^{1} e_{b_i,b_{i-1}}. \tag{3}$$

In situation iii, it is little complex than the previous two. Two paths—Path_A and Path_B, need to be concatenated. Let Path_A be the path that R_A has explored and has vertices $a_1, a_2, a_3, \cdots, a_x$. Path_B is the path that R_B has explored and has vertices $b_1, b_2, b_3, \cdots, b_y$, where a_x and b_y are the same vertices as P_M and a_1 and b_1 are P_A and P_B, respectively. The required path will be having vertices in the sequence $a_1, a_2, a_3, \cdots, a_x, b_{y-1}, b_{y-2}, b_{y-3}, \cdots, b_1$, as given in equation (3).

$$\text{Required path} = \sum_{i=1}^{x-1} e_{a_i,a_{i+1}} + \sum_{i=y}^{2} e_{b_i,b_{i-1}}. \tag{4}$$

3. Proposed Algorithm

The proposed algorithm targets only those graphs in which on two paths, from source, P_i and P_j, to fulfil the following property:

$$Len(P_i(s,w)) > Len(P_j(s,w), \tag{5}$$

$$\sum_{\substack{i=0 \\ \text{for } P_i}}^{k-1} w_{x_i,x_{i+1}} > \sum_{\substack{i=0 \\ \text{for } P_j}}^{l-1} w_{y_i,y_{i+1}}, \tag{6}$$

where $k \neq l$ and $x_0 = y_0 = s$ and $x_k = y_k = w$ and P_i and P_j are paths from s to w such that $P_i = \{x_1, x_2, x_3, \cdots, x_k\}$ and $P_j = \{y_1, y_2, y_3, \cdots, y_l\}$.

The search is executed, parallel, from both ends—source, s, and destination, t. Vertices are explored level by level. Best paths are calculated by concatenating the new edge with previously explored best path, and cost is calculated by adding the cost of all edges in that path from source.

Each vertex during the execution goes through three states, being represented by three colors—green, yellow, and red. The definitions of these colors can be seen in Table 1

Initially, all vertices, i.e., $v_1, v_2, v_3, \cdots, v_n$, are initialized with

$$\begin{cases} \text{SPCost}_{v_i} \longleftarrow \infty, \\ \pi_{v_i} \longleftarrow \varnothing, \\ C_{v_i} \longleftarrow \text{green}. \end{cases} \tag{7}$$

Starting from s and t, independently and simultaneously, the current vertex, let us call it v_{cur}, will explore its neighbors. Every neighbor, i.e., v_{nbr}, is updated with shortest-path cost, and let us call it SPCost, parent, i.e., π and color, i.e., C.

$$\text{SPCost}_{v_{\text{nbr}}} = \min\left(\text{SPCost}_{v_{\text{cur}}} + w_{v_{\text{cur}},v_{\text{nbr}}}, \text{SPCost}_{v_{\text{nbr}}}\right), \tag{8}$$

$$\pi_{v_{\text{nbr}}} = v_{\text{cur}} \text{ if } \text{SPCost}_{v_{\text{nbr}}} \text{ is updated}, \tag{9}$$

$$\text{Clr}_{v_{\text{nbr}}} = \text{yellow if } \text{SPCost}_{v_{\text{nbr}}} \text{ is updated}. \tag{10}$$

Update will be effective if new $\text{SPCost}_{v_{\text{nbr}}}$ reflects improvement over previously stored one.

Initially all vertices have green status. As soon as a vertex is explored during the search, its status is updated to yellow. As soon as any vertex completes its exploration (i.e., all its neighbors are explored, and their status is changed to yellow), its status is updated to red.

Exploration takes place, strictly, level by level. No next level vertex is explored until all the previous level yellow vertices' status is change to red. As soon as one level is completed, the control is switched to the other part of the algorithm to proceed and it also performs the identical steps.

During the search of any side if some red (light or dark) vertex is explored by the counterpart search, then two shortest paths are calculated and stored. If some shortest path has already been stored, then the path will less cost will be kept.

The algorithm terminates when the status of all vertices is updated to red (light or dark). In worst case, depending upon the type of the graph, it finishes in $O(V+E)$ time. If graph is totally disconnected, even then loop in line number 4 of *ShortestPath* will execute for $O(V)$ times. If graph is connected and is having E, then combined effect of line number 4 of *ShortestPath* and line number 2 of *SPfromSource* or *SPfromDestination* would be "all adjacents of all vertices" will cost $O(E)$ which will make it $O(V+E)$. Detailed algorithm can be seen in Algorithm 1.

Table 1: Definitions of colors that represent the state.

Color	Definition
Green	Unseen: the vertex is yet virgin, no neighboring vertex has yet seen/explored it and considered it to be the part of the shortest path
Yellow	Seen: some neighboring vertex has seen/explored it and considered it to the part of shortest path
Red	Visited: this vertex has played its part and has explored all its neighbors, making them yellow

```
ShortestPath(s, t, G)
1. Queue1.enQueue(s)
1. Queue2.enQueue(t)
2. InitializeDoubleSource(SPCost, Pi, Clr, s, t, G)
3.      Flag ← 0
4. While (Queue1 <> Null OR Queue2 <> Null) AND Flag = 0
5.      Do SPfromSource(SPCost, Pi, Clr, G)
6.         SPfromDestination(SPCost, Pi, Clr, G)
SPfromSource(G)
1. while u ← Queue1.deQueue() <> Green
2. for each v belongs to Adj[u]
3.      do if Clr_v = Green OR LightYellow OR LightRed
4.            then Clr_v ← LightYellow
5.                  SPCost_v ← min(SPCost_u +W_(u,v), SPCost_v)
6.                  If SPCost is updated
7.                     Then Pi_v ← u
8.                  If Clr_v = Green OR LightRed
9.                     Then Queue.enQueue(v)
10.            If Clr_v = DarkRed
11.               Then Concatinate two paths and store
12.                     CPCost ← min(CPCost_u + CPCost_v + w_(u,v), CPCost)
13.            If Clr_v = DarkYellow
14.               then skip
15. Clr_u ← LightRed
SPfromDestination(G)
1. while u ← Queue2.deQueue() <> Green
2. for each v belongs to Adj[u]
3.   do if Clr_v = Green OR DarkYellow OR DarkRed
4.      then Clr_v ← DarkYellow
5.      SPCost_v ← min(SPCost_u +W_(u,v), SPCost_v)
6.      If SPCost_v is updated
7.         Then Pi_v ← u
8.      If Clr_v = Green OR DarkRed
9.         Then Queue.enQueue(v)
10.   If Clr_v = LightRed
11.      Then Concatinate two paths and store
12.            SPCost_v ← min(SPCost_u + SPCost_v + w_(u,v), SPCost_v)
13. If Clr_v = LightYellow
14.            then skip
15. Clr_u ← DarkRed
```

Algorithm 1.

4. Working Example of the Algorithm

All vertices are initialized using equation (7) (see Figure 1). Both sources—*A* and *P*, are colored yellow and their SPCosts are marked 0 (see Figure 1(a)). All vertices are enqueued. Main subroutine (named ShortestPath) of the algorithm calls the two replicated subroutines (named SPfromSource and SPfromDestination) one by one until queue has any vertex.

The subroutine SPfromSource dequeues a vertex, it would be source, and here in this example, it is *A*. Algorithm then explores all its neighbors—*B*, *F*, and *E*, one by one.

Equations (8)–(10) are used to update the values of SPCost, π, and Clr. In the same way, SPfromDestination dequeues a vertex, it would be destination, and here in this example, it is *P*. Algorithm then explores all its neighbors—*L* and *K*, one by one. Equations (8)–(10) are

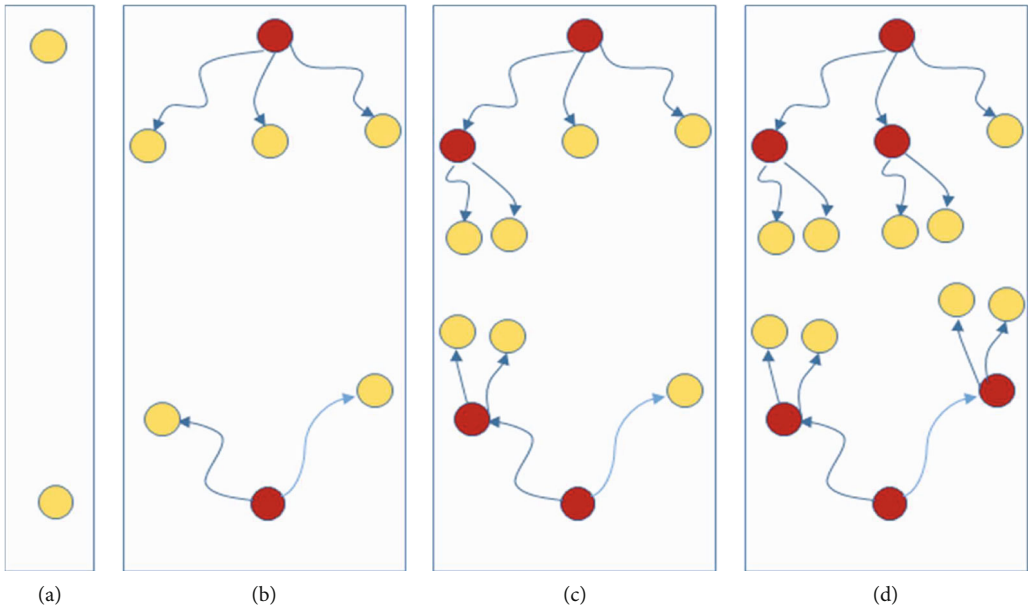


Figure 1: Explaining the process of the algorithm for simplification weights and other details are skipped: (a) initialization: (b) source and destination explored their neighbors; (c) neighbors of neighbors of source and destinations are explored one by one; (d) neighbors of neighbors of source and destinations are explored one by one.

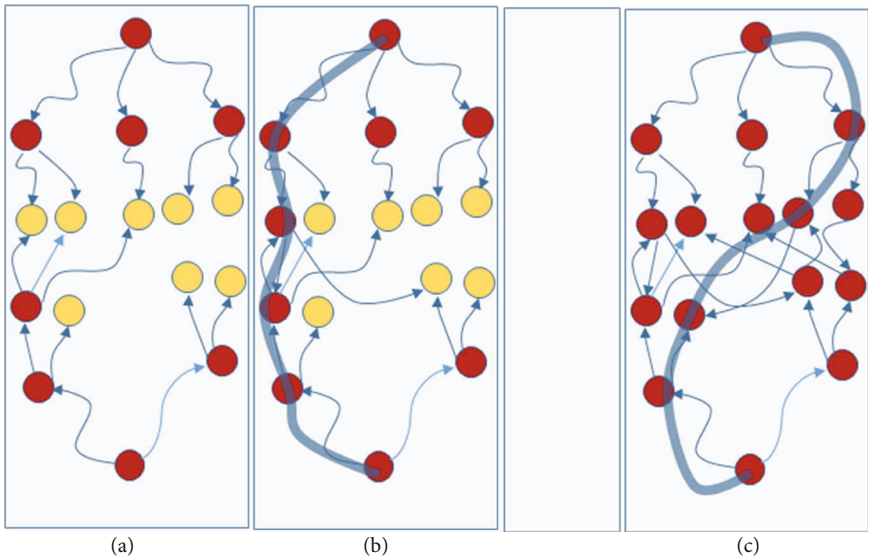


Figure 2: (a) Extract one vertex from queue from both sides and explore their neighbors; (b) explore a red vertex; applying case 3, path is calculated and stored; (c) repeat until all vertices change their color to red, and if new path is shorter, then it is stored and old one is skipped. Finally, the shortest path is obtained.

used to update the values of SPCost, π, and Clr. A and P are colored red after all explorations. New updated status can be seen in Figure 1(b).

In the same way, SPfromDestination dequeues a vertex, and it would be path, equations (9) and (10). Figures 1(c), 1(d), and 2(a) explain these iterations one by one. Figure 2(b) shows that a red vertex is explored from the other side. It generates case 3 (i.e., "R_A and R_B meet at position, P_M"). Path from both sides along with currently updating edge is concatenated to form a path using equation (7). This path is stored.

This process of dequeuing, exploring, and generating new path continues until both queues have no vertices. New paths are generated and compared with the stored one and is discarded if old one is better Figure 2(c). But if new path is better than the old one, it is replaced with the old one. The process ends with the shortest possible path.

5. Discussion

The main advantage of this algorithm is due to the nature of the networks in which it is not possible that indirect distance

can be smaller than direct distance between two vertices. Since Dijkstra invented his SPT algorithm, the research community has worked mostly in the similar direction: choose a node to propagate in nondecreasing order and manage and update neighboring vertices through it. To select vertices in a nondescending order, the half-explored vertices must be sorted, or on each selection, the minimum distance must be calculated from the remaining vertices, which again is equivalent to sort [29].

The main bottleneck of these algorithms is the intermediary sorting algorithm. Most researchers have worked to surmount this bottleneck. New data structures like tree, heap, and priority queue were designed to improve the time bounds. One of these is Fibonacci heap [30] that resulted in the best known asymptotic bound of Dijkstra's algorithm [31]. However, this bound is also criticized by researchers due to the heavy processes of Fibonacci heap [32–37]. This algorithm is beating this bound of Dijkstra's algorithm using Fibonacci heap, where sorting is not being attempted in any form. On one side, it is thrashing this time bound with simple data structure and no sorting algorithm; on the other side, it has parallelization in its nature that is beyond the scope of this paper.

At the same time, this algorithm is also better than the one presented in [29] with worst-case time bound of $O(E' + V' \lg V')$ time algorithm where $E' = E + C^{-}_{\max} C^{+}_{\max}$ and $V' = V(C^{-}_{\max} + C^{-}_{\max})$. The algorithm proposed in this research has worst-case time bound of $O(V + E)$, which is linear in time and hence is better than the one presented in [29] for routing in intelligent networks [38–42].

6. Conclusion

The proposed SPT algorithm executes with $O(V + E)$ time which is relatively better than the state of art work. It is simple to understand and easy to implement and does not require high computational resources. It has additional benefit that after some specific stage, it can always hold a path from source to destination if compromise on quality of path is possible to address the latency issues in nondelay tolerant networks that include military and tactical systems, vehicular communication networks, underwater acoustic networks, and intelligent sensor networks.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. K. Lee, J. W. Jang, S. J. Jang, and J. Y. Shin, "—Development and performance analysis of ABR-DBA algorithm for improve network performance‖," *International Journal of Future Generation Communication and Networking*, vol. 1, pp. 1–6, 2008.

[2] R. K. Ahuja, K. Melhorn, J. B. Orlin, and R. E. Tarjan, "Faster algorithms for the shortest path problem," *Journal of the ACM*, vol. 37, no. 2, pp. 213–223, 1990.

[3] M. S. Sam Coy, A. Czumaj, M. Feldmann et al., "Near-shortest path routing in hybrid communication networks," *Distrib. Parallel, Cluster Computing*, vol. 217, pp. 1–26, 2022.

[4] S. S. J. Augustine, K. Choudhary, A. Cohen, D. Peleg, and S. Sourav, "Distributed graph realizations," in *In Proc. of the 34th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2020)*, pp. 158–167, New Orleans, Louisiana, USA, 2020.

[5] O. De Jonckère and J. A. Fraire, "A shortest-path tree approach for routing in space networks," *China Communications*, vol. 17, no. 7, pp. 52–66, 2020.

[6] A. Andersson, P. B. Miltersen, and M. Thorup, "Fusion trees can be implemented with AC^0 instructions only," *Theoretical Computer Science*, vol. 215, no. 1-2, pp. 337–344, 1999.

[7] B. Zhang, J. Zhang, and L. Qi, *The Shortest Path Improvement Problems under Hamming Distance*, Springer Science+Business Media, LLC, 2006.

[8] E. W. Dijekstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[9] M. Thorup, *Undirected single-source shortest paths with positive integer weights in linear time*, vol. 46, no. 3, 1999, AT&T Labs Research, Florham Park, New Jersey, 1999.

[10] J. W. J. Williams, *Heapsort. Commun*, ACM, 1998.

[11] J. Hershberger, S. Suri, and A. Bhosle, "On the difficulty of some shortest path problems," *ACM Trans. Algorithms*, vol. 3, no. 1, pp. 1–15, 2007.

[12] M. L. Fredman and D. E. Willard, "Fibonacci heaps and their uses in improved network optimization algorithms," *Journal of the ACM*, vol. 34, no. 3, pp. 596–615, 1987.

[13] P. Van Mieghem, "The Shortest Path Problem," in *Performance Analysis of Communications Networks and Systems*, pp. 347–386, Cambridge University Press, Cambridge, 2006.

[14] Z. Y. Xie, Y. R. He, Y. T. Jiang, and C. C. Chen, "Improved and/or tree search algorithm in analysis of stochastic and time-dependent shortest path problem," *Scientific Programming*, vol. 2021, Article ID 9922466, 19 pages, 2021.

[15] W. J. Guerrero, N. Velasco, C. Prodhon, and C. A. Amaya, "On the generalized elementary shortest path problem: a heuristic approach," *Electronic Notes in Discrete Mathematics*, vol. 41, pp. 503–510, 2013.

[16] K. Magzhan and H. Jani, "A review and evaluations of shortest path algorithms," *International Journal of Scientific and Technology Research*, vol. 2, no. 6, pp. 99–104, 2013.

[17] I. A. Carvalho, T. Erlebach, and K. Papadopoulos, "An Efficient Algorithm for the Fast Delivery Problem," in *Fundamentals of Computation Theory*, L. Gąsieniec, J. Jansson, and C. Levcopoulos, Eds., vol. 11651 of Lecture Notes in Computer Science, Springer, Cham, 2019.

[18] D. Medhi and K. Ramasamy, *Routing Algorithms: Shortest Path, Widest Path, and Spanning Tree*, Network Routing, 2018.

[19] M. Therup, "On RAM priority queues," in *Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 59–67, New York, 1996.

[20] M. Thorup, "Floats, integers, and single source shortest paths," in *In Proceedings of the 15th Symposium on Theoretical Aspects*

of Computer Science. Lecture Notes on Computer Science, pp. 14–24, 1373. ,Springer-Verlag, New York, 1998.

[21] R. Raman, "Priority queues: small monotone, and trans-dichotomous," in *In Proceedings of the4th Annual European Symposium on Algorithms. Lecture Notes on Computer Science*, pp. 121–137, 1136, Springer-Verlag, New York, 1996.

[22] B. V. Cherkassky, A. V. Goldberg, and C. Silverstein, "Buckets, heaps, lists, and monotone priority queues," in *In Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 83–92, New York, 1997.

[23] R. Raman, "Recent results on the single-source shortest paths problem," *SIGACT News*, vol. 28, no. 2, pp. 81–87, 1997.

[24] S. Pettie, V. Ramachandran, and S. Sridhar, "Experimental evaluation of a new shortest path algorithm_ (extended abstract)," in *ALENEX 2002, LNCS 2409*, D. Mount and C. Stein, Eds., pp. 126–142, Springer-Verlag, Berlin Heidelberg, 2002.

[25] F. S. I. Himmich, H. B. Amor, and I. El Hallaoui, "A primal adjacency-based algorithm for the shortest path problem with resource constraints," *Transportation Science*, vol. 54, no. 5, pp. 1153–1169, 2020.

[26] S. Majumder, M. B. Kar, S. Kar, and T. Pal, "Uncertain programming models for multi-objective shortest path problem with uncertain parameters," *Soft Computing*, vol. 24, no. 12, pp. 8975–8996, 2020.

[27] P. T. Hanh, P. D. Thanh, and H. T. Binh, "Evolutionary algorithm and multifactorial evolutionary algorithm on clustered shortest-path tree problem," *Information Science*, vol. 553, pp. 280–304, 2021.

[28] C. Hansknecht, I. Joormann, and S. Stiller, "Dynamic shortest paths methods for the time-dependent TSP," *Algorithms*, vol. 14, no. 1, p. 21, 2021.

[29] R. Lewis, "Algorithms for finding shortest paths in networks with vertex transfer penalties," *Algorithms*, vol. 13, no. 11, p. 269, 2020.

[30] M. Qureshi, D. Hassan, and B. Fadzil, "AO (E) time shortest path algorithm for non negative weighted undirected graphs," *International Journal of Computer Science and Information Security*, vol. 6, no. 1, pp. 40–46, 2009.

[31] V. S. Pal and Y. R. Devi, "A survey on IP fast rerouting schemes using backup topology," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 4, pp. 213–218, 2013.

[32] I. Qadeer and M. Khurram Ehsan, "Improved channel reciprocity for secure communication in next generation wireless systems," *Computers, Materials & Continua*, vol. 67, no. 2, pp. 2619–2630, 2021.

[33] E. Gutiérrez and A. L. Medaglia, "Labeling algorithm for the shortest path problem with turn prohibitions with application to large-scale road networks," *Ann. Oper. Res.*, vol. 157, no. 1, pp. 169–182, 2007.

[34] M. Weigl, B. Siemiàatkowska, K. A. Sikorski, and A. Borkowski, "Grid-based mapping for autonomous mobile robot," *Robotics and Autonomous Systems*, vol. 11, no. 1, pp. 13–21, 1993.

[35] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, The MIT Press, 2nd edition, 2001.

[36] C. S. T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, MIT press, 2022.

[37] S. S. Skiena, *The Algorithm Design Manual*, Springer, 2008.

[38] M. Ehsan, "Performance analysis of the probabilistic models of ISM data traffic in cognitive radio enabled radio environments," *IEEE Access*, vol. 8, no. 1, pp. 140–150, 2020.

[39] M. Ehsan, A. A. Shah, M. R. Amirzada et al., "Characterization of sparse WLAN data traffic in indoor opportunistic environments as a prior for coexistence scenarios of modern wireless technologies," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 347–355, 2021.

[40] M. Ehsan and D. Dahlhaus, "Statistical modeling of ISM data traffic in indoor environments for cognitive radio systems," in *IEEE Digital Information, Networking, and Wireless Communication (DINWC), 2015 Third International Conference*, pp. 88–93, Moscow, Russia, 2015.

[41] A. Mahmood, A. Ahmed, M. Naeem, M. R. Amirzada, and A. Al-Dweik, "Weighted utility aware computational overhead minimization of wireless power mobile edge cloud," *Computer Communications*, vol. 190, pp. 178–189, 2022.

[42] N. Naz, M. Ehsan, M. R. Amirzada, M. Y. Ali, and M. A. Qureshi, "Intelligence of autonomous vehicles: a concise revisit," *Journal of Sensors*, vol. 2022, Article ID 2690164, 11 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3763555, 9 pages
https://doi.org/10.1155/2022/3763555

Research Article

Bidirectional CPW Fed Quad-Band DRA for WLAN/WiMAX Applications

Assad Iqbal,[1] Owais Owais,[1] Mohammed K. A. Kaabar,[2,3] Mustafa Shakir,[4] Inam Ullah,[5] Aftab Ahmad Khan,[1] Abd Ullah Khan,[6] and Ateeq Ur Rehman[7]

[1]*Department of Electrical and Computer Engineering, COMSATS University, Abbottabad, K.P.K, Pakistan*
[2]*Gofa Camp, Near Gofa Industrial College and German Adebabay, Nifas Silk-Lafto, 26649 Addis Ababa, Ethiopia*
[3]*Institute of Mathematical Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia*
[4]*Department of Electrical Engineering, Superior University, Lahore 54000, Pakistan*
[5]*College of Internet of Things (IoT) Engineering, Hohai University, Changzhou Campus, 213022, China*
[6]*Department Computer Science, National University of Science and Technology, Quetta Balochistan, Pakistan*
[7]*Department of Electrical Engineering, Government College University Lahore, Lahore 54000, Pakistan*

Correspondence should be addressed to Mohammed K. A. Kaabar; mohammed.kaabar@wsu.edu, Inam Ullah; inam.fragrance@gmail.com, and Ateeq Ur Rehman; ateeq.rehman@gcu.edu.pk

Received 5 January 2022; Revised 9 May 2022; Accepted 20 May 2022; Published 3 June 2022

Academic Editor: Chunqiang Hu

Due to the size constraint, antennas with compact size are required that can support multibands with single radiating element. Dielectric resonator antenna (DRA) can support multiple modes with good efficiency and gain using single radiating element. In this paper, a quad-band DRA is proposed for WLAN/WiMAX bands using a single CPW feed. Out of four bands, WLAN (2.37-2.55) GHz, (5.15-5.35) GHz, and (5.6-5.83) GHz are achieved by the excitation of fundamental and higher-order modes whereas WiMAX (3.58-3.72) GHz band is achieved by dielectric loading of the DR. Dielectric loading makes it possible to operate the antenna at desired frequency band. Due to the partial ground plane and single coplanar waveguide (CPW) feed radiation pattern of the proposed antenna is bidirectional. The design is efficient in terms of radiation efficiency. Simulated and measured results show that it can be a suitable candidate for WLAN and WiMAX applications.

1. Introduction

The advance wireless communication technologies are expected to support massive devices with low energy consumption and high spectral efficiency [1, 2]. These technologies are backscatter communications, multiple input and multiple output, device-to-device communication, vehicular communication, non-orthogonal multiple access (NOMA), Internet of things (IoT), cognitive radio, artificial intelligence, smart drones, reconfigurable intelligent surfaces, and so on. More specifically, the authors of [3] have introduced NOMA to Internet of vehicles network to maximize the total achievable energy efficiency of the system. Moreover, authors in [4] provided a new framework for spectrum and energy efficiency of NOMA network and investigated energy efficiency of vehicular network. The authors of [5] proposed reinforcement learning for blockchain networks. The work in [6] has investigated the energy efficiency of IoT networks. In addition, the work of [7] has studied the performance of drones while [8] has presented device-to-device communication in NOMA network. Of late, [9, 10] have investigated spectral efficiency and energy efficiency using single objective and multiobjective approaches. However, there exist several challenges in these technologies that need to be investigated. For example, designers are focusing on the miniaturization of devices as technology advances. Besides that, there is also an increase in the number of applications, which necessitates the use of various spectrums such as microwave, millimeter wave, and terahertz bands. Moreover, multiple antennas can be used to meet high data rate demands and provide reliable communication. As mutual coupling increases, it becomes

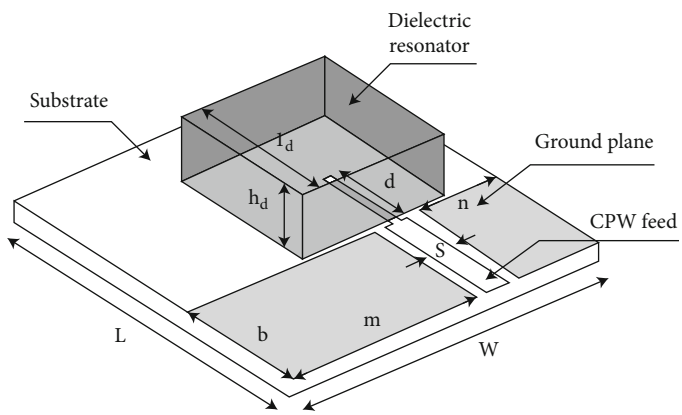


Figure 1: Geometry of the proposed DRA showing all elements of the DRA.

Table 1: Dimension of the proposed antenna design.

Variable used	Value (mm)	Variable used	Value (mm)	Variable used	Value (mm)
W	33	L	34	h_s	1.5
h_g	0.017	h_d	6	a	1
m	20	n	8.5	s	2.5
g	1	b	13	w_d	15
l_d	15	c	0.5	d	9
k	5.5	ε_s	4.4		

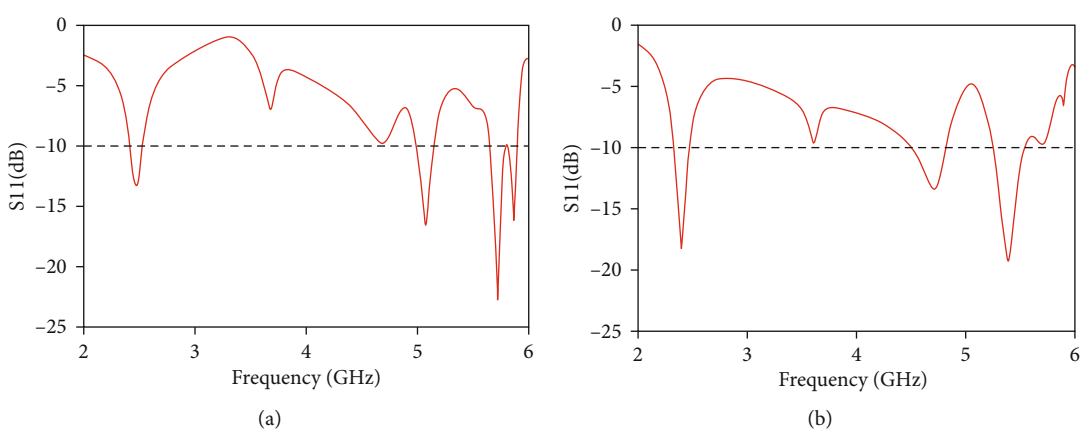


Figure 2: S-parameters due to variation in CPW strip length. (a) Strip length increased. (b) Strip length decreased.

more difficult for design engineers to incorporate a large number of antennas in small devices. Although compact single reconfigurable metallic elements cannot provide enough bandwidth, DRA has a wider impedance bandwidth at high frequencies than metallic antennas [11].

Since the last few years, the DRAs have gained more consideration due to various features like high radiation efficiency, wider impedance bandwidth, high gain, and low metallic losses as compared to metallic antennas such as microstrip patch antennas (MPAs) and planar inverted F-antenna (PIFAs). DRA has been recommended and developed for use in microwave and millimeter-wave frequency bands. Compared with microstrip patch antennas (MPA) that have discovered a wide range of applications [12]. Moreover, DRA can easily be excited by multiple feeding techniques such as coplanar waveguide, coaxial probe, hybrid feeding, and aperture coupling [13]. As compared with MPA, the DRA has a much more extensive impedance bandwidth (~10% for dielectric constant $\varepsilon r \sim 10$). This is because the microstrip antenna radiates just through two narrow radiation slots; however, the DRA radiates through its entire surface except the grounded part. Avoidance of the surface waves is one of the advantages of the DRA over the microstrip antenna. In any case, numerous attributes of the DRA and microstrip antenna (MA) are common because they

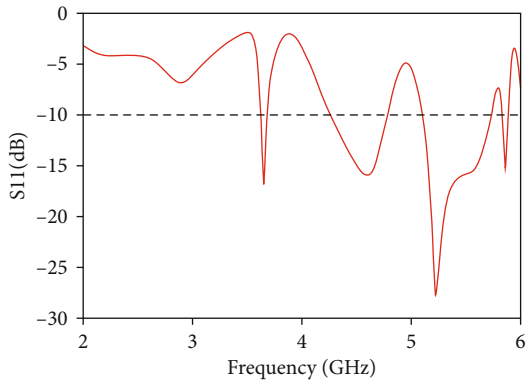
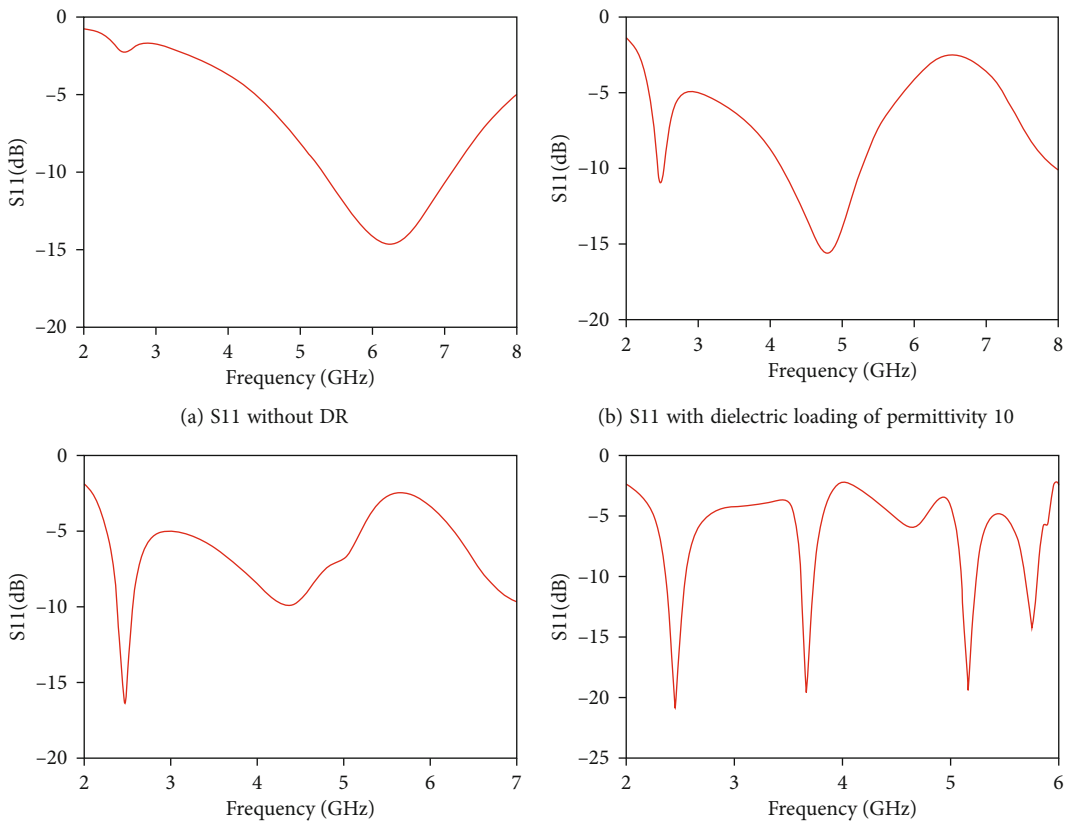


Figure 3: Effect of variation in ground plane width on S-parameters.



(a) S11 without DR

(b) S11 with dielectric loading of permittivity 10

(c) S11 with dielectric loading of permittivity 20

(d) S11 with dielectric loading of permittivity 30

Figure 4: Effect of loading with different permittivity material on S-parameters.

behave like resonant cavities. Besides, practically all the excitation methods that are applied to the MA can be utilized for the DRA. A detailed explanation of DRA with almost all aspects like its shapes, feeding mechanism, radiation patterns, and modes can be found in [14].

Recently, it has been investigated that various applications are working at different frequency bands; so, it is not viable to design an antenna for each application separately because of the increase in cost and size of the device [15]. Therefore, it is advantageous to design a multiband antenna with a single structure to reduce the cost, size, and complexity of the system [16–19]. It is the property of DRA that it can excite multiple modes using a single DR structure when excited by a suitable feeding structure [20]. Therefore, the feeding mechanism plays an important role to excite required mods in the DRA. In different papers, different feeding techniques such as microstrip line, coaxial probe, and CPW feed are used to excite multiple modes in DRA [21–24]. In some cases, pentagon and hexagon shapes of feeding structures have been used to achieve multiple bands [25]. In [26], Hamsakutty et al. achieved a triple band by using two segments of rectangular DRA, which make the overall structure bulky. In [27], the authors used segmented DRA to

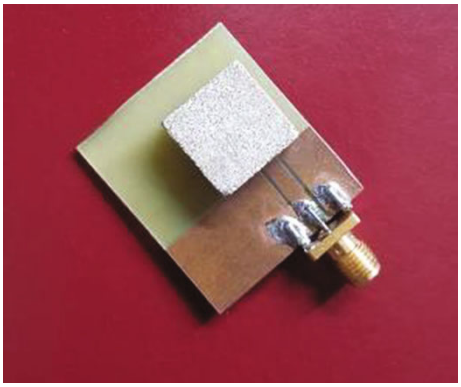


Figure 5: The prototype of the proposed quad-band DRA.

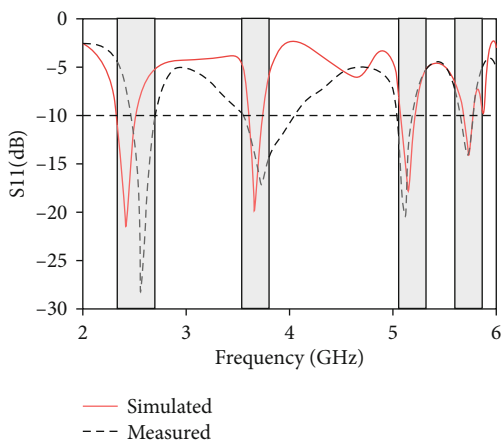


Figure 6: Measured and simulated S11 at all four bands.

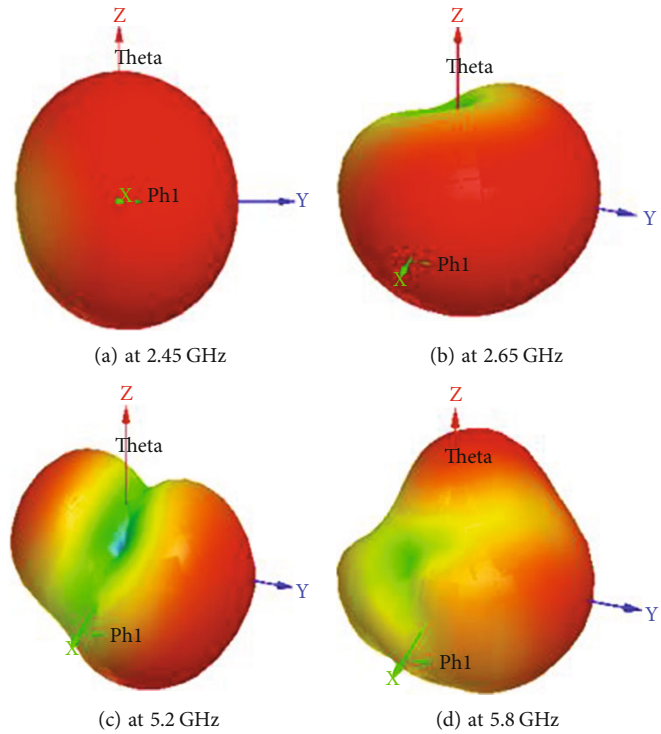
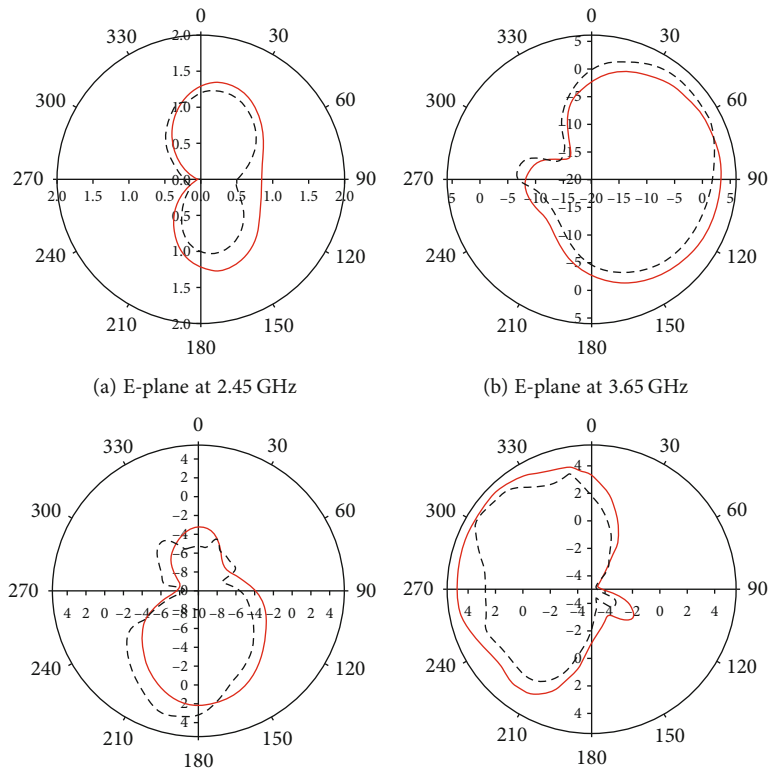


Figure 7: Three dimensional radiation patterns of the DRA.

achieve triple bands; however, they used complex feeding which ultimately complicates the fabrication process. In [28], Sharma and Gangwar proposed a triple-band hybrid DRA for 4G LTE application. In this work, two bands are achieved through feed line and higher-order modes. A unidirectional radiation pattern at all three bands is achieved in this design. For particular services where users move in a straight path like highways, tunnels and railways bidirectional antennas are more useful to maximize the link efficiency [29, 30]. In [31], two unidirectional antennas in the opposite direction have been used for bidirectional radiation pattern; however, the design was complicated and directivity was also low. A bidirectional radiation pattern has also been achieved with antennas having more than one feeding mechanism [32].

Multiple techniques have been used to design multiple-band DRAs, which are the subject of this section's investigation. The author of paper [33] achieved a triple-band by simultaneously exciting the four DR arrays. The proposed design operates at frequencies ranging from 0.86 to 1.04 GHz, 2.37 to 2.79 GHz, and 4.17 to 5.21 GHz. The array of four DR increases the overall size of the antenna while also making it bulky, making installation of such antennas in modern wireless gadgets difficult. In paper [33], the authors designed a triple-band dielectric resonate antenna by exciting two DR segments. The first two bands are achieved by exciting the two DR via feed line, and the third band is achieved through reconfigurability. Because switches are used for reconfigurability, the design of this DRA is extremely complicated. Because two DR segments are used, the overall design of this DRA is bulky. A star-shaped DR was excited by a simple microstrip feed line in paper [22], and the proposed design covered three frequency bands of 5.04 −6.13 GHz, 6.87 −7.97 GHz, and 8.58 −9.63 GHz. The construction of such an antenna is extremely difficult. This design's radiation pattern is unidirectional. Furthermore, the authors of the paper [34] designed a multiband DRA. The multiband is achieved in this design by exciting two cylindrical DR and two C shape patches. Because of the lack of a bottom ground structure, the radiation pattern of this design is bidirectional, which increases the size of the antenna of the whole communication system. A triple-band dielectric resonates antenna with operating frequencies of 1.8 GHz, 2.2 GHz, and 2.6 GHz was also designed and investigated [35]. There is no fabrication involved in this design, and the results are calculated. Only simulated results were shown by the author. The design has a substantial overall size.

Most of the DRAs listed above have multiple resonating elements, which are either very complicated to design or mostly support dual-band or triple-band operation. As a result, in this paper, a single square-shaped DR was excited via a CPW feed line. The proposed design has no bottom ground plane due to which the radiation pattern of the antenna becomes bidirectional, the operating band of the proposed design is WLAN (2.37-2.55) GHz, (5.15-5.35) GHz, (5.6-5.83) GHz, and WiMAX (3.58-3.72). Both WiMAX and WLAN bands can serve hundreds of users with a speed of 10 Mbps at 10 km with a line of sight. The use of these bands has been increased with



(a) E-plane at 2.45 GHz

(b) E-plane at 3.65 GHz

(c) E-plane at 5.2 GHz

(d) E-plane at 5.8 GHz

Figure 8: Simulated and measured E-planes.

the advent of applications with multiband support. From the comparison table, the size of the proposed design is much smaller than other DRAs, which compact and miniaturize the size of modern wireless communication devices.

In this paper, bidirectional radiation patterns have been achieved for four bands using single radiating elements with a single feeding structure. The rectangular DRA has no ground plane at the bottom which makes its radiation pattern bidirectional. It covers three bands including WiMAX and WLAN. Moreover, a quad-band DRA is proposed and achieved for WLAN/WiMAX bands using a single CPW feed. Bidirectional radiation pattern is achieved at all radiating element.

The remaining paper is organized as below: the antenna design parameters and analysis are described in Section 2. Section 3 presents the parametric study, while Section 4 presents the results and discussion. Finally, Section 5 concludes this work.

2. Antenna Design and Analysis

Figure 1 shows the geometry of the proposed DRA. Dimensions of the DR is $[l_d \times (w_d = l_d) \times h_d]\,\text{mm}^3$ placed on FR4 substrate with permittivity ε_s and height h_s. The DRA is excited through CPW feed line with optimized impedance matched at 50Ω. In order to improve the impedance matching of DR with CPW feed, a stub of width 1 mm is also added at the end of the coplanar feed. Table 1 summarizes the optimized values of the proposed DRA.

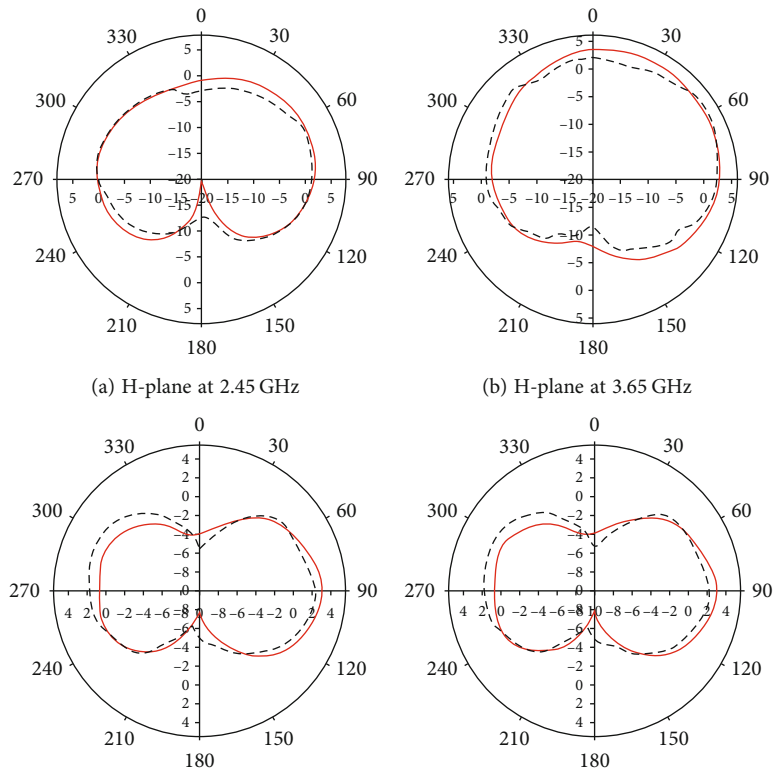
3. Parametric Study

The parametric study comprised of variation of CPW feeds on resonant frequencies with and without loading of DR on the substrate.

3.1. Variation without DR Loading. Resonant frequencies of the proposed design depend on the length of the CPW strip. If the length of the strip is increased beyond 9 mm, it is clear from Figure 2(a) that the impedance matching and bandwidth impedance of the lower WLAN band (2.45 GHz) are decreased. Impedance matching and bandwidth of the remaining two bands remain the same. If the length of CPW feed is decreased below 9 mm; the resonance of the lower band remains unaltered but resonances of the upper two bands from 5.1 GHz to 4.7 GHz and from 5.8 GHz to 5.6 GHz. This phenomenon is shown in Figure 2(b). However, it still does not fulfill the requirements of WLAN bands as it has very poor impedance matching as well.

Resonance frequencies also depend upon the width of the ground plane. When the width of the ground plane is decreased below 13 mm, it is clear from Figure 3 that the proposed design has no resonance at 2.45 GHz, and the impedance bandwidth of the WiMAX band is also decreased.

The resonance frequencies are also reliant on the width of the ground plane, and if the width of the ground plane is decreased from 13 mm, it is clear from Figure 3 that the proposed design has no resonance at 2.45 GHz, which has been achieved by DR loading.



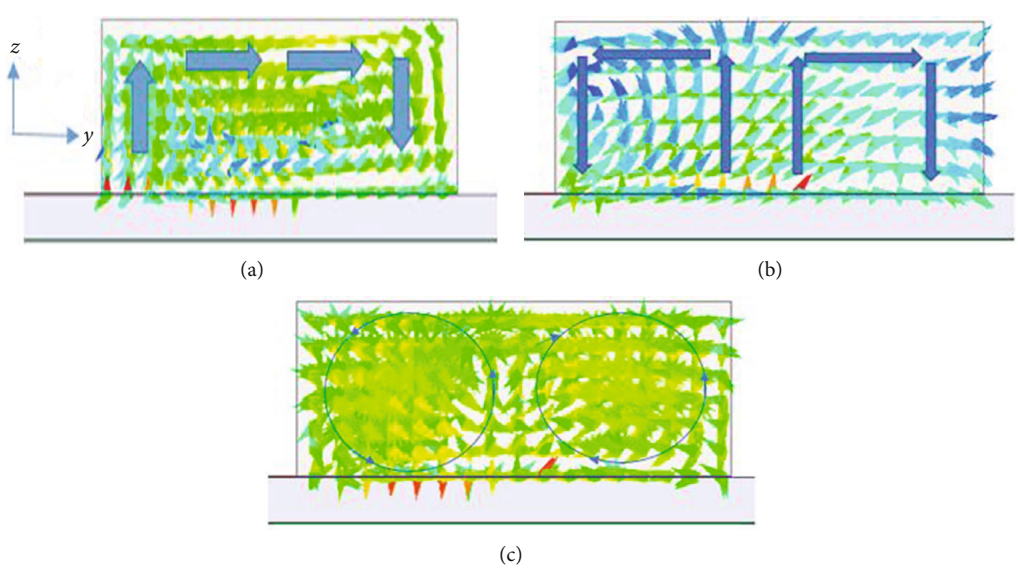
(a) H-plane at 2.45 GHz

(b) H-plane at 3.65 GHz

(c) H-plane at 5.2 GHz

(d) H-plane at 5.8 GHz

Figure 9: Simulated and measured H-planes.



(a)

(b)

(c)

Figure 10: TE modes. (a) TE_{111}. (b) TE_{121}. (c) TE_{142}.

3.2. Variation with DR Loading. The dielectric loading effect is inversely proportional to the resonance frequency. It is clear from Figure 4(a) that without dielectric loading, DRA resonates at 6.2 GHz. Due to the dielectric loading of DR with permittivity 10, the resonance point shifts back to 4.8 GHz, as shown in Figure 4(b). This is fundamental mode excitation in the DR, and other modes also shit accordingly. Figures 4(c) and 4(d) reveal the fundamental mode excitation for DR loading with permittivity 20 and 30, respectively. For permittivity 20, it shifts to 4.3 GHz, and for 30, shifting takes place on 3.6 GHz.

When the permittivity of the material is increased to 20, the resonance frequency of the antenna is shifted to 4.3 GHz from 4.8 GHz which is shown in Figure 4(c). When the

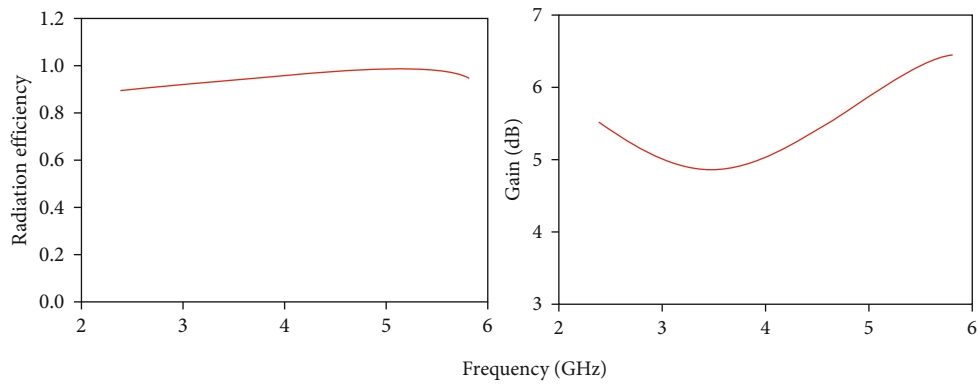


Figure 11: (a) Radiation efficiency vs. frequency. (b) Gain vs. frequency.

Table 2: Value radiation efficiency and gain are listed Table 2.

References	No. of DR element	No. of band	Gain (dBi)	Size ($L \times w \times h$) mm^3
[33]	4	3	7.2	$96 \times 113 \times 15.6$
[33]	2	3	5.2	$230 \times 134 \times 5.6$
[22]	1	3	8.10	$50 \times 50 \times 7.60$
[34]	2	4	6.5	$50 \times 50 \times 11.2$
[35]	1	2	5.3	$100 \times 70 \times 11.4$
[39]	1	3	6.86	$140 \times 140 \times 11.6$
Proposed work	1	4	6.45	$33 \times 34 \times 7.4$

Table 3: A comparison with related multiband DRAs.

Frequency (GHz)	2.45	3.65	5.25	5.81
Gain (dBi)	5.5	4.9	6.1	6.41
Rad. efficiency	0.9	0.85	0.91	0.94

material has a permittivity of 30 is loaded on CPW feed, the resonance frequency of the antenna is shifted from 4.3 GHz to 3.65 GHz, which is the required WiMAX band, the proposed design gives the resonance at 2.45 GHz is due to the excitation of the fundamental mode in DR, while resonance at 5.15 GHz and 5.8 GHz is due to the excitation of higher mode in DR which is clear from Figure 4(d).

4. Result and Discussions

This section includes the antenna and its performance parameters. First, the antenna performance parameters such as S-parameters with respect to design are described; followed by performance matrices like radiation pattern, gain, and efficiency.

Figure 5 shows the model of the proposed designed antenna. The DR has a permittivity of 30 and a tangent loss of 0.002, which is made from Eccostock® Hik500f. The ground plane is fixed on an FR4 substrate having a permittivity of 4.4 and a thickness/breadth of 1.5 mm. The DR is placed on the substrate through a glove of permittivity close to 2.

Simulated and measured S-parameters are shown in Figure 6. Shaded region in the figure shows the required band of interest. Results show that antenna is well matched at all the bands except the last band. The reason may be the ground plane size at this band because it is the higher mode excited in the DRA. Mismatch in simulated and measured results is due to the fabrication tolerances.

Figure 7 shows 3D patterns at each frequency band. It is clear from figure that direction of radiation pattern at each band is different, covering each direction at each band. This is one of the advantages of the partial ground plane. However, partial ground plane lowers the gain of the antenna due to the backlobe which is one of its drawback. A bidirectional radiation pattern with a simple structure and single feeding mechanism is achieved at all four ports. Simulated gain at all four bands is up to 6.45 dBi. It can be increased by extending the ground plane; however, this action degrades the bidirectional performance of the patterns as discussed earlier.

Figures 8 and 9 show that the simulated and measured E- and H- plane of proposed DRA, respectively. Solid line shows the simulated, and the dotted line represents the measured results. Both E and H plane of the patterns shows bidirectional radiation patterns but at the cost of backward radiations. It is clear from all the figures that the measured and simulated results are in good arrangement.

DRA can support multiple modes, and its fundamental mode depends on its aspect ratio dielectric permittivity. Aspect ratio is the ratio of length: width and length: height [36]. In the proposed work, fundamental mode in the DR is excited at 2.45 GHz as shown in Figure 10(a). This mode is TE_{111} mode, represented by a half wave of electric field pattern, showing the wave-number $= \pi/2$.

Similarly at 5.2 GHz, TE_{121} mode is excited, represented by two half waves as shown in Figure 10(b). At 5.8 GHz, TE_{142} mode has been excited. As there is no ground plane under the DR element; therefore, even mode can be excited [36]. Four half-waves in the y-axis and two half-waves in the z-axis can be seen in Figure 10(c). A complete circle represents two half-waves.

WiMAX (3.58-3.72) GHz band is achieved by dielectric loading of the DR therefore no specific pattern of the field is achieved as shown.

Figures 11(a) and 11(b) show the gain vs. frequency and radiation efficiency vs. frequency pots, respectively. It can be

seen from plots that both gain and radiation efficiency are small (4.9 and 0.85) as compared to the other values. It is because the band at 3.65 GHz is due to dielectric loading, and the remaining three bands are due to the excitation of modes in the DRA. Gain and radiation efficiency against each frequency band is shown in Table 2.

Table 3 shows the comparison of some recent research with the proposed work. It can be seen that in terms of number of DR elements, our design uses only one which results in reduced size as compared to other antennas. Moreover, the proposed design covers four bands with single CPW feed.

In this work, single CPW feed and DR element have been used to achieve bidirectional radiation pattern. This arrangement reduces the size of the antenna. This work can be extended for quad-band MIMO design for 5G applications. For this purpose, another orthogonal CPW feed can be added symmetrically to achieve quad-band MIMO in the same frequency. However, there may be the issue of isolation at each degenerate mode which can be enhanced using different techniques available in literature.

5. Conclusion

In this paper, a CPW feed bidirectional quad-band DRA is proposed for WLAN and WiMAX bands. The proposed design covers WLAN (2.37-2.55) GHz, (5.15-5.35) GHz, (5.6-5.83) GHz, and WiMAX (3.58-3.72) GHz bands. The proposed design with single feed and single DR element provides four bands with the capability of bidirectional radiation pattern. Bidirectional radiation at all four bands has been achieved using a simple design technique. The radiation efficiency of the proposed design is 94% and gains up to 6.45 dBi. This design can be used in wireless gadgets and other related applications. All the measured and simulated results of the proposed design show that it will be an appropriate candidate for WLAN and WiMAX applications.

As a future task, the proposed work can be extended for MIMO design that can be used for 5G applications. This can be done by using another CPW feed, orthogonally oriented along the same DR element. Moreover, the design can be used for quad-band array design for 5G applications.

Data Availability

The dataset used for the experiments in this study is available upon reasonable request from the corresponding author.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Assad Iqbal did the actualization, methodology, formal analysis, software, validation, investigation, and initial draft. Owais Owais did the actualization, validation, methodology, software, formal analysis, investigation, and initial draft. Mohammed K. A. Kaabar did the actualization, methodology, formal analysis, validation, investigation, initial draft, and supervision of the original draft and editing. Mustafa Shakir did the actualization, validation, methodology, formal analysis, investigation, and initial draft. Inam Ullah did the actualization, validation, methodology, formal analysis, investigation, and initial draft. Aftab Ahmad Khan did the actualization, validation, methodology, formal analysis, investigation, and initial draft. Abd Ullah Khan did the actualization, validation, methodology, formal analysis, investigation, and initial draft. Ateeq Ur Rehman did the actualization, validation, methodology, formal analysis, investigation, and initial draft. All authors read and approved the final version.

References

[1] W. U. Khan, X. Li, A. Ihsan, M. A. Khan, V. G. Menon, and M. Ahmed, "NOMA-enabled optimization framework for next-generation small-cell IoV networks under imperfect SIC decoding," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[2] W. U. Khan, T. N. Nguyen, F. Jameel et al., "Learning-based resource allocation for backscatter-aided vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[3] F. Jameel, S. Zeb, W. U. Khan, S. A. Hassan, Z. Chang, and J. Liu, "NOMA-enabled backscatter communications: toward battery-free IoT networks," *IEEE Internet of Things Magazine*, vol. 3, no. 4, pp. 95–101, 2020.

[4] W. U. Khan, M. A. Javed, T. N. Nguyen, S. Khan, and B. M. Elhalawany, "Energy-efficient resource allocation for 6G backscatter-enabled NOMA IoV networks," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[5] F. Jameel, U. Javaid, W. U. Khan, M. N. Aman, H. Pervaiz, and R. Jäntti, "Reinforcement learning in blockchain-enabled IIoT networks: a survey of recent advances and open challenges," *Sustainability*, vol. 12, no. 12, article 5161, 2020.

[6] W. U. Khan, F. Jameel, M. A. Jamshed, H. Pervaiz, S. Khan, and J. Liu, "Efficient power allocation for NOMA-enabled IoT networks in 6G era," *Physical Communication*, vol. 39, article 101043, 2020.

[7] S. K. Haider, A. Jiang, A. Almogren et al., "Energy efficient UAV flight path model for cluster head selection in next-generation wireless sensor networks," *Sensors*, vol. 21, no. 24, article 8445, 2021.

[8] S. Yu, W. U. Khan, X. Zhang, and J. Liu, "Optimal power allocation for NOMA-enabled D2D communication with imperfect SIC decoding," *Physical Communication*, vol. 46, article 101296, 2021.

[9] W. U. Khan, J. Liu, F. Jameel, V. Sharma, R. Jantti, and Z. Han, "Spectral efficiency optimization for next generation NOMA-enabled IoT networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 15284–15297, 2020.

[10] W. U. Khan, F. Jameel, G. A. S. Sidhu, M. Ahmed, X. Li, and R. Jantti, "Multiobjective optimization of uplink NOMA-enabled vehicle-to-infrastructure communication," *IEEE Access*, vol. 8, pp. 84467–84478, 2020.

[11] M. Abedian, S. K. A. Rahim, S. Danesh, S. Hakimi, L. Y. Cheong, and M. H. Jamaluddin, "Novel design of compact UWB dielectric resonator antenna with dual-band-rejection characteristics for WiMAX/WLAN bands," *IEEE Antennas Wireless Propagation Letters*, vol. 14, pp. 245–248, 2015.

[12] A. A. Khan, M. H. Jamaluddin, S. Aqeel, J. Nasir, J. U. R. Kazim, and O. Owais, "Dual-band MIMO dielectric resonator

antenna for WiMAX/WLAN applications," *IET Microwaves, Antennas & Propagation*, vol. 11, no. 1, pp. 113–120, 2017.

[13] S. Danesh, S. K. A. Rahim, and M. Khalily, "A wideband trapezoidal dielectric resonator antenna with circular polarization," *Progress In Electromagnetics Research*, vol. 34, pp. 91–100, 2012.

[14] K. M. Luk and K. W. Leung, *Dielectric Resonator Antenna*, Research Studies tablePress, U. K., 2003.

[15] Y. Gao, B. L. Ooi, W. B. Ewe, and A. P. Popov, "A compact wideband hybrid dielectric resonator antenna," *IEEE Microwave and Wireless Components Letters*, vol. 16, no. 4, pp. 227–229, 2006.

[16] S. Aqeel, M. R. Kamarudin, A. A. Khan et al., "A compact frequency reconfigurable hybrid DRA for LTE/Wimax applications," *International Journal of Antennas and Propagation*, vol. 2017, Article ID 3607195, 13 pages, 2017.

[17] I. Khan, Y. B. Tian, I. Ullah, M. M. Kamal, H. Ullah, and A. Khan, "Designing of E-shaped microstrip antenna using artificial neural network," *International Journal of Computing, Communications and Instrumentation Engineering*, vol. 5, no. 1, 2018.

[18] R. Khan, Q. Yang, I. Ullah et al., "3D convolutional neural networks based automatic modulation classification in the presence of channel noise," *IET Communications*, vol. 16, no. 5, pp. 497–509, 2021.

[19] W. U. Khan, N. Imtiaz, and I. Ullah, "Joint optimization ofNOMA-enabled backscatter communications for beyond5G IoTnetworks," *Internet Technology Letters*, vol. 4, no. 2, article e265, 2021.

[20] A. Ahmad Khan, R. Khan, S. Aqeel, J. Ur Rehman Kazim, J. Saleem, and M. K. Owais, "Dual-band mimo rectangular dielectric resonator antenna with high port isolation for LTE applications," *Microwave and Optical Technology Letters*, vol. 59, no. 1, pp. 44–49, 2017.

[21] V. Hamsakutty, A. V. Praveen Kumar, G. Bindu et al., "A multifrequency coaxial-fed metal coated dielectric resonator antenna," *Microwave and Optical Technology Letters*, vol. 47, no. 6, pp. 573–575, 2005.

[22] L. Huitema, M. Koubeissi, M. Mouhamadou, E. Arnaud, C. Decroze, and T. Monediere, "Compact and multiband dielectric resonator antenna with pattern diversity for multistandard mobile handheld devices," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 11, pp. 4201–4208, 2011.

[23] L. Ali, C. Wang, I. Ullah et al., "Design and optimization of microwave sensor for the non-contact measurement of pure dielectric materials," *Electronics*, vol. 10, no. 24, article 3057, 2021.

[24] I. Khan, Y. B. Tian, H. Vllah, S. U. Rahman, and M. M. Kamal, "Design annular ring microstrip antenna based on artificial neural network," in *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pp. 2033–2037, Xi'an, Shaanxi China, May 2018.

[25] M. M. Kamal, I. Ullah, A. Ashraf, and N. Ullah, "Designing band notch features in ultra-wideband antenna," in *2017 7th IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies (MAPE)*, pp. 92–95, Xi'an Shi, China, October 2017.

[26] V. Hamsakutty, A. V. Praveen Kumar, J. Yohannan, G. Bindu, and K. T. Mathew, "Coaxial fed hexagonal dielectric resonator antenna for multifrequency operation," *Microwave and Optical Technology Letters*, vol. 48, no. 5, pp. 878–880, 2006.

[27] M. Bemani, S. Nikmehr, and H. Younesiraad, "A novel small triple band rectangular dielectric resonator antenna for WLAN and WiMAX applications," *Journal of Electromagnetic Waves and Applications*, vol. 25, no. 11-12, pp. 1688–1698, 2011.

[28] A. Sharma and R. K. Gangwar, "Triple band two-segment cylindrical dielectric resonator antenna with a novel microstrip feed for WLAN/WIMAX applications," *Microwave and Optical Technology Letters*, vol. 57, no. 11, pp. 2649–2655, 2015.

[29] A. Carton, C. G. Christodoulou, C. Dyck, and C. Nordquist, "Investigating the impact of carbon contamination on RF MEMS reliability," in *2006 IEEE Antennas and Propagation Society International Symposium*, pp. 1934-1196, Albuquerque, NM, USA, July 2006.

[30] H. Arai, K. Kohzu, T. Mukaiyama, and Y. Ebine, "Bi-directional notch antenna with parasitic elements for tunnel booster system," *Antennas and Propagation Society International Symposium*, vol. 4, pp. 2218–2221, 1997.

[31] A. Batgerela, J. I. Choib, and S. Y. Eomb, "High-gain bidirectional MDAS antenna design excited by stacked-microstrip dipole," *Journal of Electromagnetic Waves and Applications*, vol. 26, no. 11-12, pp. 1412–1422, 2012.

[32] T. Jia and X. Li, "A compact stacked bidirectional antenna for dual-polarized WLAN applications," *Progress In Electromagnetics Research*, vol. 44, pp. 95–108, 2013.

[33] A. Vahora and K. Pandya, "Triple band dielectric resonator antenna array using power divider network technique for GPS navigation/bluetooth/satellite applications," *International Journal of Microwave and Optical Technology*, vol. 15, pp. 369–378, 2020.

[34] P. R. Girjashankar and T. Upadhyaya, "Surface mountable multiband dielectric resonator antenna for wireless communication systems," *Progress In Electromagnetics Research M*, vol. 103, pp. 173–183, 2021.

[35] A. Sharma and R. K. Gangwar, "Circularly polarised hybrid Z-shaped cylindrical dielectric resonator antenna for multiband applications," *IET Microwaves, Antennas & Propagation*, vol. 10, no. 12, pp. 1259–1267, 2016.

[36] X. S. Fang and K. W. Leung, "Designs of single-, dual-, wideband rectangular dielectric resonator antennas," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 6, pp. 2409–2414, 2011.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1498918, 15 pages
https://doi.org/10.1155/2022/1498918

Research Article

Empowering Reconfigurable Intelligent Surfaces for Security of Downlink NOMA

Nhan Duc Nguyen,[1] **Minh-Sang Van Nguyen**,[2] **and Munyaradzi Munochiveyi**[3]

[1]*Faculty of Mechanical-Electrical and Computer Engineering, School of Engineering and Technology, Van Lang University, 69/ 68 Dang Thuy, Tram Street, Ward 13, Binh Thanh District, 70000 Ho Chi Minh City, Vietnam*
[2]*Faculty of Electronics Technology, Industrial University of ho chi Minh City (IUH), Ho Chi Minh City, Vietnam*
[3]*Electrical and Electronics Engineering Department, University of Zimbabwe, Mount Pleasant, Harare, Zimbabwe*

Correspondence should be addressed to Munyaradzi Munochiveyi; mmunochiveyi@eng.uz.ac.zw

Received 3 March 2022; Revised 25 April 2022; Accepted 6 May 2022; Published 25 May 2022

Academic Editor: Han Wang

Reconfigurable intelligent surfaces (RIS) and non-orthogonal multiple access (NOMA) are promising techniques to develop next-generation wireless systems. While RIS has huge potential to create massive device connectivity, NOMA exhibits its spectrum efficient communication among multiple access approaches. RIS is a passive device made up of low-cost meta-surfaces which can control the propagation of radio waves, and it is easily deployable in lots of applications in the Internet of Things. The full-duplex nature of RIS has also been a major reason for its consideration of major emerging and trending technologies. In this paper, we aim to investigate the secrecy performance of the RIS-NOMA-assisted Internet of Things (IoT) systems in the presence of two legitimate users who belong to a cluster, and those devices are associated with the existence of an eavesdropper situated close to such a cluster. This paper considers the devices in the presence of RIS and an eavesdropper. As main performance metrics, the closed-form expressions for secrecy outage probability (SOP) and strictly positive secrecy capacity (SPSC) are derived to evaluate the performance of legitimate users. Simulations are performed in support of the Monte-Carlo method, and the obtained results show that in most of the cases, the number of meta-surfaces in RIS and signal-to-noise ratio (SNR) levels at the source also plays a pivotal role in influencing the secure performance of the system.

1. Introduction

RIS is a passive device that is made of low-cost meta-surfaces which can control the propagation of radio waves impinging on it, according to the position of the receiver [1]. Also, RIS is a passive device that does not rely on any external energy sources and performs the signal transmission using soft programming. Also, RIS can work in full-duplex mode communication, which makes it a perfect choice for enabling massive device connectivity. Even though the idea of RIS devices came into existence a couple of years back, there is currently immense research work and publications available in various journals on this topic. Primarily, the researchers have focused on integrating this technology with various trending and emerging technologies such as NOMA, cognitive radio (CR) systems, visible light communications (VLC), and physical layer security (PLS) [2–4]. Consequently, increased research work is being performed in enhancing secrecy efficiency and PLS, as the day-to-day privacy concerns of the users' increase. Therefore, in the following subsection, we introduce in detail various works related to our research.

1.1. Related Works. The integration of RIS with underlay CR network has been considered in [5], and the SOP performance is investigated in the presence of interference from the secondary network to the primary network. In [6], the authors have an intelligent reflecting surface– (IRS–) assisted mmWave system operating in the presence of CR networks to perform robust secure transmission under imperfect channel state information (CSI) conditions. The performance was analyzed in terms of achievable secrecy rate, and the proposed algorithm indicated that it had a better performance compared to traditional

systems. In [7], the authors considered a multiple-input-multiple-output– (MIMO–) aided RIS system with secure wireless information and power transfer (SWIPT) technique. The authors also considered multiple antennas at the sender, receiver, and energy receiver (ER), where ER is assumed to be the potential eavesdropper in the network. Moreover, the authors proposed various techniques to improve the efficiency of the secrecy performance of the system. A similar system in an orthogonal frequency division multiple access– (OFDM) aided systems was considered in [8], where passive beamforming and joint optimization with alternating optimization were proposed techniques to enhance the secrecy performance of the system. In [9], the authors considered a deep learning-based approach for RIS systems in the PLS perspective to maximize the secrecy rate performance of the legitimate user in real time, based on the reflecting elements in the system. The results showed that the proposed method achieved better performance, as well as reduced computational complexity. In [10], the authors investigated the performance of beamforming and RIS design to analyze and enhance the security of multiple-input single-output– (MISO–) assisted RIS system. Two designs were proposed, which are as follows: low-computational complexity successive design and high-security performance joint design. Both designs achieved better secrecy performance compared to the traditional existing systems.

In [11], the authors considered RIS-enabled IoT network when the Fisher-Snedecor $\mathcal{F}$ model is utilized to analyze the generalized composite fading and shadowing model effect. The authors performed the analysis in terms of average capacity, bit error rate, and outage probability. The results demonstrated that employing the proposed model is beneficial in the considered system and achieves efficient performance over the other fading models. Furthermore, RIS-assisted satellite for IoT network is proposed in [12]. The installation of reflecting surfaces on satellites is proposed to enhance the broadcasting and beamforming of the signal with significant gains. The study showed that up to 10^5 times increase in the uplink and downlink achievable rates of IoT networks can be obtained. In [13], the authors considered a RIS system with multiple antennas at the base station (BS) serving multiple users with single antennas on the ground. The authors also proposed a joint optimization at the BS and RIS to minimize the system sum mean squared error. The numerical results demonstrated that the proposed technique and algorithm outperformed traditional systems, whereas in [14], the authors considered RIS to enhance the performance of wireless power transfer (WPT) in the presence of mobile edge computing (MEC) IoT network. Various optimizations techniques were performed at the RIS, MEC, and WPT; subsequently, a reinforcement learning method is adopted to effectively overcome the nonconvex problems. The numerical results showed the effective and beneficial performance of the proposed system and method. Similarly in [15], resource allocation was studied for RIS-assisted wireless powered frequency division multiple access (FDMA) IoT networks. With the aid of passive beamforming reflectors at the RIS, the wireless energy transfer (WET) and wireless information transfer (WIT) have also been efficiently improved by applying various optimization algorithms. The performance of the system was analyzed in terms of system throughput, transmission time scheduling, and energy harvesting, and the demonstrated results proved to be effective compared to the traditional systems.

However, the aforementioned works have not considered secure RIS-NOMA systems; hence, this has motivated us to derive main equations of secure performance along with detailed evaluations of these equations.

1.2. Our Contributions and Organization. In this paper, we aim to analyze the secrecy performance of the RIS-NOMA system in the presence of two legitimate users located in a cluster and an eavesdropper, both possessing single antennas. Two cases were considered in the research, i.e., with and without a direct link between devices in the presence of a RIS device. The primary contributions of the papers are as follows:

(i) To characterize secure performance, we rely on the SNR at each intended NOMA device which is likely to experience degraded performance under the impact of eavesdropping. In addition, we intend to compute the SNR expressions and secrecy rate expressions for scenarios where users lack a direct link with the BS. Importantly, we want to answer the question of the impact of RIS-aided link transmission on secure performance. Our results provide important guidelines to design RIS for future IoT systems

(ii) Furthermore, the main contributions are deriving the asymptotic and closed-form expressions for SOP and SPSC based on the initial expressions obtained in related works. The main coefficients are determined in order to recommend future design of reliable transmissions from the source (BS or access point (AP)) to intended devices

(iii) We provide main parameters affecting secure performance in numerical simulations. We then simulate and compare the performance of the proposed model with the aid of simulations performed based on the obtained expressions between RIS and non-line of sight users

Meanwhile, the paper is structured as follows. Section 2 will introduce and describe the system model and its characteristics. In Section 4, the numerical results and simulations are presented and discussed to understand the performance of the proposed system in both cases. Finally, in Section 5, the paper will be concluded. After the conclusion, the appendix is provided for explaining the computations related to Sections 2.

2. System Model

Figure 1 shows the system model for the normal case when direct link transmission does not exist due to blockage, while both RIS-link work together to serve destinations. In this paper, we consider a RIS-assisted wireless communication

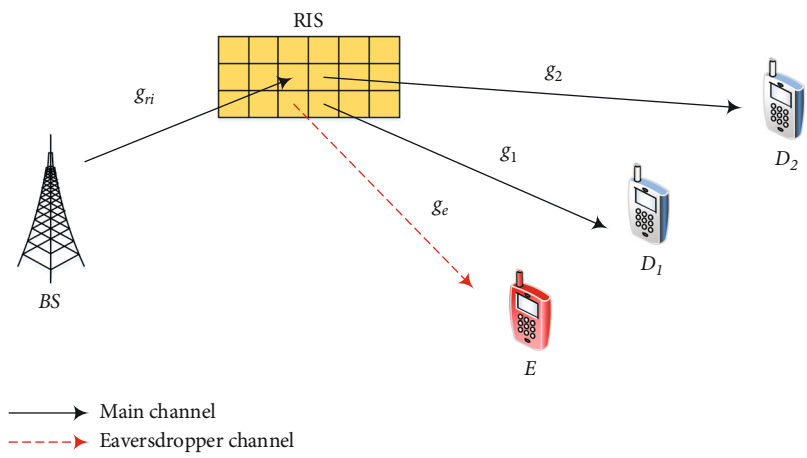


Figure 1: A sketch of IoT secure transmission relying on NOMA and RIS.

Table 1: Main notations.

Symbol	Description
$\Pr(.)$	Probability
$F_Y(.)$	The cumulative distribution function (CDF) of a random variable Y
$f_Y(.)$	The probability density function (PDF) of a random variable Y
P	The BS normalized transmission powers
x_i	The signal for D_i, $(i = 1, 2)$
τ_i	The corresponding power allocation coefficients
∂_i	The additive white Gaussian noise (AWGN) at D_i that is modeled as a zero-mean complex Gaussian distribution with variance N_0
σ	The amplitude reflection coefficient with $\sigma \in (0, 1]$
φ_n	The RIS adjustable phase applied by the n-th reflecting element
ε	The path loss exponent
R_i	The target rate at the user D_i
g_{ri}	The complex channel coefficient for the link $BS \longrightarrow$ RIS
g_1	The complex channel coefficient for the link RIS $\longrightarrow D_1$
g_2	The complex channel coefficient for the link RIS $\longrightarrow D_2$
g_e	The complex channel coefficient for the link RIS $\longrightarrow E$

network in the presence of a BS or AP, two legitimate users (D_1) and (D_2) (it can be extended to multiple users which belong to a group to implement NOMA. As reported in the literature, the performance of two NOMA users often satisfies the high requirement of services at IoT devices rather than multiple users scenarios [16]. Therefore, this paper wants to retain relevant performance by focusing on two users case), and an eavesdropper (E). The RIS is assumed to be installed with N reflecting or meta-surfaces. All the nodes are equipped with a single antenna and experience Rayleigh fading among the channels. The basic signal propagation in this model is assumed as follows: The legitimate user receives a signal from the source via RIS, in which the RIS is expected to improve the quality of the signal. The eavesdropper attempts to obtain the signal from the RIS. The CSI of the legitimate users is assumed to be known, and the RIS utilizes the CSI to maximize the received SNR at the legitimate user. The RIS will not have the CSI of the eavesdropper link. In this scenario, the signals are transmitted from the source at a rate of R_i, and secure transmission is not ensured if the secrecy rate is less than R_i. To determine the performance, we utilize the SOP and SPSC as performance metrics of interest. It is noted that the other main parameters are shown in Table 1.

In this article, the channels are assumed to be slow varying and flat fading channels. Then, the received signals reflected by the RIS at D_1 are given as follows [17–19]:

$$y_{D_1}^{no} = \sigma \sum_{n=1}^{N} \frac{g_{ri} g_1}{\sqrt{d_{ri}^{\varepsilon} d_1^{\varepsilon}}} e^{j\varphi_n} \left(\sqrt{P\tau_1} x_1 + \sqrt{P\tau_2} x_2 \right) + \partial_1, \quad (1)$$

where to ensure better fairness between the users, we assume that $\tau_2 > \tau_1$ with $\tau_1 + \tau_2 = 1$ [20]. Further, g_{ri} and g_i

Table 2: Main parameters for our simulations.

Parameters	Notation	Values
Power coefficients	$\{\tau_1, \tau_2\}$	$\{0.4, 0.6\}$
Target rates	$R_1 = R_2$; $R_1^o = R_2^o$	0.2 (bps/Hz); 1.5 (bps/Hz)
Amplitude reflection coefficient [18]	σ	0.8
Path loss exponent	ε	3
Reflecting elements	N	200
Normalized distances [20]	d_{ri}; $d_1 = d_2$; d_e	0.7; 1; 0.6
Average powers [18]	λ_{W_e}; λ_{Φ_e}	0.02; 0.07
Transmit SNR [24]	α_e	17 (dB)

represent complex Gaussian random variables (RV) with zero mean and unit variance and d_{ri} and d_i are the distances for the BS-RIS and RIS-D_i links, respectively. The small-scale fading channel coefficients are modeled as independent and identically distributed $CN(0,1)$ variables [21]. With large N, via the central limit theorem, we find that $\sum_{n=1}^{N} g_{ri} g_i \sim CN(0, N)$ [17]. When the radio frequency (RF) source transmits its signal x_i to the receiver, the RIS will also receive the same signal and then adjust the phase $\varphi_n \in [0, 2\pi)$ of reflector $n \in \{n, \cdots, N\}$ based upon CSI [17].

The resulting SNR at the legitimate user D_1 to decode x_2 can be formulated as

$$\gamma_{D_1}^{no,x_{21}} = \frac{\sigma^2 d_{ri}^{-\varepsilon} d_1^{-\varepsilon} \sum_{n=1}^{N} |g_{ri}|^2 |g_1|^2 P\tau_2}{\sigma^2 d_{ri}^{-\varepsilon} d_1^{-\varepsilon} \sum_{n=1}^{N} |g_{ri}|^2 |g_1|^2 P\tau_1 + N_0} = \frac{\chi_{W_1} \alpha \tau_2 W_1^2}{\chi_{W_1} \alpha \tau_1 W_1^2 + 1}, \tag{2}$$

where $\chi_{W_1} = \sigma^2 d_{ri}^{-\varepsilon} d_1^{-\varepsilon}$, $\alpha = P/N_0$, due to $W_1 = |\sum_{n=1}^{N} g_{ri} g_1 e^{j\varphi_n}| = \sum_{n=1}^{N} |g_{ri}||g_1|$ in the case of perfect CSI [17].

After successive interference cancellation (SIC), the resulting SNR at the legitimate user D_1 to decode x_1 can be formulated as

$$\gamma_{D_1}^{no,x_1} = \chi_{W_1} \alpha \tau_1 W_1^2, \tag{3}$$

where it assumes a slow varying and flat fading model for all the channels. The received signal reflected by the RIS at D_2 can be written as follows [17, 18]:

$$y_{D_2}^{no} = \sigma \sum_{n=1}^{N} \frac{g_{ri} g_2}{\sqrt{d_{ri}^{\varepsilon} d_2^{\varepsilon}}} e^{j\varphi_n} \left(\sqrt{P\tau_1} x_1 + \sqrt{P\tau_2} x_2 \right) + \partial_2. \tag{4}$$

The resulting SNR at the legitimate user D_2 to decode x_2 can be formulated as

$$\gamma_{D_2}^{no,x_2} = \frac{\chi_{W_2} \alpha \tau_2 W_2^2}{\chi_{W_2} \alpha \tau_1 W_2^2 + 1}, \tag{5}$$

where $\chi_{W_2} = \sigma^2 d_{ri}^{-\varepsilon} d_2^{-\varepsilon}$, $W_2 = \sum_{n=1}^{N} |g_{ri}||g_2|$ [17].

The received RIS reflected signal at E can be written as [18, 22, 23]

$$y_E^{no} = \sigma \sum_{n=1}^{N} \frac{g_{ri} g_e}{\sqrt{d_{ri}^{\varepsilon} d_e^{\varepsilon}}} e^{j\varphi_n} \left(\sqrt{P\tau_1} x_1 + \sqrt{P\tau_2} x_2 \right) + \partial_e, \tag{6}$$

where g_e is a complex Gaussian RV with zero mean and unit variance and d_e is the distances for the RIS-E links. $\partial_e \sim CN(0, N_e)$ is the AWGN at E modeled as a zero-mean complex Gaussian distribution with variance N_e.

Particularly, parallel interference cancellation (PIC) is utilized at E to differentiate the superimposed signals. The resulting SNR at E to decode x_i can be formulated as [24]:

$$\gamma_E^{no,x_i} = \chi_{W_e} \alpha_e \tau_i W_e^2, \tag{7}$$

where $\chi_{W_e} = \sigma^2 d_{ri}^{-\varepsilon} d_e^{-\varepsilon}$, $\alpha_e = P/N_e$, $W_e = \sum_{n=1}^{N} |g_{ri}||g_e|$. W_e can be approximated by an exponential random variable parameter λ_{W_e} [18].

The instantaneous secrecy rate at D_1 is written as follows [18, 24, 25]:

$$C_{D_1}^{no} = \max\left\{ \frac{1}{2}\log_2\left(1 + \min\left(\gamma_{D_1}^{no,x_{21}}, \gamma_{D_1}^{no,x_1}\right)\right) - \frac{1}{2}\log_2\left(1 + \gamma_E^{no,x_1}\right), 0 \right\}. \tag{8}$$

The instantaneous secrecy rate at D_2 can be expressed as

$$C_{D_2}^{no} = \max\left\{ \frac{1}{2}\log_2\left(1 + \gamma_{D_2}^{no,x_2}\right) - \frac{1}{2}\log_2\left(1 + \gamma_E^{no,x_2}\right), 0 \right\}. \tag{9}$$

2.1. Secure Performance Analysis. In this section, the secrecy performance in terms of SOP and SPSC metrics is determined. To gain more insights, we also provide asymptotic SOP analyses.

2.1.1. SOP Analysis. In NOMA-aided systems, signals are transmitted from the source to D_1 and D_2 with the help of a RIS, respectively. Hence, outage happens when either $C_{D_1}^{no}$ or $C_{D_2}^{no}$ falls below their own target rates. With this

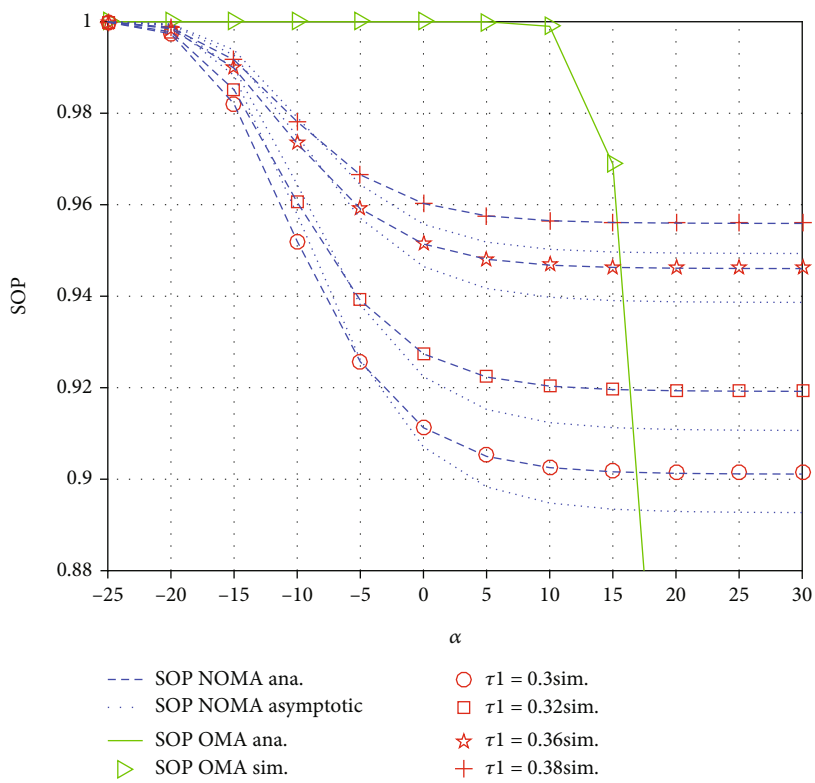


Figure 2: SOP versus α for different τ_1.

understanding, the SOP can be given as follows [24]:

$$\begin{aligned} SOP_{no} &= \Pr\left(C_{D_1}^{no} < R_1 \text{ or } C_{D_2}^{no} < R_2\right) \\ &= 1 - \underbrace{\Pr\left(\frac{1+\gamma_{D_1}^{no,x_{21}}}{1+\gamma_E^{no,x_1}} \geq C_{th_1}, \frac{1+\gamma_{D_1}^{no,x_1}}{1+\gamma_E^{no,x_1}} \geq C_{th_1}\right)}_{\theta_1^{no}} \times \underbrace{\Pr\left(\frac{1+\gamma_{D_2}^{no,x_2}}{1+\gamma_E^{no,x_2}} \geq C_{th_2}\right)}_{\theta_2^{no}}, \end{aligned} \tag{10}$$

where $C_{th_i} = 2^{2R_i}$.

Proposition 1. *The exact expression for SOP_{no} is given by*

$$\begin{aligned} SOP_{no} = 1 &- \frac{\chi_{W_1}\alpha N\zeta_1\zeta_2}{\left(C_{th_1}\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N\right)\lambda_{W_e}\lambda_{W_e}} \\ &\times \exp\left(-\frac{\mu_1}{\chi_{W_1}\alpha\tau_1 N}\right) \\ &\times \int_0^1 \exp\left(-\frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t + \mu_1}{\left(\tau_2 - \tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t + \mu_1\right)\right)\chi_{W_1}\alpha N} - \frac{\zeta_1 t}{\lambda_{W_e}}\right) dt \\ &\times \int_0^1 \exp\left(-\frac{C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 q + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 q + \mu_2\right)\right)\chi_{W_2}\alpha N} - \frac{\zeta_2 q}{\lambda_{W_e}}\right) dq, \end{aligned} \tag{11}$$

where $\zeta_1 = \tau_2 - \tau_1\mu_1/\tau_1\tau_1 C_{th_1}\chi_{W_e}\alpha_e$, $\zeta_2 = \tau_2 - \tau_1\mu_2/\tau_1\tau_2 C_{th_2}\chi_{W_e}\alpha_e$, $\mu_i = C_{th_i} - 1$.

Proof. The details are given in Appendix A. □

Remark 2. The results of (11) illustrate the SOP performance of the proposed system with no direct link between the users albeit assisted by a RIS device. The secure outage threshold is the main factor vital to SOP performance. It is intuitively seen that power allocation coefficients for the two NOMA devices are crucial factors which make a difference in evaluating the system performance for the two considered devices.

2.1.2. SOP Asymptotic. From (6), by using Gauss-Chebyshev integral [26, 27], θ_1^{no} can be computed by the following integral approximation:

$$\begin{aligned} \theta_1^{no} \approx &\frac{\chi_{W_1}\alpha N}{C_{th_1}\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N}\frac{\pi\zeta_1}{2G\lambda_{W_e}}\exp\left(-\frac{\mu_1}{\chi_{W_1}\alpha\tau_1 N}\right) \\ &\times \sum_{j=1}^{G}\Xi\exp\left(-\frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t_j + \mu_1}{\left(\tau_2 - \tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t_j + \mu_1\right)\right)\chi_{W_1}\alpha N} - \frac{\zeta_1 t_j}{\lambda_{W_e}}\right), \end{aligned} \tag{12}$$

where $t_j = 1/2(1 + \cos(2j - 1/G))$, $\Xi = |\sin(2j - 1/G)|$ and G

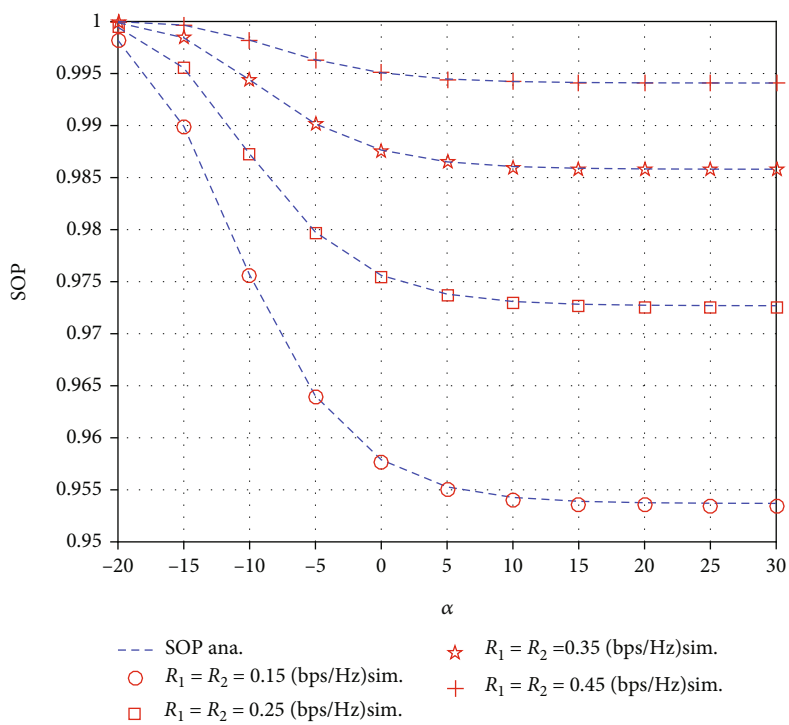


Figure 3: SOP versus α for different $R_1 = R_2$.

is the Gauss-Chebyshev integral approximated sum term [26].

From (9), by using Gauss-Chebyshev integral [26, 27], θ_2^{no} can be obtained using the following integral approximation:

$$\theta_2^{no} \approx \frac{\pi\zeta_2}{2G\lambda_{W_e}} \sum_{j=1}^{G} \Xi \exp\left(\begin{array}{c} -\dfrac{C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 t_j + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 t_j + \mu_2\right)\right)\chi_{W_2}\alpha N} \\ -\dfrac{\zeta_2 t_j}{\lambda_{W_e}} \end{array} \right). \tag{13}$$

Based on (12) and (13), the final approximate closed-form expression for SOP_{no} is given by

$$\begin{aligned} SOP_{no}^{asym} \approx 1 &- \frac{\chi_{W_1}\alpha N\pi\pi\zeta_1\zeta_2}{\left(C_{th_1}\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N\right)4GG\lambda_{W_e}\lambda_{W_e}} \\ &\times \exp\left(-\frac{\mu_1}{\chi_{W_1}\alpha\tau_1 N}\right) \\ &\times \sum_{j=1}^{G} \Xi \exp\left(\begin{array}{c} -\dfrac{C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t_j + \mu_1}{\left(\tau_2 - \tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t_j + \mu_1\right)\right)\chi_{W_1}\alpha N} \\ -\dfrac{\zeta_1 t_j}{\lambda_{W_e}} \end{array} \right) \\ &\times \sum_{j=1}^{G} \boxtimes\Xi \exp\left(\begin{array}{c} -\dfrac{C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 t_j + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 t_j + \mu_2\right)\right)\chi_{W_2}\alpha N} \\ -\dfrac{\zeta_2 t_j}{\lambda_{W_e}} \end{array} \right). \end{aligned} \tag{14}$$

2.1.3. SPSC Analysis. SPSC is one of the fundamental benchmarks for secrecy performance, and it denotes the probability of existence of secrecy capacity [24, 28]. Thus, the SPSC for a NOMA system can be written as

$$\begin{aligned} SPSC_{no} &= \Pr\left(C_{D_1}^{no} > 0, C_{D_2}^{no} > 0\right) \\ &= \underbrace{\Pr\left(\min\left(\gamma_{D_1}^{no,x_{21}}, \gamma_{D_1}^{no,x_1}\right) > \gamma_E^{no,x_1}\right)}_{\Omega_1^{no}} \\ &\quad \times \underbrace{\Pr\left(\gamma_{D_2}^{no,x_2} > \gamma_E^{no,x_2}\right)}_{\Omega_2^{no}}. \end{aligned} \tag{15}$$

Proposition 3. *The exact expression for $SPSC_{no}$ is given by*

$$\begin{aligned} SPSC_{no} = &\frac{\chi_{W_1}\alpha N\zeta_3\zeta_4}{\left(\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N\right)\lambda_{W_e}\lambda_{W_e}} \\ &\times \int_0^1 \exp\left(\begin{array}{c} -\dfrac{\chi_{W_e}\alpha_e\tau_1\zeta_3 t}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1\zeta_3 t\right)\chi_{W_1}\alpha N} \\ -\dfrac{\zeta_3 t}{\lambda_{W_e}} \end{array} \right) dt \\ &\times \int_0^1 \exp\left(\begin{array}{c} -\dfrac{\chi_{W_e}\alpha_e\tau_2\zeta_4 q}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_2\zeta_4 q\right)\chi_{W_2}\alpha N} \\ -\dfrac{\zeta_4 q}{\lambda_{W_e}} \end{array} \right) dq, \end{aligned} \tag{16}$$

where $\zeta_3 = \tau_2/\tau_1\tau_1\chi_{W_e}\alpha_e$, $\zeta_4 = 1/\tau_1\chi_{W_e}\alpha_e$.

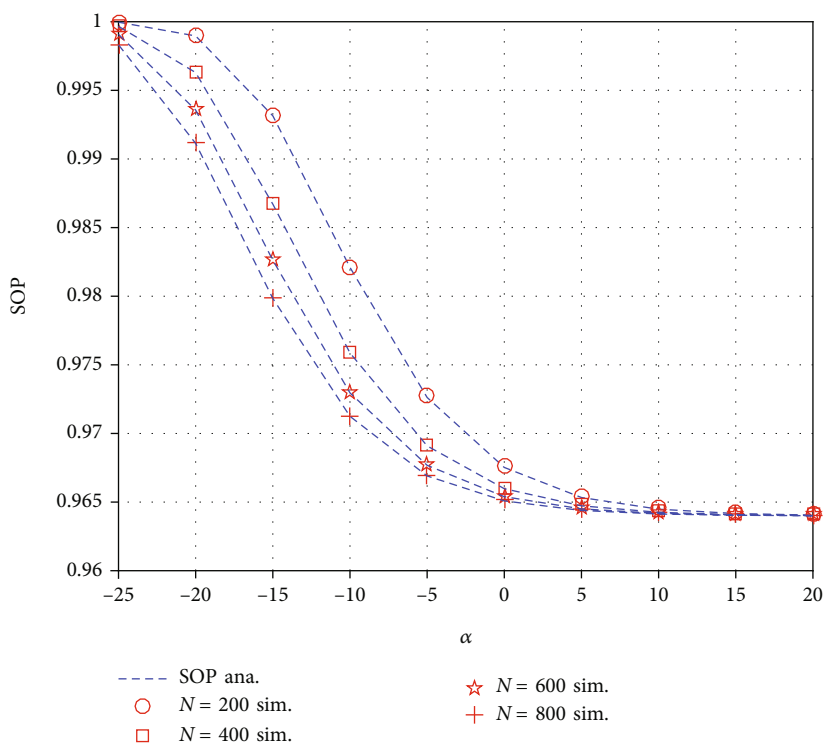


Figure 4: SOP versus α for different N.

Proof. The details are given in Appendix B. □

Remark 4. (16) illustrates the SPSC performance of the proposed system with no direct link between the users in the presence of RIS. The structure of RIS influences the SPSC performance. Therefore, by tailoring RIS, the traditional IoT gets more benefits against eavesdroppers.

2.1.4. SPSC Asymptotic. From (6), by using the Gauss-Chebyshev integral [26, 27], Ω_1^{no} can be obtained via the following integral approximation:

$$\Omega_1^{no} \approx \frac{\chi_{W_1}\alpha N\pi\zeta_3}{\left(\chi_{W_e}\alpha_e\lambda_{W_e}+\chi_{W_1}\alpha N\right)2G\lambda_{W_e}} \times \sum_{j=1}^{G}\Xi\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_1\zeta_3 t_j}{\left(\tau_2-\chi_{W_e}\alpha_e\tau_1\tau_1\zeta_3 t_j\right)\chi_{W_1}\alpha N}-\frac{\zeta_3 t_j}{\lambda_{W_e}}\right). \tag{17}$$

From (9), by using the Gauss-Chebyshev integral [26, 27], Ω_2^{no} can be formulated using the following integral approximation:

$$\Omega_2^{no} \approx \frac{\pi\zeta_4}{2G\lambda_{W_e}}\sum_{j=1}^{G}\Xi\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_2\zeta_4 t_j}{\left(\tau_2-\chi_{W_e}\alpha_e\tau_1\tau_2\zeta_4 t_j\right)\chi_{W_2}\alpha N}-\frac{\zeta_4 t_j}{\lambda_{W_e}}\right). \tag{18}$$

Based on (17) and (18), the final approximate closed-form expression for $SPSC_{no}$ is given by

$$SPSC_{no}^{asym} \approx \frac{\chi_{W_1}\alpha N\pi\pi\zeta_3\zeta_4}{\left(\chi_{W_e}\alpha_e\lambda_{W_e}+\chi_{W_1}\alpha N\right)4GG\lambda_{W_e}\lambda_{W_e}} \times \sum_{j=1}^{G}\Xi\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_1\zeta_3 t_j}{\left(\tau_2-\chi_{W_e}\alpha_e\tau_1\tau_1\zeta_3 t_j\right)\chi_{W_1}\alpha N}-\frac{\zeta_3 t_j}{\lambda_{W_e}}\right) \times \sum_{j=1}^{G}\Xi\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_2\zeta_4 t_j}{\left(\tau_2-\chi_{W_e}\alpha_e\tau_1\tau_2\zeta_4 t_j\right)\chi_{W_2}\alpha N}-\frac{\zeta_4 t_j}{\lambda_{W_e}}\right). \tag{19}$$

3. RIS-OMA Scheme

It would be better to compare with the counterpart, i.e., RIS-OMA. Similarly, the received signal reflected by the RIS at D_i are given as

$$y_{D_i}^{o} = \sigma\sum_{n=1}^{N}\frac{g_{ri}g_i}{\sqrt{d_{ri}^{\varepsilon}d_i^{\varepsilon}}}e^{j\varphi_n}\sqrt{P}x_i+\partial_i. \tag{20}$$

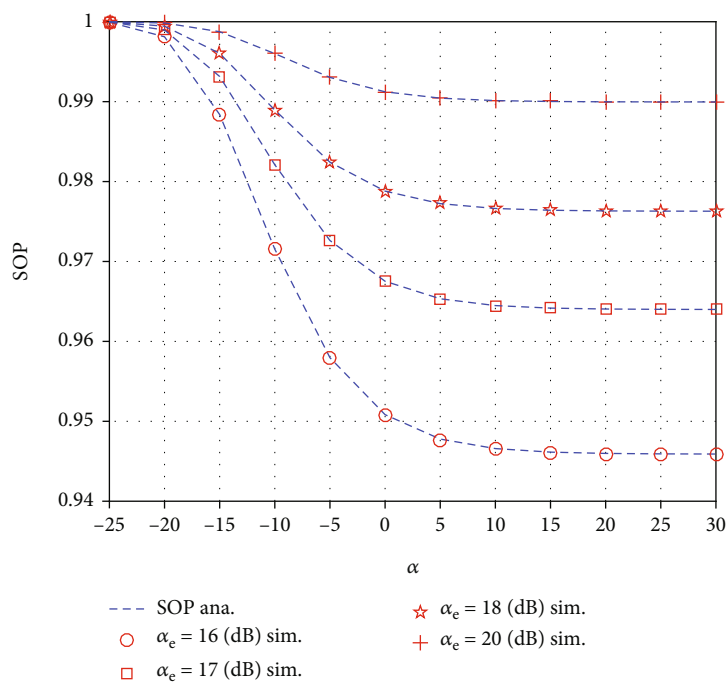


Figure 5: SOP versus α for different α_e.

We then compute SNR at the user D_i to decode x_i as

$$\gamma_{D_i}^{o,x_i} = \chi_{W_i} \alpha W_i^2, \tag{21}$$

where $\chi_{W_i} = \sigma^2 d_{ri}^{-\varepsilon} d_i^{-\varepsilon}$, $W_i = |\sum_{n=1}^{N} g_{ri} g_i e^{j\varphi_n}| = \sum_{n=1}^{N} |g_{ri}||g_i|$.

The RIS is also able to reflect signals to E and the received signal at E can be written as

$$y_E^o = \sigma \sum_{n=1}^{N} \frac{g_{ri} g_e}{\sqrt{d_{ri}^{\varepsilon} d_e^{\varepsilon}}} e^{j\varphi_n} \sqrt{P} x_i + \partial_e. \tag{22}$$

Based on the received signal at E, we then calculate SNR at E to decode x_i as

$$\gamma_E^{o,x_i} = \chi_{W_e} \alpha_e W_e^2. \tag{23}$$

The instantaneous secrecy rate at D_i in this RIS-OMA scenario is computed by

$$C_{D_i}^o = \max\left\{\frac{1}{4}\log_2\left(1+\gamma_{D_i}^{o,x_i}\right) - \frac{1}{4}\log_2\left(1+\gamma_E^{o,x_i}\right), 0\right\}. \tag{24}$$

3.1. SOP Analysis. Similarly from (10), the SOP can be given as

$$\begin{aligned} SOP_o &= \Pr\left(C_{D_1}^o < R_1^o \text{ or } C_{D_2}^o < R_2^o\right) \\ &= 1 - \underbrace{\Pr\left(\frac{1+\gamma_{D_1}^{o,x_1}}{1+\gamma_E^{o,x_1}} \geq C_{th_1}^o\right)}_{\theta_1^o} \times \underbrace{\Pr\left(\frac{1+\gamma_{D_2}^{o,x_2}}{1+\gamma_E^{o,x_2}} \geq C_{th_2}^o\right)}_{\theta_2^o}, \end{aligned} \tag{25}$$

where $C_{th_i}^o = 2^{8R_i^o}$ and R_i^o is the target rate at the user D_i for RIS-OMA case.

Firstly, θ_1^o can given by

$$\begin{aligned} \theta_1^o &= \Pr\left(\frac{1+\gamma_{D_1}^{o,x_1}}{1+\gamma_E^{o,x_1}} \geq C_{th_1}^o\right) \\ &= \Pr\left(W_1^2 \geq \frac{C_{th_1}^o \chi_{W_e} \alpha_e W_e^2 + \mu_1^o}{\chi_{W_1} \alpha}\right) \\ &= \int_0^{\infty} \left(1 - F_{W_1^2}\left(\frac{C_{th_1}^o \chi_{W_e} \alpha_e x + \mu_1^o}{\chi_{W_1} \alpha}\right)\right) f_{W_e^2}(x) dx \\ &= \frac{1}{\lambda_{W_e}} \exp\left(-\frac{\mu_1^o}{\chi_{W_1} \alpha N}\right) \\ &\quad \times \int_0^{\infty} \exp\left(-\left(\frac{C_{th_1}^o \chi_{W_e} \alpha_e}{\chi_{W_1} \alpha N} + \frac{1}{\lambda_{W_e}}\right) x\right) dx \\ &= \frac{\chi_{W_1} \alpha N}{C_{th_1}^o \chi_{W_e} \alpha_e \lambda_{W_e} + \chi_{W_1} \alpha N} \exp\left(-\frac{\mu_1^o}{\chi_{W_1} \alpha N}\right), \end{aligned} \tag{26}$$

where $\mu_i^o = C_{th_i}^o - 1$.

Next, θ_2^o is calculated similarly to θ_1^o. In particular, θ_2^o can expressed by

$$\theta_2^o = \frac{\chi_{W_2} \alpha N}{C_{th_2}^o \chi_{W_e} \alpha_e \lambda_{W_e} + \chi_{W_2} \alpha N} \exp\left(-\frac{\mu_2^o}{\chi_{W_2} \alpha N}\right). \tag{27}$$

From (26) and (27) into (25), the closed-form expression

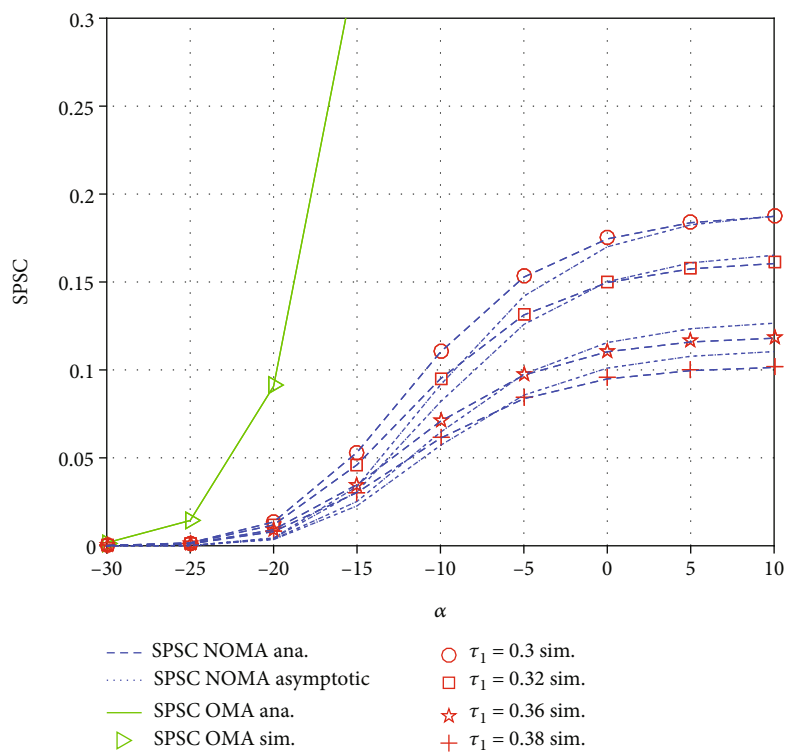


Figure 6: SPSC versus α for different τ_1.

SOP_o is given by

$$SOP_o = 1 - \frac{\chi_{W_1}\alpha N}{C_{th_1}^o \chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N} \times \frac{\chi_{W_2}\alpha N}{C_{th_2}^o \chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_2}\alpha N} \times \exp\left(-\frac{\mu_1^o}{\chi_{W_1}\alpha N} - \frac{\mu_2^o}{\chi_{W_2}\alpha N}\right). \tag{28}$$

3.2. SPSC Analysis. The SPSC for a OMA system can be written as

$$SPSC_o = \Pr\left(C_{D_1}^o > 0, C_{D_2}^o > 0\right) = \underbrace{\Pr\left(\gamma_{D_1}^{o,x_1} > \gamma_E^{o,x_1}\right)}_{\Omega_1^o} \underbrace{\Pr\left(\gamma_{D_2}^{o,x_2} > \gamma_E^{o,x_2}\right)}_{\Omega_2^o}. \tag{29}$$

From (29), Ω_1^o is written by

$$\begin{aligned}\Omega_1^o &= \Pr\left(\gamma_{D_1}^{o,x_1} > \gamma_E^{o,x_1}\right) = \Pr\left(W_1^2 \geq \frac{\chi_{W_e}\alpha_e W_e^2}{\chi_{W_1}\alpha}\right) \\ &= \int_0^\infty \left(1 - F_{W_1^2}\left(\frac{\chi_{W_e}\alpha_e x}{\chi_{W_1}\alpha}\right)\right) f_{W_e^2}(x)dx \\ &= \frac{1}{\lambda_{W_e}}\int_0^\infty \exp\left(-\left(\frac{\chi_{W_e}\alpha_e}{\chi_{W_1}\alpha N} + \frac{1}{\lambda_{W_e}}\right)x\right)dx \\ &= \frac{\chi_{W_1}\alpha N}{\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N}.\end{aligned} \tag{30}$$

Then, Ω_2^o is calculated similarly to Ω_1^o. Ω_2^o can given by

$$\Omega_2^o = \frac{\chi_{W_2}\alpha N}{\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_2}\alpha N}. \tag{31}$$

From (30) and (31) into (29), the closed-form expression $SPSC_o$ is given by

$$SPSC_o = \frac{\chi_{W_1}\alpha N}{\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N} \times \frac{\chi_{W_2}\alpha N}{\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_2}\alpha N}. \tag{32}$$

4. Numerical Results

In this section, numerical examples are presented to verify our analytical results. In the simulation results, the Rayleigh fading is assumed for all the channels [18]. The SOP and SPSC are obtained with Monte Carlo simulations. The main parameters can be considered in Table 2 except for specific cases.

Figure 2 shows the simulation of SOP versus transmit SNR for different power allocation levels. It can be determined from (10), (11) that the transmit SNR α contributes significantly to SOP performance and this can be verified in this figure. It can also be observed from the simulation that even a minute level of change in power allocation can create a noticeable difference in the performance of the system. As the transmit SNR increases beyond 15 (dB), the performance of the system goes constant. This phenomenon can be explained by the fact that SOP depends on many parameters rather than α, for example, channel gain, and data rates C_{th_i}. SOP performance is important to indicate

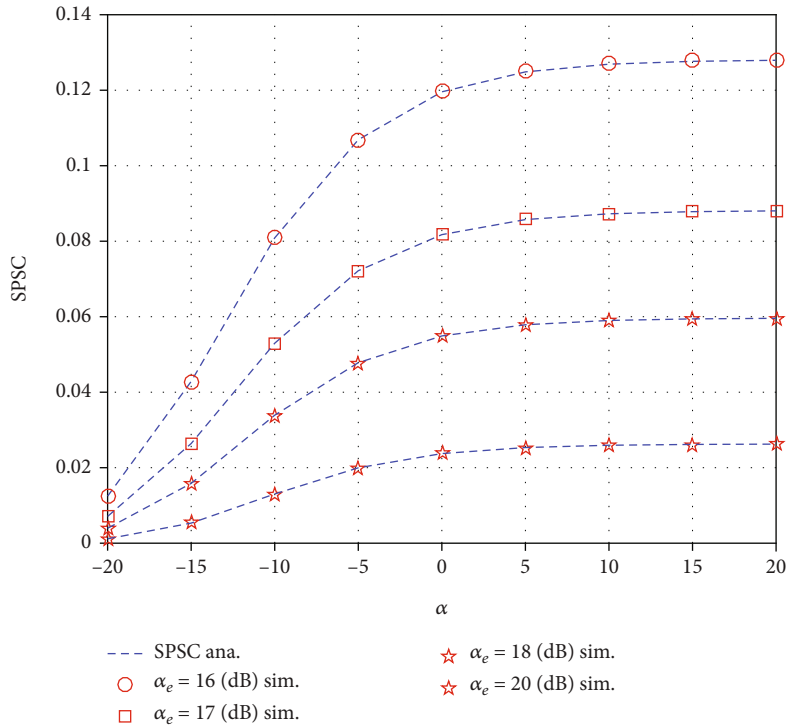


Figure 7: SPSC versus α for different α_e.

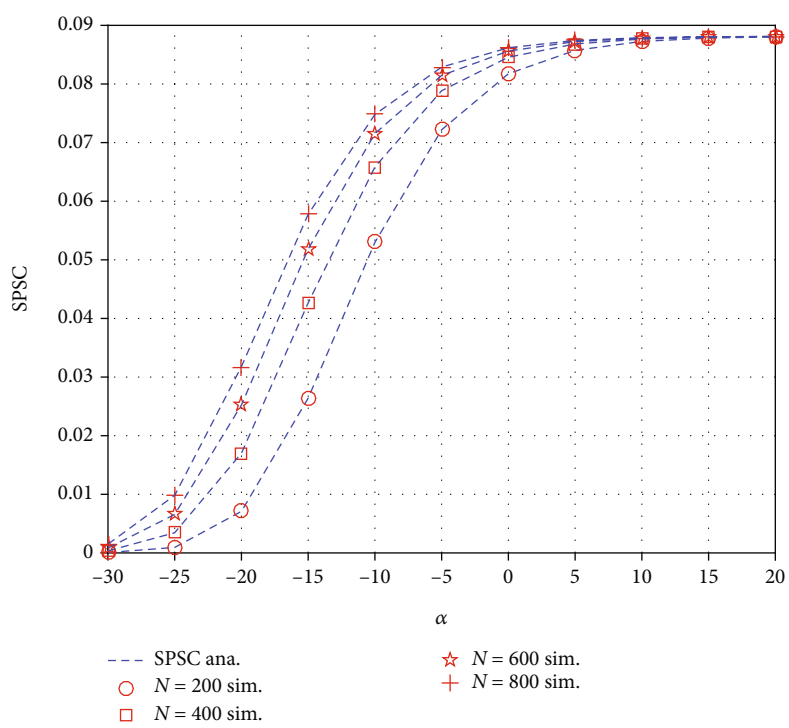


Figure 8: SPSC versus α for different N.

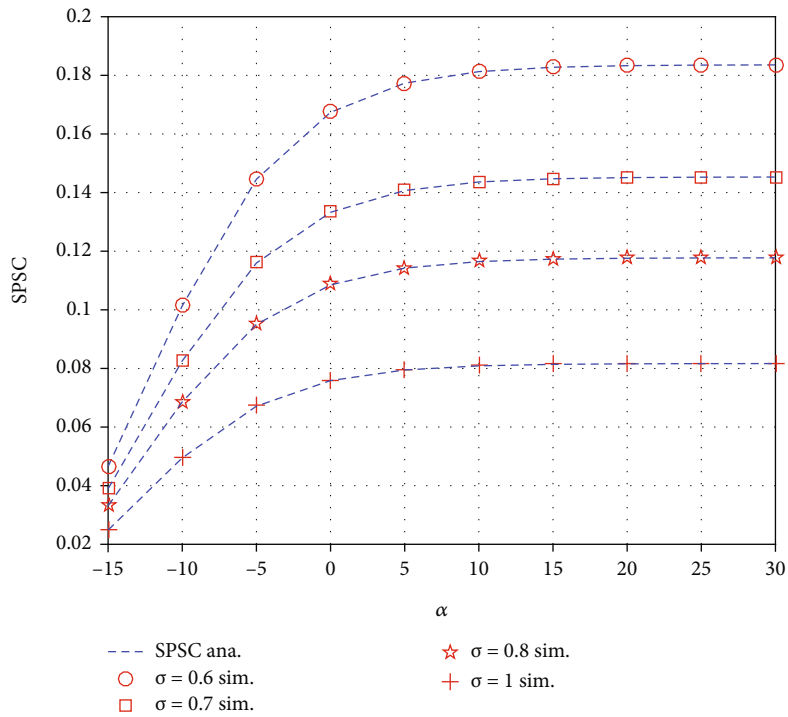


Figure 9: SPSC versus α for different σ.

how system works under particular conditions of channels. We can deal with SOP improvement if α and data rates C_{th_i} are adjusted properly. Further, we can confirm that RIS-NOMA outperforms than RIS-OMA if α is less than 17 (dB). The main reason is that RIS-NOMA provides higher spectrum efficiency.

Figure 3 shows the simulation for SOP versus transmit SNR for different levels of target rates R_1, R_2 assigned to the users. To perform a fair comparison, in each case, the target rates are assigned equally for all the users in the system. As we can observe, as the target rates increase, the performance of the system also becomes worse. The reason is that SOP performance derived in (10) and (11) are logically limited by such target rates C_{th_i}.

Figure 4 demonstrates the simulation of SOP versus transmit SNR for a different number of meta-surfaces installed at RIS. As we can see, as the number of meta-surfaces increases, the secrecy performance of the system increases. The noticeable point in this study is, even though the number of meta-surfaces is sufficiently large enough, as, with the increase in the transmit SNR, the performance of the system goes into constant mode after a certain level of α. The major reason for this situation is because, at high SNR levels, the interference between the users in the cluster becomes dominant which affects the performance of the system. SOP performance corresponds to how we design RIS with respect to the number of meta-surface. It can be explained that more meta-surface elements at RIS contributes to improve quality of received signal and the corresponding SOP can be enhanced.

Figure 5 shows the simulation of SOP versus transmit SNR for different levels of SNR at the eavesdropper. Since PIC is considered in the system, the effect of SNR at eavesdropper can be seen clearly. As the value of α_e increases, the performance of the system is decreasing rapidly. Even for the small difference between each case, the secrecy performance of the system is being affected or decreased with a huge gap between the curves.

Figure 6 shows the simulation of SPSC versus transmit SNR for different levels of power allocation at D_1. This simulation is performed similar to Figure 2 but with a different method of analysis, i.e., SPSC. As we can observe, the change in power allocation levels has shown significant change in the performance of the system. Further, it can be seen the gap between RIS-OMA and RIS-NOMA cases if we refer to SPSC case. In this case, RIS-NOMA deals with a fixed allocation scheme to assign power to two NOMA users, and hence its SPSC performance looks worse compared with RIS-OMA case.

Similarly, Figure 7 is simulated identically to Figure 5, and it shows SPSC versus transmit SNR with different levels of SNR at the eavesdropper. As mentioned, since PIC is considered, as the α_e value increases, the SPSC performance of the system decreases rapidly.

Figure 8 is simulated similarly to Figure 4. It shows that the simulation between SPSC versus transmits SNR for different numbers of meta-surfaces installed at RIS. As the number of meta-surfaces increases, the performance of the system increases comparatively.

Figure 9 shows the simulation between SPSC versus transmit SNR for different levels of amplitude reflection coefficients. We can observe that the changes in the level of

σ do not show much effect on the direct link and it shows a huge effect on the no direct link since the RIS is the only possible way of communication between the devices. As the σ value decreases, the performance of the system decreases.

5. Conclusion

In this paper, we have considered a RIS system with two legitimate users, being served in a cluster, and an eavesdropper. All users are equipped with a single antenna. The proposed model was considered no direct link between the devices in the presence of RIS. The performance of the system is analyzed from the perspective of secrecy efficiency. Asymptotic and closed-form expressions are derived for SOP and SPSC. The simulations were performed based on these expressions, and the results are verified using the Monte-Carlo method. The study provides that in most of the cases, the secrecy performance of the system was efficient in the direct link between the devices in the presence of RIS. As we previously mentioned, the number of meta-surfaces and SNR levels at the users also play a pivotal role in influencing the performance of the system.

Appendix

A. Proof of Proposition 1

From (10), θ_1^{no} is given by

$$\theta_1^{no} = \underbrace{\Pr\left(\frac{1+\gamma_{D_1}^{no,x_{21}}}{1+\gamma_E^{no,x_1}} \geq C_{th_1}\right)}_{\theta_{1a}^{no}} \times \underbrace{\Pr\left(\frac{1+\gamma_{D_1}^{no,x_1}}{1+\gamma_E^{no,x_1}} \geq C_{th_1}\right)}_{\theta_{1b}^{no}}. \tag{A.1}$$

Then, an upper bound of θ_{1a}^{no} can be obtained as

$$\begin{aligned}\theta_{1a}^{no} &= \Pr\left(\frac{\chi_{W_1}\alpha\tau_2 W_1^2}{\chi_{W_1}\alpha\tau_1 W_1^2+1} \geq C_{th_1}\chi_{W_e}\alpha_e\tau_1 W_e^2+\mu_1\right)\\ &= \Pr\left(W_1^2 \geq \frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1 W_e^2+\mu_1}{\left(\tau_2-\tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1 W_e^2+\mu_1\right)\right)\chi_{W_1}\alpha}\right),\end{aligned} \tag{A.2}$$

where $\mu_i = C_{th_i} - 1$ and θ_{1a}^{no} is calculated with the condition $W_e^2 < \tau_2 - \tau_1\mu_1/\tau_1\tau_1 C_{th_1}\chi_{W_e}\alpha_e$, we let $\zeta_1 = \tau_2 - \tau_1\mu_1/\tau_1\tau_1 C_{th_1}\chi_{W_e}\alpha_e$. It is noted that all channels follow the Rayleigh distribution with PDF and CDF $f_{|X|^2}(x) = (1/\varphi_X)\exp(-x/\varphi_X)$, $F_{|X|^2}(x) = 1-\exp(-x/\varphi_X)$, respectively [18]. Then, θ_{1a}^{no} can be calculated as

$$\begin{aligned}\theta_{1a}^{no} &= \int_0^{\zeta_1}\left(1-F_{W_1^2}\left(\frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1}{\left(\tau_2-\tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1\right)\right)\chi_{W_1}\alpha}\right)\right)\\ &\quad\times f_{W_e^2}(x)dx = \int_0^{\zeta_1}\exp\left(-\frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1}{\left(\tau_2-\tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1\right)\right)\chi_{W_1}\alpha N}\right)\\ &\quad\times\frac{1}{\lambda_{W_e}}\exp\left(-\frac{x}{\lambda_{W_e}}\right)dx\\ &= \frac{1}{\lambda_{W_e}}\int_0^{\zeta_1}\exp\left(\begin{array}{c}-\dfrac{C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1}{\left(\tau_2-\tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1\right)\right)\chi_{W_1}\alpha N}\\ -\dfrac{x}{\lambda_{W_e}}\end{array}\right)dx.\end{aligned} \tag{A.3}$$

From (A.3), $t = x/\zeta_1$. θ_{1a}^{no} can written by

$$\theta_{1a}^{no} = \frac{\zeta_1}{\lambda_{W_e}}\int_0^1\exp\left(\begin{array}{c}-\dfrac{C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t+\mu_1}{\left(\tau_2-\tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t+\mu_1\right)\right)\chi_{W_1}\alpha N}\\ -\dfrac{\zeta_1 t}{\lambda_{W_e}}\end{array}\right)dt. \tag{A.4}$$

From (A.1), θ_{1b}^{no} can be expressed as follows:

$$\begin{aligned}\theta_{1b}^{no} &= \Pr\left(W_1^2 \geq \frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1 W_e^2+\mu_1}{\chi_{W_1}\alpha\tau_1}\right)\\ &= \int_0^\infty\left(1-F_{W_1^2}\left(\frac{C_{th_1}\chi_{W_e}\alpha_e\tau_1 x+\mu_1}{\chi_{W_1}\alpha\tau_1}\right)\right)f_{W_e^2}(x)dx\\ &= \frac{1}{\lambda_{W_e}}\exp\left(-\frac{\mu_1}{\chi_{W_1}\alpha\tau_1 N}\right)\times\int_0^\infty\exp\left(-\left(\frac{C_{th_1}\chi_{W_e}\alpha_e}{\chi_{W_1}\alpha N}+\frac{1}{\lambda_{W_e}}\right)x\right)dx\\ &= \frac{\chi_{W_1}\alpha N}{C_{th_1}\chi_{W_e}\alpha_e\lambda_{W_e}+\chi_{W_1}\alpha N}\exp\left(-\frac{\mu_1}{\chi_{W_1}\alpha\tau_1 N}\right).\end{aligned} \tag{A.5}$$

From (A.4) and (A.5) into (A.1), θ_1^{no} can be written as

$$\begin{aligned}\theta_1^{no} &= \frac{\chi_{W_1}\alpha N}{C_{th_1}\chi_{W_e}\alpha_e\lambda_{W_e}+\chi_{W_1}\alpha N}\frac{\zeta_1}{\lambda_{W_e}}\exp\left(-\frac{\mu_1}{\chi_{W_1}\alpha\tau_1 N}\right)\\ &\quad\times\int_0^1\exp\left(\begin{array}{c}-\dfrac{C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t+\mu_1}{\left(\tau_2-\tau_1\left(C_{th_1}\chi_{W_e}\alpha_e\tau_1\zeta_1 t+\mu_1\right)\right)\chi_{W_1}\alpha N}\\ -\dfrac{\zeta_1 t}{\lambda_{W_e}}\end{array}\right)dt.\end{aligned} \tag{A.6}$$

Further, θ_2^{no} can be written as

$$\theta_2^{no} = \Pr\left(\frac{\chi_{W_2}\alpha\tau_2 W_2^2}{\chi_{W_2}\alpha\tau_1 W_2^2 + 1} \geq C_{th_2}\chi_{W_e}\alpha_e\tau_2 W_e^2 + \mu_2\right)$$
$$= \Pr\left(W_2^2 \geq \frac{C_{th_2}\chi_{W_e}\alpha_e\tau_2 W_e^2 + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2 W_e^2 + \mu_2\right)\right)\chi_{W_2}\alpha}\right). \tag{A.7}$$

We let $\zeta_2 = \tau_2 - \tau_1\mu_2/\tau_1\tau_2 C_{th_2}\chi_{W_e}\alpha_e$, θ_2^{no} be calculated with the condition $W_e^2 < \zeta_2$. θ_2^{no} can be expressed as follows:

$$\theta_2^{no} = \int_0^{\zeta_2}\left(1 - F_{W_2^2}\left(\frac{C_{th_2}\chi_{W_e}\alpha_e\tau_2 x + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2 x + \mu_2\right)\right)\chi_{W_2}\alpha}\right)\right) \times f_{W_e^2}(x)dx$$
$$= \int_0^{\zeta_2}\exp\left(-\frac{C_{th_2}\chi_{W_e}\alpha_e\tau_2 x + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2 x + \mu_2\right)\right)\chi_{W_2}\alpha N}\right)$$
$$\times \frac{1}{\lambda_{W_e}}\exp\left(-\frac{x}{\lambda_{W_e}}\right)dx$$
$$= \frac{1}{\lambda_{W_e}}\int_0^{\zeta_2}\exp\left(\begin{array}{c}-\dfrac{C_{th_2}\chi_{W_e}\alpha_e\tau_2 x + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2 x + \mu_2\right)\right)\chi_{W_2}\alpha N} \\ -\dfrac{x}{\lambda_{W_e}}\end{array}\right)dx. \tag{A.8}$$

From (A.8), we let $q = x/\zeta_2$. θ_2^{no} be written as

$$\theta_2^{no} = \frac{\zeta_2}{\lambda_{W_e}}\int_0^1\exp\left(\begin{array}{c}-\dfrac{C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 q + \mu_2}{\left(\tau_2 - \tau_1\left(C_{th_2}\chi_{W_e}\alpha_e\tau_2\zeta_2 q + \mu_2\right)\right)\chi_{W_2}\alpha N} \\ -\dfrac{\zeta_2 q}{\lambda_{W_e}}\end{array}\right)dq. \tag{A.9}$$

Substituting (A.6) and (A.9) into (10), we can obtain (11).

The proof is completed.

B. Proof of Proposition 3

From (15), Ω_1^{no} can be written as

$$\Omega_1^{no} = \Pr\left(\min\left(\gamma_{D_1}^{no,x_{21}}, \gamma_{D_1}^{no,x_1}\right) > \gamma_E^{no,x_1}\right)$$
$$= \underbrace{\Pr\left(\gamma_{D_1}^{no,x_{21}} > \gamma_E^{no,x_1}\right)}_{\Omega_{1a}^{no}}\underbrace{\Pr\left(\gamma_{D_1}^{no,x_1} > \gamma_E^{no,x_1}\right)}_{\Omega_{1b}^{no}}. \tag{B.1}$$

From (B.1), Ω_{1a}^{no} can be obtained as

$$\Omega_{1a}^{no} = \Pr\left(\gamma_{D_1}^{no,x_{21}} > \gamma_E^{no,x_1}\right) = \Pr\left(\frac{\chi_{W_1}\alpha\tau_2 W_1^2}{\chi_{W_1}\alpha\tau_1 W_1^2 + 1} \geq \chi_{W_e}\alpha_e\tau_1 W_e^2\right)$$
$$= \Pr\left(W_1^2 \geq \frac{\chi_{W_e}\alpha_e\tau_1 W_e^2}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1 W_e^2\right)\chi_{W_1}\alpha}\right). \tag{B.2}$$

Next, we let $\zeta_3 = \tau_2/\tau_1\tau_1\chi_{W_e}\alpha_e$ and Ω_{1a}^{no} be calculated with the condition $W_e^2 < \zeta_3$. Ω_{1a}^{no} can be rewritten as

$$\Omega_{1a}^{no} = \int_0^{\zeta_3}\left(1 - F_{W_1^2}\left(\frac{\chi_{W_e}\alpha_e\tau_1 x}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1 x\right)\chi_{W_1}\alpha}\right)\right) \times f_{W_e^2}(x)dx$$
$$= \int_0^{\zeta_3}\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_1 x}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1 x\right)\chi_{W_1}\alpha N}\right)$$
$$\times \frac{1}{\lambda_{W_e}}\exp\left(-\frac{x}{\lambda_{W_e}}\right)dx$$
$$= \frac{1}{\lambda_{W_e}}\int_0^{\zeta_3}\exp\left(\begin{array}{c}-\dfrac{\chi_{W_e}\alpha_e\tau_1 x}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1 x\right)\chi_{W_1}\alpha N} \\ -\dfrac{x}{\lambda_{W_e}}\end{array}\right)dx. \tag{B.3}$$

From (B.3), we let $t = x/\zeta_3$. Ω_{1a}^{no} be given by

$$\Omega_{1a}^{no} = \frac{\zeta_3}{\lambda_{W_e}}\int_0^1\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_1\zeta_3 t}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1\zeta_3 t\right)\chi_{W_1}\alpha N} - \frac{\zeta_3 t}{\lambda_{W_e}}\right)dt. \tag{B.4}$$

From (B.1), Ω_{1b}^{no} can be obtained as

$$\Omega_{1b}^{no} = \Pr\left(\gamma_{D_1}^{no,x_1} \geq \gamma_E^{no,x_1}\right) = \Pr\left(W_1^2 \geq \frac{\chi_{W_e}\alpha_e W_e^2}{\chi_{W_1}\alpha}\right)$$
$$= \int_0^{\infty}\left(1 - F_{W_1^2}\left(\frac{\chi_{W_e}\alpha_e x}{\chi_{W_1}\alpha}\right)\right)f_{W_e^2}(x)dx$$
$$= \frac{1}{\lambda_{W_e}}\int_0^{\infty}\exp\left(-\left(\frac{\chi_{W_e}\alpha_e}{\chi_{W_1}\alpha N} + \frac{1}{\lambda_{W_e}}\right)x\right)dx$$
$$= \frac{\chi_{W_1}\alpha N}{\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N}. \tag{B.5}$$

From (B.4) and (B.5) into (B.1), Ω_1^{no} can be written as

$$\Omega_1^{no} = \frac{\chi_{W_1}\alpha N\zeta_3}{\left(\chi_{W_e}\alpha_e\lambda_{W_e} + \chi_{W_1}\alpha N\right)\lambda_{W_e}} \times \int_0^1 \exp\left(-\frac{\chi_{W_e}\alpha_e\tau_1\zeta_3 t}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_1\zeta_3 t\right)\chi_{W_1}\alpha N} - \frac{\zeta_3 t}{\lambda_{W_e}}\right) dt. \tag{B.6}$$

From (15), Ω_2^{no} can be obtained as

$$\begin{aligned}\Omega_2^{no} &= \Pr\left(\gamma_{D_2}^{no,x_2} > \gamma_E^{no,x_2}\right) \\ &= \Pr\left(W_2^2 \geq \frac{\chi_{W_e}\alpha_e\tau_2 W_e^2}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_2 W_e^2\right)\chi_{W_2}\alpha}\right).\end{aligned} \tag{B.7}$$

Next, we let $\zeta_4 = 1/\tau_1\chi_{W_e}\alpha_e$ and Ω_2^{no} be calculated with the condition $W_e^2 < \zeta_4$. Ω_2^{no} can be rewritten as

$$\begin{aligned}\Omega_2^{no} &= \int_0^{\zeta_4}\left(1 - F_{W_2^2}\left(\frac{\chi_{W_e}\alpha_e\tau_2 x}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_2 x\right)\chi_{W_2}\alpha}\right)\right) f_{W_e^2}(x)dx \\ &= \int_0^{\zeta_4}\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_2 x}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_2 x\right)\chi_{W_2}\alpha N}\right) \times \frac{1}{\lambda_{W_e}}\exp\left(-\frac{x}{\lambda_{W_e}}\right)dx \\ &= \frac{1}{\lambda_{W_e}}\int_0^{\zeta_4}\exp\left(-\frac{\chi_{W_e}\alpha_e\tau_2 x}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_2 x\right)\chi_{W_2}\alpha N} - \frac{x}{\lambda_{W_e}}\right)dx.\end{aligned} \tag{B.8}$$

From (B.8), we let $q = x/\zeta_4$. Ω_2^{no} be given by

$$\Omega_2^{no} = \frac{\zeta_4}{\lambda_{W_e}}\int_0^1 \exp\left(-\frac{\chi_{W_e}\alpha_e\tau_2\zeta_4 q}{\left(\tau_2 - \chi_{W_e}\alpha_e\tau_1\tau_2\zeta_4 q\right)\chi_{W_2}\alpha N} - \frac{\zeta_4 q}{\lambda_{W_e}}\right)dq. \tag{B.9}$$

Substituting (B.6) and (B.9) into (15), we can obtain (16).

The proof is completed.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We are greatly thankful to Van Lang University, Vietnam, for providing the budget for this study.

References

[1] M. Munochiveyi, A. C. Pogaku, D. T. Do, A. T. le, M. Voznak, and N. D. Nguyen, "Reconfigurable intelligent surface aided multi-user communications: state-of-the-art techniques and open issues," *IEEE Access*, vol. 9, pp. 118584–118605, 2021.

[2] M. S. Van Nguyen, D. T. Do, S. Al-Rubaye, S. Mumtaz, A. Al-Dulaimi, and O. A. Dobre, "Exploiting impacts of antenna selection and energy harvesting for massive network connectivity," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7587–7602, 2021.

[3] D.-T. Do, A.-T. Le, N.-D. Xuan Ha, and N.-N. Dao, "Physical layer security for internet of things via reconfigurable intelligent surface," *Future Generation Computer Systems*, vol. 126, pp. 330–339, 2022.

[4] M. Jain, S. Soni, N. Sharma, and D. Rawal, "Performance analysis at far and near user in NOMA based system in presence of SIC error," *AEU-International Journal of Electronics and Communications*, vol. 114, article 152993, 2020.

[5] N. D. Nguyen, A. -T. Le, and M. Munochiveyi, "Secrecy outage probability of reconfigurable intelligent surface-aided cooperative underlay cognitive radio network communications," in *2021 22nd Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 73–77, Tainan, Taiwan, 2021.

[6] X. Wu, J. Ma, C. Gu, X. Xue, and X. Zeng, "Robust Secure Transmission Design for IRS-Assisted mmWave Cognitive Radio Networks," *IEEE Transactions on Vehicular Technology*, p. 1, 2022.

[7] N. Hehao and L. Ni, "Intelligent reflect surface aided secure transmission in MIMO channel with SWIPT," *IEEE Access*, vol. 8, pp. 192132–192140, 2020.

[8] W. Jiang, B. Chen, J. Zhao, Z. Xiong, and Z. Ding, "Joint active and passive beamforming design for the IRS-assisted MIMOME-OFDM secure communications," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10369–10381, 2021.

[9] Y. Song, M. R. A. Khandaker, F. Tariq, K. -K. Wong, and A. Toding, "Truly intelligent reflecting surface-aided secure communication using deep learning," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, pp. 1–6, Helsinki, Finland, 2021.

[10] B. Ning, Z. Chen, W. Chen, and L. Li, "Improving security of THz communication with intelligent reflecting surface," in *2019 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, Waikoloa, HI, USA, 2019.

[11] A. U. Makarfi, K. M. Rabie, O. Kaiwartya, O. S. Badarneh, X. Li, and R. Kharel, "Reconfigurable intelligent surface enabled IoT networks in generalized fading channels," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, Dublin, Ireland, 2020.

[12] K. Tekbıyık, G. K. Kurt, and H. Yanikomeroglu, "Energy-Efficient RIS-Assisted Satellites for IoT Networks," *IEEE Internet of Things Journal*, 2021.

[13] P. Mursia, V. Sciancalepore, A. Garcia-Saavedra, L. Cottatellucci, X. C. Perez, and D. Gesbert, "RISMA: reconfigurable intelligent surfaces enabling beamforming for IoT massive access," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1072–1085, 2021.

[14] A. Li, Y. Liu, M. Li, Q. Wu, and J. Zhao, "Joint scheduling design in wireless powered MEC IoT networks aided by reconfigurable intelligent surface," in *2021 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, pp. 159–164, Xiamen, China, 2021.

[15] Z. Chu, Z. Zhu, X. Li, F. Zhou, L. Zhen, and N. Al-Dhahir, "Resource Allocation for IRS Assisted Wireless Powered FDMA IoT Networks," *IEEE Internet of Things Journal*, 2021.

[16] L. Yang and Y. Yuan, "Secrecy outage probability analysis for RIS-assisted NOMA systems," *Electronics Letters*, vol. 56, no. 23, pp. 1254–1256, 2020.

[17] W. Zhao, G. Wang, S. Atapattu, T. A. Tsiftsis, and C. Tellambura, "Is backscatter link stronger than direct link in reconfigurable intelligent surface-assisted system?," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1342–1346, 2020.

[18] L. Yang, J. Yang, W. Xie, M. O. Hasna, T. Tsiftsis, and M. D. Renzo, "Secrecy performance analysis of RIS-aided wireless communication systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12296–12300, 2020.

[19] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[20] X. Yue, Y. Liu, S. Kang, A. Nallanathan, and Z. Ding, "Exploiting full/half-duplex user relaying in NOMA systems," *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 560–575, 2018.

[21] F. Benkhelifa, A. Tall, Z. Rezki, and M. Alouini, "On the low SNR capacity of MIMO fading channels with imperfect channel state information," *IEEE Transactions on Communications*, vol. 62, no. 6, pp. 1921–1930, 2014.

[22] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[23] X. Li, M. Zhao, Y. Liu, L. Li, Z. Ding, and A. Nallanathan, "Secrecy analysis of ambient backscatter NOMA systems under I/Q imbalance," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12286–12290, 2020.

[24] J. Chen, L. Yang, and M. Alouini, "Physical layer security for cooperative NOMA systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4645–4649, 2018.

[25] C. Yu, H. Ko, X. Peng, W. Xie, and P. Zhu, "Jammer-aided secure communications for cooperative NOMA systems," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1935–1939, 2019.

[26] Z. Yang, Z. Ding, Y. Wu, and P. Fan, "Novel relay selection strategies for cooperative NOMA," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10114–10123, 2017.

[27] E. Hildebrand, *Introduction to Numerical Analysis*, Dover, New York, USA, 1987.

[28] X. Liu, "Probability of strictly positive secrecy capacity of the Rician-Rician Fading Channel," *IEEE Wireless Communications Letters*, vol. 2, no. 1, pp. 50–53, 2013.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 5012770, 8 pages
https://doi.org/10.1155/2022/5012770

Research Article

A Cost-Effective Identity-Based Signature Scheme for Vehicular Ad Hoc Network Using Hyperelliptic Curve Cryptography

Asad Iqbal,[1] Insaf Ullah,[2] Abeer Abdulaziz AlSanad,[3] Muhammad Inam Ul Haq,[1] Muhammad Asghar Khan,[2] Wali Ullah Khan,[4] and Khaled Rabie[5]

[1]*Department of Computer Science and Bioinformatics, Kushal Khan Khattak University Karak, 27200, Karak Khyber Pakhtunkhwa, Pakistan*
[2]*Hamdard Institute of Engineering & Technology, Hamdard University, Islamabad 44000, Pakistan*
[3]*Information Systems Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University, Riyadh 11432, Saudi Arabia*
[4]*Interdisciplinary Center for Security, Reliability and Trust (SnT), University of Luxembourg, 1855 Luxembourg City, Luxembourg*
[5]*Department of Engineering, Manchester Metropolitan University, Manchester, UK*

Correspondence should be addressed to Muhammad Asghar Khan; khayyam2302@gmail.com

Received 13 March 2022; Revised 16 April 2022; Accepted 23 April 2022; Published 12 May 2022

Academic Editor: Chi-Hua Chen

A Vehicular Ad Hoc Network (VANET) is a subset of the Mobile Ad Hoc Network (MANET) that allows vehicles to communicate with each other and with roadside stations to offer efficient and safe transportation. Furthermore, when VANET is used in connection with the Internet of Things (IoT) devices and sensors, it can help with traffic management and road safety by allowing vehicles to interact with one another at any time and from any location. Since VANET's event-driven communications are carried out via an open wireless channel, there are significant security concerns. In this paper, we use Hyperelliptic Curve Cryptography (HECC) to offer a cost-effective identity-based signature for secure communication over VANET. The proposed scheme does not need certificate management, and we found that it is more secure against a variety of cryptographic threats after conducting a thorough security analysis. In addition, comparisons of communication and computational costs are made, demonstrating that the proposed scheme is more efficient in both respects than existing schemes.

1. Introduction

Vehicular Ad Hoc Networks (VANETs) have lately received a lot of attention and are now regarded as an important aspect of the automotive sector. VANET is being utilized in the Intelligent Transportation System (ITS) to aid passenger vehicles and infrastructure with issues like road safety, issuing misadventure alerts and assisting drivers, and offering other entertainment services [1]. By integrating Internet of Things (IoT) applications with intelligent transportation mechanisms, VANET creates a secure environment for vehicle communication [2]. The general architecture for VANET is shown in Figure 1, which comprises cars with built-in onboard units (OBUs), Road-Side Units (RSUs), and Trusted Authority (TA). The OBU's job is to connect with surrounding vehicles and RSUs through an open wireless channel, such as the Dedicated Short Range Communication (DSRC) protocol [3]. RSUs are antennas that are placed along the side of the road to collect traffic-related data from automobiles, while TA is a high-performance computing and storage entity in charge of numerous VANET applications including registration and key generation for OBU and RSU [4].

The VANET supports three forms of communication: vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and infrastructure-to-infrastructure (I2I). Open Dedicated Short Range Radio Signals (DSRS) are used for V2V communication, whereas secure channels are used for V2I and I2I communication [5, 6]. Each vehicle in the VANET connects with nearby vehicles and RSUs through OBU,

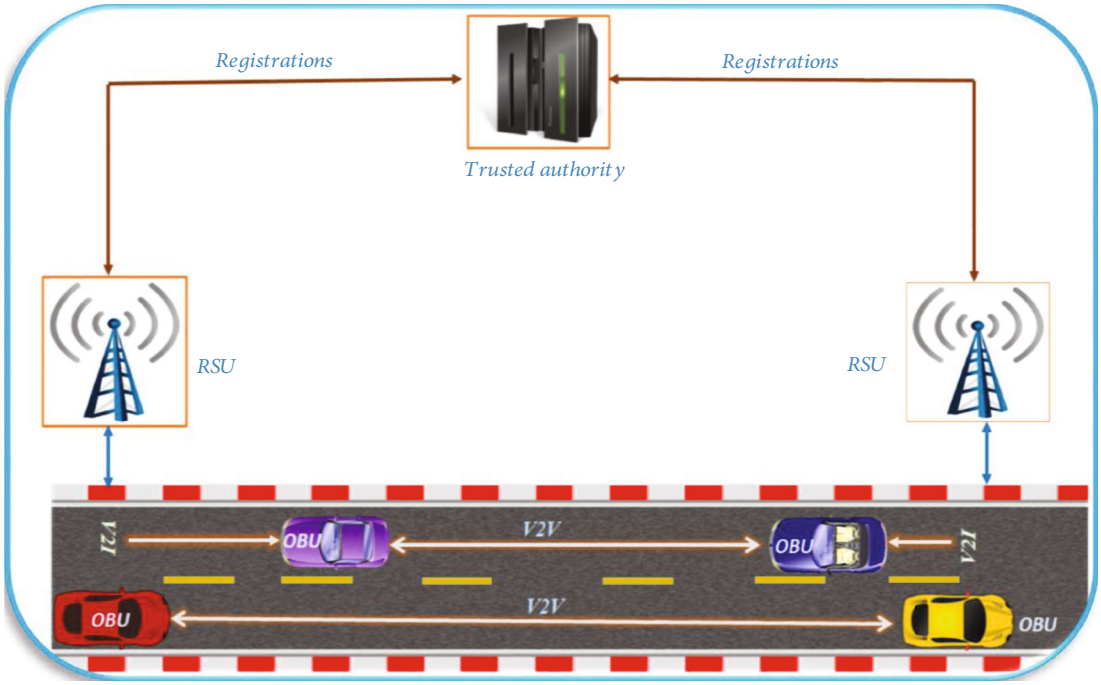


Figure 1: Flow model for VANET.

providing traffic-related information such as position, speed, current time, and traffic and road conditions [5–7].

Because VANET communication could take place over an open wireless channel, authentication is a major concern [8]. Digital signatures will be the most appropriate solution for dealing with this type of situation. It allows a VANET sender node to produce a signature on a dispatchable document using his private key and then transfer it over to the receiver node. The recipient node validates the signature using the sender's public key after getting it. Though digital signatures are based on asymmetric key cryptography, the first candidate is Public Key Infrastructure (PKI), in which the Certificate Authority (CA) presents the user with a certificate. The main disadvantage of PKI is certificate management. PKI is being phased out in favour of identity-based cryptography, which does not require certificate management. In this cryptosystem, users just provide their identities to TA, which subsequently produces the public and private keys for that identity and sends them through a secure channel.

Rivest-Shamir-Adleman (RSA), bilinear pairing (BP), and Elliptic Curve Cryptography (ECC) are commonly employed to achieve security and efficiency in security schemes. These algorithms are frequently based on computationally difficult problems. With a key size of up to 1024 bits, RSA cryptography employs enormous factorization [9]. Due to huge pairing and map-to-point function calculation, BP is 14.31 times worse than RSA. ECC, a modern cryptography method, was utilized to address the difficulties in RSA and BP with a key size up to 160 bits, reducing the computationally difficult problem to some extent, but it is still not supported by resource-constrained devices. A new cryptographic system called Hyperelliptic Curve Cryptography (HECC) was created for this purpose, and it provides the same level of security as EC [10]. While giving the security features of RSA, BP, and ECC, the HECC employs an 80-bit key size. The HECC is an excellent starting point for a VANET system.

1.1. Preliminaries. The HECC can be defined as the following: it is a generalized form of elliptic curves and state $W(Tq)$ over finite field F_q defined by equation $W : b^2 + h(a)b = t(a) \bmod q$, where $h(a) \in T[a]$ is a polynomial and degree $h(a) \leq g$ and $t(a) \in T[a]$ is a monic polynomial and degree $t(a) \leq 2g + 1$. further, it includes devisor which is a finite formal sum of points, and according to Mumford, it can be represented as $S = (x(a), y(a)) = (\sum_{i=0}^{g} x_i a^i, \sum_{i=0}^{g-1} y_i a^i)$. The divisors form an Abelian group which is called Jacobian group $J_c(T_q)$, and the order of the Jacobian group $o(J_c(T_q))$ is defined as $|(\sqrt{q} - 1)^{2g}| \leq o(J_c(T_q)) \leq |(\sqrt{q} + 1)^{2g}|$. So, the whole security of the hyperelliptic curve cryptosystem is based on the hyperelliptic curve discrete logarithm problem, which can be defined as the following: Let S be a divisor of order n in the Jacobian group $J_c(T_q)$, find an integer $a \in T_q$, such that $S_1 = a \cdot S$.

1.2. Motivation and Contributions. So, inspired by the idea of HECC, we make the following contribution to this work as a result of the preceding discussion:

(1) We propose a batch verification method based on HECC using authentication and key management mechanism

(2) We carried out a thorough security analysis and confirmed that the proposed scheme is resistant to a variety of cyberattacks

(3) By comparing the proposed scheme to a previously published scheme, we performed a cost analysis study in terms of both communication and computation, and the findings show that the proposed scheme is efficient

2. Literature Review

VANET is a network that allows vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication (V-I). VANET communication, on the other hand, uses the Internet, necessitating the need for authentication procedures to avoid rogue nodes. One of the better options is to employ digital signatures, which allow a sender to create a digital signature on data created in a VANET environment using his private key and then deliver it to the receiver. Using the sender's public key, the receiver may easily verify the signature after receiving it [11]. As a result, various academics have developed digital signature systems for traffic-related communications in VANET authentication.

Raya and Hubaux [12] proposed a Public Key Infrastructure- (PKI-) based authentication technique in which the Certificate Authority (CA) generates a large number of anonymous private/public keys and certificates in a short period of time to sign traffic-related communications. However, because of the limited storage capacity of the vehicle's OBU, it is not ideal for storing a pair of large numbers of public/private keys and certificates. On the basis of anonymous certificates, Lu et al. [13] enhanced the system used in [12] and contributed a new Conditional Privacy Preservation Authentication (CPPA) scheme.

This scheme is not appropriate for real-time communication systems due to certificate renewal issues. Freudiger et al. [14] offered another authentication approach in which they integrated the mix zone and anonymous certificate methods. However, when a high number of certificates are required, it has an impact on RSU's storage capacity. Zhang et al. [15] proposed an efficient authentication technique that included the use of Hash Message Authentication Codes (HMAC) to ensure privacy preservation. For connecting with RSU, a random public/private key pair and certificate were assigned in this scheme. However, cars are still obliged to hold the maximum amount of certificates under this approach, which solves the storage problem. Wasef and Shen [16] suggested another PKI-based approach, the Expedite Message Authentication Protocol (EMAP). By replacing the Certificate Revocation List (CRL) with keyed HMAC, they were able to speed up the revocation checking procedure.

By employing bilinear pairing, Zhang et al. [17] proposed a CPPA signature technique based on identity for VANETs. By combining the feature of group signature with batch verification and bilinear pairing, Chim et al. [18] created an identity-based technique. For V2V communication, Shim [19] presented an identity-based CPPA signature technique based on bilinear pairing. Horng et al. [20] created an ID-based signature technique for VANETs that uses bilinear pairing and also supports batch signatures. However, the techniques in [17–21] may have an impact on real-time communication since they are based on bilinear pairing, which necessitates higher channel capacity and processing resources. Sun et al. [21] designed the CPPA signature, by utilizing bilinear pairing. However, this scheme can affect real-time communication as it is based on bilinear pairing that must need greater capacity in the channel and more computational power. He et al. [22] proposed a new ID-based CPPA signature system for both V2V and V2I communication in VANET with the use of ECC. This scheme's results showed that it was successful in facilitating batch signature verification and assessing VANET in high-traffic locations. However, throughout the three-point multiplication operation, there was a delay in confirming signatures. Using ECC, Ikram et al. presented an ID-based signature technique for V2V communication on VANETs. Their technique, however, still has a significant computational overhead.

3. Network Model

Figure 2 depicts the network model for the proposed method, which includes three entities: onboard units (OBUs), roadside units (RSUs), and the Department of Transportation (DoT). The steps to take are as follows:

(1) *OBU*. It has 5G technology and can connect with other OBUs, as well as DoT and RSU. Its duty also includes registering with DoT by transmitting his identification; after DoT receives his identity, DoT generates the public and private keys for his identity and delivers them to the OBU. Then, utilizing an open network, OBU may build data signatures and transfer them to RSU.

(2) *RSU*. It is a 5G-enabled base station in charge of V-I communication management and execution. Its duty also includes registering with DoT by transmitting his identification; after DoT receives his identity, DoT generates the public and private keys for his identity and sends it back to the RSU. Furthermore, when RSU receives signed data from OBU, it performs a verification procedure; if the signature is acceptable, the message is accepted; otherwise, an error message is generated.

(3) *DoT*. The DoT is a trustworthy third party with significant computational and storage power. It establishes system parameters and makes them publicly available to other organizations. When DoT receives OBU and RSU's identities, it generates public and private keys and sends them to OBU and RSU separately.

4. Proposed Scheme

The proposed batch verification identity-based signature can be executed through the steps that are explained below. Before we start the proposed algorithm, in Table 1, the symbols used during its constructions are explained.

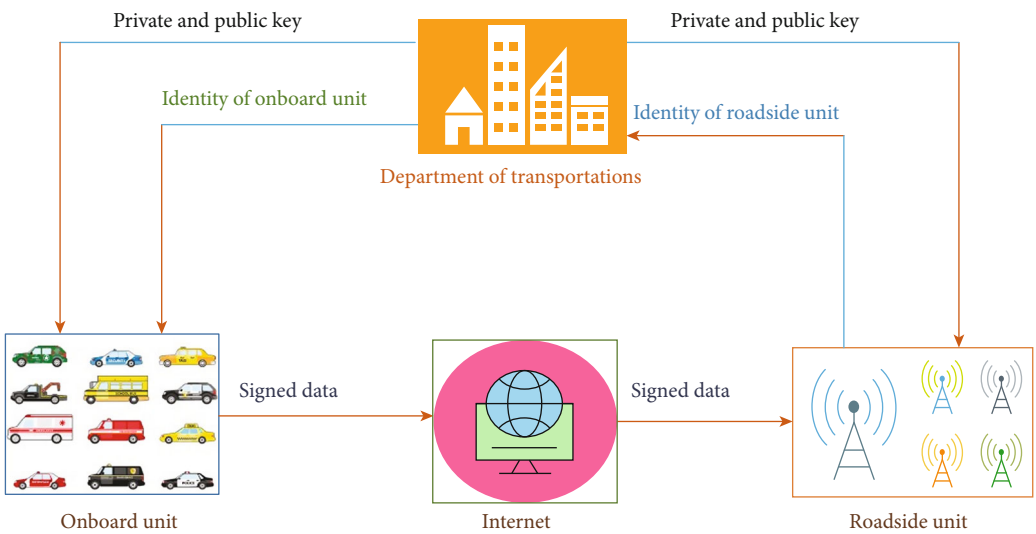


Figure 2: Network model for proposed scheme.

Table 1: Symbols used in the proposed algorithm.

Symbol	Description
λ	It is used to indicate security parameter
H	Denotes the hash function like SHA 256 which avoids the reversed manner
χ	Denotes the private key of TA
ID_{V}	Denotes the identity for each vehicle in the system
Ω_{rv}	The private key for receiver vehicle
ζ_{sv}	The public key for sender vehicle
$\mathrm{ID}_{\mathrm{rV}}$	The identity for receiver vehicle
δ	Indicates the public key of TA
$\mathbf{F}_n$	Denotes the nonfinite field of hyperelliptic curve
$\mathscr{D}$	Denotes devisor of hyperelliptic curve
Ω_{sv}	The private key for sender vehicle
$\mathrm{ID}_{\mathrm{sV}}$	The identity for sender vehicle
ζ_{rv}	The public key for receiver vehicle
M	Indicates the plaintext

Setup: this algorithm is processed by Trusted Authority (TA) when it receives the security parameter $\lambda = 80$ bits in size; further, it selects χ as his private key and executes his public key as $\delta = \chi \cdot \mathscr{D}$. Moreover, TA published the set $\gamma = \{\delta, \mathscr{D}, \mathbf{F}_n, H\}$, where $\mathbf{F}_n$ and H denotes the nonfinite field of the hyperelliptic curve and hash function like SHA 256 which avoids the reversed manner.

Key generation: for a vehicle with identity (ID_{V}), TA compute the public and private keys as follows:

(i) Compute $\zeta_{\mathrm{v}} = \alpha_{\mathrm{v}} \cdot \mathscr{D}$, where α_{v} is the secret number from $\mathbf{F}_n$

(ii) Compute $\xi = H(\zeta_{\mathrm{v}}, \mathrm{ID}_{\mathrm{V}})$ and $\beta = \xi \cdot \delta$

(iii) Compute $\Omega_{\mathrm{v}} = \alpha_{\mathrm{v}} + \xi \cdot \chi$ and send $\zeta_{\mathrm{v}}, \Omega_{\mathrm{v}}$, and β as a public key, private key, and public number to ID_{V}

Signature generation: a sender vehicle with identity ($\mathrm{ID}_{\mathrm{sV}}$) can sign the received data from OBU as follows.

(i) Compute $\ell = \upsilon \cdot \mathscr{D}$, where υ is the secret number from $\mathbf{F}_n$

(ii) Compute $\mathfrak{S} = H(M, \ell, \mathrm{ID}_{\mathrm{sV}})$ and $\mathscr{J} = \upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{sv}}$

(iii) Set $(\mathscr{J}, \ell)$ as a signature pair and send it to the receiver vehicle

Signature verifications: a receiver vehicle with identity ($\mathrm{ID}_{\mathrm{rV}}$) can verify the received signature pair $(\mathscr{J}, \ell)$ as

Table 2: Computational cost comparisons with respect to major operations and ms.

Schemes	Signing cost	Verification cost	Total cost (in ms)
Ali et al. [25]	3 ERM	1 ERM	$4 * 0.97 = 3.88$
Lo and Tsai [26]	2 ERM	2 ERM	$4 * 0.97 = 3.88$
He et al. [22]	4 ERM	3 ERM	$7 * 0.97 = 6.79$
Wang and Yao [27]	1 BPRM	3 BP	$1 * 4.31 + 3 * 14.90 = 49.01$
Bayat et al. [28]	5 BPRM	3 BP	$5 * 4.31 + 3 * 14.90 = 66.25$
Jianhong et al. [29]	5 BPRM	3 BP	$5 * 4.31 + 3 * 14.90 = 66.25$
Proposed	2HERM	2HERM	$4 * 0.48 = 1.92$

follows: it computes $\mathfrak{S} = H(M, \ell, \mathrm{ID}_{\mathrm{sV}})$ and accepts $(\mathcal{J}, \ell)$ when $\mathcal{J} \cdot \mathcal{D} = \ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \beta)$ are satisfied.

Batch signature generation: a sender vehicle with identity $(\mathrm{ID}_{\mathrm{sV}})$ can sign the batch data of OBU as follows.

(i) Compute $\mathcal{G} = \sum_{i=0}^{n} \ell$ and $\mathcal{C} = \sum_{i=0}^{n} \mathcal{J}$

(ii) Set $(\mathcal{C}, \mathcal{G})$ as a batch signature pair and send it to the receiver vehicle

Batch signature verifications: a receiver vehicle with identity $(\mathrm{ID}_{\mathrm{rV}})$ can verify the received batch signature pair $(\mathcal{C}, \mathcal{G})$ as follows: it computes $\mathfrak{S} = H(M, \ell, \mathrm{ID}_{\mathrm{sV}})$ and accepts $(\mathcal{C}, \mathcal{G})$ when $\sum_{i=0}^{n} \mathcal{J} \cdot \mathcal{D} = \sum_{i=0}^{n} \ell + \mathfrak{S} \cdot \sum_{i=0}^{n} (\zeta_{\mathrm{sv}} + \beta)$ are satisfied.

4.1. Correctness. The signature verification can be done as is $\mathcal{J} \cdot \mathcal{D} = \ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \beta)$ satisfied $\mathcal{J} \cdot \mathcal{D} = (\upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{sv}}) \cdot \mathcal{D} = (\upsilon \cdot \mathcal{D} + \mathfrak{S} \cdot \Omega_{\mathrm{sv}} \cdot \mathcal{D}) = (\upsilon . \mathcal{D} + \mathfrak{S} \cdot (\alpha_{\mathrm{sv}} + \xi \cdot \chi) \cdot \mathcal{D}) = (\ell + \mathfrak{S} \cdot ((\alpha_{\mathrm{sv}} \cdot \mathcal{D} + \xi \cdot \chi \cdot \mathcal{D})) = (\ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \xi \cdot \chi \cdot \mathcal{D})) = (\ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \xi \cdot \delta)) = (\ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \beta)) = \mathcal{J} \cdot \mathcal{D}$ hence proved.

Also, the batch signature verification can be done as is $\sum_{i=0}^{n} \mathcal{J} \cdot \mathcal{D} = \sum_{i=0}^{n} \ell + \mathfrak{S} \cdot \sum_{i=0}^{n} (\zeta_{\mathrm{sv}} + \beta)$ are satisfied.

$\sum_{i=0}^{n} \mathcal{J} \cdot \mathcal{D} = \sum_{i=0}^{n} (\upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{sv}}) \cdot \mathcal{D} = \sum_{i=0}^{n} (\upsilon \cdot \mathcal{D} + \mathfrak{S} \cdot \Omega_{\mathrm{sv}} \cdot \mathcal{D}) = \sum_{i=0}^{n} (\upsilon \cdot \mathcal{D} + \mathfrak{S} \cdot (\alpha_{\mathrm{sv}} + \xi \cdot \chi) \cdot \mathcal{D}) = \sum_{i=0}^{n} (\ell + \mathfrak{S} \cdot ((\alpha_{\mathrm{sv}} \cdot \mathcal{D} + \xi \cdot \chi \cdot \mathcal{D})) = \sum_{i=0}^{n} (\ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \xi \cdot \chi \cdot \mathcal{D})) = \sum_{i=0}^{n} \ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \xi \cdot \delta)) = \sum_{i=0}^{n} (\ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \beta)) = \sum_{i=0}^{n} \sum_{i=0}^{n} \mathcal{J} \cdot \mathcal{D}$ hence proved.

5. Security Analysis

Before going to discuss the security properties, we must discuss some properties of an attacker that can be a threat to our proposed scheme. Here, we consider the Dolev-Yao model, in which the attacker can perform interception and impersonation, break the privacy of identity, break the process of mutual authentication, and can generate a forged signature, respectively. In the following subphases, we have proved that our designed approach can resist various cyberattacks.

5.1. Authentication. For the authentication, the sender computes $\mathcal{J} = \upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{sv}}$ and sends it to the receiver. After the reception of $\mathcal{J}$, the receiver computes $\mathfrak{S} = H(M, \ell, \mathrm{ID}_{\mathrm{sV}})$ and accepts $(\mathcal{J}, \ell)$ when $\mathcal{J} \cdot \mathcal{D} = \ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \beta)$ are satisfied, so that our scheme meets the authentication security service in this way.

5.2. Identity Privacy Preservation. The identity privacy preservation can be done in the proposed scheme in a way that we are not sending any user identity in an open channel during communication between two devices in VANET. We only send $(\mathcal{J}, \ell)$ in an open network, where $\mathcal{J} = \upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{sv}}$ and $\ell = \upsilon \cdot \mathcal{D}$, so it is obvious that it does not contain any user identity.

5.3. Nonrepudiation. A vehicle or RSU in VANET should not be able to refuse any traffic-related message sent by it, because it used its private key during signature generation $(\mathcal{J} = \upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{sv}})$ which is directly associated with its public key. So, the receiver or TA can verify the received signature by using $\mathcal{J} \cdot \mathcal{D} = \ell + \mathfrak{S} \cdot (\zeta_{\mathrm{sv}} + \beta)$; if this equation holds, then TA can decide the message from the sender.

5.4. Traceability. If a malicious vehicle transmits a false traffic-related message, only the TA can trace the vehicle's original identity. In our proposed scheme, let us say if the malicious vehicle with identity $(\mathrm{ID}_{\mathrm{mV}})$ can generate a signature on a false as the following: it computes $\ell = \upsilon \cdot \ ?$, $\mathfrak{S} = H(M, \ell, \mathrm{ID}_{\mathrm{mV}})$, $\mathcal{J} = \upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{mv}}$, and sends a tuple $(\mathcal{J}, \ell)$ as a signature pair to the receiver vehicle. A receiver vehicle with identity $(\mathrm{ID}_{\mathrm{rV}})$ can verify the received signature pair $(\mathcal{J}, \ell)$ as follows: it computes $\mathfrak{S} = H(M, \ell, \mathrm{ID}_{\mathrm{mV}})$ and accepts $(\mathcal{J}, \ell)$ when $\mathcal{J} \cdot \mathcal{D} = \ell + \mathfrak{S} \cdot (\zeta_{\mathrm{mv}} + \beta)$ are satisfied. So, if the receiver found that the message signature which was sent by the malicious sender is harmful, then it reports this vehicle identity to the TA, and TA backlists this identity for the future.

5.5. Impersonation Attack. The proposed mechanism avoids this attack because it transmits only two parameters $(\mathcal{J}, \ell)$, which will be the obligatory need for the attacker to impersonate the signature, where ℓ is the public number which can easily be accessible for the attacker and $= \upsilon + \mathfrak{S} \cdot \Omega_{\mathrm{mv}}$, so for this, the attacker needs υ from $\ell = \upsilon \cdot \mathcal{D}$ which cannot be feasible because of the hard nature of the hyperelliptic curve discrete logarithm problem. Further, the attacker desires to make $\Omega_{\mathrm{v}} = \alpha_{\mathrm{v}} + \xi \cdot \chi$, which further requires α_{v} from $\zeta_{\mathrm{v}} = \alpha_{\mathrm{v}} \cdot \mathcal{D}$ and χ from $\delta = \chi \cdot \mathcal{D}$, which cannot be feasible because of the two-time hard nature of the hyperelliptic curve discrete logarithm problem.

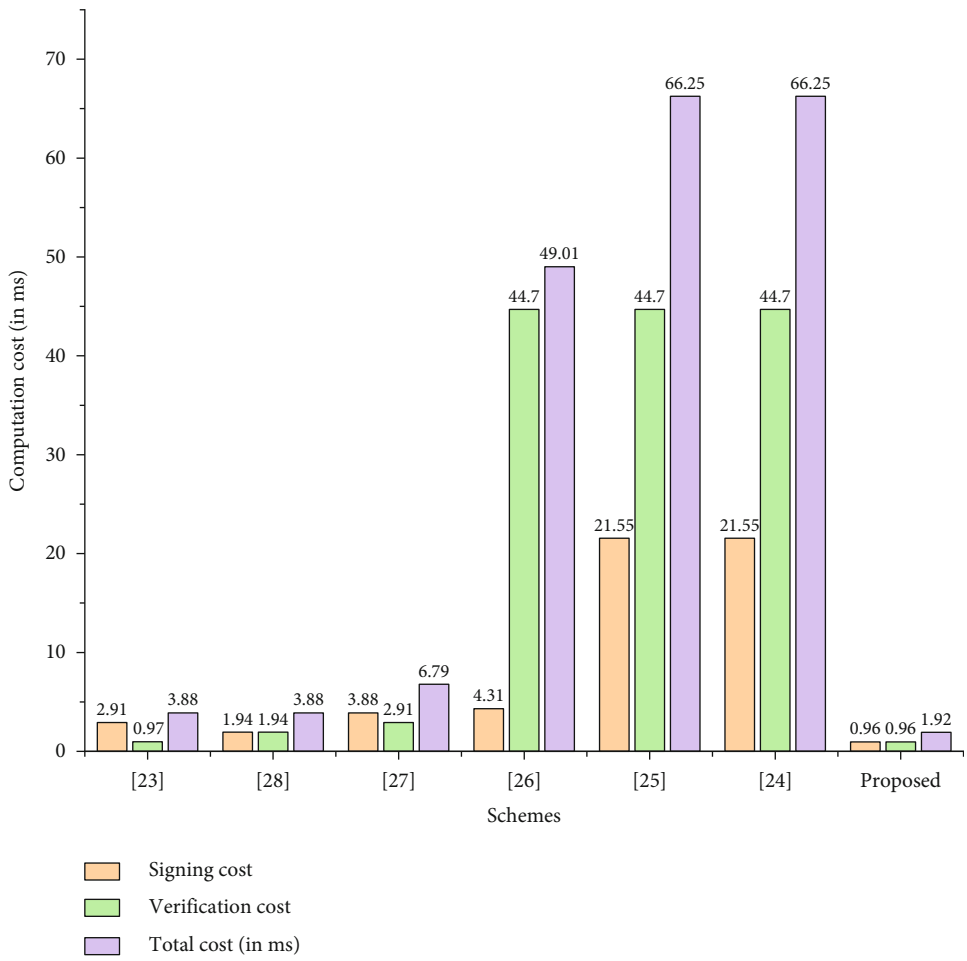


Figure 3: Comparison of computational cost.

5.6. Modification Attack. The attacker cannot modify the signature tuple; it needs υ from $\ell = \upsilon \cdot \mathscr{D}$ which cannot be feasible because of the hard nature of the hyperelliptic curve discrete logarithm problem. Further, the attacker desires to make $\Omega_v = \alpha_v + \xi \cdot \chi$, which further requires α_v from $\zeta_v = \alpha_v \cdot \mathscr{D}$ and χ from $\delta = \chi \cdot \mathscr{D}$, which cannot be feasible because of the two-time hard nature of the hyperelliptic curve discrete logarithm problem.

6. Computational Cost

In this section, we compare our proposed scheme with existing schemes in terms of computational cost. Typically, the computational cost involves heavy mathematical operations. Our scheme has been compared with three bilinear pairing (BP) schemes as well as three elliptic curves (EC), which involve heavy computation. We measure computational cost in milliseconds (ms) for comparison. The single Elliptic Curve Point Multiplication (ERM) needs 0.97 ms, Bilinear Pairing Point Multiplication (BPRM) takes 4.31 ms, and BP take 14.90 ms [23, 24]. We utilize hyperelliptic curve divisor multiplications (HERM) [10] that take 0.48 ms to process and use a half amount of key, i.e., 80 bits as compared to EC, which provides the same level of security. According to Table 2 and Figure 3, in which we have provided the comparisons of the proposed and existing schemes with the help of major operations and milliseconds, our scheme is more efficient than existing schemes.

Table 3: Communication cost analysis between proposed scheme and existing schemes on the basis of bits.

Schemes	Communication cost	Communication cost with bits
Ali et al. [25]	$\|m\| + 2\|q\|$	$850 + 2 * 160 = 1170$ bits
Lo and Tsai [26]	$\|m\| + 3\|q\|$	$850 + 3 * 160 = 1330$ bits
He et al. [22]	$\|m\| + 4\|q\|$	$850 + 4 * 160 = 1490$ bits
Wang and Yao [27]	$\|m\| + 3\|G\|$	$850 + 3 * 1024 = 3922$ bits
Bayat et al. [28]	$\|m\| + 3\|G\|$	$850 + 3 * 1024 = 3922$ bits
Jianhong et al. [29]	$\|m\| + 3\|G\|$	$850 + 3 * 1024 = 3922$ bits
Proposed	$\|m\| + 2\|n\|$	$850 + 2 * 80 = 1010$ bits

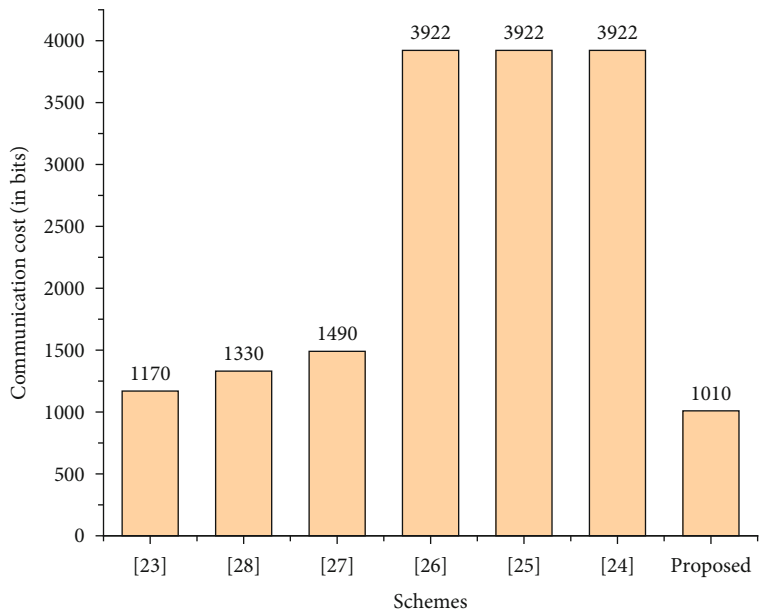


Figure 4: Comparison of communication cost.

The observation is produced from a workstation having the following specification.

(i) Intel Core i5-6300 CPU

(ii) 2.40 GHz processor

(iii) 8 GB of RAM

(iv) Windows 10 Ultimate edition

7. Communication Overhead

In this section, our proposed scheme has been compared with Ali et al. [25], Lo and Tsai [26], He et al. [22], Wang and Yao [27], Bayat et al. [28], and Jianhong et al. [29] in terms of communication overhead. For this purpose, we consider $|m|$ as the plaintext, and its size is supposed to be equal to 850 bits; $|G|$ for bilinear pairing, where its size in bits is 1024; $|q|$ for elliptic curve where its size in bits is 160; and $|n|$ for hyperelliptic curve where its size in bits is 80, respectively. Therefore, it is clear from Table 3 and Figure 4 that our scheme is superior in communicational overhead to the schemes proposed in [22, 25–29].

8. Conclusion

In this paper, we proposed a cost-effective identity-based signature for the deployment of VANET using Hyperelliptic Curve Cryptography (HECC) to lower the computational cost of verifying vehicles during message authentication. The proposed scheme supports a batch signature verification approach, which allows each vehicle in a high-traffic area to validate multiple messages at the same time. Authentication, identity privacy preservation, nonrepudiation, traceability, impersonation attack, and modification attack are all security criteria that the proposed technique meets. Our research demonstrates that the proposed scheme will be a preferable choice for VANET in terms of computational and communicational cost when compared to current similar techniques.

Data Availability

All the data is presented in this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the present study.

References

[1] J. Mahmood, Z. Duan, Y. Yang, Q. Wang, J. Nebhen, and M. N. M. Bhutta, "Security in vehicular ad hoc networks: challenges and countermeasures," *Security and Communication Networks*, vol. 2021, 20 pages, 2021.

[2] S. M. Hatim, S. J. Elias, N. Awang, and M. Y. Darus, "VANETs and Internet of Things (IoT): a discussion," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 12, no. 1, pp. 218–224, 2018.

[3] B. P. Roberts and C. Sandberg, "The role of energy storage in development of smart grids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1139–1144, 2011.

[4] M. N. Mejri, J. Ben-Othman, and M. Hamdi, "Survey on VANET security challenges and possible cryptographic solutions," *Vehicular Communications*, vol. 1, no. 2, pp. 53–66, 2014.

[5] J. Cui, J. Zhang, H. Zhong, R. Shi, and Y. Xu, "An efficient certificateless aggregate signature without pairings for vehicular ad hoc networks," *Information Sciences*, vol. 451-452, pp. 1–15, 2018.

[6] W. Khan, T. N. Nguyen, F. Jameel et al., *Learning-Based Resource Allocation for Backscatter-Aided Vehicular Networks*, TechRxiv. Preprint, 2021.

[7] W. U. Khan, A. Ihsan, T. N. Nguyen, M. A. Javed, and Z. Ali, "NOMA-enabled backscatter communications for green transportation in automotive-Industry 5.0," *IEEE Transactions on Industrial Informatics*, 2022.

[8] Q. Mei, H. Xiong, J. Chen, M. Yang, S. Kumari, and M. K. Khan, "Efficient certificateless aggregate signature with conditional privacy preservation in IoV," *IEEE Systems Journal*, vol. 15, no. 1, pp. 245–256, 2021.

[9] M. A. Khan, I. Ullah, S. Nisar et al., "An efficient and provably secure certificateless key-encapsulated signcryption scheme for flying ad-hoc network," *IEEE Access*, vol. 8, pp. 36807–36828, 2020.

[10] I. Ullah, M. A. Khan, M. H. Alsharif, and R. Nordin, "An anonymous certificateless signcryption scheme for secure and efficient deployment of internet of vehicles," *Sustainability*, vol. 13, no. 19, p. 10891, 2021.

[11] S. F. Tzeng, C. Y. Yang, and M. S. Hwang, "A new digital signature scheme based on factoring and discrete logarithms," *International Journal of Computer Mathematics*, vol. 81, no. 1, pp. 9–14, 2004.

[12] M. Raya and J. P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[13] R. Lu, X. Lin, H. Zhu, P. H. Ho, and X. Shen, "ECPP: efficient conditional privacy preservation protocol for secure vehicular communications," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, pp. 1229–1237, 2008.

[14] J. Freudiger, M. Raya, M. Felegyhazi, P. Papadimitratos, and J. P. Hubaux, "Mix-zones for location privacy in vehicular networks," in *Proc. 1st Int. Workshop on Wireless Netw. Intell. Transp. Syst.*, Vancouver, BC, Canada, 2007.

[15] C. Zhang, X. Lin, R. Lu, and P. H. Ho, "RAISE: an efficient RSU-aided message authentication scheme in vehicular communication networks," in *2008 IEEE international conference on communications*, pp. 1451–1457, 2008.

[16] A. Wasef and X. Shen, "EMAP: expedite message authentication protocol for vehicular ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 1, pp. 78–89, 2013.

[17] C. Zhang, P. H. Ho, and J. Tapolcai, "On batch verification with group testing for vehicular communications," *Wireless Networks*, vol. 17, no. 8, pp. 1851–1865, 2011.

[18] T. W. Chim, S. M. Yiu, L. C. Hui, and V. O. Li, "SPECS: secure and privacy enhancing communications schemes for VANETs," *Ad Hoc Networks*, vol. 9, no. 2, pp. 189–203, 2011.

[19] K. A. Shim, "CPAS: an efficient conditional privacy-preserving authentication scheme for vehicular sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 4, pp. 1874–1883, 2012.

[20] S. J. Horng, S. F. Tzeng, Y. Pan et al., "b-SPECS+: batch verification for secure pseudonymous authentication in VANET," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 11, pp. 1860–1875, 2013.

[21] J. Sun, C. Zhang, Y. Zhang, and Y. Fang, "An identity-based security system for user privacy in vehicular ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 9, pp. 1227–1239, 2010.

[22] D. He, S. Zeadally, B. Xu, and X. Huang, "An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2681–2691, 2015.

[23] M. A. Khan, H. Shah, S. U. Rehman et al., "Securing internet of drones with identity-based proxy signcryption," *IEEE Access*, vol. 9, pp. 89133–89142, 2021.

[24] I. Ullah, S. Zeadally, N. U. Amin, M. Asghar Khan, and H. Khattak, "Lightweight and provable secure cross-domain access control scheme for internet of things (IoT) based wireless body area networks (WBAN)," *Microprocessors and Microsystems*, vol. 81, p. 103477, 2021.

[25] I. Ali, T. Lawrence, and F. Li, "An efficient identity-based signature scheme without bilinear pairing for vehicle-to-vehicle communication in VANETs," *Journal of Systems Architecture*, vol. 103, article 101692, 2020.

[26] N. W. Lo and J. L. Tsai, "An efficient conditional privacy-preserving authentication scheme for vehicular sensor networks without pairings," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1319–1328, 2016.

[27] S. Wang and N. Yao, "LIAP: a local identity-based anonymous message authentication protocol in VANETs," *Computer Communications*, vol. 112, pp. 154–164, 2017.

[28] M. Bayat, M. Barmshoory, M. Rahimi, and M. R. Aref, "A secure authentication scheme for VANETs with batch verification," *Wireless Networks*, vol. 21, no. 5, pp. 1733–1743, 2015.

[29] Z. Jianhong, X. Min, and L. Liying, "On the security of a secure batch verification with group testing for VANET," *International Journal of Network Security*, vol. 16, no. 5, pp. 351–358, 2014.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7732029, 13 pages
https://doi.org/10.1155/2022/7732029

Research Article

A Deep Learning-Based User Selection Scheme for Cooperative NOMA System with Imperfect CSI

Zhiyuan Yin,[1] **Jin Chen**,[1] **Guoxin Li**,[1] **Haichao Wang**,[1] **Wenhui He**,[1] **and Yan Ni**[2]

[1]*College of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China*
[2]*Department of General Education, Army Engineering University of PLA, Nanjing 210007, China*

Correspondence should be addressed to Guoxin Li; guoxin@aeu.edu.cn

Received 12 February 2022; Revised 9 March 2022; Accepted 16 March 2022; Published 11 May 2022

Academic Editor: Xingwang Li

This paper studies the user selection problem for a cooperative nonorthogonal multiple access (NOMA) system consisting of a base station, a far user, and N near users. The selected near user receives its own message and assists the far user by relaying the far user's message. Firstly, we propose a user selection strategy to maximize the selected near user's data rate while satisfying the quality-of-service (QoS) requirement of the far user. Considering that the channel state information (CSI) of users in actual communication is usually imperfect, we then analyze the outage probability of the NOMA system based on the user selection strategy under imperfect CSI and obtain a closed-form expression. The theoretical analysis shows that the diversity order of the NOMA system under imperfect CSI is 0, which means the multiuser diversity order disappears. In order to improve the impact of imperfect CSI on system performance, we use the deep learning method to identify and classify channels of imperfect CSI and improve the accuracy of CSI. The simulation results show that the theoretical analysis of outage performance is consistent with the numerical results. Compared with the strategy without the deep learning method, the proposed deep learning-based user selection scheme significantly improves the system performance. Furthermore, we verify that our scheme recovers the diversity gain.

1. Introduction

With the development of communication technology, spectrum resources are increasingly scarce. To make efficient use of the limited spectrum, nonorthogonal multiple access (NOMA) technology was proposed [1]. Existing research [2–5] shows that NOMA allows multiple users to communicate simultaneously using a single channel resource, thus achieving more significant spectral efficiency than traditional orthogonal multiple access (OMA). However, there is mutual interference between users in the NOMA system. The most common way to solve this problem is to eliminate interference with successive interference cancellation (SIC) [6, 7], which gradually reduces the interference of the maximum power user in the received signal. Consequently, some NOMA users will know other users' messages; thus, they can serve as relays to help other users. In this sense, the idea of cooperative NOMA is proposed. A classical strategy was designed for cooperative NOMA in [8], where users successfully decoding other users' messages helped other users in turn. Note that existing research related to cooperative NOMA networks [9–12] mainly considered improving system performance under perfect channel state information (CSI). However, it is actually difficult to obtain perfect CSI. In NOMA networks, imperfect CSI not only causes additional interference to the expected signal but also leads to the wrong decoding sequence [13–15]. Therefore, it is particularly necessary to minimize the channel estimation error as much as possible.

With the development of artificial intelligence, deep learning (DL) technology [16] has been widely used [17–19]. In order to obtain perfect CSI, some works used

DL technology for channel estimation and achieved good performance. In [20], a channel estimation network based on DL was proposed in the high-speed mobile scenarios, reducing computational complexity and improving performance. In [21], a channel learning scheme based on the deep autoencoder was developed, which learned CSI at the energy transmitter based on the energy feedback harvested by the energy receiver. In [22], based on the fact that the propagation environment is almost identical, a DL-based CSI estimation for high mobility networks was proposed, which allowed the deep neural network to learn the nonlinear CSI relations. In [23], a five-layer deep neural network (DNN) was designed to estimate channels in orthogonal frequency-division multiplexing (OFDM) systems. In [24], the authors regarded CSI as 2D images and used DL-based image processing techniques to estimate the channel. In [25], a channel estimator using the sliding bidirectional gated recurrent unit network was designed at the receiver, which can be combined with other channel estimation techniques. In [26], a DNN was constructed for channel estimation and direction of arrival estimation, improving performance without increasing complexity. In [27], a deep image prior based DNN was proposed to improve estimation performance without training. In [28], in order to reduce the overhead in multi-input multi-output (MIMO) systems, a convolutional neural network- (CNN-) based estimator was proposed. In [29], a learned denoising-based approximate message passing (LDAMP) network was exploited for beamspace channel estimation.

Motivated by the successful research of DL in channel estimation, we introduce DL technology into cooperative NOMA networks to solve the imperfect CSI estimation and user selection problems. It is worth mentioning that the above DL schemes for channel estimation are not applicable in our NOMA system even if they have good performance on their problems. First, each transmission block is equally divided into two phases in this paper, and CSI of each hop is irrelevant. To improve learning performance, it is an effective approach to learn the CSI of each hop, respectively. Second, a key step in NOMA systems is SIC, which needs a special network design. Details of network design are shown in Section 4. The main contributions of this paper are summarized as follows.

(i) In our work, the optimal power allocation and user selection in the proposed downlink cooperative NOMA system are studied. Based on the proposed user selection strategy, the system outage probability and diversity order under imperfect CSI are analyzed. The theoretical analysis shows that the diversity order of the NOMA system under imperfect CSI is 0, which means the multiuser diversity order disappears

(ii) A channel estimation network based on CNN and long short-term memory (LSTM) is designed for our NOMA scenario, which includes offline training and online prediction. Firstly, in order to minimize the complexity of training and improve learning performance, the input data is pretrained to extract feature vectors by a three-layer one-dimensional CNN. Secondly, the learning network with two LSTM layers in parallel is built, which learns CSI and reduces the estimation error to make the appropriate user selection

(iii) The estimated performance of the proposed scheme is verified. Firstly, the LSTM network shows high learning accuracy, which largely satisfies the reliability of the mapping between user selection and input data. Secondly, the proposed scheme successfully improves the data rate of near users with imperfect CSI. Compared with the other strategy without the DL method, our proposed scheme has obvious advantages. Meanwhile, the performance of the proposed scheme is similar to the strategy with perfect CSI. Thirdly, the proposed scheme improves the outage performance and recovers the diversity gain while the diversity gain under imperfect CSI is zero without the DL method

The remainder of this paper is organized as follows. In Section 2, a downlink NOMA system with multiple near users and a far user is constructed, and the far user and one of near users are matched. In Section 3, the optimal power allocation and user selection are analyzed, and the system outage probability and diversity order under imperfect CSI are derived. In Section 4, we propose a scheme based on the CNN and LSTM for the NOMA scenario, which pretrains data and trains the network to reduce the channel estimation error. In Section 5, the performance of the proposed scheme is simulated and analyzed, which is followed by conclusions in Section 6.

2. System Model

We consider a NOMA downlink scenario consisting of a base station (BS), a far user U_f, and N near users U_1, U_2, …, U_n, …, U_N, as shown in Figure 1. The location of U_f is fixed, and near users are randomly distributed between the BS and U_f. Assume the channels between BS/U_f and near users follow Rayleigh distribution. The near users have direct links with BS, whereas there is no direct link between BS and U_f. Hence, U_f can only rely on the near users' help to communicate with BS. Specifically, nears users employ the decode-and-forward protocol to forward the far user's information.

Each transmission block is equally divided into two phases. In the first phase, supposing the near user U_n is selected to help U_f, BS sends a superimposed signal containing the information for U_n and U_f. The signal received by U_n can be expressed as

$$y_n = h_n\left(\sqrt{\alpha_n P}x_n + \sqrt{(1-\alpha_n)P}x_f\right) + \omega_n, \tag{1}$$

where x_n and x_f denote the messages for U_n and U_f, respectively, $h_n \sim CN(0, \Omega_{bn})$ is the channel coefficient from BS to U_n, $\Omega_{bn} = d_{bn}^{-\lambda}$, λ denotes the path loss exponent, d_{bn} denotes

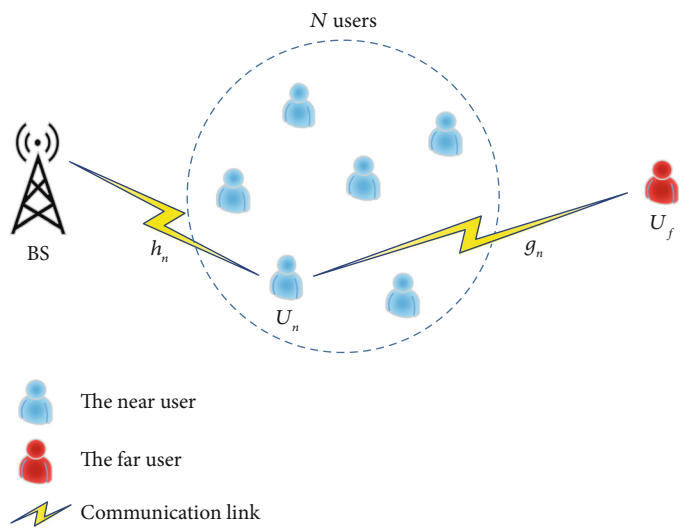


Figure 1: NOMA downlink scenario.

the distance between BS and U_n, α_n represents the power allocation factor for U_n, P represents the transmit power, and ω_n represents additive white Gaussian noise (AWGN) at U_n with mean 0 and variance σ^2.

After receiving y_n, U_n first tries to decode x_f and then decodes x_n with estimated h_n. Assuming the minimum mean square error (MMSE) channel estimator [30] is used, then we have

$$h_n = \widehat{h}_n + e_n, \quad 1 \le n \le N, \tag{2}$$

where $\widehat{h}_n$ is the complex channel coefficient estimated and e_n denotes channel estimation error following a complex Gaussian distribution, denoted by $e_n \sim CN(0, \delta_n^2)$. Hence, the received signal-to-interference-plus-noise ratio (SINR) at U_n to detect x_f and x_n can be, respectively, given by

$$\gamma_n^f = \frac{(1-\alpha_n)\rho\left|\widehat{h}_n\right|^2}{\alpha_n\rho\left|\widehat{h}_n\right|^2 + 1 + \rho\delta_n^2}, \tag{3}$$

$$\gamma_n^n = \frac{\alpha_n\rho\left|\widehat{h}_n\right|^2}{\rho\delta_n^2 + 1}, \tag{4}$$

where $\rho = P/\sigma^2$ is the transmit signal-noise ratio (SNR).

If U_n successfully decodes x_f, it forwards x_f in the second phase. The signal received by U_f can be expressed as

$$y_f = \sqrt{P} g_n x_f + \omega_f, \tag{5}$$

where $g_n \sim CN(0, \Omega_{nf})$ is the perfect complex channel coefficient from U_n to U_f, $\Omega_{nf} = d_{nf}^{-\lambda}$, d_{nf} denotes the distance between U_n and U_f, and ω_f is the AWGN at U_f with mean 0 and variance σ^2. Similar to h_n, let the estimate for the channel g_n be $\widehat{g}_n$. By assuming MMSE estimation error, it holds that

$$g_n = \widehat{g}_n + r_n, \quad 1 \le n \le N, \tag{6}$$

and the SINR at U_f to decode x_f is given by

$$\gamma_{nf} = \frac{\rho|\widehat{g}_n|^2}{\rho\varepsilon_n^2 + 1}, \tag{7}$$

where ε_n^2 is the variance of r_n.

According to (3) and (7), the achievable rate of U_f (denoted by R_f^n) can be expressed as

$$R_f^n = \frac{1}{2}\log_2\left(1 + \min\left\{\gamma_n^f, \gamma_{nf}\right\}\right). \tag{8}$$

If R_f^n exceeds the targeted data rate of U_f (denoted by R_f), the quality-of-service (QoS) requirement of U_f is satisfied. Thus, according to (4), the achievable rate of U_n (denoted by R_n^n) can be expressed as

$$R_n^n = \begin{cases} \frac{1}{2}\log_2(1+\gamma_n^n), & R_f^n \ge R_f, \\ 0, & R_f^n < R_f. \end{cases} \tag{9}$$

Note that R_n^n is achievable only when $R_f^n \ge R_f$, which means the signal x_f is decoded at U_n successfully. If $R_f^n < R_f$, the corresponding near user will not be selected and we regard its data rate as 0.

The aim of this paper is to maximize the achievable rate of near users while ensuring the QoS requirement of U_f, i.e., maximizing R_n^n, by jointly optimizing near-user selection and power allocation coefficient α_n. In the next two sections, we will study this problem without and with the DL method.

3. Power Allocation, User Selection, and Outage Performance Analysis

In this section, we study the optimal power allocation coefficient and user selection and then investigate the system outage performance without the DL method.

3.1. Optimal Power Allocation Coefficient and User Selection. Firstly, we optimize the power allocation coefficient. Based on (3), (8), and (9), the supremum of α_n can be given by

$$\alpha_n \leq \frac{\rho\left|\widehat{h}_n\right|^2 - \left(1+\rho\delta_n^2\right)\varepsilon_1}{\rho\left|\widehat{h}_n\right|^2\left(1+\varepsilon_1\right)}, \tag{10}$$

where $\varepsilon_1 = 2^{2R_f} - 1$. In order to ensure that U_n decodes x_f successfully, α_n should be larger than zero, in which case we can find that

$$\left|\widehat{h}_n\right|^2 > \frac{\left(1+\rho\delta_n^2\right)\varepsilon_1}{\rho}. \tag{11}$$

Under the condition of (11), since the aim of this paper is to maximize the performance of near users on the basis of meeting the QoS requirement of U_f, the optimal power allocation coefficient α_n^* should be taken as the upper bound, which is given by

$$\alpha_n^* = \frac{\rho\left|\widehat{h}_n\right|^2 - \left(1+\rho\delta_n^2\right)\varepsilon_1}{\rho\left|\widehat{h}_n\right|^2\left(1+\varepsilon_1\right)}. \tag{12}$$

According to (7), (8), (9), and (11), the set of effective near users which can ensure the QoS requirement of U_f can be expressed as

$$E_n = \left\{ n : \left|\widehat{h}_n\right|^2 > \frac{\left(1+\rho\delta_n^2\right)\varepsilon_1}{\rho}, \left|\widehat{g}_n\right|^2 \geq \frac{\left(1+\rho\varepsilon_n^2\right)\varepsilon_1}{\rho} \right\}. \tag{13}$$

To maximize the achievable rate of near users, the selected near users should be contained in set E_n. Thus, the best user selection is expressed as follows:

$$n^* = \arg \max_{n\in E_n} \left\{ \left|\widehat{h}_n\right|^2 \right\}. \tag{14}$$

It can be observed from (12)–(14) that channel estimation has a significant impact on power allocation and user selection.

3.2. System Outage Probability and Diversity Order. The outage probability of this paper is defined as the probability that U_f fails to decode x_f or U_n fails to decode x_n. Let R_n be the targeted rate of U_n. If U_n successfully decodes x_n, i.e., $R_n^n \geq R_n$, we can find that

$$\left|\widehat{h}_n\right|^2 \geq \frac{\left(1+\rho\delta_n^2\right)\varepsilon_2}{\alpha_n\rho}, \tag{15}$$

where $\varepsilon_2 = 2^{2R_n} - 1$. According to (12), formula (15) can be further expressed as follows:

$$\left|\widehat{h}_n\right|^2 \geq \frac{\left(1+\rho\delta_n^2\right)\left(\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2\right)}{\rho}. \tag{16}$$

Taking the constraints of (13) and (16) into consideration, the system outage probability P_{system} is given by

$$\begin{aligned} P_{\text{system}} &= \prod_{n=1}^{N} P_n = \prod_{n=1}^{N}\left\{1 - \Pr\left\{\left|\widehat{h}_n\right|^2 > \frac{\left(1+\rho\delta_n^2\right)\varepsilon_1}{\rho}, \left|\widehat{g}_n\right|^2 \geq \frac{\left(1+\rho\varepsilon_n^2\right)\varepsilon_1}{\rho}, \left|\widehat{h}_n\right|^2 \right.\right. \\ &\left.\left. \geq \frac{\left(1+\rho\delta_n^2\right)\left(\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2\right)}{\rho}\right\}\right\} = \prod_{n=1}^{N}\left\{1 - \Pr\left\{\left|\widehat{g}_n\right|^2 \geq \frac{\left(1+\rho\varepsilon_n^2\right)\varepsilon_1}{\rho}, \left|\widehat{h}_n\right|^2 \right.\right. \\ &\left.\left. \geq \frac{\left(1+\rho\delta_n^2\right)\left(\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2\right)}{\rho}\right\}\right\} \\ &= \prod_{n=1}^{N}\left\{1 - e^{-\left[\left(1+\rho\varepsilon_n^2\right)\varepsilon_1/\rho\left(d_{nf}^{-\lambda}-\varepsilon_n^2\right)+\left(1+\rho\delta_n^2\right)\left(\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2\right)/\rho\left(d_{bn}^{-\lambda}-\delta_n^2\right)\right]}\right\}, \end{aligned} \tag{17}$$

where P_n denotes the outage probability for U_n. Note that only when all near users fail, the system is completely unable to communicate. Therefore, P_{system} is the cumulative outage probability of all N near users. According to the definition of diversity order [31], when $\rho \longrightarrow \infty$, we can find that

$$\lim_{\rho\longrightarrow\infty} P_n = \lim_{\rho\longrightarrow\infty}\left\{1 - e^{-\left(k_{n2}+\rho k_{n3}/\rho k_{n1}+k_{n5}+\rho k_{n6}/\rho k_{n4}\right)}\right\} = 1 - e^{-k_{n3}/k_{n1}-k_{n6}/k_{n4}}, \tag{18}$$

where $k_{n1} = d_{nf}^{-\lambda} - \varepsilon_n^2$, $k_{n2} = \varepsilon_1$, $k_{n3} = \varepsilon_1\varepsilon_n^2$, $k_{n4} = d_{bn}^{-\lambda} - \delta_n^2$, $k_{n5} = \varepsilon_1 + \varepsilon_2 + \varepsilon_1\varepsilon_2$, and $k_{n6} = \left(\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2\right)\delta_n^2$. According to (17) and (18), we can find that

$$\lim_{\rho\longrightarrow\infty} P_{\text{system}} = \lim_{\rho\longrightarrow\infty}\prod_{n=1}^{N} P_n = \prod_{n=1}^{N}\left\{1 - e^{-k_{n3}/k_{n1}-k_{n6}/k_{n4}}\right\}, \tag{19}$$

which is a constant. Thus, the diversity gain of the cooperative NOMA system without the DL method can be expressed as

$$D_{\text{order}} = -\lim_{\rho\longrightarrow\infty}\frac{\lg P_{\text{system}}}{\lg \rho} = 0. \tag{20}$$

According to the abovementioned analysis, we can tell that the accuracy of each hop CSI has a major effect on each user's detection performance and also the user selection result. Next, we will introduce our proposed DL-based channel estimation method for the considered cooperative NOMA network.

4. CNN and LSTM-Based User Selection Scheme for the Downlink Cooperative NOMA

Due to the poor system outage performance under imperfect CSI, this section considers using the DL method to obtain more accurate CSI to improve system performance. In this section, a scheme adopting the LSTM network to solve the proposed optimization problem is considered. The proposed problem needs a large amount of data for learning when using DL, which leads to the vanishing gradient problem. The memory cells in the LSTM network can save previously extracted information for later use, solving the vanishing gradient problem [32]. Thus, LSTM is employed in this paper to learn the CSI of each hop. In addition, considering the large amount of data for training and the SIC method is complex, CNN is added in front of the LSTM network to extract channel features and improve the learning performance.

4.1. Design Concept of Neural Network. Our aim is to minimize channel estimation error and get accurate CSI to select the optimal user and achieve the best performance. In order to learn the CSI of each hop, a neural network with multiple LSTM layers in parallel is designed, and each LSTM layer corresponds to one-hop CSI. In this paper, the network uses two LSTM layers as each transmission block has two-hop CSI. In order to fit the SIC process, multiple hidden layers are added to the network, and the number of hidden layers corresponds to the times of decoding in each transmission. Thus, two hidden layers are added after the LSTM layers in this paper. Meanwhile, a CNN, which pretrains data to make the data easier for learning, is added in front of the LSTM network. It is worth mentioning that according to our design, if the NOMA system becomes more complex, the neural network can be adjusted to adapt to the new NOMA scenario.

The network design is summarized as follows. The first three layers of the network are one-dimensional CNN, pretraining data to extract feature vectors and simplifying learning parameters. The extracted feature vectors are passed to the fourth layer. Then, the fifth layer learns the feature vectors of the imperfect CSI. Note that feature vectors may vanish; thus, the fifth layer is the LSTM layer. After that, two hidden layers are added. Finally, the eighth layer is used to handle the network output. Details are shown in the following subsection.

4.2. Proposed Channel Estimation Framework. The channel estimation framework proposed is shown in Figure 2, which includes offline training and online prediction. The learning network is trained via using a mass of data which is pretrained well for offline training. In the online prediction part, the feedback of the downlink NOMA system is the input of the learning network and the estimated channel coefficient is the output. The following parts introduce the data pretraining and the proposed learning network in detail, respectively.

4.2.1. Data Pretraining. In order to improve the generalization ability and accelerate the convergence speed of the proposed network, the original input data needs to be pretrained. Note that data is converted into a sequence of transmitted symbols in the communication system. Since the data is in the plural form, the real part and the imaginary part need to be separated first and then reconstitute a real number sequence for pretraining. Thus, let the transmitted signal vector $I_n = \{s_1^n, s_2^n, \cdots\infty, s_q^n\}$ be a sequence with all data of U_n, where q is the number of transmitted symbols in the sequence at U_n. If the number is not q in the process of pretraining, the data is invalid. Note that the transmitted signal vector is one-dimensional. Therefore, the CNN used for pretraining data is one-dimensional here. In the task of extracting feature vectors of sequences, one-dimensional CNN shows excellent performance. It can extract the input data sequence into a shorter sequence composed of high-level features, so as to shorten the training time of the neural network and reduce the load of each neuron. After feature extraction, a different sequence O_n will be output.

The CNN constructed in this paper is the superposition of convolution layers and a pooling layer. The specific structure is as follows. The first layer is a one-dimensional convolution layer, which calculates the convolution sum of the input data through the sliding window and obtains the output of the convolution layer after weighting processing. After I_n passes through the convolution layer, the original information sequence is convoluted and transformed, and all convolution outputs form a characteristic matrix, and the matrix data is output to the next layer of the network. The second layer is a one-dimensional maximum pooling layer, which searches the maximum output value of the previous convolution layer through the pooling window, and extracts the characteristic vectors from the characteristic matrix to reduce the number of characteristic parameters. The third layer again stacks a one-dimensional convolution layer, performs the same operation as the first layer, and completes the extraction of feature vectors. The pretrained data O_n can be input into the LSTM network for channel estimation, because the data still represents CSI. The activation function used by the CNN in this paper is the hyperbolic tangent (tanh) activation function, and its formula is expressed as follows:

$$f_{\tanh}(Z) = \frac{1 - e^{-2Z}}{1 + e^{-2Z}}. \tag{21}$$

The pretraining method is shown in Algorithm 1. The pretraining improves the performance of the network, avoids the situation that the network cannot converge, and speeds up the convergence speed.

4.2.2. Proposed Learning Network. The structure of offline training is shown in Figure 3. In order to improve the learning capacities of the network, we have built the LSTM network behind the 3-layer CNN. For each layer of the network, the output is the weighted sum of neurons the layer is equipped with. Since data has been pretrained and the dimension of input data has been reduced, the number of

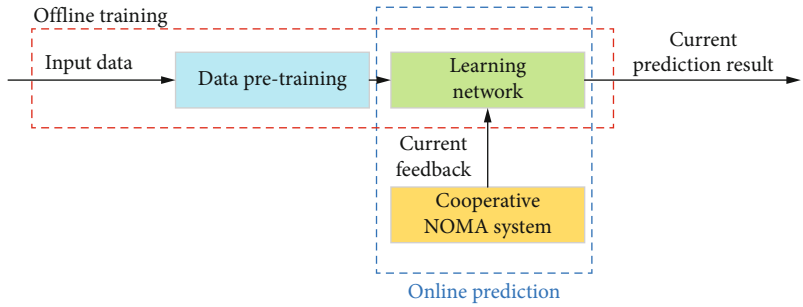


Figure 2: The channel estimation framework.

Input: Data I_n.
Output: Data O_n.
1: **for** I_n is a sequence with all data of U_n **do**
2: Separate the real part and the imaginary part of the plural data.
3: Set $I_n = \{s_1^n, s_2^n, \cdots, s_q^n\}$, where q is the number of transmitted symbols in the sequence in theory.
4: The actual number of transmitted symbols is r.
5: **if** $r = q$ **then**
6: I_n is trained by the proposed one-dimensional CNN.
7: Extract feature vectors.
8: **else**
9: Discard the data;
10: Update O_n.
11: Output data O_n.

Algorithm 1: Data pretraining.

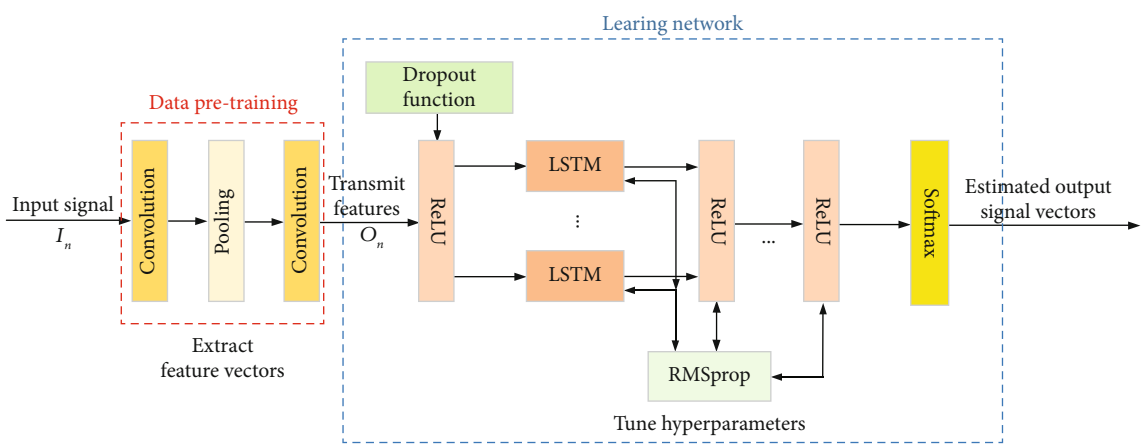


Figure 3: The structure of offline training.

features used for training will not be very large. When the pretrained data enters the LSTM network, these features will perform some kinds of action and propagate some kinds of symbols to the node of the next layer. The dimension of the fourth layer, which is a dense layer with 32 neurons conveying features, is set to the length l of the sequence used for training. Afterwards, the fifth layer consists of multiple parallel LSTM layers. In this paper, two LSTM layers are used to learn the two-hop CSI without correlation. The sixth and seventh layers have 64 and 32 neurons, respectively, which are dense layers, corresponding to two decoding in the NOMA system. If the number of hidden layers here is less than the decoding times, the learning performance will decrease, which we will verify in Section 5. When the number of decoding times in the NOMA system increases, the hidden layers should be added to maintain the learning capacities. Additionally, the output layer processed by the softmax function provides the estimated output signal vectors. The fourth, sixth, and seventh layers are processed by the Rectified Linear Unit (ReLU) function. The above two

Input: Data O_n.
Output: The estimated output signal vectors.
1: **while** There is a training sequence of length l**do**
2: Transmits features to the fourth layer.
3: Use the dropout function on the network.
4: Two parallel LSTM layers learn uncorrelated CSI respectively.
5: Two dense layers correspond to the two decoding in SIC process.
6: **while** The loss function does not converge **do**
7: Use the RMSprop optimizer.
8: **if** Network converges **then**
9: Feed data for the current slot into LSTM network.
10: **if** CSI is successfully identified **then**
11: Select the best near user.
12: **else**
13: Retrain data.
14: **else**
15: Return to step 1.
16: **while** Find the optimal selection **do**
17: Output the estimated output signal vectors.

Algorithm 2: Proposed learning network.

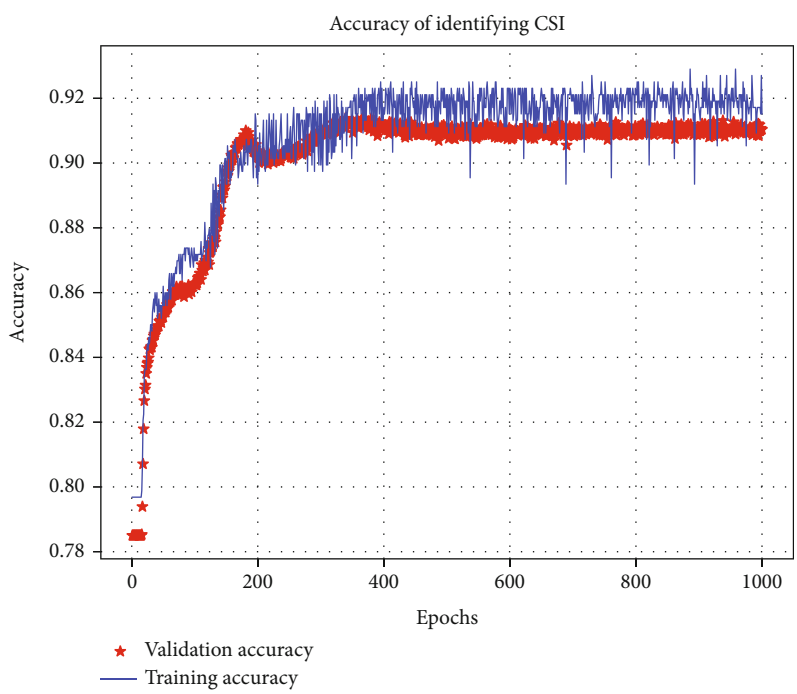


Figure 4: Identification accuracy of the proposed learning network.

functions are expressed as follows:

$$f_{\text{softmax}}(Z_i) = \frac{e^{Z_i}}{\sum_{t=1}^{T} e^{Z_t}}, \qquad (22)$$
$$f_{\text{ReLU}}(Z) = \max\{0, Z\},$$

where Z_i is the output value of the ith node and T is the number of output nodes, namely, the number of categories of classification.

Note that the proposed scheme attempts to minimize the channel estimation error, approximating imperfect CSI to perfect CSI. Whether the judgment of CSI is accurate or not has a great impact on user selection. Based on the successful identification of accurate CSI, the most appropriate near user can be selected.

The principle of the proposed LSTM network is shown in Algorithm 2. The goal of our model is to minimize the channel estimation error, which is equivalent to minimizing the difference between the input data and the output of the network. Therefore, the loss function we use is the

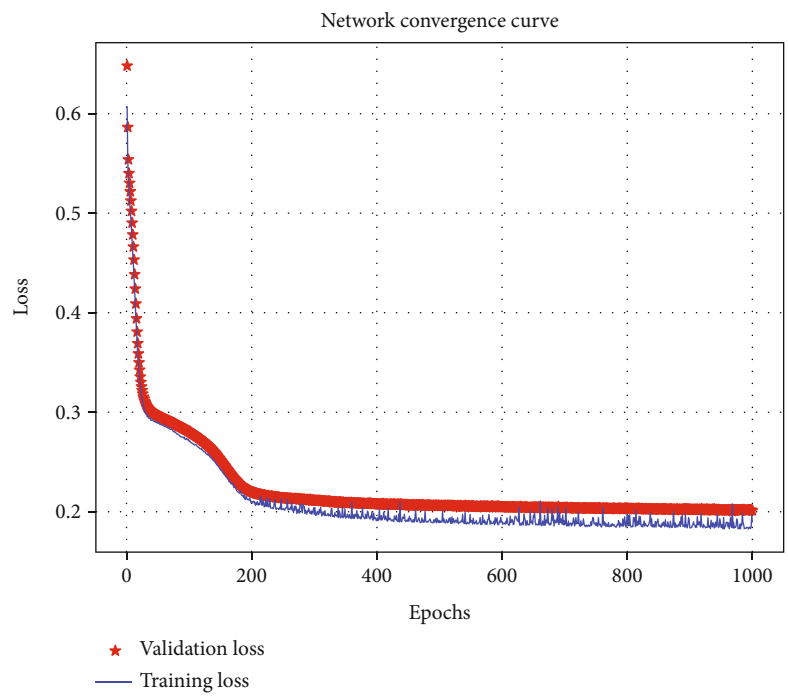


Figure 5: Loss value of the proposed learning network.

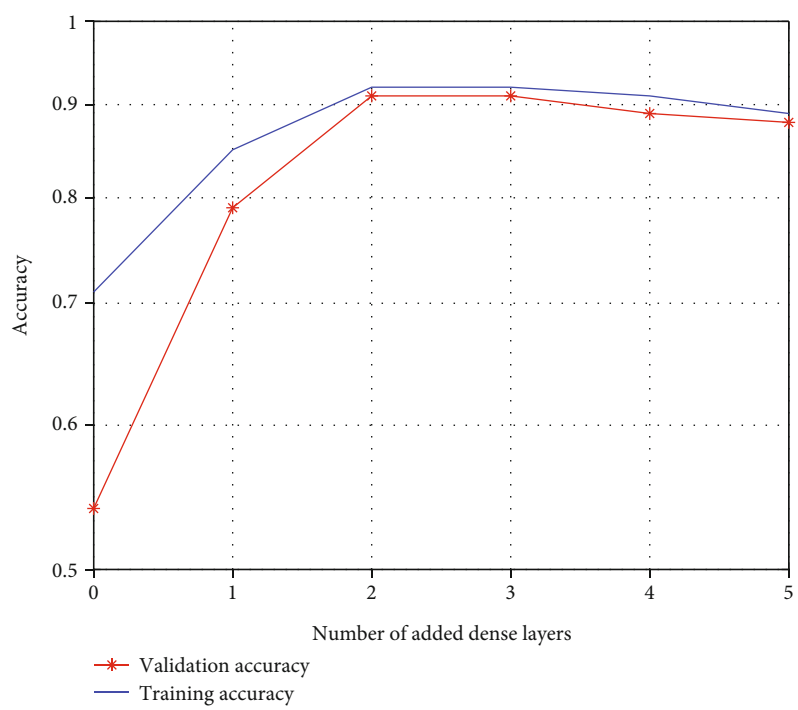


Figure 6: Comparison of different numbers of dense layers behind LSTM layers.

categorical cross-entropy, and its formula is given by

$$f_{\text{loss}} = -\frac{1}{D_{\text{out}}} \sum_{i=1}^{D_{\text{out}}} Y_i \log \widehat{Y}_i + (1 - Y_i) \log \left(1 - \widehat{Y}_i\right), \quad (23)$$

where D_{out} is the dimension of output vectors, Y_i is the expected output, and $\widehat{Y}_i$ is the actual output. It is used to measure the distance between the probability distribution of the network output and the distribution of the label we specify. By minimizing the distance between the two distributions, a well-trained network can make the output as close to the desired result as possible. Meanwhile, in order to reduce the computational complexity and improve the

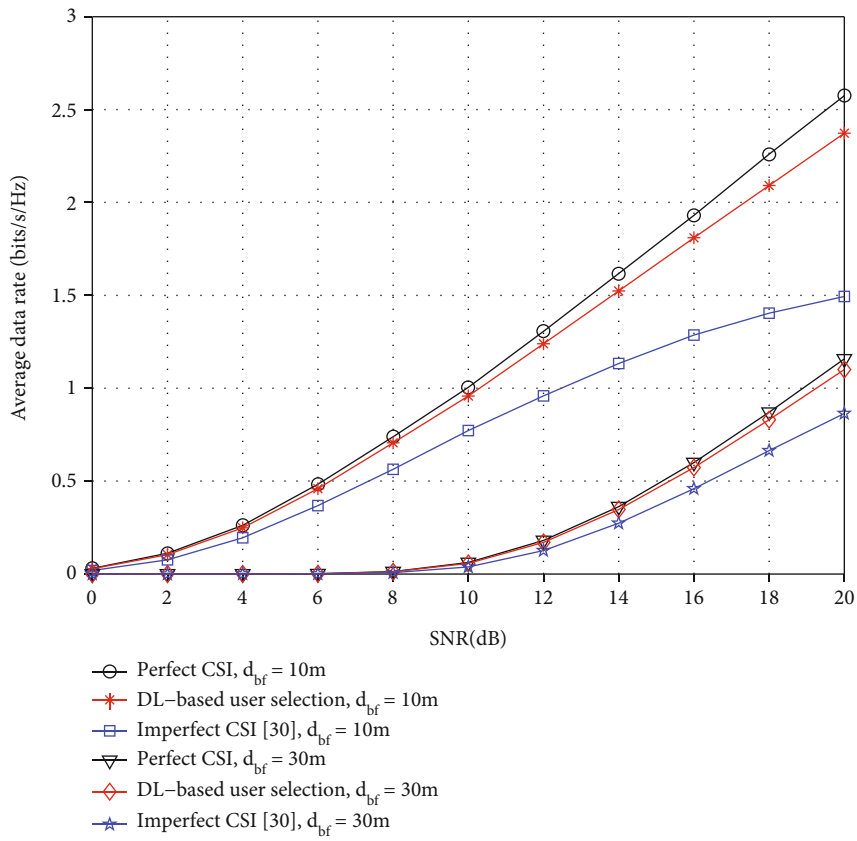


Figure 7: Average data rate of selected near users with different distances from BS to U_f.

learning performance of the network, we choose the RMSprop as the optimizer, which tunes hyperparameters to minimize the value of loss function. Based on the LSTM network, we use the recurrent dropout to improve the generalization ability of the neural network. We set the input unit of a certain layer of the network to 0 randomly, in order to break the accidental correlation in the training data of this layer.

5. Simulation Results and Discussion

In this section, we check the performance of the proposed deep neural network with LSTM framework when it is used in the NOMA system. Specifically, Python 3.7 and Keras are used for programming neural networks, and MATLAB is used for simulation of NOMA data. Keras is a DL framework that makes it easy to define and learn almost any type of DL model. Simulation parameters are set as follows. Every 100000 simulations are averaged to eliminate the randomness caused by channel fading. The distance between BS and U_f (denoted by d_{bf}) is set as 10 m or 30 m, and the near users are randomly distributed between U_f and BS.

Our simulation is divided into the following three parts. Firstly, we verify the performance of the proposed neural network. Specifically, the learning accuracy of the neural network is analyzed, and the convergence is reflected by the loss value. Secondly, the average data rate of the proposed scheme is simulated. We take the proposed DL-based user selection as our learning strategy under imperfect CSI and apply it to the simulation of near-user data rate in this paper. We also compare this strategy with the other two user selection strategies without DL, which are contrasted strategy with imperfect CSI [30] and contrasted strategy with perfect CSI. The user selection with imperfect CSI utilizes the imperfect instantaneous CSI and selects the user with $\widehat{h}_n$, and the user selection with perfect CSI selects the user with h_n. Thirdly, the system outage probability of the proposed scheme is simulated and compared with the other two strategies without DL to verify that the proposed scheme has good outage performance and diversity gain. The numerical results of outage performance are consistent with the theoretical analysis. Details are shown in the following subsections.

5.1. Performance of the Proposed Learning Network. In this subsection, we analyze the performance of our deep neural network. As shown in Figure 4, the constructed LSTM network can finally achieve 91% accuracy to identify CSI. Such precision can largely satisfy the reliability of the user selection. The loss value of the proposed network is shown in Figure 5, in which the network gradually converges after 400 epochs. The above simulation results show that the trained neural network can accurately judge the CSI of the input data, which means the appropriate near users can be selected.

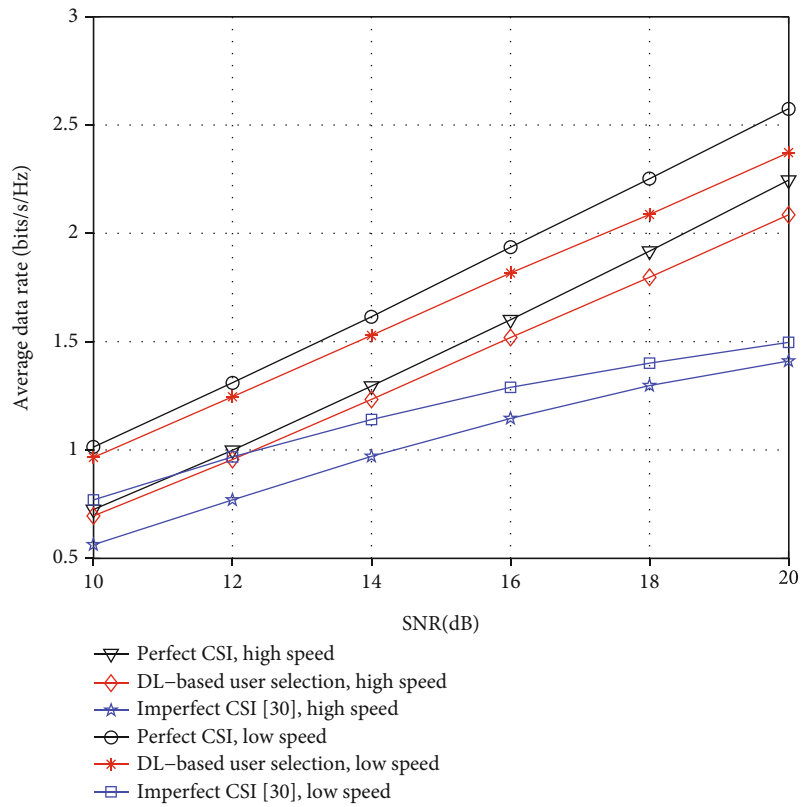


Figure 8: Average data rate of selected near users at different speeds. The distance between BS and U_f is 10 m.

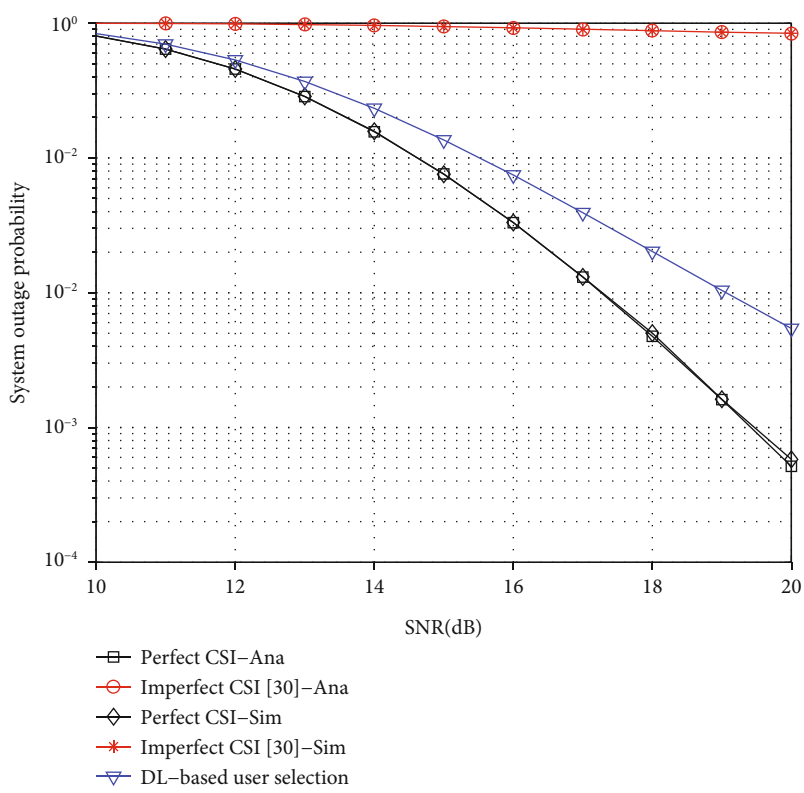


Figure 9: Outage performance comparison and diversity gain of different strategies on cooperative NOMA system.

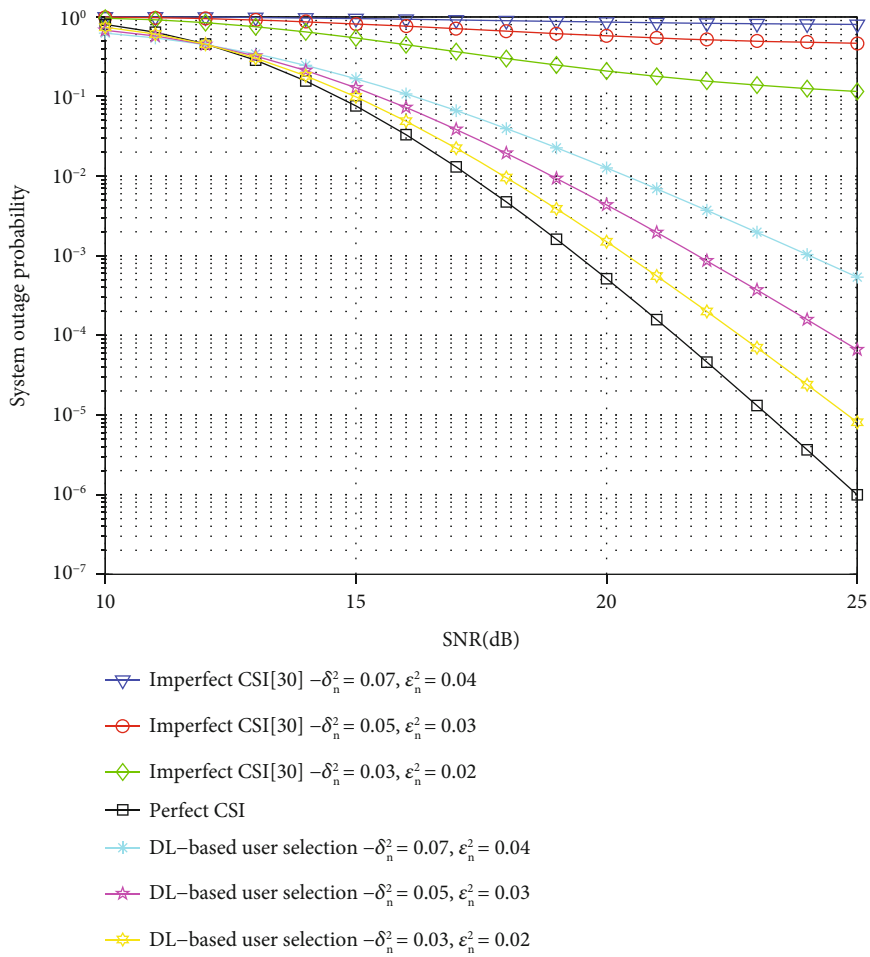


Figure 10: Effect of different estimation errors on system outage performance and diversity gain.

In order to verify the impact of the number of dense layers behind LSTM layers on the performance of neural networks, the learning accuracy of different networks is simulated in Figure 6. When the number of added dense layers is 2, the learning accuracy of the neural network is higher. If the number is less, performance will drop dramatically. In addition, continuing to add dense layers will not achieve greater gains and will increase learning costs. Therefore, we can conclude that when the number of added dense layers is the same as decoding times, the neural network has the best effect.

5.2. Average Data Rate of the Proposed User Selection Scheme. Based on the high accuracy of the learning network, we analyze the average data rate of the proposed user selection scheme when the transmit SNR changes in this subsection. Figure 7 simulates the average data rate of selected near users when d_{bf} is 10 m and 30 m, respectively, and Figure 8 compares the average data rate of selected near users at low speed of 1 m/s and at high speed of 20 m/s, respectively.

As shown in Figure 7, when d_{bf} is 10 m, our proposed strategy has a great advantage over the strategy without the DL method under imperfect CSI. The performance improvement in the average data rate is so significant that it is not far behind even the strategy with perfect CSI. When the transmit SNR increases, the average data rate of near users also increases, and the proposed strategy maintains its superiority. We change the random initial distribution of near users, and the performance advantage remains the same. In order to verify the performance of each strategy under a poor channel state, we add d_{bf} and randomly deploy near users. The average data rate under this circumstance is simulated. It is obvious from Figure 7 that the average data rate of selected users has dropped dramatically. Due to the great distances between users, the channel quality is poor; thus, high power is required to communicate successfully and average data rate is low. Although the channels are very bad, the proposed strategy can still play a good role and is even closer to the perfect CSI strategy than the one with the good channel state. This benefits from the advantages of DL. The more obvious the features are, the better the fitting effect will be, and it is easier to find a better choice in a poor communication environment.

The moving speeds of near users are 1 m/s and 20 m/s in Figure 8, respectively. As shown in Figure 8, the proposed scheme still has advantages, and the curve trend remains constant when the speed changes, which proves that the proposed learning strategy can adapt to user mobility well. By

comparing the performance at low speed and high speed, it can be seen that the average data rate at low speed is higher than that at high speed, which is in line with common sense. According to the simulation results above, we can draw a conclusion that the proposed DL-based user selection can maintain good performance advantages on an average data rate regardless of the speed of near users.

5.3. Outage Performance and Diversity Gain of the Proposed User Selection Scheme. In this subsection, we simulate the system outage performance and discuss the diversity gain. The analytical result in (19) is validated in Figure 9 that when $\rho \longrightarrow \infty$, the system outage probability of the strategy without DL under imperfect CSI is a constant, which means the diversity gain is zero. The outage performance of our proposed DL-based user selection scheme is shown in Figure 9. Obviously, the proposed scheme effectively reduces the system outage probability. Although the full diversity is not obtained, the proposed scheme successfully recovers the diversity gain, which is no longer 0 under imperfect CSI. In addition, the system outage probability under perfect CSI is studied in [33], which has full diversity. The theoretical analysis results of the two contrasted strategies are consistent with the Monte Carlo simulations in Figure 9.

As shown in Figure 10, with the decrease of channel estimation error, the system outage probability decreases and gradually approaches the system outage probability under perfect CSI. Note that higher transmit SNR is needed to make the outage probability of the user selection under imperfect CSI tend to a constant value when the estimation error decreases. However, due to the fact that the channel estimation error exists, the user selection without the DL method under imperfect CSI always has no diversity gain, even if the error is small. For our proposed user selection scheme, a smaller estimation error can help the network obtain higher diversity gain. Thus, our proposed scheme has great advantages over the user selection without the DL method.

6. Conclusion

In this paper, a DL-based user selection scheme for cooperative NOMA is proposed, and its performance is well investigated. Different from most studies, this paper considers the NOMA system with imperfect CSI and analyzes the outage performance and designs a special neural network to accurately identify CSI. In particular, the proposed scheme improves outage performance and recovers diversity gain, showing the superiority of the scheme. We believe that the DL method has more room for development in the domain of NOMA research.

Data Availability

The simulation data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors would like to declare that there are no potential conflicts of interest.

Acknowledgments

The work of Guoxin Li was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1801103, in part by the National Science Foundation of China under Grant 62101595, and in part by the Jiangsu Province Natural Science Foundation under Grant BK20200580.

References

[1] J. M. Meredith, "Study on downlink multiuser superposition transmission for LTE," in *TSG RAN Meeting*, vol. 67, Shanghai, China, March 2015.

[2] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.

[3] X. Li, Z. Xie, Z. Chu, V. G. Menon, S. Mumtaz, and J. Zhang, "Exploiting benefits of IRS in wireless powered NOMA networks," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 175–186, 2022.

[4] G. Li, H. Liu, G. Huang, X. Li, B. Raj, and F. Kara, "Effective capacity analysis of reconfigurable intelligent surfaces aided NOMA network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, 16 pages, 2021.

[5] Z. Ding, Y. Liu, J. Choi et al., "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, 2017.

[6] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 611–615, London, U.K., Sep 2013.

[7] V. Ozduran, "Advanced successive interference cancellation for non-orthogonal multiple access," in *2018 26th Telecommunications Forum (TELFOR)*, pp. 1–4, Belgrade, Serbia, Nov 2018.

[8] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1462–1465, 2015.

[9] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 938–953, 2016.

[10] L. Yang, J. Chen, Q. Ni, J. Shi, and X. Xue, "NOMA-enabled cooperative unicast-multicast: design and outage analysis," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 7870–7889, 2017.

[11] L. Yang, H. Jiang, Q. Ye, Z. Ding, L. Lv, and J. Chen, "On the impact of user scheduling on diversity and fairness in cooperative NOMA," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 11296–11301, 2018.

[12] Q. Y. Liau and C. Y. Leow, "Successive user relaying in cooperative NOMA system," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 921–924, 2019.

[13] X. Li, Q. Wang, M. Liu et al., "Cooperative wireless-powered NOMA relaying for B5G IoT networks with hardware impairments and channel estimation errors," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5453–5467, 2021.

[14] S. Arzykulov, T. A. Tsiftsis, G. Nauryzbayev, and M. Abdallah, "Outage performance of cooperative underlay CR-NOMA with imperfect CSI," *IEEE Communications Letters*, vol. 23, no. 1, pp. 176–179, 2019.

[15] J. Cui, Z. Ding, and P. Fan, "Outage probability constrained MIMO-NOMA designs under imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8239–8255, 2018.

[16] G. E. Hinton, S. Osindero, and Y.-W. The, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[17] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2063–2079, 2018.

[18] L. Shao, D. Wu, and X. Li, "Learning deep and wide: a spectral method for learning deep networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 12, pp. 2303–2308, 2014.

[19] M. U. Ghani and W. C. Karl, "Fast enhanced CT metal artifact reduction using data domain deep learning," *IEEE Trans. on Computational Imaging*, vol. 6, pp. 181–193, 2020.

[20] Y. Liao, Y. Hua, X. Dai, H. Yao, and X. Yang, "ChanEstNet: a deep learning based channel estimation for high-speed scenarios," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, pp. 1–6, Shanghai, China, May 2019.

[21] J. Kang, C. Chun, and I. Kim, "Deep-learning-based channel estimation for wireless energy transfer," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2310–2313, 2018.

[22] H. Khan, M. M. Butt, S. Samarakoon, P. Sehier, and M. Bennis, "Deep learning assisted CSI estimation for joint URLLC and eMBB resource allocation," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Dublin, Ireland, June 2020.

[23] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2018.

[24] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Communications Letters*, vol. 23, no. 4, pp. 652–655, 2019.

[25] Q. Bai, J. Wang, Y. Zhang, and J. Song, "Deep learning-based channel estimation algorithm over time selective fading channels," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 125–134, 2020.

[26] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8549–8560, 2018.

[27] E. Balevi, A. Doshi, and J. G. Andrews, "Massive MIMO channel estimation with an untrained deep neural network," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2079–2090, 2020.

[28] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. Naderi Alizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE Journal on Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 989–1000, 2019.

[29] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, 2018.

[30] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 654–667, 2016.

[31] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1389–1398, 2003.

[32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[33] Z. Yang, Z. Ding, Y. Wu, and P. Fan, "Novel relay selection strategies for cooperative NOMA," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10114–10123, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8455065, 13 pages
https://doi.org/10.1155/2022/8455065

Research Article

An Effective Wireless Sensor Network Routing Protocol Based on Particle Swarm Optimization Algorithm

Mohammed Zaid Ghawy,[1] Gehad Abdullah Amran,[2] Hussain AlSalman,[3] Eissa Ghaleb,[4,5] Javed Khan,[6] Ali A. AL-Bakhrani,[7] Ahmed M. Alziadi,[8] Abdulaziz Ali,[8] and Syed Sajid Ullah[9]

[1]*Department of Computer Science, Faculty of Computer Sciences and Information Systems, Thamar University, Dhamar, Yemen*
[2]*Department of Management Science and Engineering Faculty of Management and Economics, Dalian University of Technology, China Liaoning, Dalian*
[3]*Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia*
[4]*College of Computer and Information Science, Southwest University, Chongqing 400715, China*
[5]*Computer Science Department, Aden Community College, Aden, Yemen*
[6]*Department of Electrical Engineering, University of Science and Technology Bannu, Pakistan*
[7]*Department of Computer Science, Technique Leaders College, Sana'a, Yemen*
[8]*Software College, Northeastern University, No. 195, Chuangxin Road, Hunnan District, Shenyang, 110169 Liaoning, China*
[9]*Department of Information and Communication Technology, University of Agder, Norway*

Correspondence should be addressed to Mohammed Zaid Ghawy; mohammdghawi@gmail.com and Gehad Abdullah Amran; jehad.westran@gmail.com

Received 11 March 2022; Revised 6 April 2022; Accepted 11 April 2022; Published 5 May 2022

Academic Editor: Xingwang Li

Improving wireless communication and artificial intelligence technologies by using Internet of Things (Itoh) paradigm has been contributed in developing a wide range of different applications. However, the exponential growth of smart phones and Internet of Things (IoT) devices in wireless sensor networks (WSNs) is becoming an emerging challenge that adds some limitations on Quality of Service (QoS) requirements. End-to-end latency, energy consumption, and packet loss during transmission are the main QoS requirements that could be affected by increasing the number of IoT applications connected through WSNs. To address these limitations, an effective routing protocol needs to be designed for boosting the performance of WSNs and QoS metrics. In this paper, an optimization approach using Particle Swarm Optimization (PSO) algorithm is proposed to develop a multipath protocol, called a Particle Swarm Optimization Routing Protocol (MPSORP). The MPSORP is used for WSN-based IoT applications with a large volume of traffic loads and unfairness in network flow. For evaluating the developed protocol, an experiment is conducted using NS-2 simulator with different configurations and parameters. Furthermore, the performance of MPSORP is compared with AODV and DSDV routing protocols. The experimental results of this comparison demonstrated that the proposed approach achieves several advantages such as saving energy, low end-to-end delay, high packet delivery ratio, high throughput, and low normalization load.

1. Introduction

Wireless sensor networks (WSNs) have recently become among the most important techniques in the world. WSN is a wireless network system that self-organizes, recognized through numerous energy- restricted micro sensors and at least one sink base station (BS) [1–3]. A WSN is made up of large, low-power sensor nodes that are intelligent with a

high-power sink that is personally responsible for building pathways between themselves while adhering to particular transmission rules. The development of wireless sensor technologies sparked Internet of Things (IoT) is required in today's world applications, ushering in a new era of pervasive and intelligent IoT applications [2]. The IoT comprises numerous inventions that provide a consistent network, everywhere, and about everything [4]. The IoT is being used in various applications, including military, monitoring systems, smart buildings, agriculture, and smart cities [5, 6]. WSNs may now build spontaneous connections betwixt devices with or without any infrastructures thanks to recent advancements in ad hoc wireless technology [7]. Furthermore, WSNs have also evolved the most important part of the smart city with IoT environments, utilizing intelligent devices that might be tailored and automatic constructed by the WSN sending, receiving, and sharing data on a very limited region. Wireless technologies of the next generation are on expecting to enable large numbers of connections and produce the highest data rates with decreased energy consumption and transmitting latency [8–16]. With the growth of IoT and WSN, IoT devices are widely used in intelligent industries [17–28]. IoT devices, on the other hand, have limited computational resources and energy [29–44]; therefore, the advancement of IoT devices has created enormous challenges for the IoT system. The sensor node sends the sensor information to the (BS) for further processing in several IoT applications [23]. This is accomplished by using efficient routing algorithms in WSNs to improve the network and make it more profitable by improving data transmission, scalability, and energy efficiency, making the network more viable [2]. On the other hand, there are numerous obstacles in building an effective communication that is resourceful routing protocol for WSNs/IoT, such as the unreliability of low-power wireless networks and restricted resources, which are typically lacking in terms of QoS needs. Due to the dynamical and self-organizing nature of WSNs, connection between devices with one another via wireless channel and the typical routing technique is inapplicable due to the topology's high dynamic characteristic. Thus, these models were only concerned with basic routing requirements. As a result, when developing routing protocols, researchers face numerous challenges [45–49]. Still, there is now an urgent need for developing routing protocols to maximize network lifetime, improve essential QoS requirements, and enable additional features that include a high-packet delivery ratios, low latency, information security, energy effectiveness, and flexible network topology [45]. Furthermore, the aforementioned requirements in the preceding line, establishing a multipath-aware routing protocol based on intelligent optimizations techniques, is still an essential prerequisite for many IoT applications [51–53].

Swarm intelligence [46] is a relatively new field well defined as the work we put effort into developing a distributed problem-solving algorithm that is based on social insect collective behavior. Currently, sensor appliances actually require networking solutions that are low in cost and complexity while simultaneously increasing reliability. Alternatively, the route problem can commonly be represented including a multidimensional optimization issue in which throughput must be maximized while latency is minimized. In recent years, many bioinspired evolutionary techniques (swarm intelligence) have piqued interest, as an instance Ant Colony Optimization (ACO) and PSO, for determining optimal paths in WSN-based IoT applications [55–59]. The basic idea underlying these algorithms is to calculate the cost/fitness iteratively and calculate a function over a population initially and the cost or the fitness on a fresh population, usually acquired by performing operations to populations. Finally, we can determine the optimal solution by weighing all options against the best option. The most effective solution provides us with the population's optimal values, which may be utilized to determine the best routing path.

A unique multipath routing protocol is given in this work, which can utilize the Particle Swarm Optimization (PSO) approach to generate several pathways. The protocol is named as Multipath-based Particle Swarm Optimization Routing Protocol (MPSORP) to improve QoS in WSN for IoT applications. Figure 1 depicts the structure of a WSN for IoT applications. The diagram depicts how sensed data was delivered from the source to the end-user via a multihop path and gateway. This allows end users to access data anytime and from any location. Based on the calculation of a fitness function that may be used to locate the optimal routing path [47], the protocol enables for a source node to choose the best path to data transmission from a collection possible path identified.

The proposed protocol MPSORP is an improvement over traditional Destination- Sequenced Distance Vector (DSDV) routing protocol to consider multiple paths instead of a single path. Furthermore, MPSORP provides the optimization mechanism select optimal path to the destination nodes with achieve more load balancing and high QoS while maintaining power conservation. Generally, the structure of Mobile Ad Hoc Network (MANET) is not self-organized [48]. Still, we can make it as self-organized as possible with the assist of computational intelligence-based methods for instance swarm intelligence. A new optimized routing solution known as MPSORP was proposed to address these issues, which can improve performance QoS measurements in WSN-based IoT applications, such as average end-to-end delay, packet delivery ratio, bandwidth, and packet loss. WSN-based IoT applications have additional hurdles, for instance, high bandwidth requirement, minimal delay, real-time delivery, suitable jitter, and low packet loss ratio, in addition to the challenges posed by WSNs in term of energy limitation, deployment, coverage, and reliability. These features impose much more resource limits, including energy usage, memory, bandwidth, and computing capability [49].

In this paper, we chose this method since it is well suited to our problem. This is known as a population-based search method inspired by flocking bird social behavior or schooling fish. The performance of the suggested approach was evaluated by comparing them with the DSDV and Ad Hoc On-Demand Distance Vector (AODV) protocols using the Network Simulator NS-2 [50, 51].

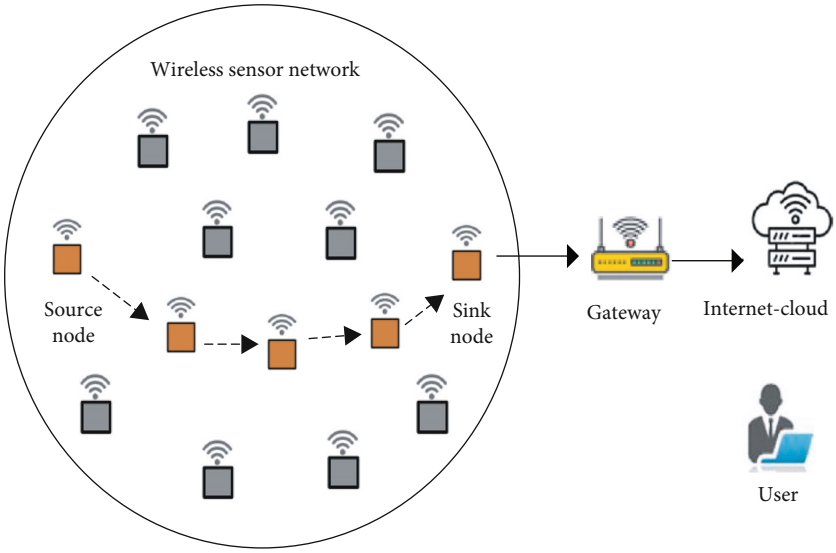


Figure 1: Structure of WSN-based IoT applications.

phase one :[Initialization Phase]
for (x =0 to number of solutions or populations). for (n =0 to number of sensor nodes).
Arbitrarily solutions are selected. Calculate new route using solution. End for .
Calculate fitness value of initialized solution. Calculate global best and Local best.
End for .
phase two :[Update Phase] while criteria does not match
for (x =1 to number of solutions)
for (n =1 to number of sensor nodes). update solution using PSO update equation. Generate new path based on update solution. End for.
Calculate fit-ness value for updated route. Calculate global best and local best.
End for.
Note the global best End while.

Pseudocode 1: Pseudocode- (PSO-) based routing algorithm.

The PSO algorithm was chosen for this investigation for a variety of reasons, including the following: to begin, its operating principle is based on swarm intelligence, which is applicable to the problem we are attempting to solve. Second, the idea of PSO is to combine the collective memory of the entire swarm with individual recollections. Moreover, PSOs can search very large spaces for candidate solutions and use these on optimization problems that are partially irregular and noisy and change over time. Also, it is a fast algorithm and requires less memory compared with other optimization algorithms and is easy to implement. Third, among the mentioned approaches, in PSO, the solution of a particular problem is being represented with the aid of a multidimensional role of a particle and a swarm of particles is working collectively to search the first-class role which corresponds to the satisfactory problem solution, so that the PSO approach is applicable to the disciplines of discrete and combinatorial optimizing. Fourth, PSO is a population-based stochastic search approach that does not use the gradient of the issue being optimized; hence, it does not require the optimization problem to be substantially differentiable, as standard optimization techniques do. As a result, it can also be used in partially irregular, time-variable, and noisy optimization situations. Fifth, the simulation results for several recent researches show that PSO is a reliable method to optimize the WSN for multipath routing, because this method achieves faster convergence and gives more accurate results than nonoptimized Hopfield Neural Network (HNN), noisy HNN, shortest path (SP) algorithm, Genetic Algorithm, ACO, and new research in this field [47, 49, 52–54]. Sixth, PSO is an optimization method that is extremely simple, effective, and efficient. PSO is often used to investigate the search location. It is fairly simple to construct and can be used for both scientific study and engineering purposes. It has a higher optimization capability and can accomplish tasks quickly. PSO is a more robust and simple solution for serious environmental monitoring and data aggregating challenges [55]. The section goes over the literature review of multipath routing approaches and the methods' drawbacks as the reason for this study. This section includes an investigation of existing methodologies, which motivated the researcher to develop an effective and optimal multipath routing system.

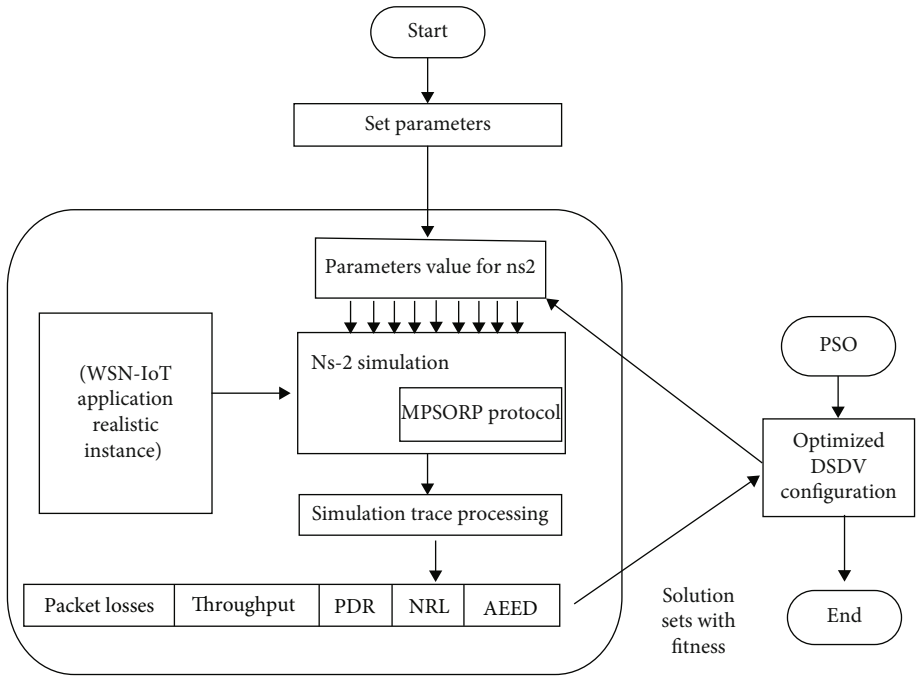


Figure 2: The optimization framework multipath using PSO.

Table 1: Simulation of parameters.

Parameter	Value
Mobility model	Random way-point
Simulator	NS 2.34
Number of nodes	25
Area of simulation	500 k m ∗ 500 km, wide
Simulation time	3600 second
Speed	Varying from 0.006 m/s to 0.47 m/s (ocean env)
Pause time	300 second
Traffic type	Constant bit rate (CBR)
Data packet length	70 bytes
Connection rate	0.05 to 4 packets per second
Routing protocol	MPS-ORP, DSDV, AODV
Radio-propagation model	Two-ray-ground
Interface queue type	Drop-tail/Pri-queue
Maximum packet in queue	150
Network interface type	WirelessPhy/802_15_4
MAC type	802_15_4
Antenna type	OmniAntenn
Transmitting power (pt_)	0.001 Watts
Receiver threshold (RXThresh_)	$3.981e-13$ Watts
Capture power threshold (CPThresh_)	10 dB
Carrier sensing threshold (CSThresh_)	$3.981e-13$ Watts
Operating Freq (freq_)	$2.4e+9$ GHz
System loss factor (L_)	1.0

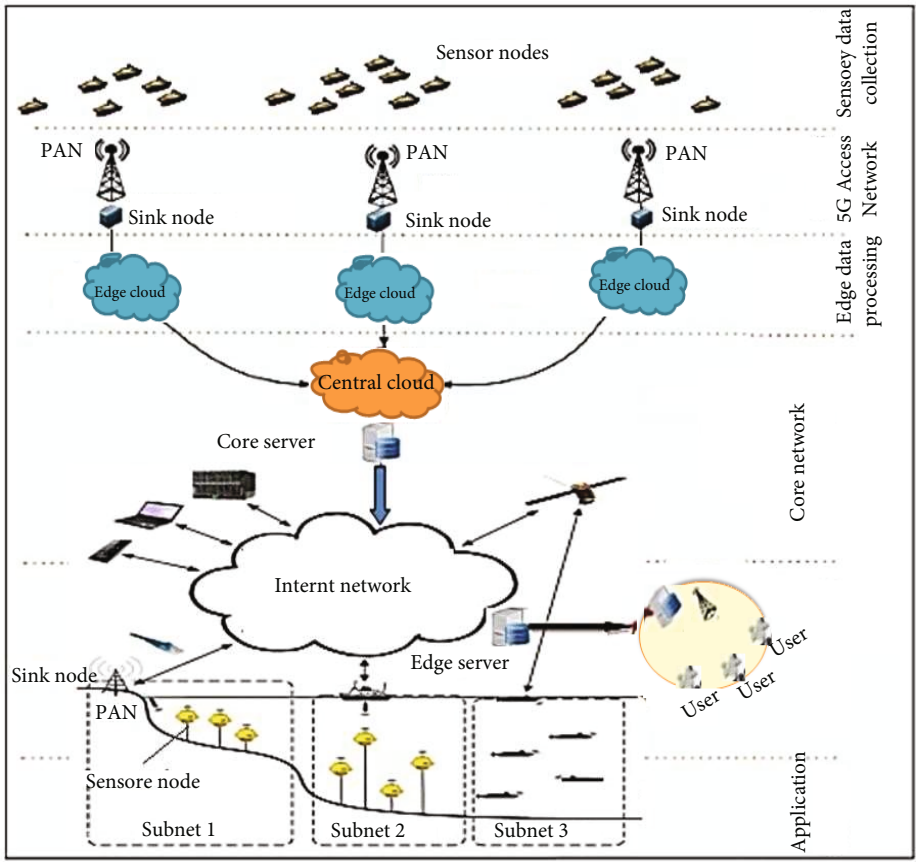


Figure 3: Proposed system model of Internet oceanic of things.

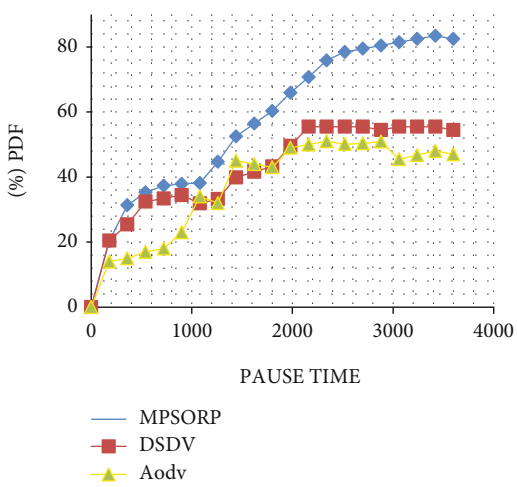


Figure 4: Packet of delivery ratio versus pause time.

2. Background

The basic target of routing protocol structure is to decrease energy usage by determining the most efficient paths between source and destination. As a result, the researcher's task has been to create an energy effective routing protocol with better QoS parameters, as shown in [55–57]. As a result, several studies have been proposed by academics and academicians in order to determine the best way while taking into account QoS metrics like throughput, Packet Delivery Fraction (PDF), and delay [58].

Liu et al. proposed an accurate multipath routing protocol with two-path selection matrices assumed link stability and the value of time constraints into account, resulting in low link interruption probability and delay [59].

The authors in [60] propose an MDW (multipath DSDV-based routing protocol) for WIA-PA (Wireless Networks for Manufacturing Automation-Process Automation) network. It adopts the routing choosing criterion based on the link stability and the disjoint node strategy. Their method showed enhanced results compared to DSDV multipath (DSDVM) over QoS metrics: PDF, the average end-to-end delay, and the average remaining energy the route, but this is limited by static network and no consideration of mobility or other hard environmental condition in the evaluation of protocol.

The authors in [61] propose the Neighbor Coverage Multipath DSDV (NCMDSDV). Implementing two fields called "Second-hop" and "Link-id" in the routing table generates nonlinked pathways betwixt the source and destination nodes. According to findings, the multipath DSDV offers a higher throughput and packet delivery ratio. It is faster in comparison because it has a shorter end-to-end delay and less packet loss.

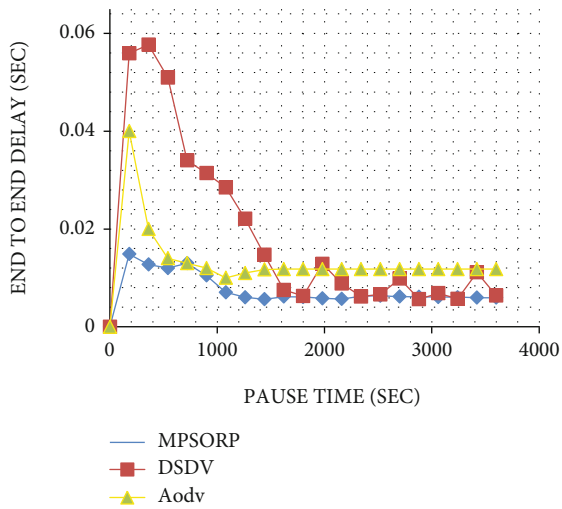


Figure 5: Aver-age end-to-end delay versus pause time.

Two multipath routing protocols (M-AODV and MDSDV) were examined in [62] in order to keep improving QoS for real-time multimedia applications in MANET. The outcomes showed that MDSDV performs well in certain situations (network load and reliability for large-scale networks). Furthermore, regardless of the number of nodes in medium mobility, this protocol provides appropriate and good quality with low jitter.

Design an energy efficient routing system for an IoT application based on a WSN that has inequity in the network due to high traffic load as in [47]. To choose the best path, the three factors are considered in the proposed protocol: longevity, reliability, and traffic intensity at the next-hop node. In comparison to other protocols, the proposed protocol saves more energy, has larger packet delivery ration, has a shorter end-to-end delay, and has a longer network lifetime.

In [63], the authors present an energy-aware, distance-based model. In addition, the proposed solution has the ability to achieve the best balance energy of nodes and increases the IoT lifetime.

In [64], the authors discuss wireless sensor networks, routing strategies, and hierarchy architecture, as well as a brief discussion of wireless sensor networks' energy harvesting. Similarly, in [65], the authors proposed an enhanced QoS-based on clustering with multipath routing protocol (OQoS-CMRP) for WSNs that reduces energy consumption in the sink exposed zone by forming clusters and selecting cluster heads in the sink exposed zone using a modified PSO-based clustering algorithm to address the energy hole issue. Designing energy-effectual multipath routing protocol is critical for WSN applications. The authors of [66] present a multipath routing system that accounts for wireless interference and network energy savings. The path discovery method of the proposed protocol selects the next-hop node, potentially lowering the linking cost.

In [67], the authors suggested an energy-efficient routing system based on PSO. They used to delay and energy as two limitations to reduce PSO problem to its simplest form. Their simulation outcomes and a comparison study with the Genetic Algorithms (GA) revealed that PSO outperforms GA in terms of finding the best path with the least amount of energy usage.

In [4], the authors suggested a bioinspired routing method for constructing, recovering, and selecting k-disjoint pathways that tolerate failure while meeting QoS requirements.

In [49], the authors conduct a survey that focuses on highlighting and explaining available swarm intelligence-based routing techniques for WSNs. Also included are instructions on constructing and building smart routes to support QoS-aware applications.

The general technique in earlier works has been to use multipath routing to reduce rates of data packet dropouts and increase network lifetime. Furthermore, previous studies did not consider optimizing the network's performance of multipath routing protocols. Nevertheless, in this study, we investigate the widespread use of mobile WSNs to meet various IoT application environments.

After reviewing numerous strategies utilized in the background in order to address the issue of optimal path finding in routing protocols, we discovered that there is still a study topic that has yet to be investigated. For example, WSN-based IoT applications minimize energy usage in discovering multipath to transport data from source to destination.

The following is the paper's primary contribution:

(i) Designing and developing an efficient routing method for finding the optimum path

(ii) Choosing the best path based on the most effective effect

(iii) To increase network-related QoS metrics for instance bandwidth, packet delivery rate, and end-to-end delay

(iv) Using the usual NS-2 simulations tool, evaluate the suggested protocol's performance

This study is aimed at proposing an effective routing recovery technique that allows the optimal transmitting path to be carefully reviewed and updated while the mobile sink moves. We develop fault-tolerant routing models to ensure network stability, improve network performance, minimize network energy consuming, lengthen network lives, and keep improving network robustness and reliability.

The main aims of this paper are to boost the effectiveness of multipath routing protocols in the network, and we differ from previous research in the following ways:

(1) We present the method of an enhanced PSO-based multipath routing algorithm for mobile WSNs, which improves network efficiency and reliability in comparison to the most existing work

(2) In contrast to optimal routing with a static sink, our experiment employs sink mobility to save energy, increase network throughput, reduce packet loss, and extend network lifetime

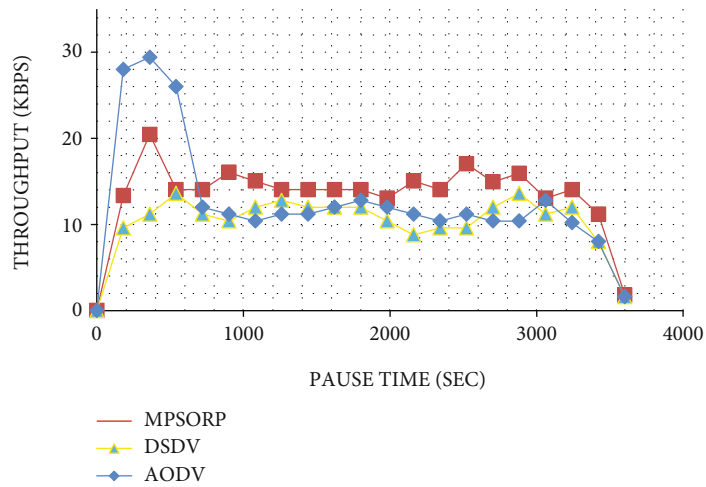


Figure 6: Throughput versus movement in time.

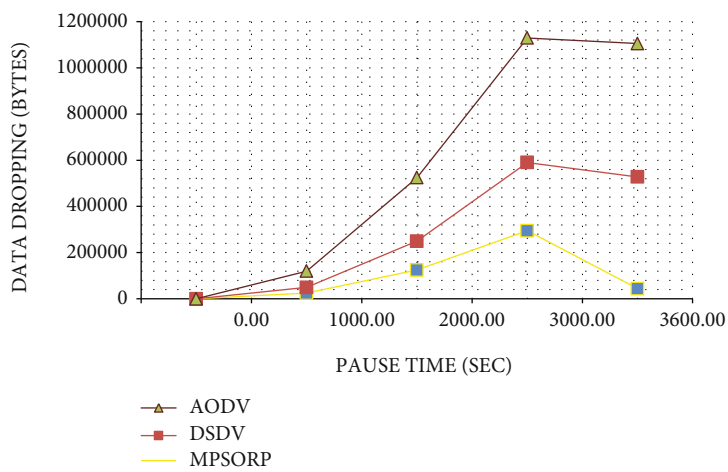


Figure 7: Packet loss versus pause time.

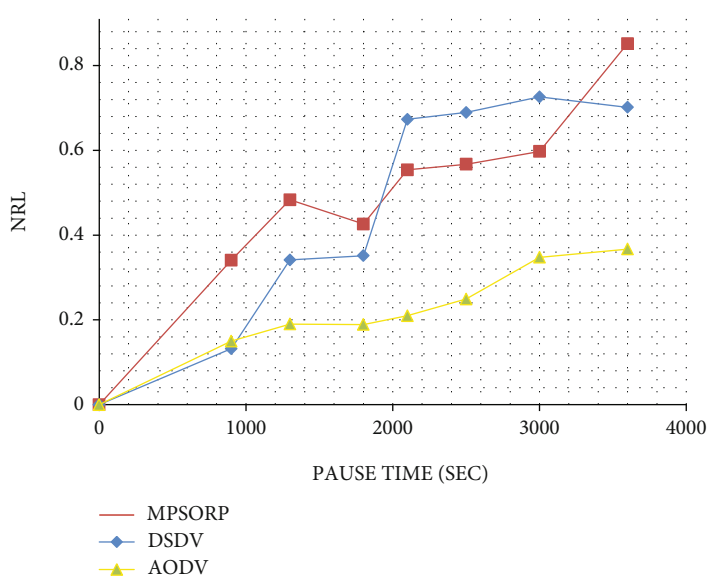


Figure 8: Normalized routing load versus movement in time.

(3) In order to address routing concerns, we impose reliability and delay constraints on sink mobility additionally to the energy constraint imposed by the network's lifetime

3. Proposed Model

We formulate the issue of the proposed routing methods for WSN with mobile sink in this section. The routing problem might be alternatively expressed as a multidimensional optimization problem in which the throughput must be maximized, and the latency must be minimized. The proposed approach is aimed at efficiently transmitting data betwixt sources and sinks across a wireless medium with enhanced throughput, lower delay, and routing overhead, as much as energy consuming.

The mobility in mobile WSNs is used as a mobile sink node for data collecting. Meanwhile, it is built on the traditional clustering model, as well as the proactive routing environment for establishing a mobile sensor network environment. To improve the dependability of data transmission, fault tolerance technology for multipath routing is utilized to build numerous transmitting paths betwixt the source node and the destination. Although this strategy boosts the network's assist to load balancing and transmission bandwidth, it also appears to increase the power consumption as well as complexity of the route structure. Furthermore, it improves the data transmission's stability and reliability. It is a widely used technique for achieving fault tolerance at the network layer.

3.1. PSO Algorithm. PSO techniques replicate a group of birds searching for food and sharing information using the cognitive processing as well as experience gained from food foraging and sharing of information. Individual partnerships will be used to find the first best answer, with a wide range of applications [26]. The particle mobility is influenced by two elements that use information as of iteration to iteration and particle to particle. The particle saves the best solution visited thus far, designated pbest, in its memory due to iteration-to-iteration information and experiences and attractiveness to this solution as it journeys across the solution into search space. The perfect solution visited by any particle is stored in the particle's memory due to particle-to-particle interaction, and the particle becomes attracted to this solution, denoted gbest. The cognitive and social constituents are the first and second factors. If a better or more dominant solution exists (in aspects of fitness) after iteration, the pbest and gbest for each particle are changed. This approach is repeated iteratively until the demand result is either obtained or decided that an acceptable solution is not possible within the computing constraints. The ith particle of the swarm is denoted by an n-dimensional vector in an n-dimensional search space, $Xj = (xj1, xj2\, xjn)$. One more n-dimensional vector $Vi = ({}_{vi1}\,; v_{i2} v_{\text{in}})\, T$represent the particle's velocity.

The previously best decided place to visit of the ith particle is indicated as $P_i = (p_{i1}, p_{i2} \cdots p_{\text{in}})\, T$. The best particle in the swarm is represented by "b." Using the velocity update equation (1), the ith particle's velocity is updated.

The equation of velocity update is

$$vik = vik + c1r1(pik - xik) + c2r2(pbk - xik), \tag{1}$$

and the position is updated by

$$xik = xik + vik, \tag{2}$$

where $k = 1, 2\, n$; $I = 1\,; 2\, S$, where S is the swarm's size; $c1$ and $c2$ are constants, referred to as scaling parameters for social cognitive and behavior, respectively (typically, $c1 = c2$); $r1$ and $r2$ are set up random values uniformly distributed in [0, 1]. The basic form of the PSO algorithm is shown in equations (1) and (2). $V_{\max}$ is a constant that used arbitrarily limit particle velocities and improve the search resolution. Furthermore, to improve the governance of exploration and exploitation, the idea of an inertia weight was devised. The purpose was to do away with the necessity for $V_{\max}$. In 1998 (Shi and Eberhart, 1998) [27], the addition of an inertia weight (w) in the PSO algorithm had been indicated in background. As a result, velocity update equation is

$$vik = w * vik + c1r1(pik\text{-}xik) + (pbk\text{-}xik). \tag{3}$$

According to Eberhart and Shi [28, 68], the best technique is to start with w at 0.9 and linearly reduce it to 0.4, enabling for initial exploration accompanied by acceleration as for a better global optimum.

3.2. Proposed Multipath-Based PSO Routing Protocol (MPSORP). The main attentiveness of previous work of WSN was on dropping energy utilization in improved PSO protocol with cluster energy optimization algorithm. Furthermore, delay initiated and the packet dropped in communication use a lot of energy in the substitute key, which diminishes the life span of the network in an extensive sense. As a result, network life is extended at the expense of large packet losses and increased delay with low throughput. The importance of providing superior performance metrics has received some attention. It is necessary to concentrate on such constraints while expending less energy. This dissertation carries out the following activities in order to implement PSO on the DSDV protocol with cluster energy optimization.

PSO algorithm is primarily a computational technique for optimizing a problem by improving a candidate solution iteratively in relation to a given quality measure. To solve optimization problems, we must to do first formulate the issue in terms of the optimization problem. The optimum path is picked in this suggested algorithm based on fitness value, which could be determined by the lowest distance to be traveled through data up to the base node and the node's energy.

3.2.1. Fitness Function. To use PSO to select the best path, you must first determine the fitness value of each path.

$$\text{Fitnessval} = \text{dis}(i, j) + \text{dist}(j, \text{base}) + \text{remaining energy}(i) + \text{remainingenergy}(j). \quad (4)$$

This fitness value has been utilized to choose the best local and global candidates for PSO. The best optimal solution will be the path with the lowest fitness value.

3.2.2. Calculation of Route Selection of Parameters. Source node, destination node, transmitting range, and power loss are used as an input to calculate the best selection parameter. The route choice parameter is evaluated as follows:

$$\text{TP}(tx)d - \alpha(i, j)Rsp = N\sum k = 1\text{TP}(tx)d - \alpha(i, j) + \alpha 2, \quad (5)$$

where TP (tx) is transmission power, $d - \alpha\,(i, j)$ is distance betwixt nodes i and j, σ is power level of noise, and α is path loss components.

The proposed approach has the following advantages: end-to-end delay is minimized, and energy utilization is lowered due to the small number of routes discovered. The technique uses a routing selection parameter based on the route's power and bandwidth requirements to choose the forwarding nodes or forwarding connection, resulting in a high throughput; consequently, the route chosen is bandwidth aware.

3.2.3. Proposed Algorithm. We have set PSO to find out the superior optimal path while using the least amount of energy. We end up choosing a random variety of solutions from of the set of x solutions. To create an initial solution, x is the total number of solutions. Following the random of initial solutions, we also evaluate the fitness value of each solution using equation (4). After that, we compute the best solution from among them and make it the global and local best starting point. To update the PSO, the update equation is used. For previous solutions, generate new ones, and calculate the nodes of those solutions. These solutions and their nodes are then utilized to calculate each solution's fitness value. The process will be continued until the specified iteration is complete. The superior solution to the other alternatives is replaced based on this continuing iteration as well as fitness value. Pseudocode 1 lists the pseudocode of PSO-based routing algorithm.

4. Framework of Optimization

There are two aspects to our optimization framework. As indicated in Figure 2, one is an optimization method, and the other is called evaluation of solution evaluation. PSO generates sets of solution vectors (also known as new populations) in the optimization process, which are then employed in simulation for performance evaluation. PSO is aimed at identifying the best route through the search space (solution vector). The simulation model runs each solution after receiving a population from PSO into MDSDV one by one. Now, NS-2 takes the WSN realistic instance and the simulation settings listed in Table 1 and they are configured appropriately.

We changed NS-2 in such a way that it now accepts PSO population automatically. Then, in the simulation, NS-2 generates global data known as the trace of simulation. We evaluated the resulting network performance QoS by considering the following five measures in order to compare default and optimized MPSORP:

PDR (packet delivery ratio): the ratio of the amount of complete and accurate packets received by the actual destination to the number of data packets generated by the source.

Average end-to-end delay (AEED): the average amount of time a data packet takes to travel from source to final destination.

NRL (normalized routing load): this represents the average number of routing packets sent to deliver a single data packet.

Throughput: is defined as the ratio of the total amount of data received by a receiver from a sender to the time it takes the receiver to begin receiving the last packet. Packet loss is the number of data packets that are dropped by a protocol.

5. Simulation Model

This research uses the Network Simulator NS2.34 [29] with the Ubuntu10.10 (Linux) operating system to develop the proposed MPSORP. MPSORP, AODV, and DSDV three different routing protocols are compared in this study. The trials are carried out with 25 randomly distributed nodes in rectangular fields measuring 500 km by 500 km for simulation purposes. A rectangular area is chosen to allow for the evaluation of transmission for far-flung nodes. The de facto standard for accessing the communication medium WSNs is IEEE 802.15.4, which describes Medium Access Control (MAC) sublayers and the physical. The size for each data packet is set to 70 bytes. Table 1 shows the parameters chosen as well as their specifications.

5.1. Simulation Setup of IoT Application. In this study, the IoT-App is offered as a nautical data collecting and mapping system through Ship Ad hoc Sensor Networks (SASNET). Ships will be communicating via Very High Frequency (VHF), which is now available on most of the ships and will be outfitted by a variety of sensors such as seat depths, temperature, and wind speed as well orientation. On the ground, 5G base station nodes act as sinks for the collecting data and outfitted with Mobile Edge Computing (MEC) data collection and processing abilities. The sensory data is eventually compiled in a central cloud on the Internet to create publicly accessible, present cartography systems. We explore the suggested system's deployment constraints and advantages and its performance, utilizing our proposed routing protocol MPSORP.

The proposed Internet ocean of things system model is shown in Figure 3. The evolving IoT promises a fully connected world with numerous linked devices and services.

VHF communication is identified as the IoT enablers with the possibility that can provide an entirely new set of previously unavailable apps and services. As a result, overcome has introduced VHF to the radio spectrum for 5G, paving the way for 5G-VHF connection.

As in study, we take advantage of the new 5G era to propose a cartography application in which a SANET is utilized to collect several nautical sensual data from ships and vessels then send it back to onshore sinks that are collocated with 5G base stations that would include dedicated storage as a component of MEC services. As shown in Figure 3, we suggest a new MEC application in which a portion of the edge computing resources is used as an edge repository (cloud) of acquired sensory data that consistently reaches the shore. The edge clouds eventually merge to a central internet cloud. All of the sensory data is consolidated, filtered, and analyzed to provide real-time map of ground and underwater environmental info for beneficiary consumers. The acquired sense data and map information are frequently cached at network edges near users when the network detects significant request. The cartography system shown in Figure 3 can be collected data, for instance, sea state, dimensions, temperature, wind speed, and salinity among other things. The results of the testing and simulations show that the design effectively enables IoT interoperability. We effectively transmitted and received a message utilizing acoustic signals from the deployed underwater sensor node to the border router, which was relayed by a 4G/5G broadband Internet to a laptop and a smartphone linked to the traditional Internet.

As is well known, devices in any WSN system that uses the IEEE 802.15.4 standard can perform one of two functions: Full-Function Device (FFD) or Reduced Function Device (RFD). Coordinator, Device, and Personal Area Network (PAN) coordinator are the three modes of the former (PAN coordinator). The PAN coordinator must be configured on the destination node (sink node) to which the collected information will be sent. The rest of the nodes (those that detect the environment) are normally set up to work in RFD mode. An RFD node cannot send or receive from another RFD node; an FFD node can transmit and receive from all other modes.

6. Discussion and Results

In the proposed MPSORP's performance, compared with DSDV and AODV, five performance metrics is determined by calculating. These performance metrics are PDR, AEED, NRL, throughput, and packet loss (PL).

6.1. Packet Delivery Ratio. Figure 4 illustrates a compared of packet delivery rates for MPSORP, AODV protocol, and DSDV. DS-DV produced the much lowest packet delivery ratio. The maximum PDR reached quite 83% in the MPSORP. In the ocean environment, the speed of the sensor node slowly moving has only 0.47 m/sec based on ocean speed, and subsequently, MPSORP method provides a higher packet delivery rate as in Figure 4.

6.2. End-to-End Delay Average. The graphical outcomes shown in Figure 5 are a measurement of delay, which is a very power full tool. It demonstrates that the end-to-end delay in IEEE 802.15.4 with MPSORP is the shortest of all the routing protocols. This demonstrates that the MPSORP, which is compliant IEEE 802.15.4 standard, is remarkably well suited for delay-sensitive applications. This characteristic can be explained. MPSORP is a proactive routing protocol, and with these protocols, many pathways to destination are accessible right away. To put it another way, routing discovery does not cause any delays. Furthermore, if it is impossible to transport the packets and the MPDSDV, as a result, the routing protocol attempts to drop them and tends to outcome in less delay. It incorporates the delay as a result of route detection procedure and the queue during data packet transmission when calculating the time it takes for a packet to go from the CBR source to the destination. Only the successful data packets—fully delivered to the destinations that the DSDV and MDSDV protocols—show the biggest initial delay required for route setup, of roughly 0.055 seconds. While MPSORP has the shortest latency, at only 0.045 seconds, this might be ascribed to the facts that MPSORP is a multipath protocol based on tables which needs to maintain route tables to reduce the time required for route detection. In contrast, all of other protocols rely on on-demand routing discovery.

6.3. Throughput. In general, the network throughput grows continuously during the simulation time. In all mobility circumstances, MPSORP achieves the best throughput and displays efficient behavior.

The following are some main reasons for this high throughput: first, when the initial data packet comes, it is held until the best routing for a specific destination is discovered. Second, a choice may postpone advertising of routes that are set to change, reducing oscillations in the route tables. By postponing the advert of unsterilized routes, the number of rebroadcasts of routes with the same sequence number is reduced. As shown in Figure 6, this improves the precision of valid routes and tends to result in higher MPSORP throughput at a mobility rate of 4 packets per second.

6.4. The Packet Loss. The comparison of the three methods in terms of data discarded by each procedure is shown in Figure 7. In all circumstances, AODV is the worst in terms of data loss since it broadcasts a route request in the event of a connection failure and then waits for a route reply to receive fresh information; during this time, AODV queues the packets. AODV queue timeout refers to when packets are discarded due to expiration in the AODV queue. DSDV, on the other hand, waits for a length of time to obtain information; if no route is available and a node using DSDV wishes to transmit data, DSDV must queue the packets, which will be dropped if the queue is full.

Because MPSORP uses an alternate path in the event of a link failure, it drops slightly less data than other protocols in sparse networks (networks with 25 nodes).

Packet drop rate happens in mobile sensor networks due to transmission failures, mobility, and congestion. In this setting (marine network), the number of dropped packets is extremely high. In this context, the influence of sparseness in the network is the primary cause of dropped packets (marine network).

6.5. Normalizing Routing Load. The numeral of routing packets transferred through the emulation is known as the routing load normalized. For packets delivered over multiple hops, every transmission of a packet (every hop) counts as such transmission. In other words, it is a comparison of the total number of transferred data packets to the network's total number of control packets.

In order to save energy, it is preferable to have a few control packages as feasible. The normalized routing load vs. pause time is depicted in Figure 8. Due to the large number of route requests launched, the MPSORP has the highest routing load values, while AODV and DSDV have nearly identical figures, in contrast to the AODV protocol, which would have the lowest routing load. As a result, DSDV in concepts of routing cost has the highest routing efficiency, while MPDSDV has the best routing performance in packet delivery. This clearly shows that the optimum packet delivery metrics is frequently associated with a highest routing load.

7. Conclusion and Future Work

An energy-efficiency multipath routing protocol based on the PSO technique is proposed. This method is utilized to choose the best way among all the shortest paths found that uses the least amount of energy from the nodes and allows the wireless network to last for a long period. The suggested technique uses Particle Swarm Optimization to discover the best pathways by taking into account the distance between nodes as well as their energy. Only one optimum way is chosen from these selected paths, decreasing the network's energy consumption. The proposed MPSORP improves the performance of WSN-based IoT applications by increasing the packet delivery ratio, decreasing average delay, and reducing overall energy usage during execution. Other QoS measures, including such normalized routing load, throughput, and reduced data packet losses, have improved as a result of this effort. The proposed algorithm works by calculating the route first and then optimizing it. The simulation outcomes indicate that the proposed method is effective and gives better performance in extending the wireless sensor network lifetime. The outcomes demonstrate that the MPSORP has a higher QoS and better performance than the other routing protocols (AODV and DSDV). Finally, multipath routing algorithms are an effective way to enable real-time and dependability applications with proper QoS over IoT. In future work, PSO can be combined with an existent technique to handle with huge-scale routing challenges utilizing hybrid sensor networks.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Mohammed Zaid Ghawy and Gehad Abdullah Amran contributed equally as first authors. All authors contributed in acquisition, analysis, or interpretation of data and results and review and rewriting of the manuscript.

Acknowledgments

This research was supported by the Researchers Supporting Project number (RSP-2021/244), King Saud University, Riyadh, Saudi Arabia.

References

[1] A. Nandi, B. Sonowal, D. Rabha, and A. Vaibhav, "Centered sink LEACH protocol for enhanced performance of wireless sensor network," in *International Conference on Automation, Computational and Technology Management (ICACTM)*, pp. 436–440, London, United Kingdom, 2019.

[2] T. Zhang and J. Zhang, "A kind of effective data aggregating method based on compressive sensing for wireless sensor network, EURASIP," *Journal on Wireless Communications and Networking*, vol. 15, no. 9, pp. 1–15, 2018.

[3] D. G. Zhang, G. Li, K. Zheng, X. Ming, and Z. H. Pan, "An energy-balanced routing method based on forward-aware factor for wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 766–773, 2014.

[4] M. Z. Hasan and F. Al-Turjman, "Optimizing multipath routing with guaranteed fault tolerance in Internet of Things," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 6463–6473, 2017.

[5] D. G. Zhang, X. Wang, and X. D. Song, "A novel approach to mapped correlation of ID for RFID anti-collision," *IEEE Transactions on Services Computing*, vol. 7, no. 4, pp. 741–748, 2014.

[6] X. D. Zhang, "Design and implementation of embedded uninterruptible power supply system (EUPSS) for web-based mobile application," *Enterprise Information Systems*, vol. 6, no. 4, pp. 473–489, 2012.

[7] R. Jha and S. Ghosh, "Energy efficient particle swarm optimization based multipath routing in WSN," *International journal online of science*, vol. 10, no. 4, 2018.

[8] F. Jameel, W. U. Khan, N. Kumar, and R. Jantti, "Efficient power-splitting and resource allocation for cellular V2X communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3547–3556, 2021.

[9] W. U. Khan, F. Jameel, X. Li, M. Bilal, and T. A. Tsiftsis, "Joint spectrum and energy optimization of NOMA-enabled small-cell networks with QoS guarantee," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8337–8342, 2021.

[10] M. S. Ali, M. S. Islam, M. Asif, W. U. Khan, F. Lin, and O. Waqar, "On efficient DCT type-I based low complexity CE for uplink NB-IoT systems," *IEEE Access*, vol. 9, 2021.

[11] L. Cao, "Task offloading method of edge computing in Internet of Vehicles based on deep reinforcement learning," *Cluster Computing*, vol. 1, pp. 1–15, 2022.

[12] W. U. Khan, E. Lagunas, A. Mahmood, S. Chatzinotas, and B. Ottersten, "Integration of backscatter communication with multi-cell NOMA: a spectral efficiency optimization under imperfect SIC," 2021, http://arxiv.abs/org/2109.11509.

[13] W. U. Khan, T. N. Nguyen, F. Jameel et al., *Learning-based resource allocation for backscatter-aided vehicular networks*, IEEE Transactions on Intelligent Transportation Systems, 2021.

[14] W. U. Khan, F. Jameel, A. Ihsan, O. Waqar, and M. Ahmed, "Joint optimization for secure ambient backscatter communication in NOMA enabled IoT networks," 2021, http://arxiv.abs/org/2111.10872.

[15] W. U. Khan, F. H. Memon, K. Dev, M. A. Javed, D. T. Do, and N. M. F. Qureshi, "Ambient BacCom in beyond 5G NOMA networks: a multi cell resource allocation framework," *TechRxiv*, vol. 1, pp. 1–10, 2021.

[16] W. U. Khan, M. A. Jamshed, E. Lagunas, S. Chatzinotas, X. Li, and B. Ottersten, "Energy efficiency optimization for backscatter enhanced NOMA cooperative V2X communications under imperfect CSI," 2022, http://arxiv.abs/org/2202.01592.

[17] B. Prachin and S. Parul, "Communication technologies and security challenges for internet of things: a comprehensive review," *AEU-International Journal of Electronics and Communications*, vol. 99, no. 2, pp. 81–99, 2019.

[18] D. G. Zhang and X. D. Song, "Extended AODV routing method based on distributed minimum transmission (DMT) for WSN," *AEU-International Journal of Electronics and Communications*, vol. 69, no. 1, pp. 371–381, 2015.

[19] T. Zhang, "Novel self-adaptive routing service algorithm for application of VANET," *Applied Intelligence*, vol. 49, no. 5, pp. 1866–1879, 2019.

[20] K. Zheng and A. Zhang, "A novel multicast routing method with minimum transmission for WSN of cloud computing service," *Soft Comput*, vol. 19, no. 7, pp. 1817–1827, 2015.

[21] T. Zhang, "Novel optimized link state routing protocol based on quantum genetic strategy for mobile learning," *Netw Comput Appl*, vol. 12, no. 2, pp. 37–49, 2018.

[22] D. G. Zhang, "New multi-hop clustering algorithm for vehicular ad hoc networks," *IEEE Trans Intell Transp Syst*, vol. 20, no. 4, pp. 1517–1530, 2019.

[23] S. Liu, "Novel unequal clustering routing protocol considering energy balancing based on network partition & distance for mobile education," *Journal of Network and Computer Applications*, vol. 88, no. 15, pp. 1–9, 2017.

[24] S. Liu, "Novel dynamic source routing protocol (DSR) based on genetic algorithm-bacterial foraging optimization (GA-BFO)," *International Journal of Communication Systems*, vol. 31, no. 18, pp. 1–20, 2018.

[25] Y. M. Tang, "Novel reliable routing method for engineering of internet of vehicles based on graph theory," *Engineering Computations*, vol. 36, no. 1, pp. 226–247, 2019.

[26] C. L. Gong and K. W. Jiang, "A kind of new method of intelligent trust engineering metrics (ITEM) for application of mobile ad hoc network," *Engineering Computations*, vol. 11, pp. 1–13, 2019.

[27] X. H. Liu and Y. Y. Cui, "A kind of novel RSAR protocol for mobile vehicular ad hoc network," *CCF Trans Netw*, vol. 2, no. 2, pp. 111–125, 2019.

[28] S. Liu, "Adaptive repair algorithm for TORA routing protocol based on flood control strategy," *Comput Commun*, vol. 151, no. 1, pp. 437–448, 2020.

[29] X. H. Peng, J. Ren, and L. She, "BOAT: a block-streaming app execution scheme for lightweight IoT devices," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1816–1829, 2018.

[30] X. Wang and X. D. Song, "New medical image fusion approach with coding based on SCD in wireless sensor network," *Journal of Electrical Engineering and Technology*, vol. 10, no. 6, pp. 2384–2392, 2015.

[31] D. G. Zhang, "A new approach and system for attentive mobile learning based on seamless migration," *Applied Intelligence*, vol. 36, no. 1, pp. 75–89, 2012.

[32] K. Zheng and D. X. Zhao, "Novel quick start (QS) method for optimization of TCP," *Wireless Netw*, vol. 22, no. 1, pp. 211–222, 2016.

[33] Y. Y. Zhu, "A new constructing approach for a weighted topology of wireless sensor networks based on local-world theory for the Internet of Things (IOT)," *Computers & Mathematics with Applications*, vol. 64, no. 5, pp. 1044–1055, 2012.

[34] S. Zhou, "A low duty cycle efficient MAC protocol based on self-adaption and predictive strategy," *Mobile Networks and Applications*, vol. 23, no. 4, pp. 828–839, 2018.

[35] H. L. Niu, "Novel PEECR-based clustering routing approach," *Soft Computing*, vol. 21, no. 24, pp. 7313–7323, 2017.

[36] X. Wang and X. D. Song, "New clustering routing method based on PECE for WSN," *EURASIP Journal on Wireless Communications and Networking*, vol. 16, no. 2, pp. 1–13, 2015.

[37] S. Liu, "Dynamic analysis for the average shortest path length of mobile ad hoc networks under random failure scenarios," *IEEE Access*, vol. 7, no. 1, pp. 21343–21358, 2019.

[38] J. X. Gao, "Novel approach of distributed & adaptive trust metrics for MANET," *Wireless Networks*, vol. 25, no. 6, 2019.

[39] T. Zhang, "A kind of novel method of power allocation with limited cross-tier interference for CRN," *IEEE Access*, vol. 7, no. 1, pp. 82571–82583, 2019.

[40] X. H. Liu, "A new algorithm of the best path selection based on machine learning," *IEEE Access*, vol. 7, no. 1, pp. 126913–126928, 2019.

[41] P. Z. Zhao and Y. Y. Cui, "A new method of mobile ad hoc network routing based on greed forwarding improvement strategy," *IEEE Access*, vol. 7.l, no. 1, pp. 158514–158524, 2019.

[42] P. B. Duan, "A unified spatio-temporal model for short-term traffic flow prediction," *Transportation Systems*, vol. 20, no. 9, pp. 3212–3223, 2019.

[43] J. Q. Chen and G. Q. Mao, "Capacity of cooperative vehicular networks with infrastructure support: multi-user case," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1546–1560, 2018.

[44] J. Q. Chen, "A topological approach to secure message dissemination in vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 135–148, 2020.

[45] M. Radi, B. Dezfouli, K. A. Bakar, and M. Lee, "Multipath routing in wireless sensor networks: survey and research challenges," *Sensors*, vol. 12, no. 1, pp. 650–685, 2012.

[46] M. Sharawi, I. A. Saroit, H. El-Mahdy, and E. Emary, "Routing wireless sensor networks based on soft computing paradigms: survey," *International Journal on Soft Computing, Artificial*

Intelligence and Applications (IJSCAI), vol. 2, no. 4, pp. 21–36, 2013.

[47] K. Jaiswal and V. Anand, *EOMR: an energy-efficient optimal ulti-path routing. Wireless Personal Communications*, Springer Science+Business Media, LLC, part of Springer Nature, 2019.

[48] D. Sinwar, N. Sharma, S. K. Maakar, and S. Kumar, "Analysis and comparison of ant colony optimization algorithm with DSDV, AODV, and AOMDV based on shortest path in MANET," *Journal of Information and Optimization Sciences*, vol. 41, no. 2, pp. 621–632, 2020.

[49] F. L. Benmansour and N. Labraoui, "A comprehensive review on swarm intelligence-based routing protocols in wireless multimedia sensor networks," *International Journal of Wireless Information Networks*, vol. 28, no. 2, pp. 175–198, 2021.

[50] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *ACM SIGCOMM computer communication review*, vol. 24, no. 4, pp. 234–244, 1994.

[51] S. Mccanne, S. Floyd, and K. Fall, "Network Simulator 2 (NS-2) version 2.28," 1998, http://www-nrg.ee.lbl.gov/ns/,http://www.isi.edu/nsnam/ns.

[52] K. Vijayalakshmi and P. Anandan, "A multi objective Tabu particle swarm optimization for effective cluster head selection in WSN," *Cluster Computing*, vol. 22, no. S5, pp. 12275–12282, 2019.

[53] A. Raychaudhuri and D. De, *Bio-inspired algorithm for multi-objective optimization in wireless sensor network*, Nature Inspired Computing for Wireless Sensor Networks, Springer, 2020.

[54] S. Singh and P. Kumar, "MH-CACA: multi-objective harmony search-based coverage aware clustering algorithm in WSNs," *Enterprise Information Systems*, vol. 14, no. 9-10, pp. 1325–1353, 2020.

[55] J. Wang, C. Ju, Y. Gao, A. K. Sangaiah, and G. J. Kim, "A PSO based energy efficient coverage control algorithm for wireless sensor networks," *Computers, Materials \& Continua*, vol. 56, no. 3, pp. 433–446, 2018.

[56] J. Wang, Y. Gao, C. Zhou, R. S. Sherratt, and L. Wang, "Optimal coverage multi-path scheduling scheme with multiple mobile sinks for WSNs," *Computers, Materials & Continua*, vol. 62, no. 2, pp. 695–711, 2020.

[57] A. A. M. Rahat, R. M. Everson, and J. E. Fieldsend, "Evolutionary multi-path routing for network lifetime and robustness in wireless sensor networks," *Ad Hoc Networks*, vol. 52, pp. 130–145, 2016.

[58] B. Guruprakash, C. Balasubramanian, and R. Sukumar, "An approach by adopting multi-objective clustering and data collection along with node sleep scheduling for energy efficient and delay aware WSN," *Peer-to-Peer Networking and Applications*, vol. 13, no. 1, pp. 304–319, 2020.

[59] S. H. Liu, W. Zeng, Y. Lou, and J. Zhai, "A reliable multi-path routing approach for medical wireless sensor networks," in *International Conference on Identification, Information, and Knowledge in the Internet of Things (IIKI)*, Beijing, Oct. 2015.

[60] T. Wang, J. Liu, and L. Cheng, "Robust collaborative mesh networking with large-scale distributed wireless heterogeneous terminals in industrial cyber-physical systems," *International Journal of Distributed Sensor Networks*, vol. 13, Article ID 1550147717729640, 2017.

[61] R. Datla, Y. Mai, and N. Wang, *Neighbor coverage multipath DSDV*, California State University, Fresno Fresno. CA. USA, 2018.

[62] M. Sedrati and A. Benyahia, "Multipath routing to improve quality of service for video streaming over mobile ad hoc networks," *Wireless Personal Communications*, vol. 99, no. 2, pp. 999–1013, 2018.

[63] T. Qiu, R. Qiao, M. Han, A. K. Sangaiah, and I. Lee, "A lifetime-enhanced data collecting scheme for the internet of things," *IEEE Communications Magazine*, vol. 55, no. 11, pp. 132–137, 2017.

[64] M. M. Warrier and A. Kumar, "An energy efficient approach for routing in wireless sensor networks," *Procedia Technology*, vol. 25, pp. 520–527, 2016.

[65] O. Deepa and J. Suguna, *An optimized QoS-based clustering with multipath routing protocol for wireless sensor networks*, Journal of King Saud University-Computer and Information Sciences, 2017.

[66] P. Maratha, K. Gupta, and P. Kuila, "Energy balanced delay aware multi-path routing using particle swarm optimisation in wireless sensor networks," *International Journal of Sensor Networks*, vol. 35, no. 1, pp. 10–22, 2021.

[67] A. Sajedi, V. Derhami, L. Mohammad, and A. Mohammad, "Energy-aware multicast routing in manet based on particle swarm optimization," *Procedia Technology*, vol. 1, pp. 434–438, 2012.

[68] R. C. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proceedings of the 2000 Congress on Evolutionary Computation*, vol. 1 of *CEC00 (Cat. No.00TH8512), IEEE*, pp. 84–88, La Jolla, CA, USA, July 2000.

[69] D. Zhang, G. Li, and K. Zheng, "An energy balanced routing method based on forward-aware factor for wireless sensor network," *IEEE transactions on industrial informatics*, vol. 10, no. 1, pp. 766–773, 2014.

[70] B. Mans and N. Shrestha, "Performance evaluation of approximation algorithms for multipoint relay selection," *Proceedings of the 3rd Annual Mediterranean Ad Hoc Networking Workshop*, vol. 1, no. 1, pp. 50–59, 2004.

[71] A. Moussaoui, F. Semchedine, and A. Boukerram, "A link-state QoS routing protocol based on link stability for mobile ad hoc networks," *Journal of Network and Computer Applications*, vol. 39, no. 3, pp. 117–125, 2014.

[72] P. Sondi, "Design guidelines for quality of service support in optimized link state routing-based mobile ad hoc networks," *Ad Hoc Networks*, vol. 11, no. 3, pp. 298–323, 2013.

[73] X. Wang and X. Song, "A kind of novel VPF-based energy-balanced routing strategy for wireless mesh network," *International Journal of Communication Systems*, vol. 30, no. 6, pp. 1–15, 2017.

[74] G. Cervera, "Multipath routing strategy to prevent flooding disruptionattacks in link state routing protocol for MANET," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 744–755, 2013.

[75] L. Chen and J. Zhang, "A multi-path routing protocol based on link lifetime and energy consumption prediction for mobile edge computing," *IEEE Access*, vol. 8, no. 1, pp. 69058–69071, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7353080, 18 pages
https://doi.org/10.1155/2022/7353080

Research Article

Driver's Face Pose Estimation Using Fine-Grained Wi-Fi Signals for Next-Generation Internet of Vehicles

Zain Ul Abiden Akhtar,[1] Hafiz Faiz Rasool,[1] Muhammad Asif,[2] Wali Ullah Khan,[3] Zain ul Abidin Jaffri,[4] and Md. Sadek Ali[5]

[1]*Faculty of Engineering, Department of Information and Communication Engineering, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan*
[2]*College of Electronics and Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China*
[3]*Interdisciplinary Centre for Security, Reliability and Trust (SnT)/ SigCom, University of Luxembourg, Luxembourg*
[4]*College of Physics and Electronic Information Engineering, Neijiang Normal University, Neijiang 641100, China*
[5]*Communication Research Laboratory, Department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh*

Correspondence should be addressed to Md. Sadek Ali; sadek@ice.iu.ac.bd

Received 25 November 2021; Revised 26 March 2022; Accepted 15 April 2022; Published 5 May 2022

Academic Editor: Antonio De Domenico

Driver's behavior and gesture recognition are most significant in the emerging next-generation vehicular technology. Driver's face may provide important cues about his/her attention and fatigue behavior. Therefore, driver's face pose is one of the key indicators to be considered for automatic driver monitoring system in next-generation Internet of Vehicles (IoV) technology. Driver behavior monitoring is most significant in order to reduce road accidents. This paper aims to address the problem of driver's attentiveness monitoring using face pose estimation in a nonintrusive manner. The proposed system is based on wireless sensing, leveraging channel state information (CSI) of WiFi signals. In this paper, we present a novel classification algorithm that is based on the combination of support vector machine (SVM) and K nearest neighbor (KNN) to enhance the classification accuracy. Experimental results demonstrate that the proposed device-free wireless implementation can localize a driver's face very accurately with an average recognition rate of 91.8%.

1. Introduction

With the rapid growing automobile industry, Internet of Vehicles (IoV) has attracted many researchers due to its enormous commercial applications [1–4]. For the next generation fast-paced intelligent vehicles, driver's face estimation may provide important cues to solve many human-centered problems, e.g., driver behavior recognition and driver attention analysis for safe driving. Face localization is a special case of head pose estimation that is widely used in various applications, e.g., saliency prediction, facial expression analysis, and video conferencing [5]. The nominal face orientation while driving is frontal. If the driver's face orientation is in other directions (e.g., tilting down or sideway), this is either due to inattention or fatigue. From the literature, it is evident that driver's head analysis generally points out the attention level of a driver, as well as his/her distraction and fatigue behavior [6, 7].

During the previous decays, various driver assistance systems are designed to avoid accidents. These systems could significantly provide essential information at an early stage to avoid possible accidental risks from occurring. The driver's attention can be analyzed from several visual or nonvisual parameters, i.e., heart rate variability [8], motion of the hands [9] or the feet [10], and gaze tracking [11] or eye blinking [12]. Among the others, face is an important indicator to measure driver's attention that deserves further consideration.

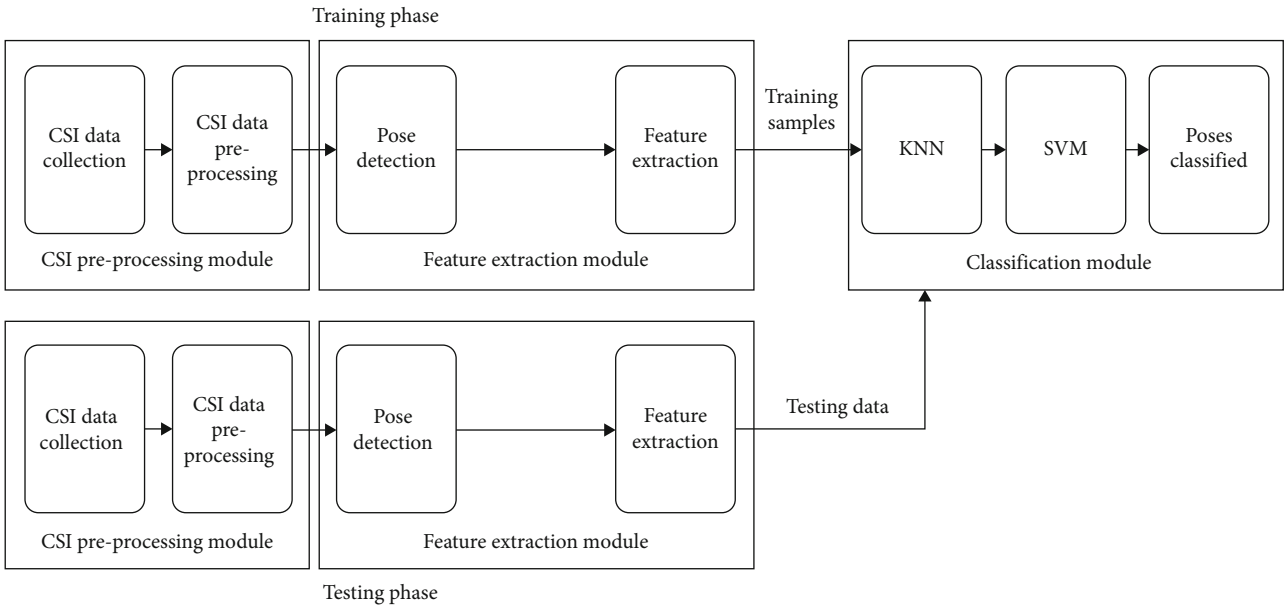


Figure 1: System architecture.

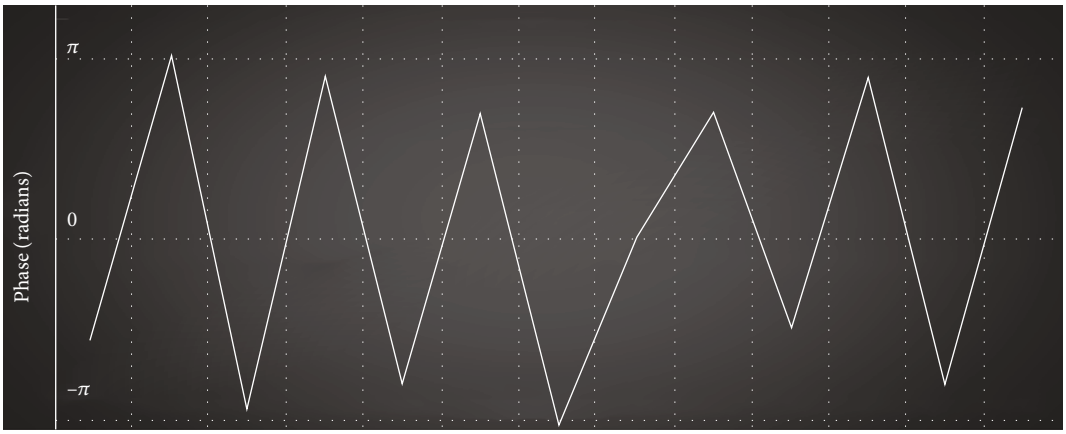


Figure 2: Raw CSI phases.

In general, the choice of the most suitable driver's inattention monitoring system is crucial. For optimal performance, the sensing system should be noninvasive and able to perform accurately under various driving conditions, e.g., nights, clouds, sunrises, and sunsets. During the recent years, WiFi-based activity and gesture recognition systems have been emerged with remarkable performance [13–27], leveraging channel state information (CSI). Motivated by the desire, a WiFi CSI-based wireless sensing framework is presented that is simple yet accurate face localization system to overcome the difficulties of existing methods. Our proposed WiFi CSI-based wireless sensing solution is nonintrusive to user, calibration-free, and can work well in smoke, darkness, line-of-sight, and nonline-of-sight. As low-density parity-check codes (LDPC) scheme has become the significant choice of WiFi (802.11n/ac/ad) [28–31], this innovative idea may also accelerate the benefits of 5G in IoV.

In this research work, we propose a novel hybrid classification technique that is based on the combination of support vector machine (SVM) classifier with K nearest neighbor (KNN), to enhance the recognition performance. Both SVM and KNN have been effectively used for various WiFi-based activity and gesture recognition systems [32–36]. The performance of KNN is dependent on the size of the training samples. As a result, if the size of training samples is very less, it often cannot predict very accurately due to the problem of high variance. Therefore, nearest neighbor classifiers do not have a good generalization capability. Meanwhile, SVM classifier has a good generalization capability because it is based on finding the optimal hyperplane for nonseparable input data. On the other hand, SVM uses all training samples which may cause time-consuming computation. In order to overcome the issues of the high computational burden of SVM and requirement of large sample size of KNN, a combination of SVM and KNN is proposed. The combination of KNN and SVM algorithms (KSVM) yields excellent results and deals in the multiclass setting with reasonable computational complexity in

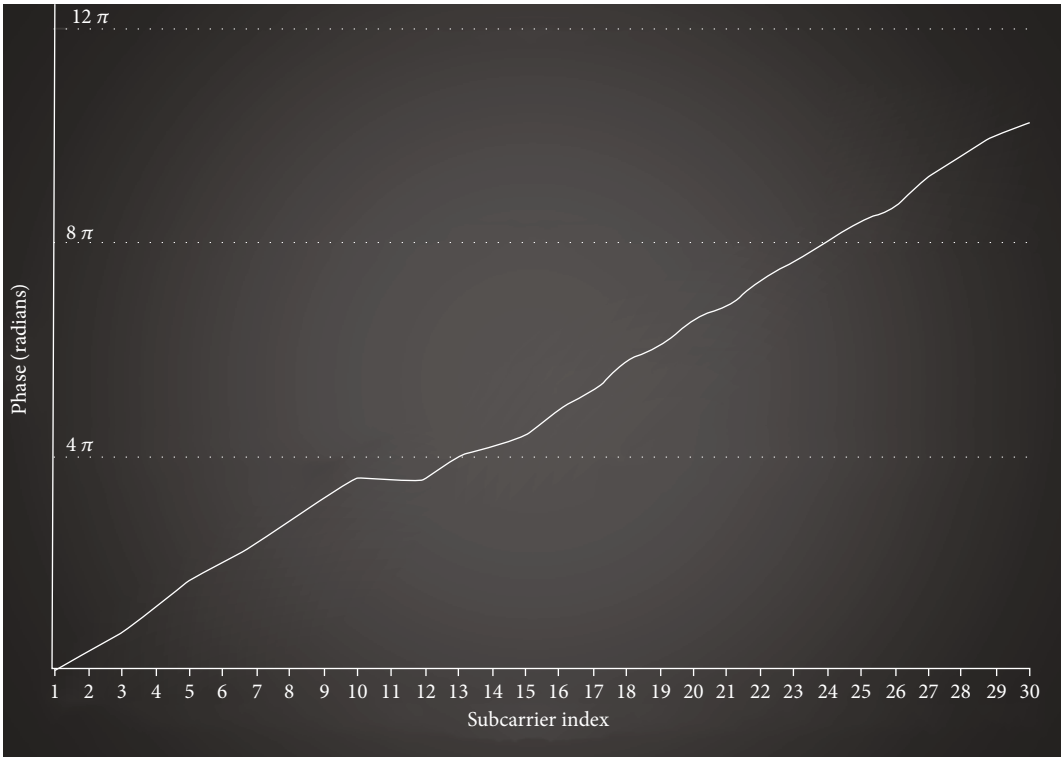
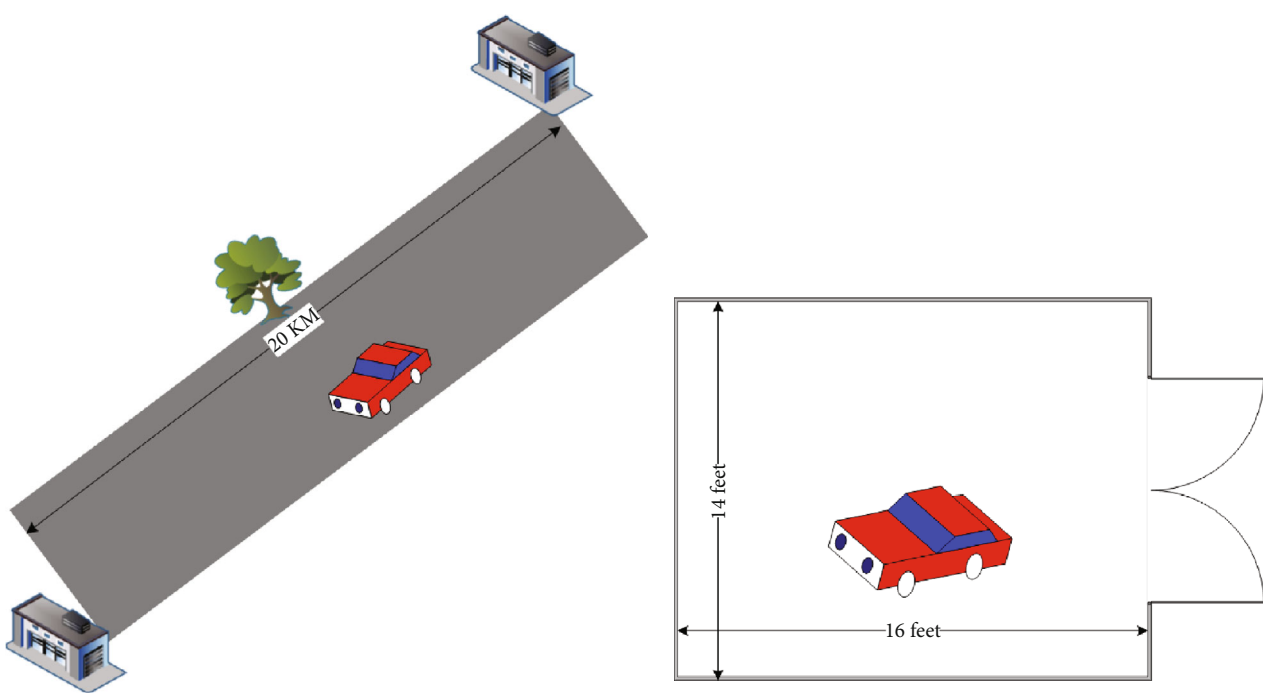


Figure 3: Unwrapped CSI phases.



(a) Experimentation route

(b) Garage layout

Figure 4: Experimentation settings.

practice. The proposed KSVM technique follows two simple steps [37]. First, we find close neighbors to a query sample. Second, we train a local SVM classifier which preserves the minimal values on the set of collected neighbors. The proposed method outperforms nearest neighbor and support vector machines for large and multiclass data sets.

The presented scheme utilizes commercially available WiFi devices to record and acquire channel information,

Table 1: Driver face poses.

Activity type	Activity performed	Activity label
Distraction	Looking over right shoulder	RS
	Looking over left shoulder	LS
	Looking in right rear view mirror	RM
	Looking in left rear view mirror	LM
	Looking at road ahead	RA
	Looking down at infotainment system	LI
	Looking at center rear view mirror	CM
Fatigue	Head tilting down	HD
	Head tilting right	HR
	Head tilting left	HL

which are readily available in the form of CSI measurements [38]. The proposed mechanism leverages the variations in WiFi channel information caused by driver's poses in WiFi coverage area. As per author knowledge, this is the pioneer work for device-free WiFi CSI-based driver's face localization system using the combination of K nearest neighbor with SVM.

Our remarkable contributions are summarized below:

(i) We present a wireless device-free driver's face localization system utilizing CSI of WiFi signals

(ii) We propose a novel hybrid classification method based on the combination of SVM and KNN to increase the recognition accuracy of the system with less computational burden

(iii) To validate the reliability of our presented scheme, comprehensive experiments are performed in cluttered scenarios

(iv) To evaluate the performance of our proposed classification method, we compare our experimental results with conventional classification algorithms

The remaining part of the paper is organized as Section 2 briefly reviews the traditional methods relevant to our presented research work. Section 3 gives the highlights of suggested framework. In Section 4, the detailed implementation of our proposed system is discussed. In Section 5, we explain the experimentation settings and results. Section 6 demonstrates the main limitations of the presented technique. Finally, Section 7 provides the conclusion with future suggestions.

2. Related Work

This section will review the existing WiFi CSI-based device-free gesture and activity recognition systems relevant to our research work. WiFi-based fine-grained physical layer CSI has attracted many scientists because of its high localization accuracy. With the pervasiveness in wireless sensing technology, WiFi-based device-free indoor localization has been entered into a modern era of life [14–16]. The emerging device-free activity recognition takes the advantages of WiFi CSI for the characterization of human activities [39–42]. WiFi-based localization and recognition has been extended to Wi-COVID [26], a WiFi-based COVID-19 detection and patient monitoring system.

Recently, WiPass [25] introduced WiFi CSI-based smartphone keystroke recognition. The WiFi CSI-based intrusion detection [33, 43, 44] and microactivity recognition [45] systems have been presented with remarkable recognition performance. In recent years, ubiquitous WiFi-based training-free localization system has been suggested with good recognition results [46]. The authors of [17] presented a multiuser gesture recognition system using WiFi signals. DeepSeg [22] and Wihi [23] worked on WiFi-based activity recognition using deep learning approach. During the recent years, WiFi-based posture recognition system has been developed with good recognition results [21]. WiAct [19] proposed a device-free passive activity recognition system exploiting the correlations between WiFi CSI amplitude information and human body movement. The authors of [20] demonstrated the concept of temporal frequency for WiFi-based human activity recognition.

WiFi vision [13] introduced the idea of indoor positioning, WiFi imaging, daily activities recognition, gesture recognition, gait recognition, human identification, fall detection, and human detection using WiFi devices. WiGer [47] presented a WiFi-based hand gestures recognition system via a fast dynamic time warping algorithm. In recent years, the idea of writing in air [18, 48] is introduced for virtual reality devices using WiFi signals which is much complex in comparison to simple gestures recognition. DF-WiSLR [27] is a sign language recognition model exploiting WiFi signals.

Traditionally, support vector machine (SVM) and K nearest neighbor (KNN) methods have been widely used in numerous device-free WiFi CSI-based localization, activity, and gesture recognition systems, as stand-alone classifier. In this context, WiCatch [49] presented a WiFi CSI-based hand gesture recognition system leveraging SVM classification method. Wi-Key [34] is WiFi CSI-based system to recognize keystrokes using KNN classifier. WiFall [35] is WiFi CSI-based abnormal behavior detection system leveraging local outlier factor. In this model, one-class SVM is used to successfully classify the features.

In recent decades, WiFi-based driver's in-vehicle activity and gestures recognition systems have been introduced with good recognition performance [50–52]. WiFind [36] presented a WiFi CSI-based driver fatigue detection system. This model is based on one-class SVM technique. WiDriver [53] is dependent on driver's hand movements to recognize driver actions using CSI of WiFi signals. Different from existing systems, we use hybrid classification approach, i.e., combination of KNN and SVM algorithms for WiFi CSI-based driver's face localization.

Actual activity performed	RS	LS	RM	LM	RA	LI	CM	HD	HR	HL
RS	0.91	0.01	0.06	0	0	0	0.01	0	0.01	0
LS	0.02	0.90	0.02	0.03	0.01	0.01	0	0	0	0.01
RM	0.05	0.02	0.90	0	0	0	0	0.01	0.01	0.01
LM	0	0.03	0.01	0.91	0	0.01	0	0.01	0	0.03
RA	0.01	0	0	0.01	0.89	0.04	0.03	0.01	0.01	0
LI	0	0	0.01	0	0.02	0.92	0.01	0.03	0	0
CM	0	0	0	0.01	0.04	0.01	0.93	0.01	0	0
HD	0.01	0	0	0.01	0.01	0.06	0.01	0.90	0	0
HR	0.01	0	0.01	0	0.02	0.01	0	0.02	0.92	0.01
HL	0	0.02	0	0.01	0	0.01	0	0.02	0.01	0.93

Activity classified

(a) Scenario-I

Actual activity performed	RS	LS	RM	LM	RA	LI	CM	HD	HR	HL
RS	0.93	0	0.04	0	0	0.01	0.01	0	0.01	0
LS	0.02	0.90	0.02	0.03	0	0.01	0	0	0	0.02
RM	0.04	0.02	0.91	0.01	0.01	0	0	0	0.01	0
LM	0	0.02	0.01	0.92	0.01	0.01	0	0.01	0	0.02
RA	0	0	0	0	0.91	0.01	0.03	0.04	0.01	0
LI	0	0	0.01	0	0.01	0.94	0.01	0.03	0	0
CM	0	0	0	0.01	0.03	0.02	0.92	0.02	0	0
HD	0	0	0	0.01	0.01	0.03	0.01	0.92	0.01	0.01
HR	0.02	0	0.02	0	0.02	0.01	0	0.01	0.91	0.01
HL	0	0.01	0	0.02	0	0.02	0	0.02	0.01	0.92

Activity classified

(b) Scenario-II

Figure 5: Confusion matrix of activity recognition using our proposed algorithm.

3. System Overview

In this section, we will demonstrate the important facts about WiFi CSI, basic system architecture, and overview of our presented classification mechanism.

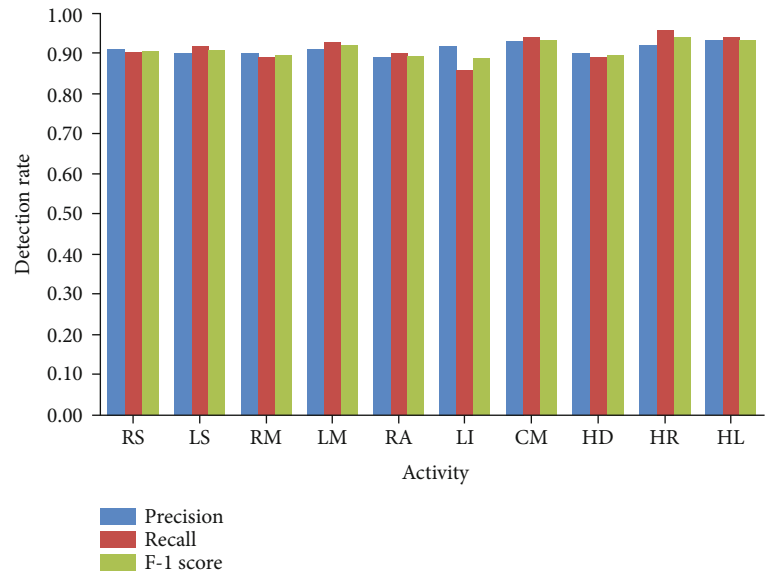
3.1. CSI Overview. The proposed system is based on WiFi devices having IEEE 802.11n/ac enabled protocols. The channel state information (CSI) of WiFi signal is used as information source. Commercially available off-the-shelf WiFi devices that exploit IEEE 802.11n/ac usually support multiple-input multiple-output (MIMO) technology and thus comprises of multiple transmitter (Tx) and receiver (Rx) antennas. The CSI of WiFi signal is fine-grained information containing physical layer data that is based on widely used orthogonal frequency division multiplexing (OFDM) technology.

In this work, a WiFi router or access-point is used as a transmitter. The receiver is an Intel 5300 NIC that is used to collect CSI information from physical layer. Both transmitter and receiver are enabled with IEEE 802.11n protocol. The channel properties are usually available in the form of CSI measurements on commercial WiFi devices [38]. Intel 5300 network interface card (NIC) supports 30 subcarriers to records and acquire the channel variations of each CSI Tx-Rx antenna pair in orthogonal frequency division multiplexing (OFDM) system.

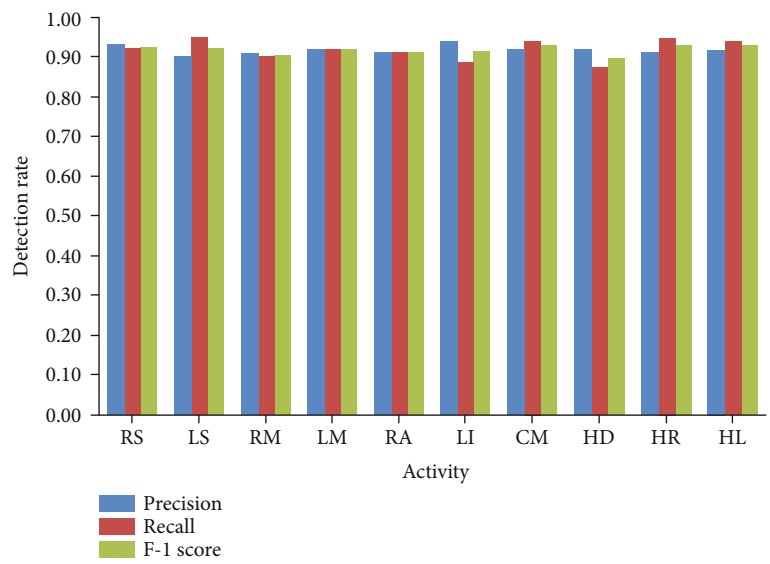
The commonly used narrowband flat-fading channel for i^{th} packet, leveraging OFDM and MIMO technique is formulated as

$$Y_i = H_i X_i + \mathcal{N}_i \; i \in [1, N], \tag{1}$$

where Y_i is the received signal, and X_i is transmitted signal.



(a) Scenario-I



(b) Scenario-II

Figure 6: Precision, recall, and $F-1$ score.

Table 2: Overall detection rate with precision, recall, and F_1-score.

Experiment	Average rate (%)					
	Precision		Recall		F_1-score	
	Min.	Max	Min.	Max.	Min.	Max.
Scenario-I	89	93	86	96	89	94
Scenario-II	90	94	88	95	90	93

N is the total number of packets received. H_i refers to the CSI channel matrix for i^{th} packet, and $\mathcal{N}_i$ represents the Gaussian noise vector.

Let $\mathbb{N}_{Rx}$ and $\mathbb{N}_{Tx}$ refer to the total number of receiving and transmitting antennas, respectively. For each stream, the CSI matrix comprises of $\mathbb{N}_{Tx} \times \mathbb{N}_{Rx} \times 30$ complex values. For each Tx-Rx antenna pair CSI matrix, H is presented as

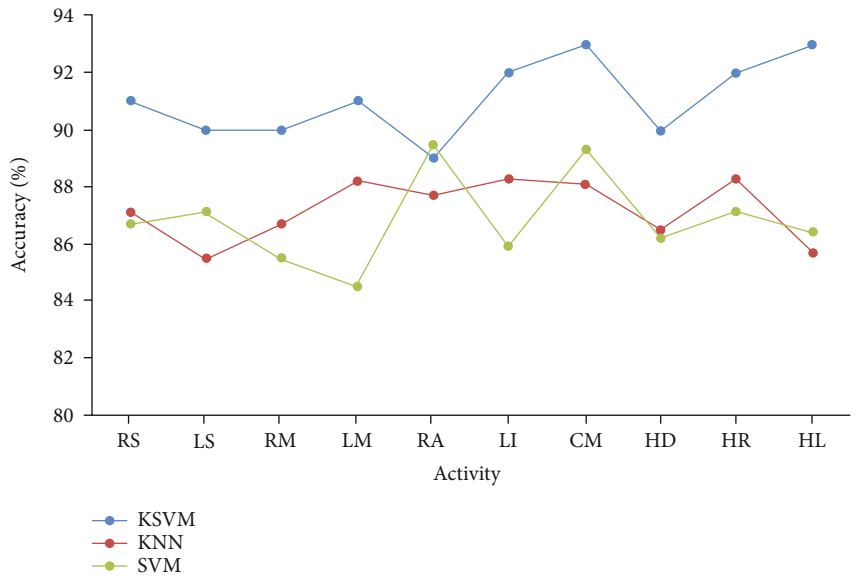
$$H_i = [h_1, h_2, \cdots, h_{30}]\ i \in [1, N], \tag{2}$$

where h carries both phase and amplitude measurements in the form of complex number; calculated as

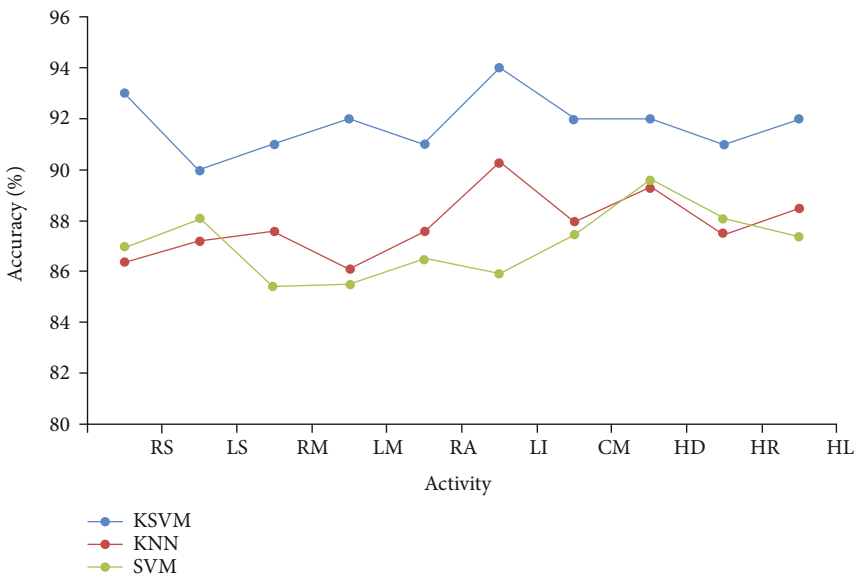
$$h = |h| e^{j\sin\{\angle h\}}, \tag{3}$$

where $|h|$ stands for the amplitude while $\angle h$ denoted the phase information.

3.2. System Architecture. Our WiFi CSI-based driver's face localization system consists of following three main modules: (1) CSI preprocessing module, (2) feature extraction module, and (3) classification module, as shown in Figure 1.



(a) Scenario-I



(b) Scenario-II

Figure 7: Comparison of KSVM with stand-alone KNN and SVM classifiers.

Table 3: Comparison of KSVM with stand-alone KNN and SVM classifiers.

Experiment	Average recognition accuracy (%)		
	SVM	KNN	KSVM
Scenario-I	86.8	87.2	91.1
Scenario-II	87.1	87.9	91.8

CSI preprocessing module collects and acquires the CSI data from physical layer. Using basic filtering techniques, the CSI data is preprocessed to remove unwanted noises. Feature extraction module detects the relevant face poses and extracts the meaningful features from pre-processed CSI data. The classification module relies on hybrid classification technique to recognize different poses. In the next section, each module is explained briefly.

4. Methodology

In this section, we will explain the complete flow of our system methodology, i.e., CSI preprocessing, pose detection, features extraction, and classification.

4.1. CSI Preprocessing. The acquired CSI data is a composite signal which comprises of useful information and embedded unwanted noises from the surroundings. First, the CSI received information is filtered using basic filtering techniques. From the literature, it is clear that the human

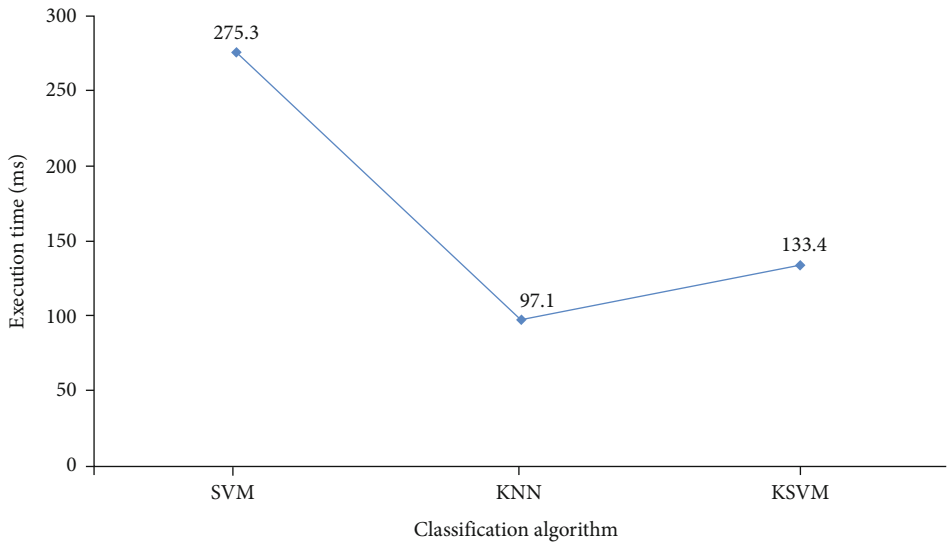


Figure 8: Comparison of execution time.

activities and poses have less frequency in comparison to the frequency of noise [43]. Therefore, we need to remove high-frequency noise. For the purpose, a second order low pass Butterworth filter is implemented. In our experiments, the packets sampling rate (F_s) is adjusted at 80 packets/second that is equal to the normalized cutoff frequency $w_n = 2\pi f / F_s = 0.025\pi$ rad/sec. The received CSI data may be affected by some static path components. Therefore, we subtract the corresponding constant offsets from the streams to alleviate these static path components. The raw CSI data cannot be directly used and can give wrong information because it is wrapped between $-\pi$ and π [54]. Therefore, it is required to unwrap the measured CSI data. The raw and unwrapped CSI data are shown in Figures 2 and 3, respectively.

Due to unsynchronized time clock between transmitter and receiver, the CSI raw phase data behaves extremely random. The relation between true phase and measured phase can be formulated as

$$\angle \widehat{h}_i = \angle h_i + 2\pi \frac{n_i}{N} \Delta t + \beta + z, \tag{4}$$

where $\angle \widehat{h}_i$ represents the measured phase of ith subcarrier while $\angle h_i$ shows the actual phase, Δt is the time lag, n_i stands for the subcarrier index, N is used for the size of FFT, β represents the unknown phase offset, and z indicates the random noise.

The phase error $2\pi n_i / N\Delta t + \beta$ is linear function of subcarrier index n_i. We can formulate two calibration parameters A and B for phase represented as

$$A = \frac{\angle \widehat{h}_{i_N} - \angle \widehat{h}_{i_1}}{i_N - i_1}, \tag{5}$$

$$B = \frac{1}{N} \sum_{k=1}^{N} \angle \widehat{h}_{i_k}. \tag{6}$$

We subtract $A_{i_k} + B$ from raw phase $\angle \widehat{h}_{i_k}$ and get the sanitized phase $\angle \tilde{h}_{i_k}$ as

$$\angle \tilde{h}_{i_k} = \angle \widehat{h}_{i_k} - A_{i_k} - B. \tag{7}$$

The phase sanitization is performed on all the subcarriers and reassembled according to the corresponding amplitudes.

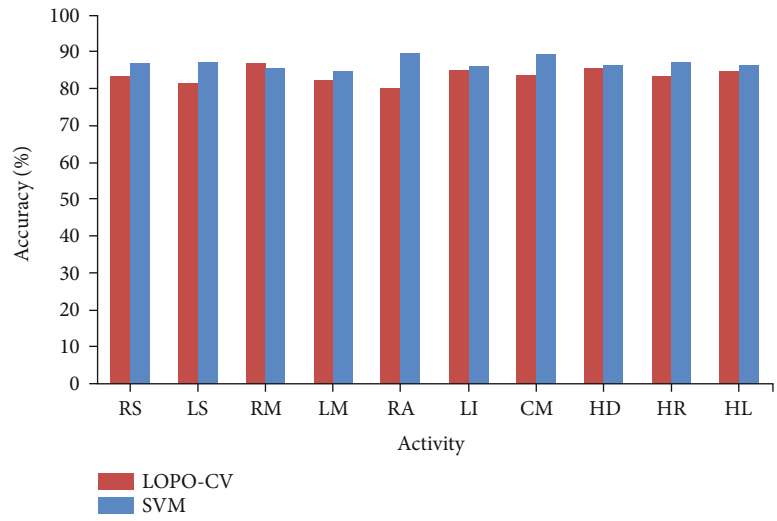
4.2. Pose Detection. In this section, we discuss how to detect the driver's face-related activities or poses using the received CSI data and how we can distinguish it from other in-vehicle activities. The first step towards driver's face localization using CSI data is to detect whether a driver has performed some activity or not. For this purpose, the meaningful CSI streams caused by human motion are initially segmented using step-by-step moving variance technique [41]. In each step, a sliding window of length L is used across neighboring CSI packets. The high values in moving variance are caused by dynamic motions while, on the other hand, low values mean slight fluctuation due to surroundings. The moving variance is given as

$$\mathscr{V}(\text{CSI}) = \sum_{m=1}^{n} \left[\frac{1}{W-1} \sum_{k=1}^{W} |\text{CSI}_{k \in W} - \mu|^2 \right], \tag{8}$$

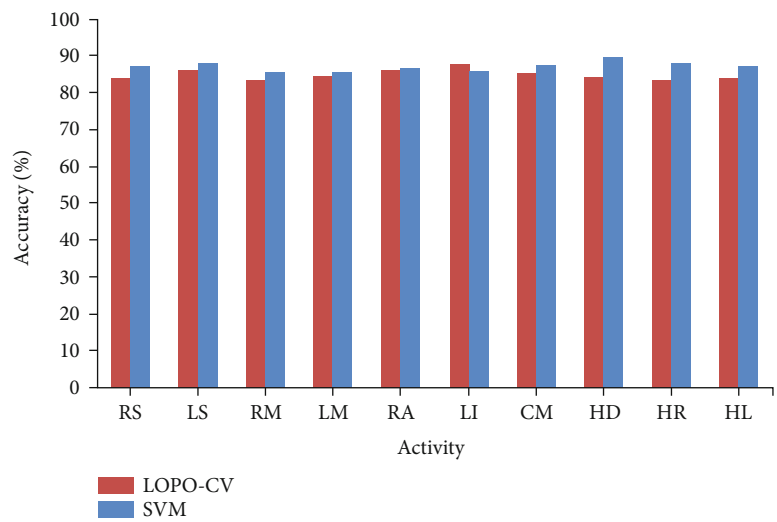
here, μ is the mean and can be defined as

$$\mu = \frac{1}{W} \sum_{k=1}^{W} \text{CSI}_k, \tag{9}$$

where n is the number of packets. k is the packet number in sliding window while m is the packet number in CSI stream. Afterward, we calculate the variance for 30 CSI filtered streams and obtain a $30 \times t$ matrix, where t is the sliding window moving times. After detecting the moving part of



(a) Scenario-I



(b) Scenario-II

Figure 9: Accuracy test using LOPO-CV scheme with SVM.

CSI sequence, we accurately detect the face-related activities. Through experimental investigations, we find that the face activities cause large fluctuation in CSI streams. Therefore, we apply a threshold based technique. We set a threshold τ_{th} and compare the maximum value of $\mathscr{V}(CSI)$ with τ_{th} for activity related component selection. Mathematically,

$$\Psi = \begin{cases} 1, & \max\left(\mathscr{V}(\text{CSI})\right) \tau_{\text{th}}, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

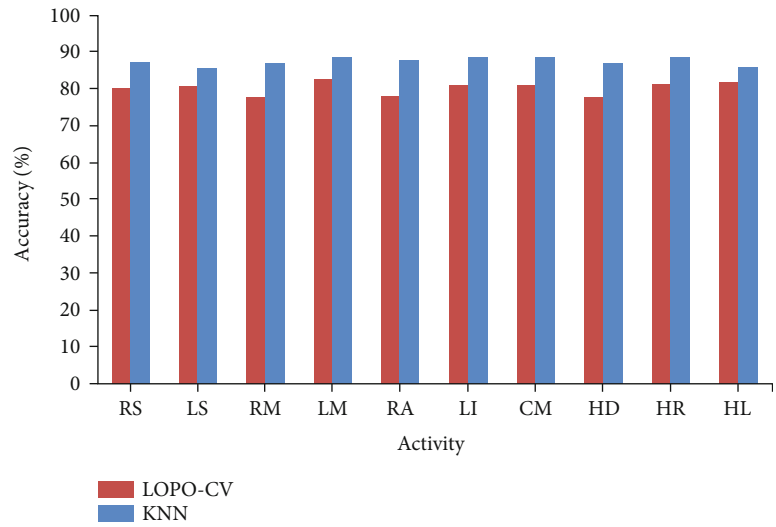
where Ψ represents the presence of face-related activity. If Ψ is set to 1, it means face-related activity is detected. The value of threshold τ_{th} is empirically selected based on the variance of our preliminary measurements, which varies with the activities. To reduce the dimensionality of the acquired data, we perform principle component analysis (PCA) as demonstrated in [55]. After getting the required profile, we can apply extract feature extraction method.

4.2.1. Feature Extraction. Based on the detailed analysis of extracted moving average data and our preliminary experimental investigations, we specifically choose following six statistical features: (i) mean, (ii) standard deviation, (iii) median absolute deviation, (iv) maximum value, (v) 25th percentile, and (vi) 75th percentile. In order to differentiate multiple poses features, we integrate all obtained features into a tuple $\mathscr{F}$ and can be defined as

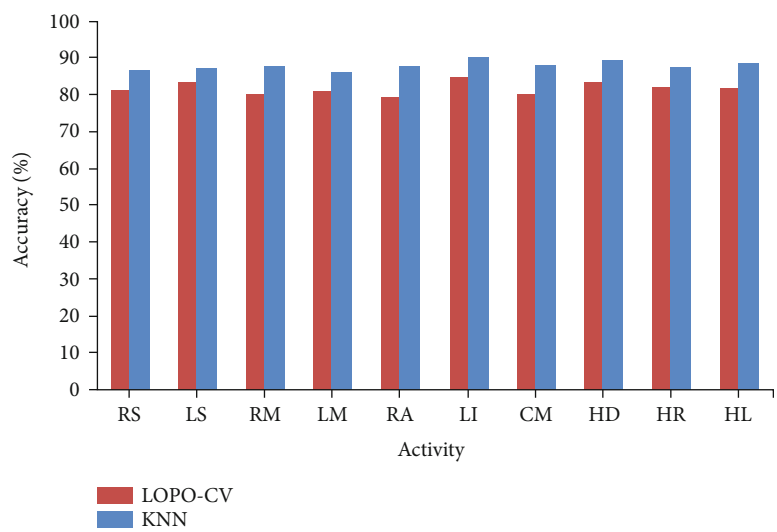
$$\mathscr{F} = \{f_1, f_2, \cdots, f_n\} \subset \mathbb{F}, \tag{11}$$

where $\mathbb{F}$ represents the set of all features, and f_i stands for a feature.

4.3. Classification Module. Our presented KSVM classification algorithm combines the benefits of both KNN and SVM. It finds the K nearest neighbors of query sample x and then trains an SVM classifier to perform the recognition [37].



(a) Scenario-I



(b) Scenario-II

Figure 10: Accuracy test using LOPO-CV scheme with KNN.

4.3.1. KNN Classifier. The K nearest neighbor (KNN) is a very simple machine learning algorithm that is based on statistical data for the classification of features. The KNN method chooses the most relevant class of a testing sample from the available K nearest samples. Suppose we have a training dataset D can be represented as

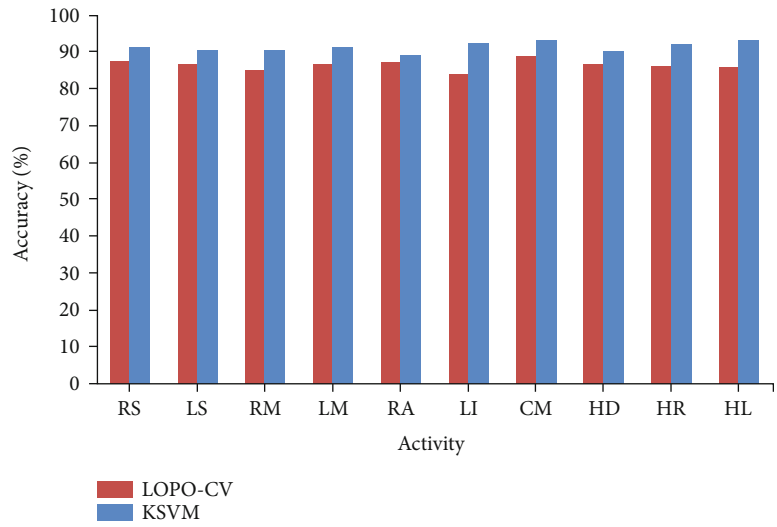
$$D = (c_1, l_1), (c_2, l_2), \cdots \cdots (c_n, l_n), \tag{12}$$

where n is the number of samples $c_i (i = 1, 2, \cdots, n)$, and l is the corresponding class label $l_i (i = 1, 2, \cdots, n)$.

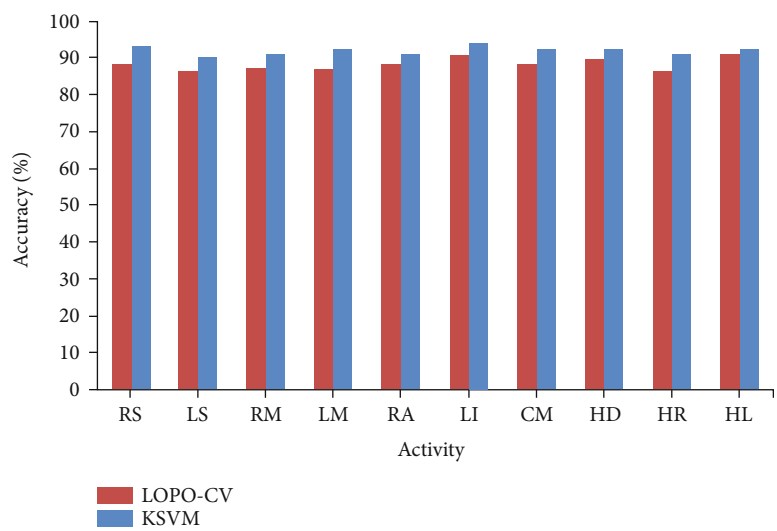
According to K nearest neighbor rule, any unknown sample c is assigned to the class obtained by majority voting of its K nearest neighbors in dataset D. For binary classification problems, the decision rule of the KNN classifier is mathematically represented as

$$\mathrm{KNN}(c) = \mathrm{sign}\left(\sum_{i=1}^{K} c_i\right). \tag{13}$$

4.3.2. SVM Classifier. The support vector machine (SVM) is a linear discriminant classification algorithm. This machine learning method is based on the principle of maximizing the classification effect by establishing a hyperplane. The classification hyperplane provides the decision surface and maximizes the isolation boundary between different types of samples. A support surface is drawn on both sides of hyperplane containing the samples that are close to classification interface. The training samples on the support surface are called support vectors.



(a) Scenario-I



(b) Scenario-II

Figure 11: Accuracy test using LOPO-CV scheme with KSVM.

Table 4: Overall generalization performance.

Experiment	Average accuracy (%)					
	SVM		KNN		KSVM	
	Original	LOPO-CV	Original	LOPO-CV	Original	LOPO-CV
Scenario-I	86.8	83.7	87.2	80.2	91.1	86.4
Scenario-II	87.1	84.9	87.9	81.7	91.8	88.2

The training dataset D is defined as

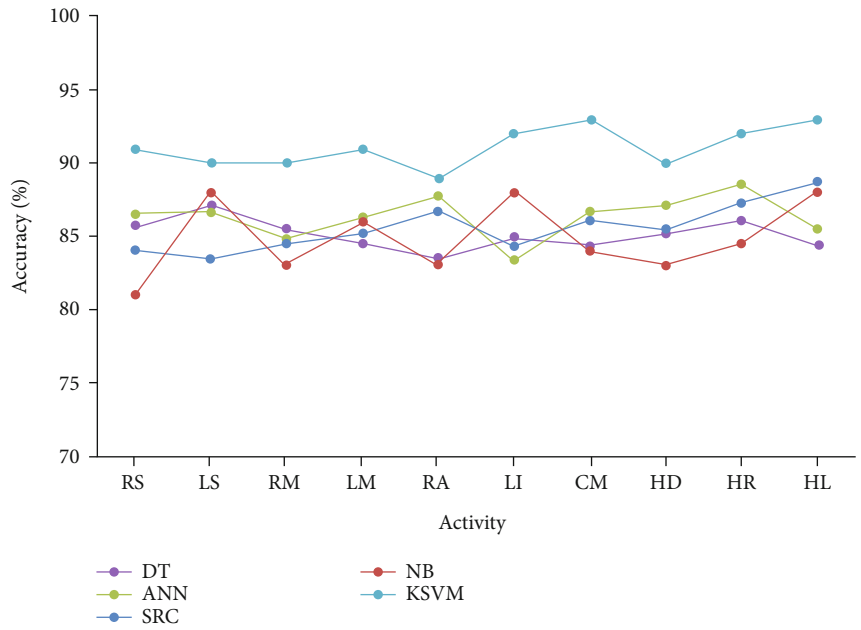
$$D = [(c_1, d_1), (c_2, d_2), \cdots, (c_n, d_n)], \tag{14}$$

where n is the number of samples. Each sample c_i is a d -dimensional vector while class label d_i is either 1 or -1.

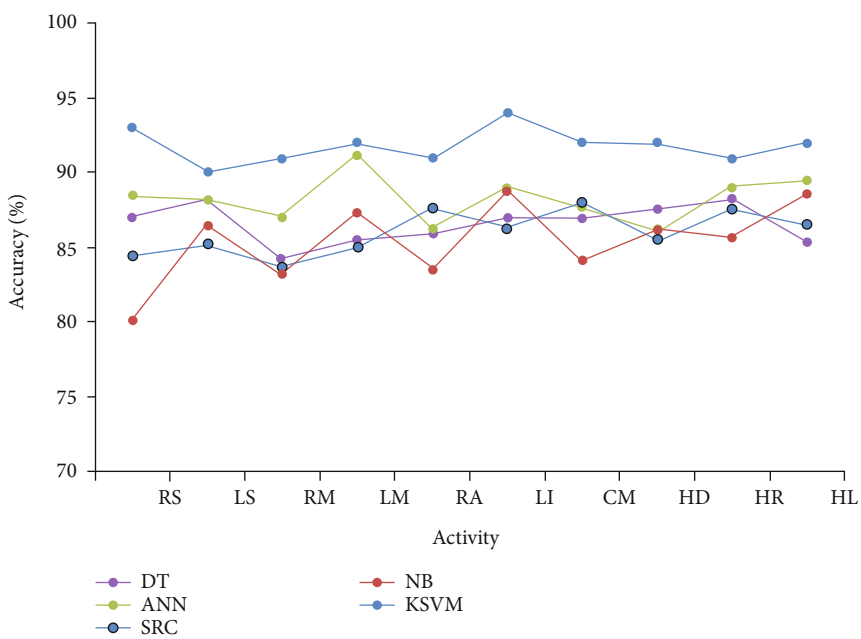
We can map the samples to a feature space of higher dimensions using a nonlinear mapping transformation function Φ. The SVM classifier takes the following decision rule:

$$\mathrm{SVM}(c) = \mathrm{sign}\left(\sum_{i=1}^{L} a_i d_i K(c_i c) + e\right), \tag{15}$$

where $K(a, e) = \Phi(a).\Phi(e)$ is the kernel function. The value of a_i and e are adjusted to maximize the marginal distance of separating hyperplane. The sign function gives the results in binary form. To obtain results in nonbinary form, we



(a) Scenario-I



(b) Scenario-II

Figure 12: Comparison of KSVM with state-of-the-art classifiers.

Table 5: Comparison of accuracy with state-of-the-art classification methods.

Experiment	Average recognition accuracy (%)				
	DT	ANN	SRC	NB	KSVM
Scenario-I	85.3	86.3	85.1	84.9	91.1
Scenario-II	86.5	88.4	86	85.4	91.8

remove the sign function:

$$\mathrm{SVM}(c) = \left(\sum_{i=1}^{L} a_i d_i K(c_i c) + e \right). \tag{16}$$

4.3.3. KSVM Classifier. The basic idea of KSVM method follows a very simple procedure to classify the features. In the first step, it finds the K nearest neighbors to query sample c. Afterward, it trains an SVM that is used to perform the recognition. Finally, the samples with minimum values of $K(c_i, c_i) - 2K(c_i, c)$ are considered to be the closest samples

Table 6: Classification accuracy (%) of KSVM with varying K values.

K value	Gesture class									
	RS	LS	RM	LM	RA	LI	CM	HD	HR	HL
$K = 1$	87.5	88.2	88.4	88.7	89.1	86.9	89.2	88.1	90.3	89.7
$1 < K < 5$	<89.2	<88.9	<90.2	<89.7	<88.5	<89.3	<90.6	<87.7	<90.4	<90.4
$K = 5$	90.4	91.1	92.6	92.5	89.3	90.7	91.4	89.9	92.2	92.1
$5 < K < 10$	<90.5	<89.3	<91.7	<92.1	<89.7	<90.3	<89.9	<90.2	<91.6	<91.0
$K = 10$	89.3	91.0	90.5	90.6	89.4	89.6	90.1	88.9	90.6	90.1
$10 < K < 15$	<89.6	<91.3	<91.7	<88.5	<88.9	<90.3	<90.7	<88.4	<91.0	<90.4
$K = 15$	88.3	90.7	90.6	91.1	89.6	90.1	88.9	89.0	90.5	88.6
$15 < K < 20$	<90.5	<89.1	<89.5	<90.2	<89.8	<89.6	<88.7	<89.0	<90.3	<90.6
$K = 20$	87.5	90.0	88.7	87.3	88.1	87.7	89.3	88.4	89.0	87.6
$20 < K < 25$	<87.3	<89.4	<89.1	<88.2	<89.0	<86.9	<88.1	<87.5	<86.8	<88.2
$K = 25$	89.1	87.9	90.0	88.8	89.5	86.7	88.3	87.5	88.1	86.2
$25 < K < 30$	<86.3	<89.5	<89.1	<89.0	<87.7	<86.5	<88.2	<86.9	<90.0	<89.5
$K = 30$	86.5	86.3	87.6	89.0	86.3	88.9	88.2	90.1	86.2	87.8

to sample c. In any transformed feature space, the following inequality can be used for the said purpose:

$$\|\Phi(c_i) - \Phi(c)\|^2 = \Phi^2(c_i) + \Phi^2(c) - 2\Phi(c_i)\Phi(c) = K(c_i, c_i) + K(c, c) - 2K(c_i, c). \tag{17}$$

The final decision function follows

$$\mathrm{KSVM}(c) = \mathrm{sign}\left(\sum_{i=1}^{K} a_i d_i K(c_i c) + e\right). \tag{18}$$

5. Experimentation and Evaluation

This section describes the experimentation settings and evaluation performance of our proposed KSVM scheme.

5.1. Experimentation Settings. In our presented framework, all experiments are performed with 802.11n enabled WiFi devices as described in [50–52]. Specifically, a laptop is used as a receiver that is equipped with Intel 5300 NIC and three receiving antennas, i.e., $\mathbb{N}_{Rx} = 3$. To record and acquire CSI data, we run 802.11n CSI Tool [38] on the receiver with Ubuntu 11.04 LTS operating system. A single antenna TP-Link router, i.e., $\mathbb{N}_{Tx} = 1$ is used as a transmitter or access-point that operates at frequency of 2.4 GHz. In this experiment, the receiver pings the access-point that is set at 80 packets/s. The system is generating 3 CSI streams of 30 subcarriers each forming a 1×3 MIMO system with channel bandwidth of 20 MHz. We have used MATLAB R2016a to perform signal processing. We setup our testbed in a locally manufactured vehicle which is not equipped with preinstalled WiFi devices. Due to unavailability of WiFi access-point in our test vehicle, we installed a commercially available router (TP-Link) as access-point, placed on the dashboard in front of driver's seat. To acquire CSI information, the receiver (laptop) is configured at copilot's seat. To evaluate the performance of our presented method, the following two scenarios are chosen:

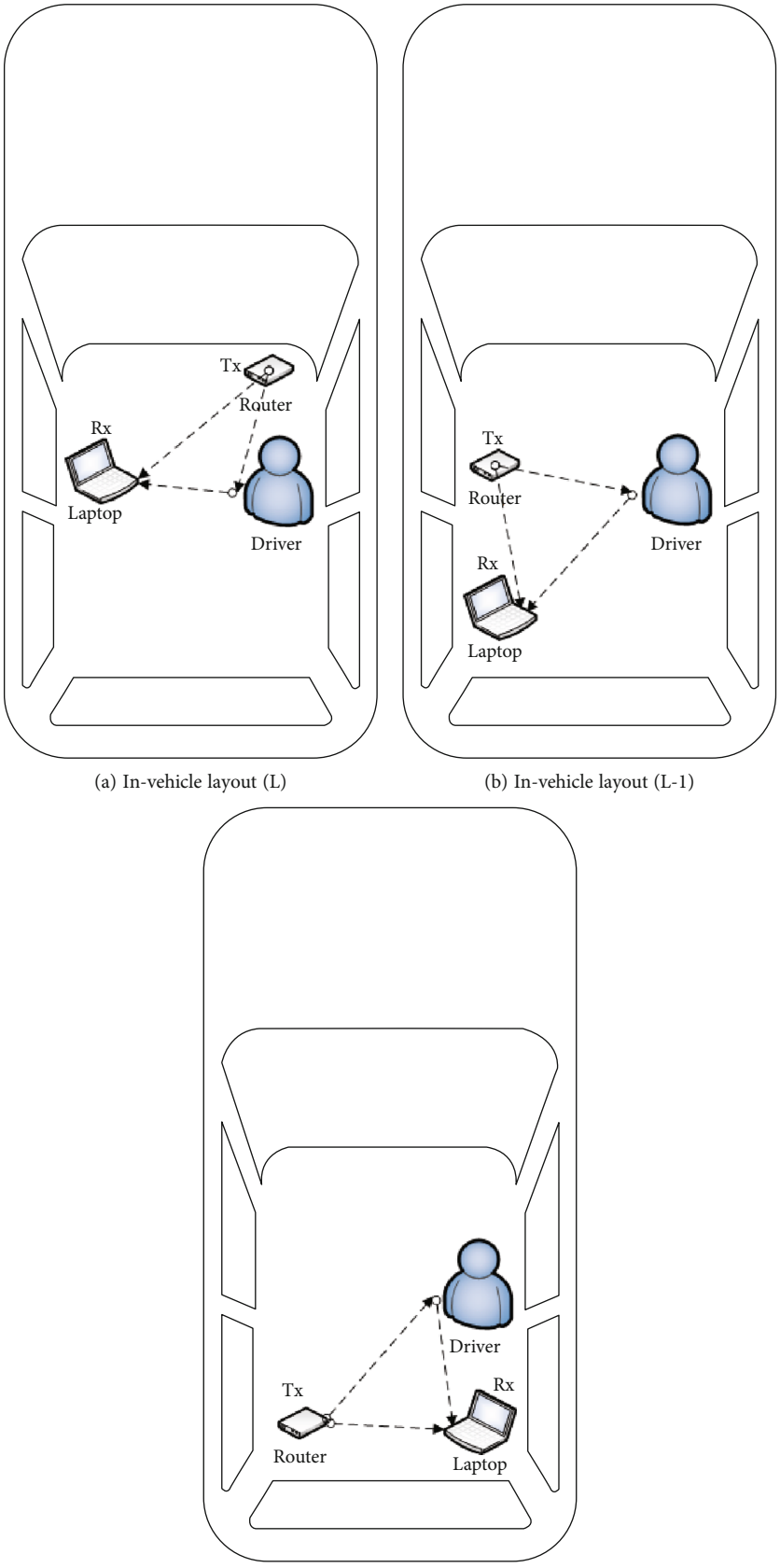
(i) *Scenario-I.* This is the actual driving scenario where all prescribed poses are performed while driving a vehicle. The vehicle is driven at an average speed of 20 km/hr on a straight road of 20 km long, as shown in Figure 4(a). To avoid the interference of other in-vehicle activities, during pose performance, no other activity is performed

(ii) *Scenario-II.* In this scenario, all prescribed poses are performed in a vehicle standing in a garage. The garage size is 14×16 feet, as shown in Figure 4(b)

In each experiment, 10 in-vehicle human poses, as shown in Table 1, are performed by five volunteers (3 male and 2 female university students). For each experiment, each volunteer repeated all poses 20 times. Therefore, the data set consists of total 1000 samples ($5 - \text{volunteers} \times 10 - \text{poses} \times 20 - \text{times repeated}$) for each experiment. The 50% of the total samples are used for training purpose while 50% for testing purpose. For cross-validation, we keep the testing samples out, i.e., the training data do not have the samples from testing data.

5.2. Performance Evaluation. For performance evaluation, we specifically choose recognition accuracy and confusion matrix. The actual pose occurred is shown on the column of confusion matrix while the pose classified is shown by the rows of confusion matrix. As shown in Figure 5, the presented scheme can recognize 10 different in-vehicle poses with an average accuracy of 91.1% and 91.8% for scenarios I and II, respectively.

To evaluate the efficacy and reliability of presented scheme, the obtained results are analyzed by choosing different evaluation metrics, i.e., precision, recall, and F_1-score, defined as



(a) In-vehicle layout (L)

(b) In-vehicle layout (L-1)

(c) In-vehicle layout (L-2)

Figure 13: Varying layout.

Table 7: Performance evaluation with varying layouts.

Experiment	Average recognition accuracy (%)		
	L-1	L-2	L
Scenario-I	89.9	90.7	91.1
Scenario-II	91.3	91.5	91.8

(1) Precision is the positive predictive measurement, mathematically represented as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{19}$$

where TP stands for true positive, and FP is false positive. True positive (TP) is defined as the probability that correctly predicts the positive class of any model. On the other hands, false positive (FP) is the probability that incorrectly predicts the positive class of model.

(2) Recall measures the sensitivity of a model and described as the true positive rate (TPR). Recall is represented as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{20}$$

where FN stands for false negative. FN is defined as the probability that incorrectly predicts the negative class.

(3) F-measure or F_1-score is the weighted average of recall and precision, represented as

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{21}$$

Figure 6 shows the results related to precision, recall, and F_1-score. The average minimum and maximum values are summarized in Table 2. It is clear from the results that for both scenarios, all the prescribed poses are recognized with reasonable limits of precision, recall, and F_1-score.

The accuracy of KSVM is compared with stand-alone KNN and SVM, as shown in Figure 7. The overall results are summarized in Table 3. From the obtained results, it is concluded that KSVM has high recognition accuracy as compared to stand-alone KNN or SVM.

To examine the computational complexity of our proposed KSVM algorithm, we compare the execution time of KSVM with KNN and SVM as illustrated in Figure 8. It is notified that the execution time of KSVM is less as compared to SVM but little higher in comparison to KNN. It can be compromised because the recognition accuracy of KSVM is comparatively far better than KNN.

User independence test is performed to evaluate the generalization of proposed KSVM scheme. For the purpose, leave-one-participant-out cross-validation (LOPO-CV) method is adopted. In LOPO-CV mechanism, the training data is not familiar with testing data, i.e., the whole data set is used as training data set except a specific users' data that is used as the testing data. This method is repeated for each individual user until all users are treated as testing data. We have separately applied LOPO-CV on SVM, KNN, and KSVM to compare the results, as shown in Figures 9–11. The overall results are concluded in Table 4. From the obtained results, it is obvious that the proposed mechanism has generalization capability with an average accuracy of 86.4% and 88.2% for scenario-I and II, respectively. Although the stand-alone SVM algorithm has a comparatively good generalization capability but overall performance is far better using KSVM.

The performance of proposed KSVM algorithm is evaluated by comparing its recognition accuracy with conventional classification methods, i.e., Naive Bayes (NB) [56], artificial neural networks (ANN) [57, 58], decision tree (DT) [59], and sparse representation-based classification (SRC) [60, 61]. The overall comparison is shown in Figure 12, and results are summarized in Table 5. One can notice that the recognition performance of KSVM scheme is higher as compared to conventional classification methods.

Extensive experiments are performed to observe the effect of nearest neighbors with varying K values, as described in Table 6. One can observe that best results are obtained at $K = 5$; therefore, optimal values $K = 5$ are used throughout the experiments.

We have examined the robustness of our proposed KSVM scheme with different locations of transmitter (router) and receiver (laptop). Our actual layout is "L" as shown in Figure 13(a), while two varying layouts are "L-1" and "L-2," as shown in Figures 13(b) and 13(c), respectively. The results with varying layouts are described in Table 7. From the results, it is clear that our presented mechanism has acceptable recognition performance at all in-vehicle layouts.

6. Results Analysis and Discussion

In this section, we will discuss about the obtained results with prominent limitations. We observed that all poses are recognized with very good accuracy using our proposed KSVM classification method, however, the accuracy may be degraded due to different limiting factors. In this context, we observed that the CSI of WiFi signal is highly influenced by moving objects. Therefore, other vehicles and people on outside road may influence the recognition performance of the system [36]. Furthermore, the presented mechanism is designed by considering only a single person, i.e., the driver in the vehicle. However, in practical, more than one people may exist inside the vehicle which can make the recognition system more complex. To overcome these issues, we suggest to perform some additional signal processing which will be considered in our future work. Moreover, the effect of driver's orientation and personalized driving habits is needed to be considered in future study.

Despite these limitations, the proposed device-free WiFi CSI-based driver pose localization system is easy to deploy

and more scalable. It is clear that the presented classification algorithm KSVM is a general solution. It can be implemented to solve any device-free WiFi-based localization and gesture or activity recognition problem. In this research work, we have used this scheme for driver's face localization. The overall performance of our proposed KSVM is far better as compared to existing methods. There are still several aspects which need be considered in future.

7. Conclusion

In this research work, we have proposed a novel classification scheme for WiFi CSI-based device-free driver's face localization. We have presented a hybrid classification algorithm, i.e., KSVM that is based on the combination of traditional KNN and SVM classification methods. From the experimental results, it can be concluded that recognition performance has been remarkably improved by utilizing our proposed KSVM algorithm. This hybrid classification scheme opens a new window for diverse scale of potential applications. For future, we are interested to explore more complex driving scenarios and intend to observe the impact of roadway types based on the findings presented in this research work.

Data Availability

The data that support the research findings are available on request. The data are not publicly available due to privacy of research participants.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] W. U. Khan, X. Li, A. Ihsan, M. A. Khan, V. G. Menon, and M. Ahmed, *NOMA-Enabled Optimization Framework for Next-Generation Small-Cell IoV Networks under Imperfect SIC Decoding*, IEEE Transactions on Intelligent Transportation Systems, 2021.

[2] W. U. Khan, T. N. Nguyen, F. Jameel et al., *Learning-Based Resource Allocation for Backscatter-Aided Vehicular Networks*, IEEE Transactions on Intelligent Transportation Systems, 2021.

[3] W. U. Khan, M. A. Javed, T. N. Nguyen, S. Khan, B. M. Elhalawany, and B. M. Elhalawany, *Energy-Efficient Resource Allocation for 6G Backscatter-Enabled NOMA IoV Networks*, IEEE Transactions on Intelligent Transportation Systems, 2021.

[4] W. U. Khan, F. Jameel, N. Kumar, R. Jäntti, and M. Guizani, "Backscatter-enabled efficient V2X communication with non-orthogonal multiple access," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1724–1735, 2021.

[5] A. Dahmane, S. Larabi, I. M. Bilasco, and C. Djeraba, "Head pose estimation based on face symmetry analysis," *Signal, Image and Video Processing*, vol. 9, no. 8, pp. 1871–1880, 2015.

[6] K. Diaz-Chito, A. Hernández-Sabaté, A. M. López, and A. MLópez, "A reduced feature set for driver head pose estimation," *Applied Soft Computing*, vol. 45, pp. 98–107, 2016.

[7] P. WKithil, R. D. Jones, and J. McCuish, "Driver alertness detection research using capacitive sensor array," *Technical report, SAE Technical Paper*, 2001.

[8] M. Patel, S. K. L. Lal, D. Kavanagh et al., "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7235–7242, 2011.

[9] S. Martin, E. Ohn-Bar, A. Tawari, and M. M. Trivedi, "Understanding head and hand activities and coordination in naturalistic driving videos," in *In 2014 IEEE Intelligent Vehicles Symposium Proceedings*, IEEE, 2014.

[10] C. Tran, A. Doshi, and M. M. Trivedi, "Modeling and prediction of driver behavior by foot gesture analysis," *Computer Vision and Image Understanding*, vol. 116, no. 3, pp. 435–445, 2012.

[11] F. Vicente, Z. Huang, X. Xiong, F. de la Torre, W. Zhang, and D. Levi, "Driver gaze tracking and eyes off the road detection system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2014–2027, 2015.

[12] H. Jo and M. Lee, "In-attention state monitoring for a driver based on head pose and eye blinking detection using one class support vector machine," in *In International Conference on Neural Information Processing*, pp. 110–117, Springer, 2014.

[13] Y. He, Y. Chen, Y. Hu, and B. Zeng, "WiFi vision: sensing, recognition, and detection with commodity MIMO-OFDM WiFi," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8296–8317, 2020.

[14] H. Fei, F. Xiao, J. Han, H. Huang, and L. Sun, "Multi-variations activity based gaits recognition using commodity WiFi," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2263–2273, 2020.

[15] F. Wang, W. Gong, J. Liu, and W. Kui, "Channel selective activity recognition with wifi: a deep learning approach exploring wideband information," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 181–192, 2020.

[16] J. Zhang, F. Wu, B. Wei et al., "Data augmentation and dense-LSTM for human activity recognition using WiFi signal," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4628–4641, 2021.

[17] R. H. Venkatnarayan, S. Mahmood, and M. Shahzad, "WiFi based multi-user gesture recognition," *IEEE Transactions on Mobile Computing*, vol. 20, no. 3, pp. 1242–1256, 2021.

[18] Z. Guo, F. Xiao, B. Sheng, H. Fei, and Y. Shui, "WiReader: adaptive air handwriting recognition based on commercial WiFi signal," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10483–10494, 2020.

[19] H. Yan, Y. Zhang, Y. Wang, and X. Kangle, "WiAct: a passive WiFi-based human activity recognition system," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 296–305, 2020.

[20] X. Yang, R. Cao, M. Zhou, and L. Xie, "Temporal-frequency attention-based human activity recognition using commercial WiFi devices," *IEEE Access*, vol. 8, pp. 137758–137769, 2020.

[21] J. Zuo, X. Zhu, Y. Peng, Z. Zhao, X. Wei, and X. Wang, "A new method of posture recognition based on WiFi signal," *IEEE Communications Letters*, vol. 25, no. 8, pp. 2564–2568, 2021.

[22] C. Xiao, Y. Lei, Y. Ma, F. Zhou, and Z. Qin, "DeepSeg: deep-learning-based activity segmentation framework for activity recognition using WiFi," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5669–5681, 2021.

[23] J. Ding, Y. Wang, and F. Xiangcong, "Wihi: WiFi based human identity identification using deep learning," *IEEE Access*, vol. 8, pp. 129246–129262, 2020.

[24] X. Ma, Y. Zhao, Z. Liang et al., "Practical device-free gesture recognition using WiFi signals based on metalearning," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 228–237, 2020.

[25] X. Shen, Z. Ni, L. Liu, J. Yang, and K. Ahmed, "WiPass: 1D-CNN-based smartphone keystroke recognition using WiFi signals," *Pervasive and Mobile Computing*, vol. 73, no. 101393, 2021.

[26] F. Li, M. Valero, H. Shahriar, R. A. Khan, and S. I. Ahamed, "Wi-COVID: a COVID-19 symptom detection and patient monitoring framework using WiFi," *Smart Health*, vol. 19, no. 100147, 2021.

[27] H. F. T. Ahmed, H. Ahmad, K. Narasingamurthi, H. Harkat, and S. K. Phang, "DF-WiSLR: device-free Wi-Fi-based sign language recognition," *Pervasive and Mobile Computing*, vol. 69, no. 101289, 2020.

[28] M. Asif, W. Zhou, M. Ajmal, Z. u. A. Akhtar, and N. A. Khan, "A construction of high performance quasicyclic LDPC codes: a combinatoric design approach," *Wireless Communications and Mobile Computing, vol.*, vol. 2019, pp. 1–10, 2019.

[29] M. Asif, W. Zhou, Q. Yu, X. Li, and N. A. Khan, "A deterministic construction for jointly designed quasicyclic LDPC coded-relay cooperation," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[30] M. Asif, W. Zhou, Q. Yu et al., "Jointly designed quasi-cyclic LDPC-coded cooperation with diversity combining at receiver," *International Journal of Distributed Sensor Networks*, vol. 16, no. 7, 2020.

[31] M. Asif, W. U. Khan, H. M. Afzal et al., "Reduced-complexity LDPC decoding for next-generation IoT networks," *Wireless Communications and Mobile Computing, vol.*, vol. 2021, 2021.

[32] Y. Wang, X. Jiang, R. Cao, and X. Wang, "Robust indoor human activity recognition using wireless signals," *Sensors*, vol. 15, no. 7, pp. 17195–17208, 2015.

[33] T. Wang, D. Yang, S. Zhang, Y. Wu, and X. Shugong, "Wi-Alarm: low-cost passive intrusion detection using WiFi," *Sensors*, vol. 19, no. 10, p. 2335, 2019.

[34] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Recognizing keystrokes using WiFi devices," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1175–1190, 2017.

[35] Y. Wang, K. Wu, L. M. Ni, and L. M. Ni, "WiFall: device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2017.

[36] H. Peng and W. Jia, "WiFind: driver fatigue detection with fine-grained Wi-Fi signal features," *IEEE Transactions on Big Data (Early Access)*, pp. 1–1, 2018.

[37] X. Yongliang and X. Limin, "Shot boundary detection based on supervised locality preserving projections and KNN-SVM classifier," in *In 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics*, p. 11262926, IEEE, 2010.

[38] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[39] M. Liu, L. Zhang, P. Yang, L. Lu, and L. Gong, "Wi-Run: device-free step estimation system with commodity Wi-Fi," *Journal of Network and Computer Applications*, vol. 143, pp. 77–88, 2019.

[40] F. Li, M. A. A. Alqaness, Y. Zhang, B. Zhao, and X. Luan, "A robust and device-free system for the recognition and classification of elderly activities," *Sensors*, vol. 16, no. 12, p. 2016, 2043.

[41] Z. Dong, F. Li, J. Ying, and K. Pahlavan, "Indoor motion detection using Wi-Fi channel state information in flat floor environments versus in staircase environments," *Sensors*, vol. 18, no. 7, p. 2177, 2018.

[42] Y. Xu, F. Xiong, Y. Shao, and Q. Niu, "WmFall: WiFi-based multistage fall detection with channel state information," *International Journal of Distributed Sensor Networks*, vol. 14, no. 10, 2018.

[43] J. Lv, D. Man, Y. Wu et al., "Robust device-free intrusion detection using physical layer information of WiFi signals," *Applied Sciences*, vol. 19, no. 1, p. 175, 2019.

[44] E. Ding, X. Li, T. Zhao, L. Zhang, and H. Yanjun, "A robust passive intrusion detection system with commodity WiFi devices," *Journal of Sensors, page*, vol. 2018, pp. 1–12, 2018.

[45] M. A. A. Al-qaness, "Device-free human micro-activity recognition method using WiFi signals," *Geo-spatial Information Science*, vol. 22, no. 2, pp. 128–137, 2019.

[46] L. Zhang, Q. Gao, X. Ma, J. Wang, T. Yang, and H. Wang, "DeFi: robust training-free device-free wireless localization with WiFi," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8822–8831, 2018.

[47] M. A. A. Al-qaness and F. Li, "WiGer: WiFi-based gesture recognition system," *International Journal of Geo-Information*, vol. 5, no. 6, p. 92, 2016.

[48] F. Zhangjie, J. Xu, Z. Zhu, A. X. Liu, and X. Sun, "Writing in the air with WiFi signals for virtual reality devices," *IEEE Transactions on Mobile Computing*, vol. 18, no. 2, pp. 473–484, 2019.

[49] Z. Tian, J. Wang, X. Yang, and M. Zhou, "WiCatch: a Wi-Fi based hand gesture recognition system," *IEEE Access*, vol. 6, pp. 16911–16923, 2018.

[50] Z. U. A. Akhtar and H. Wang, "WiFibased driver's activity monitoring with efficient computation of radio-image features," *Sensors*, vol. 20, no. 5, p. 1381, 2020.

[51] Z. U. A. Akhtar and H. Wang, "WiFibased gesture recognition for vehicular infotainment system-an integrated approach," *Applied Sciences*, vol. 9, no. 24, p. 5268, 2019.

[52] Z. U. A. Akhtar and H. Wang, "WiFi-based driver's activity recognition using multi-layer classification," *Neurocomputing*, vol. 405, pp. 12–25, 2020.

[53] S. Duan, T. Yu, and J. He, "WiDriver: driver activity recognition system based on WiFi CSI," *International Journal of Wireless Information Networks*, vol. 25, no. 2, pp. 146–156, 2018.

[54] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–21, 2018.

[55] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and L. Sanglu, "Device-free human activity recognition using commercial WiFi devices," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1118–1131, 2017.

[56] M. Kholoud, G. Norjihan, H. S. Abdul, and R. M. Babar, "Human activity classification using decision tree and Naïve Bayes classifiers," *Multimedia Tools and Applications*, vol. 80, pp. 21709–21726, 2020.

[57] T. Kanokoda, Y. Kushitani, M. Shimada, and J.-i. Shirakashi, "Gesture prediction using wearable sensing systems with neural networks for temporal data analysis," *Sensors*, vol. 19, no. 3, p. 710, 2019.

[58] P. Goswami, S. Rao, S. Bharadwaj, and A. Nguyen, "Real-time multi-gesture recognition using 77 GHz FMCW MIMO single chip radar," in *In 2019 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–4, IEEE, 2019.

[59] T. Zhang, T. Song, D. Chen, T. Zhang, and J. Zhuang, "WiGrus: a WiFi-based gesture recognition system using software-defined radio," *IEEE Access*, vol. 7, pp. 131102–131113, 2019.

[60] B. Tan, Q. Chen, K. Chetty, K. Woodbridge, W. Li, and R. Piechocki, "Exploiting WiFi channel state information for residential healthcare informatics," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 130–137, 2018.

[61] M. A. A. Al-qaness, F. Li, X. Ma, Y. Zhang, and L. Guo, "Devicefree indoor activity recognition system," *Applied Sciences*, vol. 6, no. 11, p. 329, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3032997, 18 pages
https://doi.org/10.1155/2022/3032997

Research Article

An Ant Colony Optimization-Based Routing Algorithm for Load Balancing in LEO Satellite Networks

Xia Deng,[1] Shouyuan Zeng,[1] Le Chang,[2] Yan Wang,[1] Xu Wu,[3] Junbin Liang,[3] Jiangtao Ou,[4] and Chengyuan Fan[4]

[1]*School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China*
[2]*School of Automation, Guangdong University of Technology, Guangzhou 510006, China*
[3]*School of Computer and Electronics Information, Guangxi University, Nanning 530004, China*
[4]*AI Sensing Technology, Chancheng District, Foshan 528000, China*

Correspondence should be addressed to Le Chang; lechang@gdut.edu.cn

Received 26 January 2022; Revised 3 March 2022; Accepted 16 March 2022; Published 23 April 2022

Academic Editor: Xingwang Li

Satellite networks can provide a wider service range and lower delay than traditional terrestrial optical fiber networks. However, due to the bursty characteristic of the Internet traffic and the distributive feature of satellite links, traffic-intensive areas often suffer from link congestion while links in other areas are underutilized, i.e., the traffic imbalance problem in LEO satellite networks. In this paper, an ant colony optimization routing algorithm with window reduction for LEO satellite networks, ACORA-WR, is proposed to achieve load balancing. ACORA-WR limits the movement of the ant colony to a specific range and comprehensively considers the path distance, transmission direction, and link load to find a path with low delay and overhead. Simulation results verify that the proposed ACORA-WR scheme demonstrates high data delivery ratio and network throughput, while ensuring low average delay and network transmission overhead.

1. Introduction

With the development of on-board processing and intersatellite link technology, satellite communication networks have become a reality and attracted great attention. Low Earth Orbit (LEO) satellites are often used to form satellite networks, due to the relatively low operational height, low communication delay with ground, low link loss, and low transmission power. Iridium, Globalstar, Teledesic, Starlink, and other satellite networks are all built on LEO satellites.

Compared with terrestrial 4G/5G cellular networks, the biggest advantage of LEO satellite networks is the ability of providing seamless Internet services to the whole world [1–5]. LEO satellites can cover remote areas that are not covered by 4G/5G cellular networks and are of great value in ocean operations, scientific and technological broadband, aeronautical broadband, and disaster emergency communication. As shown in Figure 1, LEO satellites extend 4G/5G cellular networks to oceans, air, and other remote areas, which allows people to transfer data from urban areas to airplanes, cruises, and other vehicles in remote areas. However, because the land area accounts for less than 30% of the total area of the earth and the population is concentrated in large cities, the satellite data volume regarding populated areas is significantly greater than other places, so the distribution of data flow is unbalanced, i.e., the traffic imbalance problem.

The routing algorithms in LEO networks must be carefully designed to meet high-quality and large-capacity business under limited satellite resources and imbalanced traffic distribution. Compared with traditional terrestrial networks [6–8], the load balancing design for satellite communication networks is more complex. Terrestrial wired networks have

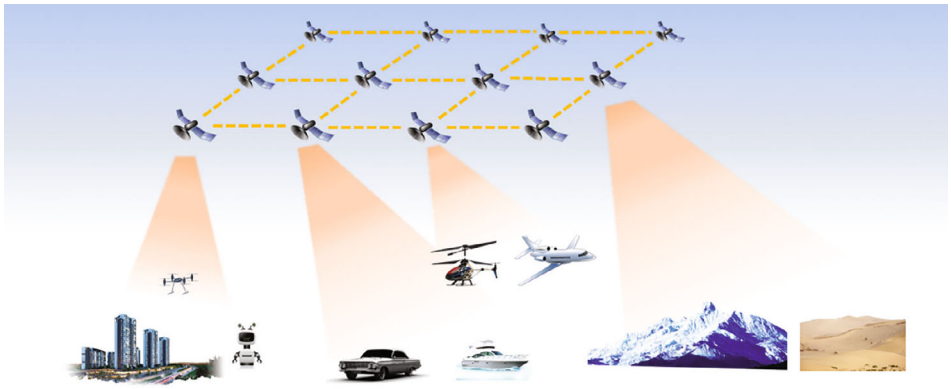


Figure 1: Application of LEO satellite networks in the global Internet.

relatively fixed topology, while the topology of the satellite communication networks is highly dynamic. The continuous movement of satellites leads to frequent change of the connections between satellites, which poses a great challenge to balancing the traffic load in such satellite networks.

Existing centralized load balancing routing schemes need to obtain path information for the entire network before calculating the optimal path, which consumes great network resources [9–11]. In contrast, distributed load balancing schemes can provide fast reaction to traffic changes [12–17]. In such category, the ant colony algorithm is a class of simulated evolutionary algorithms good at finding the optimal path under dynamic network environments [18]. Therefore, in this paper, we aim at the traffic imbalance problem in LEO satellite networks and propose an ant colony optimization routing algorithm with window reduction (ACORA-WR). Our contributions are summarized as follows.

(i) First, we have proposed an ant colony-based routing algorithm to tackle the load balancing problem in LEO satellite networks. We design congestion-avoiding heuristic information and combine it with pheromone to improve the ability of ants to enhance the local search ability and avoid stagnation. The pheromone update rule is also optimized based on path length and buffer status to improve the convergence speed to the global optimal solution

(ii) Second, we design a window-reduction mechanism to restrict the explore range of ants, which helps improve search efficiency and reduce the complexity

(iii) Last but not the least important, we implement ACORA-WR on NS2 and conduct extensive simulations to evaluate its performance. The results show that ACORA-WR achieves desirable performance on data delivery ratio and network throughput, while keeping the delay and overhead low

The remainder of this paper is organized as follows. We review the related works in Section 2. In Section 3, we explain the satellite network architecture. In Section 4, we describe the new routing algorithm. Section 5 presents the performance results of our ACORA-WR algorithm, and Section 6 summarizes the paper.

2. Related Works

Existing load-balancing routing strategies for LEO satellite networks can be mainly classified into two categories: centralized and distributed strategies.

Centralized strategies usually use a global optimization algorithm to optimize the system traffic allocation and compute the routing table. Li et al. proposed a mechanism to quantitatively estimate the link state and dynamically adjusted the weight of queuing delay [9]. The occupancy rate of each queue was divided into multiple levels, and the queuing delay was determined according to the corresponding level. HGL [10] proposed by Liu et al. combined the global and local strategy to optimally allocate IoT traffic flows. Wang et al. proposed a load balancing scheme based on Stackelberg game algorithm to make more effective use of the satellite storage space [11].

Distributed schemes are also developed because centralized strategies are difficult to implement on satellites. Na et al. proposed a distributed routing strategy for LEO satellite networks based on machine learning [12]. Taleb et al. proposed the Explicit Load Balancing (ELB) scheme which predicted the current transmission congestion by explicitly exchanging the queue usage on adjacent satellites [13]. Song et al. extended the idea of ELB and brought forward the traffic light-based routing (TLR) mechanism [14]. Papapetrou et al. proposed Location-Assisted On-demand Routing (LAOR) [15], where the satellite actively searched for a path to the destination satellite after receiving a user communication request. Karapantazis et al. extended the idea of LAOR and proposed Multiservice On demand Routing (MOR) [16]. Rao and Wang proposed Agent-based Load Balancing Routing (ALBR) [17], which used mobile agents to explore link states to randomly selected destinations.

Currently, ant colony algorithms are widely used in networks, such as avoiding the interflow and intraflow interference or balancing the loads in backbone networks [19], calculating the best route for vehicles in vehicle networks [20, 21], designing dynamic source routing algorithm

meeting high QoS requirements in ad hoc networks [22], balancing energy efficiency for WSNs and IPv6 networks [23, 24], and solving the scheduling (SBS) problem of satellite broadcasting between satellites and ground stations [25]. Wang et al. proposed a load balancing scheme based on the ant colony algorithm (LBRA-CP) [18]. In LBRA-CP, the satellites sought out the congestion areas and shared the information of their positions to predict congestion, and the ant colony algorithm was used to find an optimal path for every connection request. However, ant colony algorithms still need to tackle two challenges. First, they may result in local optimal solutions. Second, the search range may be too wide, making the search efficiency lower with the increase of the number of satellites.

In this paper, we consider the total cost of the path and the size of satellite buffer and also set the upper limit of pheromone to avoid the ant colony algorithm to get the local optimal solution. Meanwhile, we propose a window reduction mechanism to increase the search speed of the ant colony, which can make full use of intersatellite links to improve data delivery ratio and throughput, while keeping a low transmission overhead and delay.

3. System Model

3.1. Satellite Networks. We model a LEO satellite network as a graph $G = (V, E)$, where V and E represent the set of satellite nodes and directed ISLs (intersatellite links), respectively. In the context of such network, we use the term "node" and "satellite" interchangeably. The types of ISLs include intraplane ISLs and interplane ISLs. Intraplane ISLs refer to the links between satellites in the same orbit, and interplane ISLs refer to those between satellites in different orbits. The intraplane ISLs remain constant with fixed length. Satellites move predictably and regularly with a constant period. The topology of satellite networks is treated as constant at an instant or within a small time duration following [26]. As shown in Figure 2, each satellite typically has two intraplane ISLs and two interplane ISLs. An intraplane ISL is used to connect a satellite with its predecessor or successor satellites in the same orbit, and interplane ISLs are used to connect two adjacent satellites in two adjacent orbits. We represent the logical address of a satellite using $N_{i,j}$, where i is the orbit number and j is satellite number within its orbit.

3.2. End-to-End Path Cost. The ISL link cost is based on the end-to-end delay, which is the summation of queuing delay and propagation delay [17]. The cost of the link between satellite i and j, i.e., on transmission link (i, j), is thus calculated as

$$\mathrm{Cost}(i,j)_t = \mathrm{PDelay}(i,j)_t + \mathrm{QDelay}(i,j)_t, \tag{1}$$

where the propagation delay $\mathrm{PDelay}(i,j)_t$ is calculated as

$$\mathrm{PDelay}(i,j)_t = \frac{d(i,j)_t}{c}. \tag{2}$$

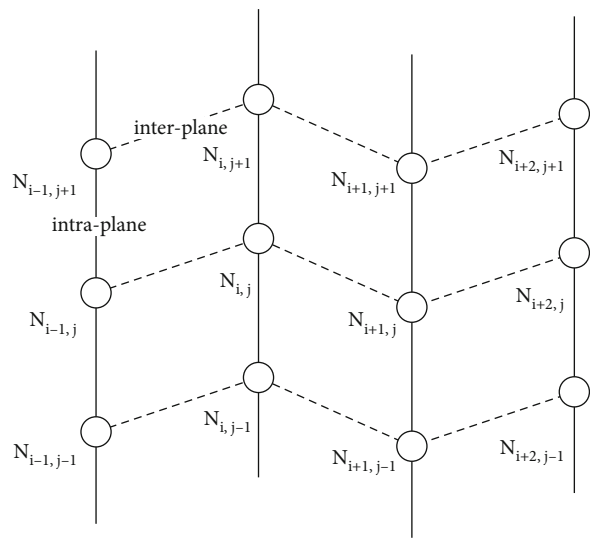


Figure 2: LEO satellite network topology.

The propagation delay $\mathrm{PDelay}(i,j)_t$ is the physical distance $d(i,j)$ of satellite link (i,j) at time t divided by velocity c, the optical transmission speed in vacuum, i.e., 3×10^8 m/s. The queuing delay $\mathrm{QDelay}(i,j)_t$ is calculated using the same method in [17]. Assuming that both the packet arrival and service pattern are Poisson processes, then the arrival and service intervals follow exponential distributions. We adopt the M/M/1 queuing model proposed in reference [27] and estimate the average number of packets in a queue using

$$\mathrm{Num}_q = \frac{\mu}{1-\mu}, \tag{3}$$

where μ represents the mean ISL utilization following the calculation in [28].

μ_h is the mean ISL utilization when the h-th packet reaches the interface queue, calculated as

$$\mu_h = \mathrm{LinkState} + e^{-\Delta t_h} \times (\mu_{h-1} - \mathrm{LinkState}), \tag{4}$$

where LinkState is set to 0 if there are no packets in the interface queue or in transmission, and 1, otherwise. $e^{-\Delta t_h}$ represents the forgetting rate, and Δt_h is the time interval between the arrival of the h-th and $(h-1)$th packets in the queue. Therefore, the queuing delay can be estimated as

$$\mathrm{QDelay} = \mathrm{Num}_q \times \frac{S_{\mathrm{avg}}}{B}, \tag{5}$$

where S_{avg} is the average packet size in the queue and B is the link bandwidth.

$H(s,d)$ represents the set of all links from the source node s to the destination node d. The total cost from the source to the destination is

$$\mathrm{Cost}(s,d)_t = \sum_{\substack{i\in H(s,d)\\ j\in H(s,d)}} \mathrm{Cost}(i,j)_t. \tag{6}$$

We set $C(0 < C < 1)$ as the threshold and let $\varrho_{ij}(t)$ represent the waiting queue occupancy rate of link (i, j) at time t. If $\varrho_{ij}(t) > C$, $\mathrm{LinkAvailable}(i, j) = 0$ indicates that link (i, j) is congested and the ISL cannot be used. Otherwise, $\mathrm{LinkAvailable}(i, j) = 1$ indicates that ISL is used normally.

$$\mathrm{LinkAvailable}(i,j) = \begin{cases} 0 & \varrho_{ij}(t) > C \\ 1 & \text{otherwise} \end{cases}. \tag{7}$$

The goal of the end-to-end path cost model is to find the optimal path with the smallest cost for each packet from the source to destination under the above constraints. Therefore, the optimization model is formulated as

$$\begin{gathered} \text{Minimize}: \mathrm{Cost}(s,d)_t \\ \text{s.t.}\forall \mathrm{link}(i,j)\mathrm{LinkAvailable}(i,j) = 1 \\ (i,j) \in H(s,d) \end{gathered} \tag{8}$$

Our goal is to find the path with the lowest total cost from source to destination. Moreover, to ensure reliable data transmission, $(i, j) \in H(s, d)$ represents each ISL on $H(s, d)$ should be usable.

4. Ant Colony Optimization Routing Algorithm with Window Reduction

Ant colony algorithm is a probabilistic algorithm used to find the optimal path [29]. It has the characteristics of distributed calculation, positive information feedback, and heuristic search. Ant colony algorithm uses a control message (also known as ant) to collect path information for routing. The packet selects the next hop according to the probability formula. However, in the LEO satellite network environment, the ant colony algorithm may stagnate for a period of time, which makes the traffic congestion more serious. Therefore, we propose the congestion-avoiding heuristic information to alleviate this problem. On the other hand, due to the long distance and global coverage of satellite links (ISLs), in order to make the ant colony more efficient, we propose the window reduction mechanism to limit the search range of ants. In this section, we explain how we design an ant colony algorithm to fit the LEO scenarios.

4.1. Ant Colony Algorithm-Based Routing. We use data packets to simulate ants, which are divided into two kinds: forward ants and backward ants. The forward ants represent the messages from the source node to the destination node, which collect the path information from the source node to the destination node including end-to-end delay and number of hops. The backward ants represent the messages returned from the destination node to the source node, which change the routing information of each satellite node.

Table 1: The probability table structure of node i.

Flow	Adjacent node	Probability
f	j	P_{ijf}
f	k	P_{ikf}
g	l	P_{ilg}

Pheromone is the information left by ants when they pass through nodes. The forward ant calculates and determines the next hop using pheromone. When the backward ant passes through a node, it will update the pheromone of the node. With a larger number of ants passing through the node, the pheromone density will be greater. Intuitively, ants prefer to choose those nodes with greater pheromone density. The routing table of each node in the network is represented by a probability table, and the probability value is related to the pheromone density. Table 1 shows the structure of the probability table of node i, which includes three kinds of information: flow, adjacent node, and probability. The flow number is determined by source-destination pairs, and the adjacent node is the node directly adjacent to node i. When a data packet of flow f reaches satellite node i, the probability of selecting the next node j is P_{ijf}.

In the initial state, the pheromones on all paths are evenly distributed. As shown in Figure 3, a group of forward ants are sent from source S, and the forward ants are forwarded in the networks according to the pheromone. If the pheromone density is 0, it is forwarded randomly. The ants can remember the information of the path, including hops and delay, and finally reach the destination node receiver D. Obviously, the ant passing through the shortest path will reach the destination node at the fastest speed. If there are multiple identical ants, the destination node only receives the first arriving ant. When the ant arriving at the destination first reaches receiver D, receiver D will send the backward ant to the source point S, and the backward ant will return to S following the original path of the forward ant and modify the pheromone when passing through the node. The backward ant determines whether to increase or decrease the pheromone according to the communication quality of the whole path and the increment value of the pheromone according to the congestion degree of the node. As denoted by the yellow links in Figure 3, when the path is busy, the backward ant will reduce the pheromone of the node on the yellow path and thus reduce the probability of the forward ant choosing this path. At the same time, the backward ant passing through the blue path will increase the pheromone of this path and allow more data to be forwarded through the blue path. In addition, we set the red link as a congested link, which is temporarily unavailable. In this way, the iterative operation is carried out continuously to select the best path for the data packets.

4.1.1. Forward Ants. We let the source node periodically send out forward ants in the process of data transmission. A forward ant uses unicast to select the next hop node according to the state transition rules and records the travel

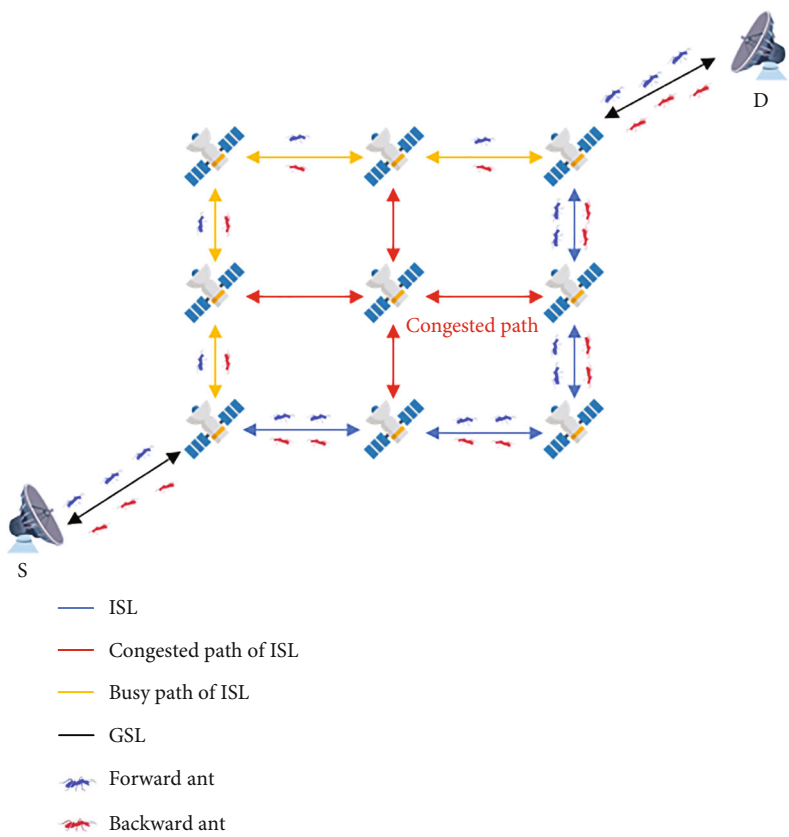


Figure 3: Ant colony algorithm in LEO satellite networks.

path and connection information along the way as the pheromone, so that the data passing through the satellite node can select the next hop node according to such pheromone. When it finally reaches the destination node, the forward ant dies and a backward ant is then generated. In addition, each forward ant also maintains a routing table independently to record the traversed satellite nodes, as to avoid loops in the forwarding process.

The state transition rule is the core rule of all ant colony algorithms; each ant selects the next node according to the state transition rule. Through the state transition rules, the algorithm can use the network link information to find the optimal path from the source node to the destination node.

When the forward ant is at satellite node i, the probability that the adjacent node j of node i is selected as the next hop is

$$P_{ijf} = \begin{cases} \dfrac{\left[p_{ijf}(t)\right]^{\alpha}\left[\eta_{ijf}(t)\right]^{\beta}}{\sum_{k\epsilon K}\left[p_{ikf}(t)\right]^{\alpha}\left[\eta_{ikf}(t)\right]^{\beta}} & q < q_0 \text{and LinkAvailable}(i,j) = 1, \\ \dfrac{1}{A} & q \geq q_0 \text{ and LinkAvailable}(i,j) = 1, \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where $p_{ijf}(t)$ and $\eta_{ijf}(t)$ are the pheromone density and congestion-avoiding heuristic information of link (i, j) with data flow f at time t, respectively, and α and β are the control parameters. q is a random number between 0 and 1. The value of q_0 determines the weights of exploring new paths and using prior knowledge. A is the number of candidate nodes for the next hop, and K is the set of adjacent nodes. If the link is available, $\text{LinkAvailable}(i, j) = 1$ and $q < q_0$, we calculate the probability according to the congestion-avoiding heuristic information and pheromone. If the link is available and $q \geq q_0$, we randomly select the next hop. If the link is unavailable, the probability value is set to 0, and the link is blocked.

The influence factors α and β represent the weights of the pheromone and congestion-avoiding heuristic information, respectively, where $\alpha + \beta = 1$. In order to accelerate the convergence speed in the initial stage and avoid rapid local convergence, we design the influence factor α at moment t as

$$\alpha(t) = \begin{cases} e^{-\sigma t} & \dfrac{\text{Num}_{\Delta v} \bullet S_{\text{avg}}}{\Delta v} > B \\ \varepsilon\left(1 + e^{-\sigma t}\right) & \text{otherwise} \end{cases}, \tag{10}$$

where ε and σ are constants and $\varepsilon, \sigma \epsilon (0, 1]$. t is the searching times. The larger the value of t is, the smaller the value of $\alpha(t)$ is. The weight of the pheromones decreases with the increase of the search times, while the congestion-avoiding heuristic information value β increases. $\text{Num}_{\Delta v}$ indicates the number of packets reaching the current node i in the past Δv time,

S_{avg} is the average packet size in the queue, and $\mathrm{Num}_{\Delta v} \bullet S_{avg} / \Delta v$ indicates the bandwidth required for successful forwarding of $\mathrm{Num}_{\Delta v}$ packets per second.

When the intersatellite link bandwidth B cannot support transmitting the data in time, we increase the weight of congestion-avoiding heuristic information to forward packets with idle ISLs to avoid adjacent links entering the congestion state quickly. When the congestion pressure of this node is light, we mainly make routing choices based on pheromone density to dynamically find the global optimal solution.

4.1.2. Congestion-Avoiding Heuristic Information. The pheromone only considers the transmission quality of the whole path from the source to the destination node but fails to handle local conditions in time. For example, after the ant colony algorithm runs for a period, there will be a stagnation phenomenon, i.e., all the individuals find the same solution and the search space converge, and no better solutions can be found. Stagnation may cause local node congestion. To tackle this problem, we propose the congestion-avoiding heuristic information, which estimates the busy degree of the link by considering the cache occupancy of the link. As the length of ISL changes from time to time, we also consider the length of the communication link. We design the congestion-avoiding heuristic information $\eta_{ijf}(t)$ of link (i, j) at moment t as

$$\eta_{ijf}(t) = \frac{1}{\mu d(i,j)_t + (1-\mu)\varrho_{ij}(t)}, \tag{11}$$

where $d(i,j)_t$ is the length of link (i,j) at moment t, $\varrho_{ij}(t)$ represents the waiting queue occupancy rate of link (i,j) at time t, and μ is the weight of the link length. If the length of link (i,j) is shorter or the cache occupancy is lower, the congestion-avoiding heuristic information value is greater. $\varrho_{ij}(t)$ is calculated as

$$\varrho_{ij}(t) = \frac{\delta_{ij}(t)}{Q}, \tag{12}$$

where Q is capacity of the buffer. $\delta_{ij}(t)$ is the space occupied by packets in the cache queue of current link (i,j) at time t.

4.1.3. Backward Ants. The backward ant returns to the source node from the destination according to the path recorded by the forward ant and updates the pheromone value of each node along the way according to the link information collected by the forward ant. With the increase of the number of backward ants passing through, the more pheromones accumulated, and the forward ants tend to choose the path with higher pheromone density. However, if the number of ant colonies is large and each individual ant starts to find the path from the source to the destination, the pheromone difference on each path will be small, and it is possible to obtain the stable solution after a certain number of iterations. However, an ant colony system algorithm is prone to stagnation, i.e., the stable solution is not the global optimal solution, and ants lose the ability to explore new paths. To solve these problems, we optimize the pheromone update rules to tackle the problems of slow convergence and untimely information update. We use p to represent the density of pheromones and design the pheromone update as

$$p_{ijf}(t+1) = \begin{cases} (1-\rho)p_{ijf}(t) + \sum_{k=1}^{m} \Delta\tau_{ijf}^{k}(t) & p_{ijf}(t+1) < \Gamma \\ \Gamma & \text{otherwise} \end{cases}, \tag{13}$$

where m is the number of ants at each iteration and $\rho \in (0,1)$ is the pheromone evaporation coefficient. The main purpose of pheromone evaporation is to avoid stagnation. Each ant will experience pheromone update, and the best ant will place more pheromones on the node with the best solution. It will lead ants to explore the optimal solution in subsequent iterations. In order to prevent the infinite increment of pheromone, let Γ represent the upper limit of the pheromone density of link (i,j). The pheromone increment $\Delta\tau$ is related to the total cost of the current path and the congestion of the current link, and $\Delta\tau$ is calculated as

$$\Delta\tau_{ijf}^{k}(t) = \begin{cases} \varphi \dfrac{(1-\varrho_{ij}(t))}{\mathrm{Cost}_{sdf}} & \mathrm{Cost}_{sdf} \le \mathrm{Cost}D_f \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

where Cost_{sdf} is the total cost of the data flow f from the source node s to the destination d and $\mathrm{Cost}D_f$ is the maximum tolerated cost. If the path length explored by the forward ant exceeds $\mathrm{Cost}D_f$, there will be a high delay in transferring data on this path, so we set the pheromone increment $\Delta\tau$ to 0. ϱ represents an immediate quantitative value related to the queue waiting time, and the pheromone increment $\Delta\tau$ is negatively correlated with the value of ϱ. We use φ to correct pheromone increment $\Delta\tau$, i.e., the relative relationship between the degree of link congestion and the cost of link.

4.2. Window Reduction. Ant colony algorithms use forward ants to explore new paths heuristically. A greater number of forward ants can find more available paths but also increase the transmission overhead. To reduce such overhead, we thus propose a window reduction mechanism to limit the routes to a specific range. It specifies a rectangular range according to the current satellite and the logical address of the destination satellite. With the movement of the forward ant, the size of the restricted area, referred to as the "window," is shrinking, and this process resembles reducing the size of the window. Each satellite node has a unique number, and the logical address can be calculated using this number. If there are m orbits in the constellation system and n satellites in each orbit, we use $<i, j>$ to represent the j-th satellite in the i-th orbit and $N_{i,j}$ to represent the logical address of the satellite, where i and j are calculated as

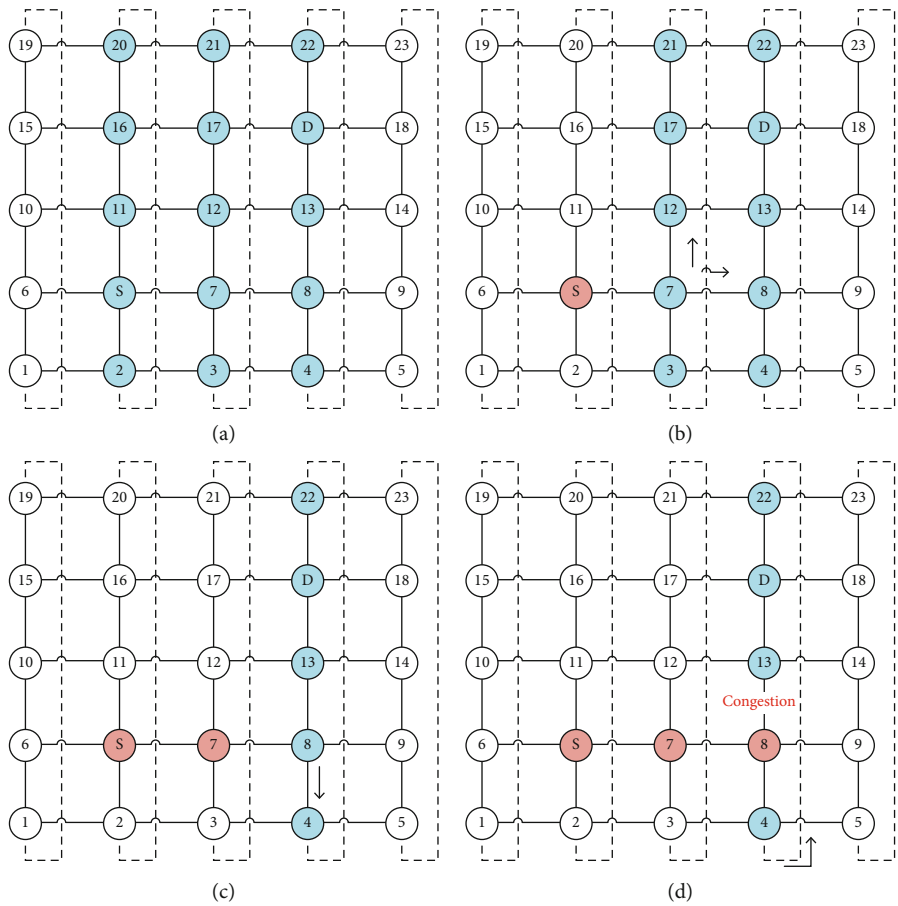


Figure 4: Example of path selection for window reduction mechanism.

$$\begin{cases} i = \dfrac{N_{i,j}}{n} \\ j = N_{i,j} \bmod n \end{cases}. \tag{15}$$

In satellite networks, the four adjacent satellites of satellite $<i_c, j_c>$ are $<(i_c+1) \bmod \mathrm{m}, j_c>$, $<(i_c-1) \bmod \mathrm{m}, j_c>$, $<(i_c, (j_c+1) \bmod \mathrm{n}>$, and $<(i_c, (j_c-1) \bmod \mathrm{n}>$. $<i_e, j_e>$ is the destination when packets are forwarded to the destination node:

$$\begin{cases} i_c = i_e, \\ j_c = j_e. \end{cases} \tag{16}$$

Then, $<i_e, j_e>$ sends the data to the ground destination node directly.

If the current satellite node is not the destination satellite, we define $\mathrm{Direction}_X$ as the x-axis relative distance between the specified node and the destination satellite. The x-axis relative distance is calculated using the orbit number of the logical address as

$$\mathrm{Direction}_X(i_e - i) = \begin{cases} m - |i_e - i| |i_e - i| > \dfrac{m}{2} \\ |i_e - i| \text{ otherwise} \end{cases}. \tag{17}$$

In the data forwarding process, the window reduction mechanism requires that the next-hop satellite node K must meet the following conditions:

$$\mathrm{Direction}_X(i_e - i_k) \le \mathrm{Direction}_X(i_e - i_c). \tag{18}$$

The orbits of the current satellite and the destination satellite divide the satellite networks into two ranges. Equations (17) and (18) mean that the search area of an ant colony is restricted to the smaller one between two ranges. If $\mathrm{Direction}_X(i_e - i) = 0$, it means that the current satellite and the destination satellite are in one orbit, and the packet will only be forwarded on that orbit.

The purpose of the window reduction mechanism is to limit the routing to the range formed by the current node and the destination satellite. As shown in Figure 4(a), a group of data is forwarded from satellite S to satellite D. The window reduction mechanism allows ants to be forwarded to all blue nodes. The adjacent satellites of S are 2, 6, 7, and 11, and satellite nodes 2, 7, and 11 meet the conditions of the window reduction mechanism, so nodes 2, 7, and 11 are added to the candidate set, K. When a packet is about to be forwarded, only the nodes in set K can be selected. Similarly, as shown in Figure 4(b), when the data reaches satellite 7, satellites 3, 8, and 12 are added to set K according to the above rules. Figure 4(c) shows that when

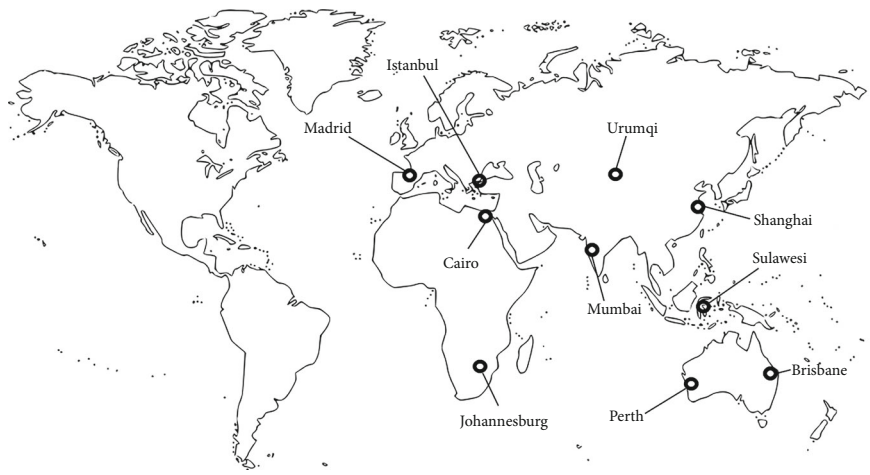


Figure 5: Distribution of ground destination cities.

the data reaches satellite 8, a satellite from satellites 4 and 13 will be selected as the next hop since satellite 8 is in the same orbit as the destination satellite D. Passing satellite 13 will make it faster to reach destination satellite D, so more ants will choose node 13 to forward. With the increase of packets passing through link $(8, 13)$, according to Equation (7), when $\varrho_{ij}(t) > C$, link$(8, 13)$ is not available. Satellite 4 is selected to forward the data. As shown in Figure 4(d), then the data of satellite 4 is forwarded to satellite 22 and finally to the destination satellite. The window reduction mechanism uses the x-axis coordinates of the satellite nodes to limit the range of forward ants. The size of this range is determined by the x-axis coordinates of the current and destination satellite nodes. We combine the optimized ant colony algorithm with the window reduction mechanism above to reduce the additional overhead caused by ants exploring new paths.

ACORA-WR is a distributed algorithm in which each satellite node updates its routing table independently according to Equation (13). The computational complexity at node is proportional to the number of neighbors. For example, in Iridium, there are no more than 4 satellite nodes adjacent to each satellite. In Starlink, there are about 4 adjacent satellite nodes, and in other LEO networks, there are usually no more than 8 neighbors. In addition, ground nodes are not counted in ACORA-WR. If the ground node is the destination, the satellite will forward the data to it directly. Therefore, the number of the neighbors of each node is a constant, so the computational complexity is O(1). Each satellite node stores only the pheromones, cache occupancy, and distance of its neighbors, so the storage complexity is O(1), too.

5. Results and Discussion

5.1. Simulation Setup. We use the NS2 simulation platform to simulate an Iridium-like system and evaluate the performance of our proposed ACORA-WR routing algorithm. The Iridium satellite network consists of 66 satellites with 11 satellites in each orbit. Each satellite has four intersatellite ISLs, including intraplane ISLs and interplane ISLs. We set the bandwidth of each ISL to 2 Mbps. The uplink and downlink bandwidth of the satellite is 2 Mbps. Other parameters of Iridium networks are as follows: altitude = 780 km and inclination = 86.4°. We use constant traffic with a fixed packet size of 512 B. The simulation time is 100 s. There are seven flows distributed among the cities shown in Figure 5, which are distributed between 50 degrees north and 50 degrees south latitudes. The flows, i.e., source-destination pairs, are shown in Table 2. We compare our ACORA-WR with three popular routing protocols, LBRA-CP [18], SPR, and LCRA [30].

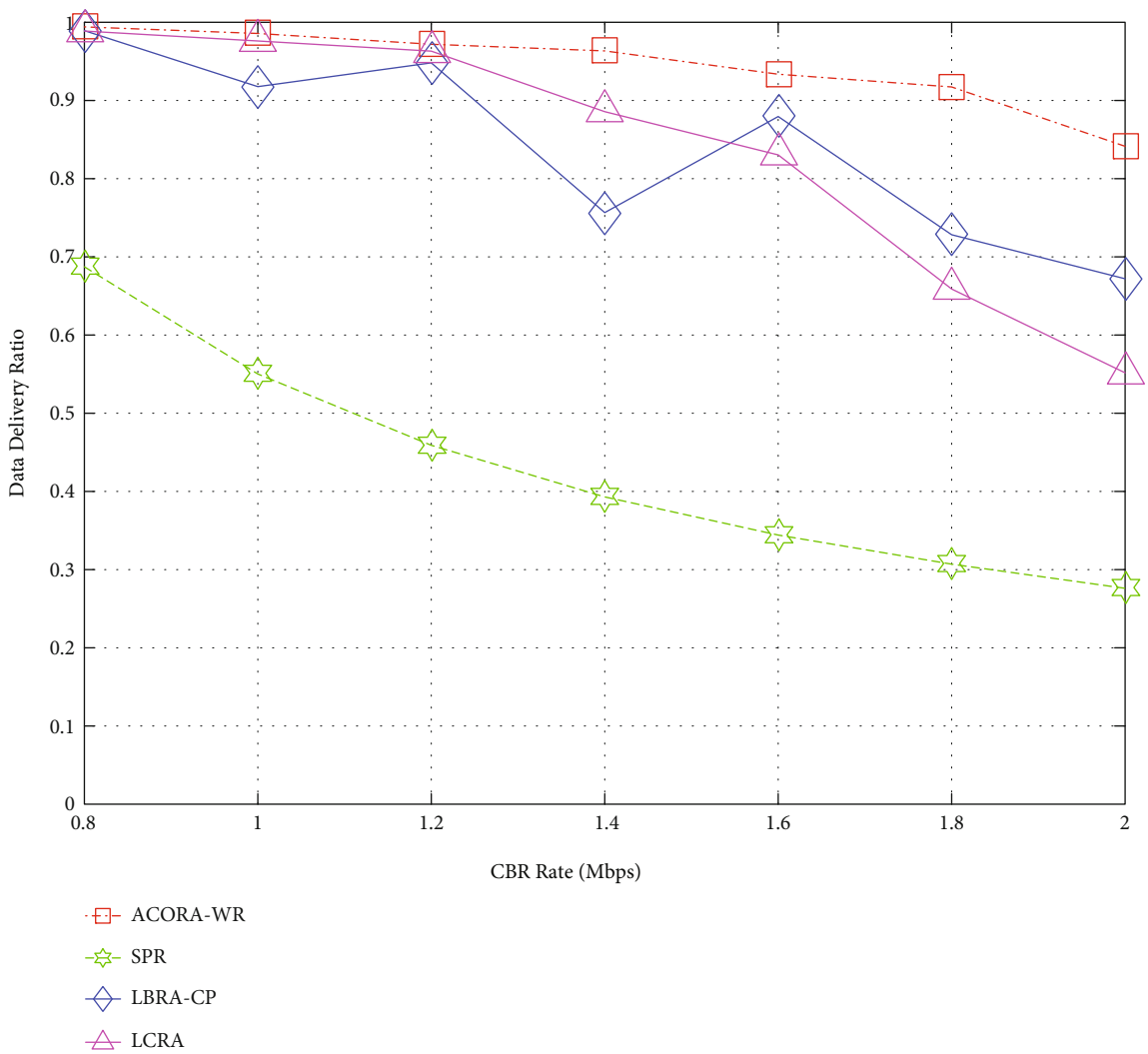
Table 2: Sending and receiving cities of simulated data flow.

Source	Destination
Madrid	Brisbane
Istanbul	Sulawesi
Mumbai	Perth
Urumqi	Johannesburg
Shanghai	Cairo
Johannesburg	Shanghai
Perth	Istanbul

(1) LBRA-CP is a load-balanced routing algorithm based on congestion prediction for LEO satellite networks. LBRA-CP sets up hot spot areas to avoid congestion by transferring traffic from hot spot areas to nonhot spot areas

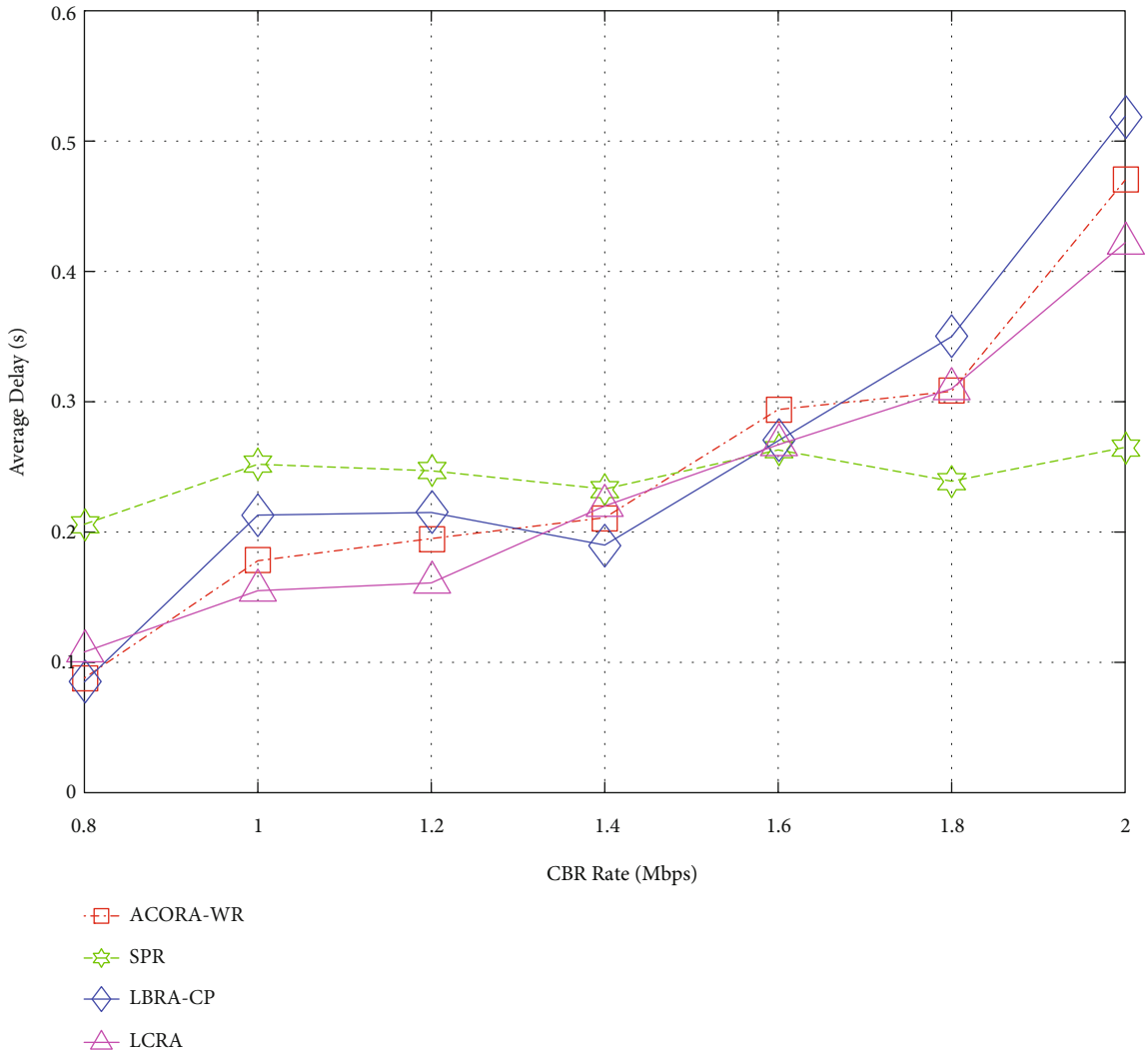
(2) SPR is the shortest path routing protocol based on Dijkstra

(3) LCRA is the cost-balanced routing protocol in LEO satellite networks. It selects the shortest path to the destination and uses the congestion information of neighboring nodes to alleviate the network congestion. It has low computational complexity and can effectively reduce network overhead



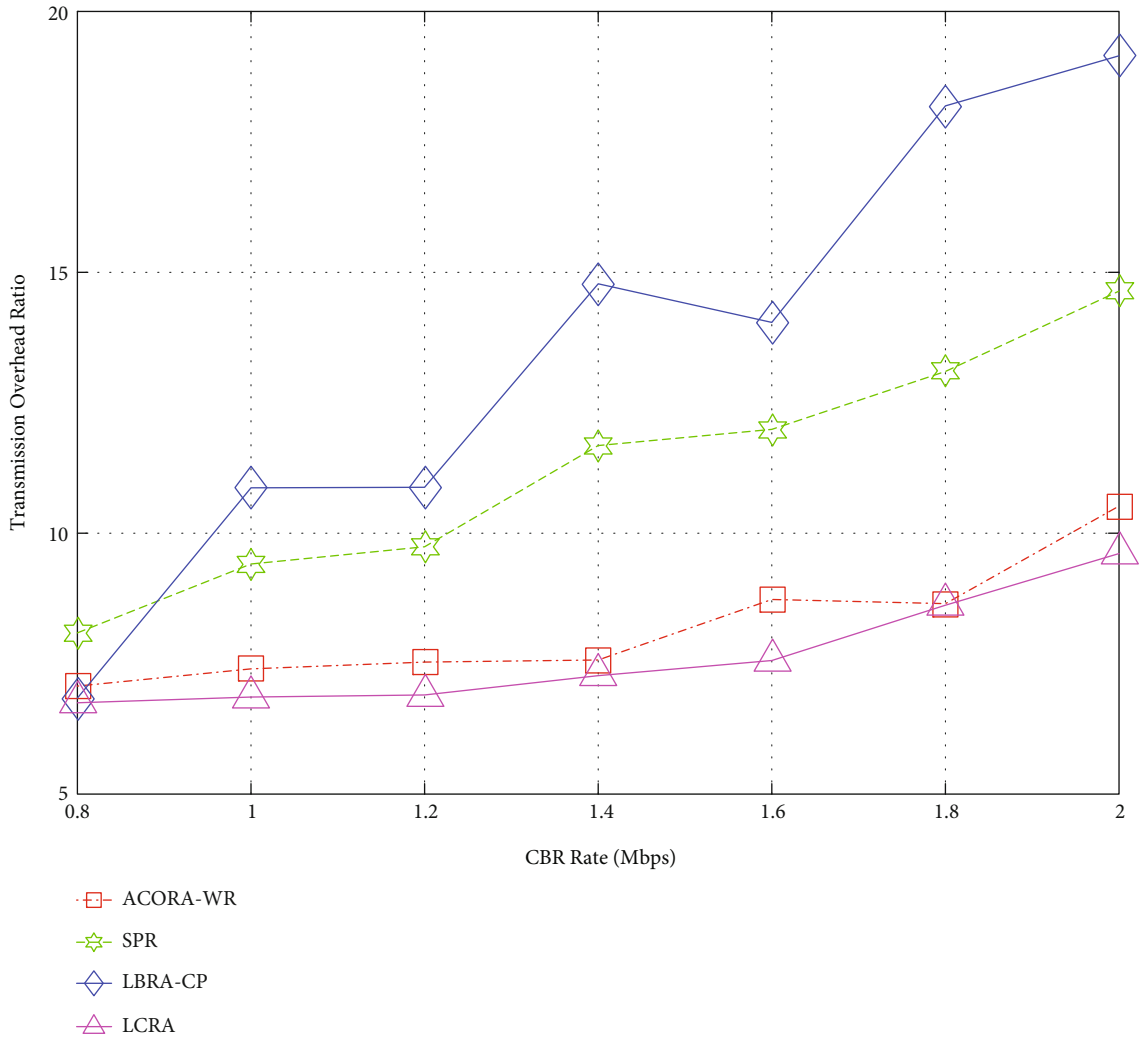
(a) Data delivery ratio

Figure 6: Continued.



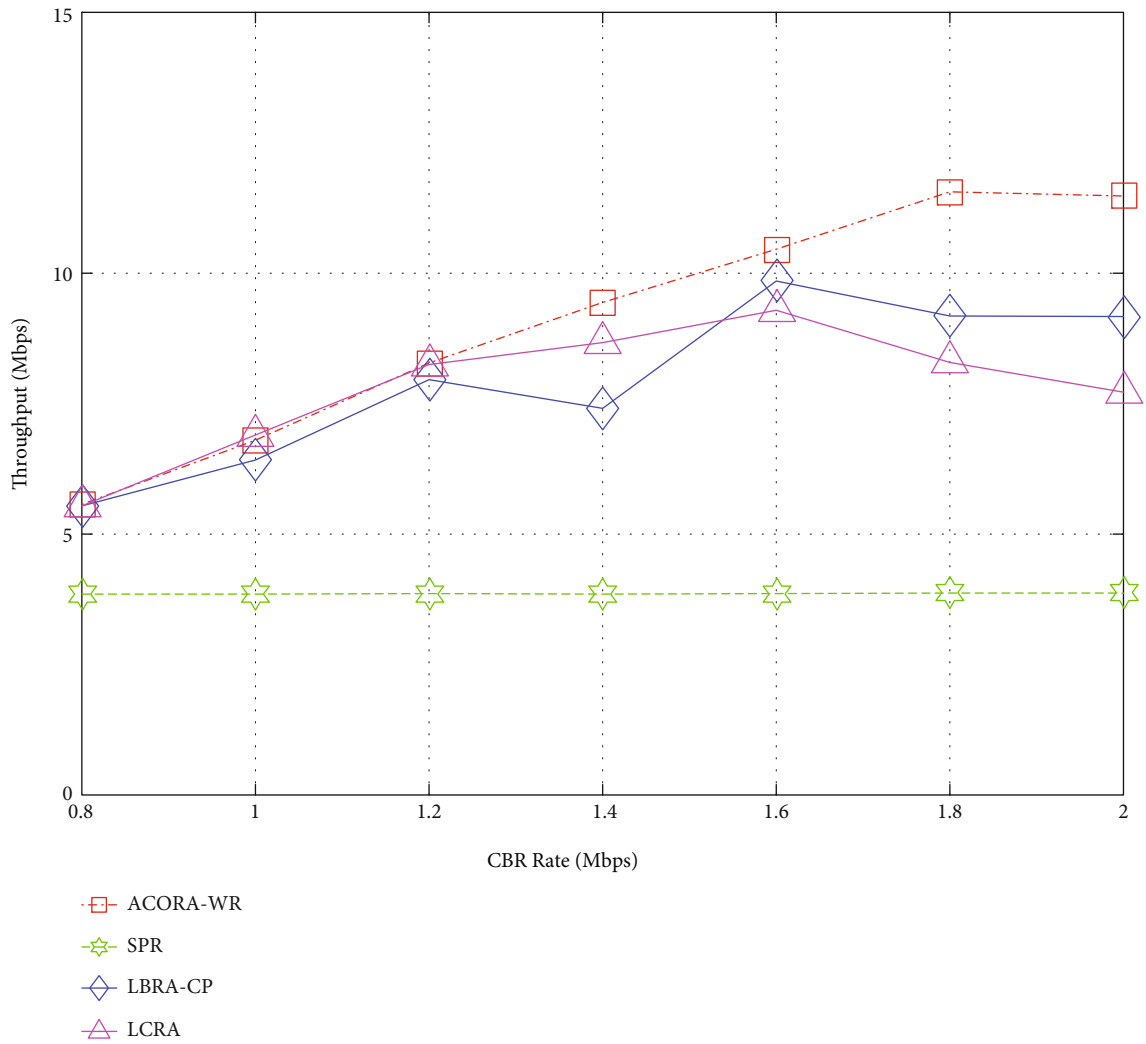
(b) Average delay

Figure 6: Continued.



(c) Transmission overhead ratio

Figure 6: Continued.



(d) Throughput

Figure 6: The performance of ACORA-WR with varying CBR rate.

The performance is evaluated using following metrics:

(1) *Data Delivery Ratio*. The ratio of the number of messages arrived at destinations to the number of messages expected to arrive at destinations

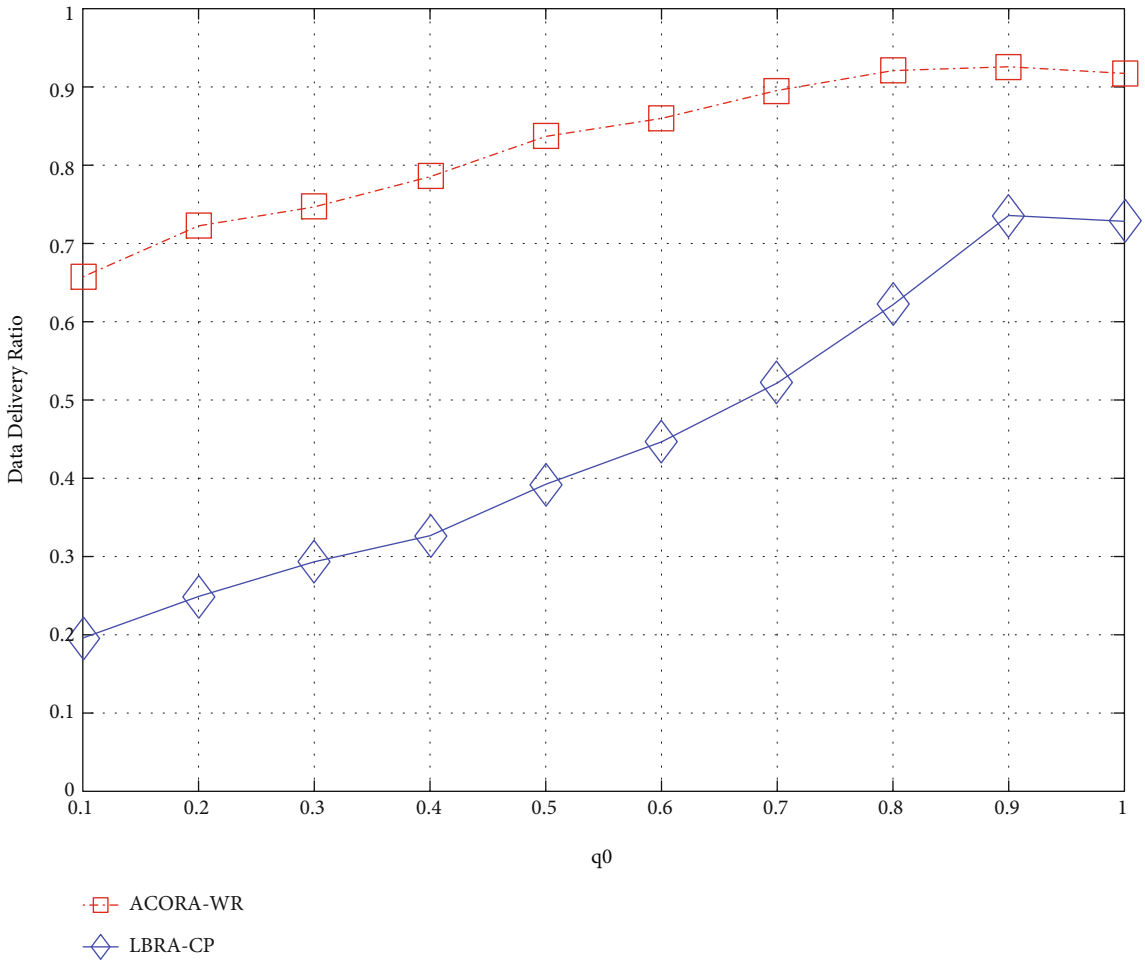
(2) *Average Delay*. The average time spent by all messages from the sources to the destinations

(3) *Transmission Overhead Ratio*. The total number of relayed messages divided by the number of messages arriving at the destination

(4) *Throughput*. The total successful message delivery rates on the destinations

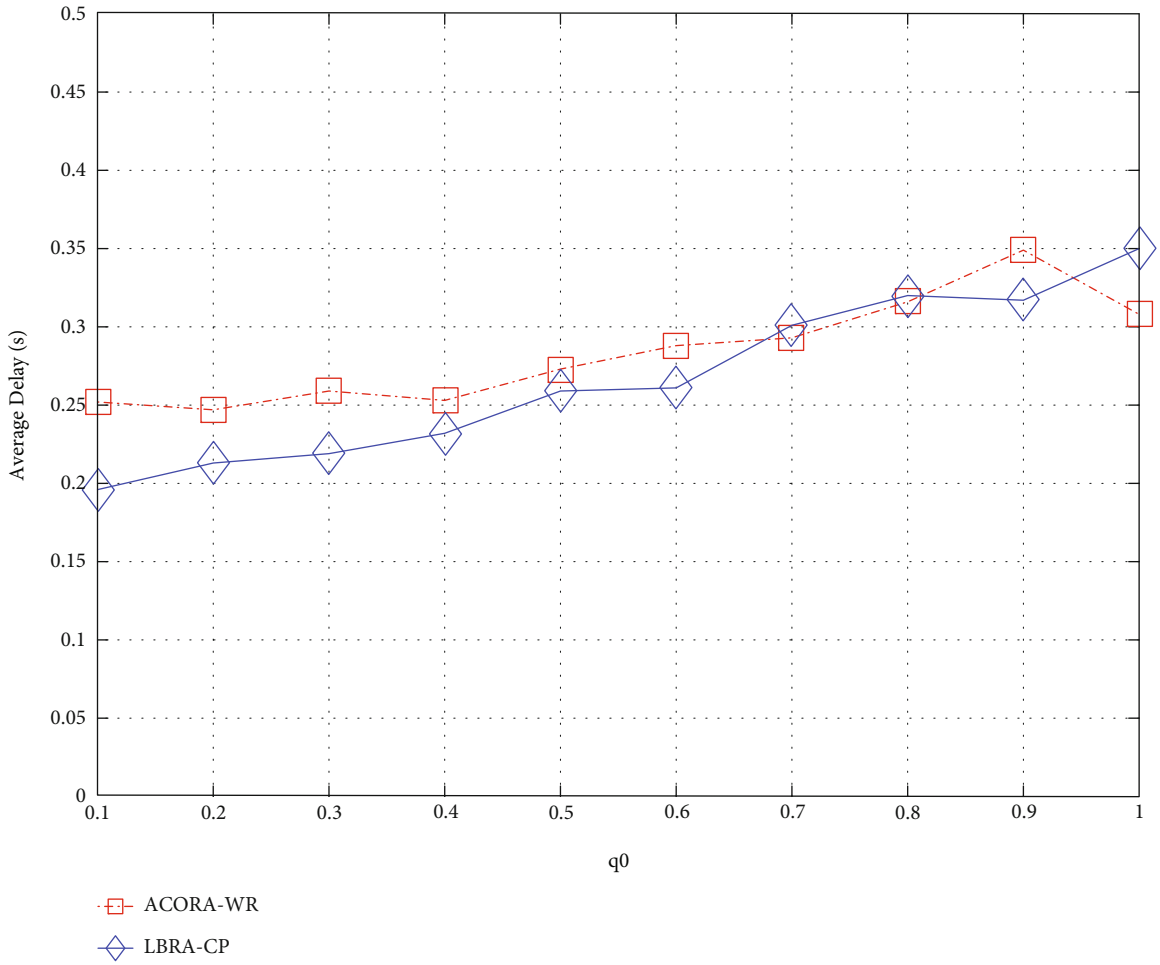
5.2. The Performance of ACORA-WR with Varying CBR Rate. To evaluate the effects of the transmission rate of CBR (constant bit rate) on ACORA-WR performance, we increase the total CBR generation rate in the networks from 0.8 Mbps to 2 Mbps. The buffer size of the node is 50 packets. The total simulation time is 100 s. For comparison fairness, we set the q_0 of ACORA-WR and LBRA-CP to be the same.

In Figure 6(a), as the CBR rate increases, the data delivery ratio in all protocols decreases due to the heavy transmission load. ACORA-WR shows the highest data delivery ratio and traffic balancing capability in the range of 0.8 Mbps to 2 Mbps of CBR rate. The shortest path algorithm SPR has the worst performance, and the average data delivery ratio of ACORA-WR is nearly twice as high as that of SPR. The LCRA algorithm considers the congestion information of adjacent nodes and chooses the shortest congestion path to the destination. However, LCRA only considers the congestion situation of adjacent nodes in routing, so it cannot optimize the routes from the global perspective. Due to the limitation of the transmission range, LCRA cannot provide enough paths to dynamically adapt to the traffic load. In contrast, ACORA-WR considers both global and local aspects through the ant colony algorithm. Compared with LCRA, the average data delivery ratio of ACORA-WR



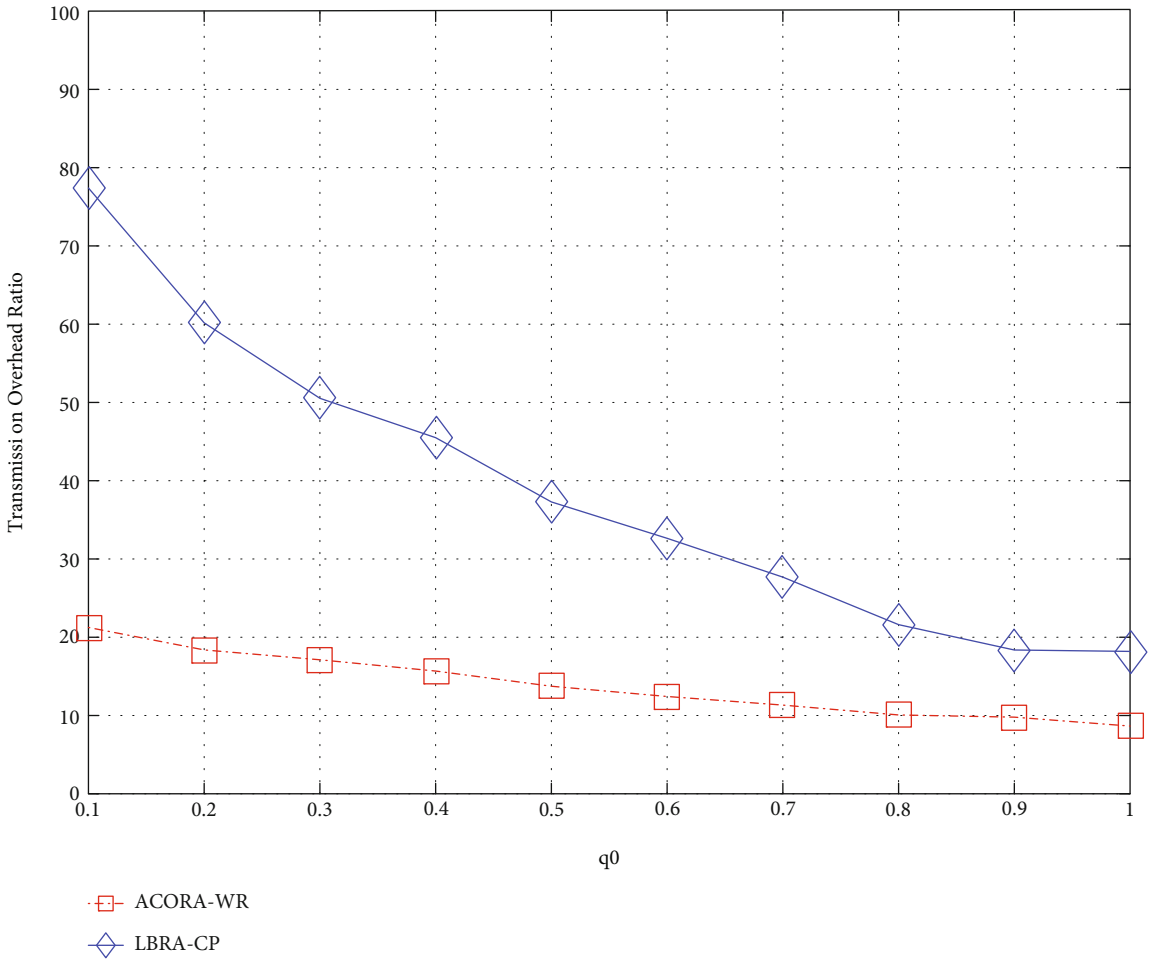
(a) Data delivery ratio

Figure 7: Continued.



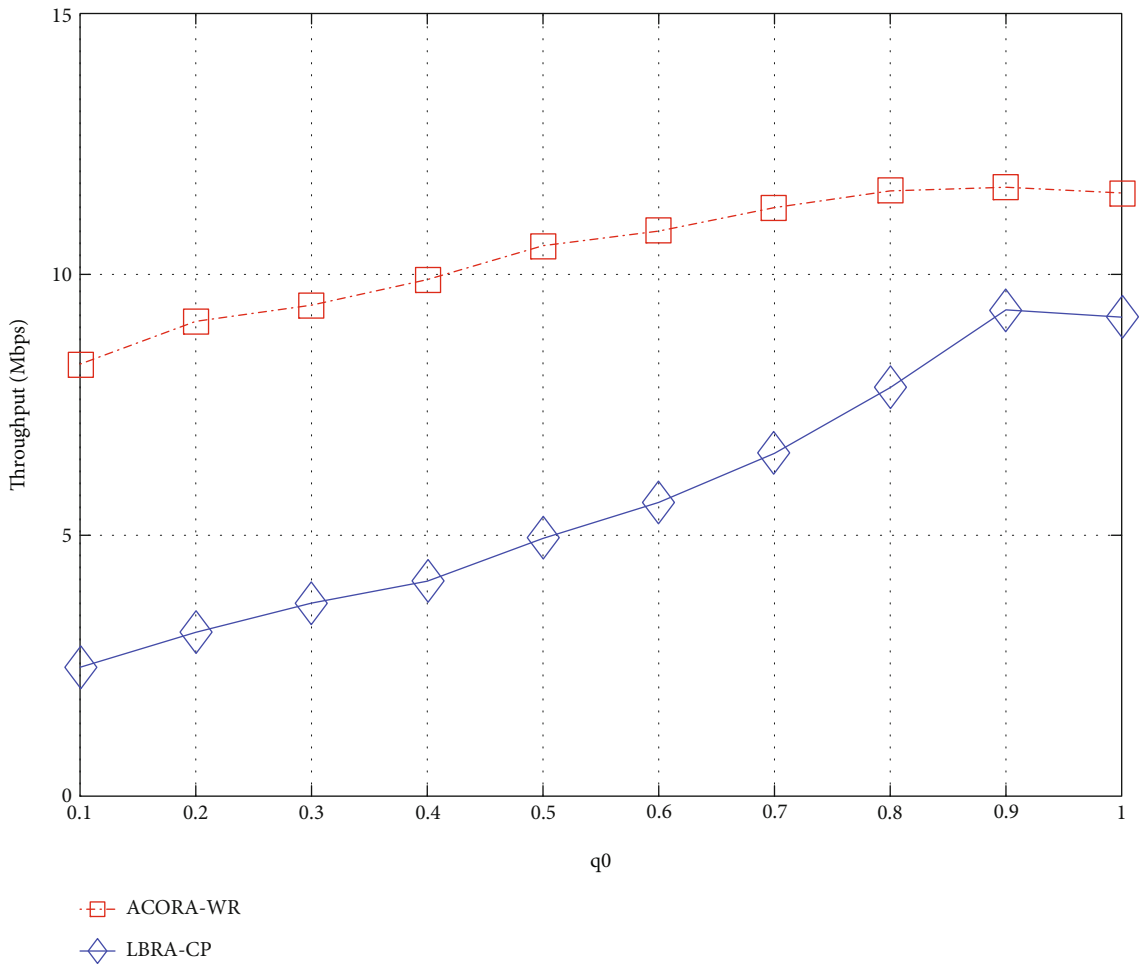
(b) Average delay

Figure 7: Continued.



(c) Transmission overhead ratio

Figure 7: Continued.



(d) Throughput

Figure 7: The performance of ACORA-WR with varying q_0.

increases by about 10%, and the LCRA performance decreases significantly when the CBR rate is at 2 Mbps. The data delivery ratio of ACORA-WR is nearly 50% higher than LCRA when the CBR rate is at 2 Mbps. LBRA-CP is also based on ant colony algorithms and combines congestion prediction to find the best route; however, its overhead is higher. ACORA-WR can speed up ant colony search and modify routing table in time. The experimental results show that the data delivery ratio of ACORA-WR is about 10% higher than LBRA-CP.

From Figure 6(b), we can see that the performance of our ACORA-WR is between LCRA and LBRA-CP in terms of the average delay. LCRA has the best average delay because the length of the path chosen by LCRA is generally smaller than ACORA-WR and LBRA-CP. However, a large amount of data is discarded while waiting in queues, and the percentage of packets discarded increases with the increase of the CBR rate. Similarly, SPR has a good average delay, but they always choose the shortest path. When the load is heavy, many packets will be discarded. LBRA-CP and ACORA-WR require higher average delay to ensure data delivery. The results show that the average delay of ACORA-WR is better than that of LBRA-CP in the range of 0.8 Mbps to 2 Mbps CBR rate.

In terms of transmission overhead ratio, as shown in Figure 6(c), LCRA has the lowest transmission overhead ratio, and ACORA-WR is slightly higher than LCRA because both LCRA and ACORA-WR limit the range of data transmission, but ACORA-WR can choose more flexible paths than LCRA. However, compared with LBRA-CP, the average transmission overhead of ACORA-WR is about 40% lower than that of LBRA-CP. This is because LBRA-CP has a broader ant search range and a slower ant colony convergence rate than ACORA-WR. Some data choose longer and longer paths, which results in a higher transmission overhead ratio.

In Figure 6(d), throughput of all protocols increases with the increase of CBR rate. ACORA-WR always maintains the highest throughput, SPR has the lowest throughput, LCRA has slightly higher throughput than LBRA-CP when CBR rate is 0.8 Mbps to 1.4 Mbps, and LBRA-CP overwhelms LCRA when CBR rate reaches 1.6 Mbps. When CBR rates range from 1.8 Mbps to 2 Mbps, the throughput of ACORA-WR increases by about 25% compared with LBRA-CP and 40% compared with LCRA.

In summary, the experimental results show that ACORA-WR can maintain a high data delivery ratio and network throughput. It can balance the traffic under heavy network load while maintaining a small average delay and overhead.

5.3. The Performance of ACORA-WR with Varying the Value of q_0. The value of q_0 determines the relative importance of exploring new paths and using prior knowledge. According to Equation (7), the probability values are calculated using congestion-avoiding heuristic information and pheromone when $q < q_0$. Routing forwarding is chosen randomly when the link is available and $q \geq q_0$. q is a random number, q_0 is a fixed value, and q_0 determines the probability ratio of the two rules to choose a route. The larger the value of q_0 is, the smaller the probability of forward ants will explore randomly, and the lower the ability of finding a new route. The smaller the value of q_0 is, the slower the convergence speed of the ant colony algorithm will have, but it will result in higher network overhead and delay. To evaluate the impact of the value of q_0 on the performance of two kinds of ant colony algorithm ACORA-WR and LBRA-CP, we change the value of q_0 in the networks from 0.1 to 1 and set the CBR generation rate to 1.8 Mbps. The buffer size of the node is 50 packets.

In Figure 7(a), both ACORA-WR and LBRA-CP have significantly improved data delivery rates as the value of q_0 increases, and both ACORA-WR and LBRA-CP show the best performance when the value of q_0 is 0.9. When the value of q_0 is low, most ants in ACORA-WR and LBRA-CP randomly select routes in the networks, but due to the narrowing of the window, ACORA-WR data is easier to find destinations, so even when the value of q_0 is low, it still shows a significantly better data delivery rate. When the value of q_0 is in the range of 0.1 to 0.8, the data delivery rate of ACORA-WR is about 90% higher than that of LBRA-CP.

In Figure 7(b), with the increase of q_0, the average delay of ACORA-WR differs slightly from that of LBRA-CP. This is because we take into account the delay of the entire path from the sources to the destinations of the data flows in the pheromone overlay process and adjust the traffic allocation of the link locally with congestion-avoiding heuristic information to ensure a lower average delay while increasing the data delivery rate.

In Figure 7(c), LBRA-CP has a very high transmission overhead ratio when the value of q_0 is small, because most of the data is searched randomly and unrestrictedly, resulting in longer paths being selected and higher transmission overhead ratio. With the constraint of the window reduction mechanism, even if most ants randomly select routes, they can still successfully forward data to their destinations with a low transmission overhead ratio.

Figure 7(d) shows the changes in network throughput of ACORA-WR and LBRA-CP as the value of q_0 increases. Overall, the network throughput of ACORA-WR and LBRA-CP increases with the value of q_0. The networks throughput of ACORA-WR is significantly better than that of LBRA-CP when the value of q_0 ranges from 0.1 to 0.8. When q_0 is 0.9, both algorithms have the best performance. With faster convergence, LBRA-CP throughput tends to approach ACORA-WR gradually, but overall, ACORA-WR is still better than LBRA-CP.

6. Conclusions and Future Work

In this paper, we aimed at balancing traffic load in LEO satellite networks and proposed an improved ant colony routing algorithm with window reduction. When packets are forwarded between intersatellite links, congestion-avoiding heuristic information is added to the state transition rules, and forward ants detect the route through our improved state transition rules periodically, as to tackle the problem that ant colony algorithms fail to adjust the local information in time. Moreover, we proposed a window reduction mechanism to limit the transmission direction of data and remove redundant paths, which accelerates the convergence speed of ant colony and reduce network transmission overhead. Finally, the new routing algorithm is implemented on the network simulation platform NS2. The experimental results show that compared with other distributed algorithms in LEO satellite networks, the proposed ACORA-WR scheme demonstrates higher data delivery ratio and network throughput. In conclusion, ACORA-WR can effectively balance the network traffic load and improve the data delivery ratio of satellite networks.

In our future work, we will consider energy consumption in load balancing algorithms. Satellites mainly receive energy from the sun, which is not guaranteed anywhere as they move. Load balancing for LEO satellite networks under these cases requires further study.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant Nos. 61702127 and 62062006), the Science and Technology Program of Guangzhou (Grant No. 201804010461), the China Scholarship Council (Grant Nos. 201908440064 and 201908440085), the Hundred Young Talents Plan Project of Guangdong University of Technology (Grant No. 220413618), and the Guangxi Innovation-driven Development Major Project Platform (No. Guike AA20302002).

References

[1] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2714–2741, 2018.

[2] N. N. Dao, Q. V. Pham, N. H. Tu et al., "Survey on aerial radio access networks: toward a comprehensive 6G access infrastructure," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1193–1225, 2021.

[3] B. Shang, Y. Yi, and L. Liu, "Computing over space-air-ground integrated networks: challenges and opportunities," *IEEE Network*, vol. 35, no. 4, pp. 302–309, 2021.

[4] Q. Wang, X. Li, Y. Liu, L. T. Alex, S. A. Khowaja, and V. G. Menon, "UAV-enabled non-orthogonal multiple access networks for ground-air-ground communications," *IEEE Transactions on Green Communications and Networking*, p. 1, 2022.

[5] X. Deng, J. Shao, L. Chang, and J. Liang, "A blockchain-based authentication protocol using cryptocurrency technology in LEO satellite networks," *Electronics*, vol. 10, no. 24, p. 3151, 2021.

[6] X. Li, Y. Zheng, W. U. Khan et al., "Physical layer security of cognitive ambient backscatter communications for green Internet-of-Things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[7] Z. Xie, Z. Chu, V. G. Menon, S. Mumtaz, and J. Zhang, "Exploiting benefits of IRS in wireless powered NOMA networks," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 175–186, 2022.

[8] G. Li, H. Liu, G. Huang, X. Li, B. Raj, and F. Kara, "Effective capacity analysis of reconfigurable intelligent surfaces aided NOMA network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, 16 pages, 2021.

[9] X. Li, F. Tang, L. Chen, and J. Li, "A state-aware and load-balanced routing model for LEO satellite networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, Singapore, Dec. 2017.

[10] Z. Liu, J. Li, Y. Wang, X. Li, and S. Chen, "HGL: a hybrid global-local load balancing routing scheme for the Internet of Things through satellite networks," *International Journal of Distributed Sensor Networks*, vol. 13, 16 pages, 2017.

[11] E. Wang, H. Li, and S. Zhang, "Load balancing based on cache resource allocation in satellite networks," *IEEE Access*, vol. 7, pp. 56864–56879, 2019.

[12] Z. Na, Z. Pan, X. Liu, Z. Deng, Z. Gao, and Q. Guo, "Distributed routing strategy based on machine learning for LEO satellite network," *Wireless Communications and Mobile Computing*, vol. 2018, 10 pages, 2018.

[13] T. Taleb, D. Mashimo, A. Jamalipour, N. Kato, and Y. Nemoto, "Explicit load balancing technique for NGEO satellite IP networks with on-board processing capabilities," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 281–293, 2009.

[14] G. Song, M. Chao, B. Yang, and Y. Zheng, "TLR: a traffic-light-based intelligent routing strategy for NGEO satellite IP networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3380–3393, 2014.

[15] E. Papapetrou, S. Karapantazis, and F. N. Pavlidou, "Distributed on-demand routing for LEO satellite systems," *Computer Networks*, vol. 51, no. 15, pp. 4356–4376, 2007.

[16] S. Karapantazis, E. Papapetrou, and F. N. Pavlidou, "Multiservice on-demand routing in LEO satellite networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, pp. 107–112, 2009.

[17] Y. Rao and R. C. Wang, "Agent-based load balancing routing for LEO satellite networks," *Computer Networks*, vol. 54, no. 17, pp. 3187–3195, 2010.

[18] H. Wang, G. Wen, N. Liu, J. Zhang, and Y. Tao, "A load balanced routing algorithm based on congestion prediction for LEO satellite networks," *Cluster Computing*, vol. 22, no. S4, pp. 8025–8033, 2019.

[19] F. Bokhari and G. Zaruba, "Interference-aware routing using ant colony optimization in wireless mesh networks," in *Wireless Communications and Networksing Conference*, Budapest, Hungary, April 2009.

[20] M. H. Eiza, T. Owens, Q. Ni, and Q. Shi, "Situation-aware QoS routing algorithm for vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5520–5535, 2015.

[21] X. Xiang, Y. Tian, X. Zhang, J. Xiao, and Y. Jin, "A pairwise proximity learning-based ant colony algorithm for dynamic vehicle routing problems," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP(99), pp. 1–12, 2021.

[22] S. Chatterjee and S. Das, "Ant colony optimization based enhanced dynamic source routing algorithm for mobile ad-hoc network," *Information Sciences.*, vol. 295, pp. 67–90, 2015.

[23] M. Rathee, S. Kumar, A. H. Gandomi, K. Dilip, B. Balusamy, and R. Patan, "Ant colony optimization based quality of service aware energy balancing secure routing algorithm for wireless sensor networks," *IEEE Transactions on Engineering Management*, vol. 68, no. 1, pp. 170–182, 2021.

[24] W. Ni, Z. Xu, J. Zou, Z. Wan, and X. Zhao, "Neural network optimal routing algorithm based on genetic ant colony in IPv6 environment," *Computational Intelligence and Neuroscience*, vol. 2021, 3115713 pages, 2021.

[25] S. Kilic and O. Ozkan, "Ant colony optimization approach for satellite broadcast scheduling problem," in *2017 8th International Conference on Recent Advances in Space Technologies*, Istanbul, Turkey, June 2017.

[26] H. S. Chang, B. W. Kim, C. G. Lee et al., "FSA-based link assignment and routing in low-earth orbit satellite networks," *IEEE Transactions on Vehicular Technology*, vol. 47, no. 3, pp. 1037–1048, 1998.

[27] D. Bertsekas and R. Gallager, *Data networks*, Prentice-Hall, Inc, 1992.

[28] L. Bai Jianjun and X.,. L. Zexin, "Compact explicit multi-path routing for LEO satellite networks, in," in *Workshop on High Performance Switching and Routing*, pp. 386–390, Hong Kong, China, May 2005.

[29] M. Dorigo and L. M. Gambardella, "Ant colonies for the travelling salesman problem," *Biosystems*, vol. 43, no. 2, pp. 73–81, 1997.

[30] X. Liu, X. Yan, Z. Jiang, C. Li, and Y. Yang, "A low-complexity routing algorithm based on load balancing for LEO satellite networks," in *2015 IEEE 82nd Vehicular Technology Conference*, Boston, MA, USA, Sept. 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9162422, 10 pages
https://doi.org/10.1155/2022/9162422

Research Article

Empirical Matching-Based Computation Offloading Optimization for 5G and Edge Computing-Integrated EIoT

Hui Zhang, Huixia Ding, Yang Wang, Sachula Meng, Sicheng Zhu, Ling Teng, and Fangyun Dong

China Electric Power Research Institute, Beijing, China

Correspondence should be addressed to Hui Zhang; zhanghui@epri.sgcc.com.cn

Received 28 January 2022; Revised 24 February 2022; Accepted 1 March 2022; Published 15 April 2022

Academic Editor: Han Wang

Electric Internet of things (EIoT) that integrates 5G and edge computing can provide data transmission and processing guarantee for smart grid. However, computation offloading optimization including joint optimization of server selection and computation resource allocation still faces several challenges such as difficulty in tradeoff balance among various quality of service (QoS) parameters, coupling between server selection and computation resource allocation, and multi-device competition. To address these challenges, we propose an empirical matching-based computation offloading optimization algorithm for 5G and edge computing-integrated EIoT. The optimization objective is to minimize the computation offloading delay by jointly optimizing large timescale server selection and small timescale computation resource allocation. We first model the large timescale server selection problem as a many-to-one matching problem, which can be decoupled from small timescale computation resource allocation by establishing a matching preference list based on empirical performance. Then, the large timescale server selection problem is solved by pricing-based matching with a quota algorithm. Furthermore, based on the obtained suboptimal result of large timescale server selection, the small timescale computation resource allocation problem is subsequently solved by Lagrange dual decomposition, the result of which is used to update large timescale empirical performance. Finally, extensive simulations are carried out to demonstrate the superior performance of the proposed algorithm by comparing it with existing algorithms.

1. Introduction

In order to achieve the energy supply-demand balance and the safe and stable operation of the smart grid, massive electric Internet of things (EIoT) devices need to be deployed to support multiple types of real-time data collection, such as voltage, current, active/reactive power, electric energy, temperature, and humidity [1–3]. The integration of 5G and edge computing provides a viable solution for the real-time data monitoring. Specifically, 5G provides communication guarantee for real-time data collection and information interaction due to the advantages of high reliability, wide connection, and low delay [4]. Compared with ad hoc networks, 5G with central control functionality can achieve centralized and efficient resource scheduling and management for computation offloading optimization in EIoT. Edge computing provides computation guarantee for real-time data processing with the superiority of high-speed computation and low-delay transmission [5, 6]. Compared with cloud computing, edge computing alleviates the high transmission delay and network congestion by reducing transmission distance between devices and servers. In addition, it outperforms fog computing in terms of privacy, security, and computation capacity due to the unified security management and abundant computation resources.

Computation offloading plays a crucial role in 5G and edge computing-integrated EIoT, which includes server selection and computation resource allocation [7, 8]. First, devices select an appropriate server and offload the computation task to the selected edge server via 5G for real-time processing. Then, edge servers allocate computation resources based on service requirements and computation capacity to

reduce computation delay [9]. However, the computation offloading optimization for 5G and edge computing-integrated EIoT still faces several challenges.

(i) Difficulty in tradeoff balance among various QoS parameters: services in 5G and edge computing-integrated EIoT have differentiated quality of service (QoS) requirements in delay, reliability, and so on. The QoS parameters are mutually influencing; e.g., improving reliability by using extra signaling overhead may harm delay and vice versa. Hence, it is necessary to balance the tradeoff among various QoS parameters through computation offloading optimization.

(ii) Coupling between large timescale server selection and small timescale computation resource allocation: server selection in large timescale leads to changes in the characteristics and number of devices served by servers, which in turn affects the computation resource allocation in small timescale. The inefficient decisions of computation resource allocation in small timescale directly affect the data computation delay performance of devices, which in turn affects the evaluation of the empirical performance of the servers in large timescale and may result in frequent server selection switching.

(iii) Multi-device competition in server selection and computation resource allocation: the contradiction between limited computation resources and multiple devices not only leads to competition for powerful servers in large timescale server selection but also causes competition for CPU frequency in small timescale computation resource allocation. In particular, the performance of delay and reliability is severely reduced by the competition for the same server.

Computation offloading within multi-access edge computing (MEC) has drawn extensive research in academia. In [10], Zaman et al. provided a survey of mobility management-based task offloading in edge networks. The taxonomy of research work classification is based on objectives, constraints, models, scenarios, and so on, and several future research directions for offloading in edge computing were presented. Lagrange dual approach has been widely adopted for addressing computation offloading problems with convex properties. In [11], Zhou et al. studied a multi-access edge computing-based task offloading algorithm for lightweight user by combining machine learning with Lagrange dual approach to minimize offloading energy consumption. In [12], Wang et al. proposed a Lagrange dual method-based communication resource allocation algorithm to minimize the weighted sum-energy consumption in MEC. Nevertheless, Lagrange dual approach cannot handle nonconvex problems which are common in multi-access networks. As an alternative approach, game theory is explored to cope with multi-user task offloading problems in MEC. In [13], Apostolopoulos et al. proposed a noncooperative game-based distributed cognitive data offloading algorithm to maximize users' utilities. In [14], Wang et al. proposed a noncooperative game-based computation offloading algorithm to adjust the offloading probability of each user and maximize vehicular utility. The drawback of game theory lies on required prior knowledge to derive utilities of players, which are essential to solve both cooperative and noncooperative games. Compared with game theory, reinforcement learning is more suitable for addressing computation offloading problems with uncertain information. In [15], Dinh et al. studied a model-free reinforcement learning-based computation offloading mechanism to maximize the utility function of each mobile user. In [16], Gao et al. proposed a novel Q learning-based computation offloading scheme for MEC system to minimize the system loss function. However, the aforementioned works have not considered the difficulty in tradeoff balance among various QoS requirements such as delay and reliability. Moreover, they focus on single timescale computation offloading optimization, while the coupling and mutual interconnection between large timescale server selection and small timescale computation resource allocation are neglected. The elimination of multi-access competition requires cooperation among different devices, which imposes another challenge on reinforcement learning-based computation offloading.

Matching theory is an effective method to solve the optimization of mutual relationship between two sides [17]. In [18], Seng et al. proposed an efficient task virtual machine matching algorithm to coordinate the offloading problems among mobile users and edge servers. In [19], Zhou et al. proposed a low complexity and stable computation offloading mechanism to minimize the total network delay based on pricing-based matching. However, the existing works only consider the one-to-one matching, which is not suitable to the scenario that the same server can serve multiple devices at the same time. In [20], Liu et al. proposed a many-to-one matching-based channel allocation algorithm for mobile users, which achieves higher throughput performance. In [21], Wang et al. proposed a many-to-one matching theory-based subchannel allocation algorithm to reduce the energy efficiency. However, the joint optimization of reliability and delay is ignored. In [22], Zhang et al. proposed a joint optimization algorithm of computation offloading and computation resource allocation, which optimizes network stability and achieves the tradeoff between energy efficiency and average service delay. Nevertheless, the aforementioned works do not consider the coupling between small timescale computation resource allocation and large timescale sever selection.

Motivated by the aforementioned challenges including difficulty in tradeoff balance among various QoS parameters, coupling between large timescale server selection and small timescale computation resource allocation, and the multi-device competition in server selection and computation resource allocation, we propose an empirical matching-based computation offloading optimization algorithm for 5G and edge computing-integrated EIoT. The optimization objective is to minimize the computation offloading delay by jointly optimizing large timescale server selection and

small timescale computation resource allocation. We first model the large timescale server selection problem as a many-to-one matching problem, which can be decoupled from small timescale computation resource allocation by establishing matching preference lists based on empirical performance. Then, the large timescale server selection problem is solved by pricing-based matching with quota algorithm. Furthermore, based on the obtained suboptimal result of large timescale server selection, the small timescale computation resource allocation problem is subsequently solved by Lagrange dual decomposition, the result of which is used to update large timescale empirical performance. Finally, extensive simulations are carried out to demonstrate the superior performance of the proposed algorithm by comparing it with existing algorithms. The main contributions of this paper are summarized as follows:

(i) Differentiated QoS guarantee: differentiated QoS demands such as reliability and delay are guaranteed through the minimization of the long-term total delay of all devices under the constrains of server computation capacity and transmission signal-to-interference-plus-noise ratio (SINR). Specifically, reliability is enforced by eliminating servers which cannot satisfy the SINR constraint from the device preference lists.

(ii) Multi-timescale computation offloading optimization: server selection is optimized in large timescale based on empirical matching with empirical performance enabled matching preference list and pricing-based matching with quota enabled matching conflict elimination. On the basis, computation resource allocation is optimized in small timescale based on Lagrange dual decomposition to reduce computation delay.

(iii) Extensive performance evaluation: compared with two state-of-the-art algorithms, simulation results demonstrate that the proposed algorithm has superior performance in computation offloading delay. Moreover, the impacts of key parameters such as quota and SINR threshold on delay performance are revealed to provide guidance for the real-world implementation of computation offloading optimization in 5G and edge computing-integrated EIoT.

The rest of the paper is organized as follows. Section 2 demonstrates the system model of computation offloading for 5G and edge computing-integrated EIoT and formulates the multi-timescale joint optimization problem. Section 3 introduces the proposed empirical matching-based computation offloading optimization algorithm. The simulation results are shown in Section 4. Section 5 concludes the paper.

2. System Model

In this paper, we consider a scenario of computation offloading for 5G and edge computing-integrated EIoT, as shown in Figure 1. The set of M devices and N base stations (BSs) are represented as $\mathcal{M} = \{1, 2, \cdots, m, \cdots, M\}$ and $\mathcal{N} = \{1, 2, \cdots, n, \cdots, N\}$, respectively. The edge server and BS are located in the same place to provide computation and communication services. For simplicity, we denote the sets of BS and edge server as $\mathcal{N}$. An example is shown in Figure 1, where the process of computation offloading includes two stages, i.e., server selection and computation resource allocation. In the first stage, when devices 1, 2, and 3 select the same BS 1 with limited offloading quota of 2 devices, BS 1 chooses to allocate computation resources for devices 1 and 2 and reject the access request of device 3. Device 3 finally offloads the computation task to BS 2 with a larger quota. We assume that BSs receive the time synchronization signal from the 5G time synchronization network and broadcast to all edge servers and devices to ensure that computation offloading is performed on the basis of time synchronization [15, 23].

A slot model is adopted [24] where server selection is optimized in large timescale, i.e., epoch, and computation resource allocation is optimized in small timescale, i.e., slot. The total optimization period has G epochs, each of which consists of T_0 slots, i.e., $T = GT_0$. The sets of epochs and slots are denoted as $\mathcal{G} = \{1, \cdots, g, \cdots, G\}$ and $\mathcal{T} = \{1, \cdots, t, \cdots, T\}$. The slot set of the g-th epoch is given by

$$\mathcal{T}(g) = \{(g-1)T_0 + 1, (g-1)T_0 + 2, \cdots, gT_0\}. \quad (1)$$

The key notations used in this paper are summarized in Table 1.

2.1. Transmission Model. At each epoch, a device selects a server for computation offloading. Define $x_{m,n}(g)$ as the server selection variable, where $x_{m,n}(g) = 1$ represents that device m selects server n for computation offloading, and otherwise, $x_{m,n}(g) = 0$. Considering uplink data transmission, the SINR between device m and server n in the t-th slot is [25]

$$\gamma_{m,n}(t) = \frac{P_{TX}(t) h_{m,n}(t)}{B_n N_0 + \lambda_{m,n}(t)}, \quad (2)$$

where B_n is the transmission bandwidth. $P_{TX}(t)$ is the transmission power, $h_{m,n}(t)$ is the channel gain between device m and server n in the t-th slot, N_0 is the noise spectral power density, and $\lambda_{m,n}(t)$ is the electromagnetic interference power.

To ensure reliability requirement, SINR constraint is given by

$$\gamma_{m,n}(t) \geq \gamma_{\min}, \quad (3)$$

where $\gamma_{\min}$ is SINR threshold for data reliable transmission. The transmission rate between device m and server n in the t-th slot is given by

$$R_{m,n}(t) = B_n \log\left(1 + \gamma_{m,n}(t)\right). \quad (4$$

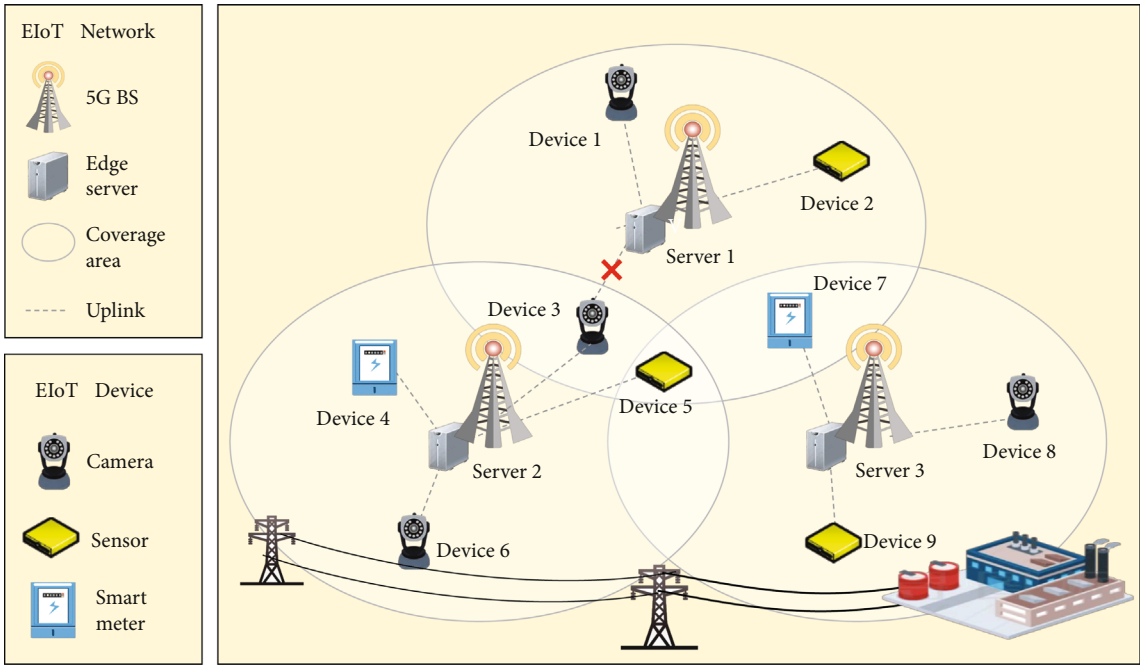


Figure 1: Computation offloading for 5G and edge computing-integrated EIoT.

Table 1: Summary of notations.

Notations	Definition	Notations	Definition
M	Number of devices	$\tau^{O}_{m,n}(t)$	Transmission delay of m to n in the t-th slot
N	Number of BSs	$\tau^{C}_{m,n}(t)$	Computation delay for processing $A_m(t)$ in the t-th slot
G	Number of epochs	$\tau_{m,n}(t)$	The total computation offloading delay
T	Number of slots	$f_{n,\max}$	The maximum computation resources
$\mathscr{G}$	Set of epochs	$f_{m,n}(t)$	Allocated computation resources by n to m
$\mathscr{T}$	Set of slots	$\mathscr{L}$	The preference list of devices
$\mathscr{M}$	Set of devices	$\gamma_{m,n}(t)$	SINR between m and n in the t-th slot
$\mathscr{N}$	Set of BSs or edge servers	$R_{m,n}(t)$	Transmission rate between m and n in the t-th slot
$x_{m,n}(g)$	Server selection variable	$\theta_{m,n}(g)$	Preference value of m towards n in the g-th epoch
B_n	Transmission bandwidth	$p_n(g)$	Server price of matching m with n
$P_{TX}(t)$	Transmission power	$\tilde{\tau}_{m,n}(g)$	Average offloading delay of m in the g-th epoch
$h_{m,n}(t)$	Channel gain between m and n	$\gamma_{\min}$	SINR threshold for data reliable transmission
N_0	Noise spectral power density	z	Iterative index of Lagrangian multiplier update
$\lambda_{m,n}(t)$	Electromagnetic interference power	$\rho_{\zeta_n(t)}(t,z)$	The update step length of $\zeta_n(t)$
$\zeta_n(t)$	Lagrangian sub vector	ε_m	Data computation density of m
Q_n	The quota of n	$A_m(t)$	The size of packets arriving at m in the t-th slot

We assume that the size of packets arriving at device m in the t-th slot is $A_m(t)$. Hence, the transmission delay of device m offloading $A_m(t)$ data to server n in the t-th slot is

$$\tau^{O}_{m,n}(t) = \frac{A_m(t)}{R_{m,n}(t)}. \tag{5}$$

2.2. Computation Model. Define $f_{m,n}(t)$ as the allocated computation resources by server n to device m. The computation delay for processing $A_m(t)$ data in the t-th slot [26] is given by

$$\tau^{C}_{m,n}(t) = \frac{\varepsilon_m A_m(t)}{f_{m,n}(t)}, \tag{6}$$

where ε_m is the data computation density of device m, i.e., the required CPU cycles to process one bit data.

The total allocated computation resource of server n should be lower than the maximum computation resource $f_{n,\max}$, i.e.,

$$\sum_{m=1}^{M} f_{m,n}(t) \le f_{n,\max}. \tag{7}$$

2.3. Problem Formulation. The total delay that device m selects server n for computation offloading is composed of the transmission delay and the computation delay, i.e.,

$$\tau_{m,n}(t) = \tau_{m,n}^{O}(t) + \tau_{m,n}^{C}(t). \tag{8}$$

The objective is to minimize the long-term total delay of all devices under the constrains of servers' computation capacity and transmission reliability, which can be expressed as

$$\begin{aligned}
\text{P1}: &\min_{\{x_{m,n}(g)\},\{f_{m,n}(t)\}} \sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{n=1}^{N} x_{m,n}(g)\tau_{m,n}(t)\\
\text{s.t.}\quad & C_1: x_{m,n}(g) = \{0,1\}, \quad \forall m \in \mathscr{M}, \forall n \in \mathscr{N}, \forall g \in \mathscr{G},\\
& C_2: \sum_{m=1}^{M} x_{m,n}(g) \le Q_n, \quad \forall n \in \mathscr{N}, \forall g \in \mathscr{G},\\
& C_3: \sum_{n=1}^{N} x_{m,n}(g) = 1, \quad \forall m \in \mathscr{M}, \forall g \in \mathscr{G},\\
& C_4: \sum_{m=1}^{M} f_{m,n}(t) \le f_{n,\max}, \quad \forall n \in \mathscr{N}, \forall t \in \mathscr{T},\\
& C_5: \gamma_{m,n}(t) \ge \gamma_{\min}, \quad \forall m \in \mathscr{M}, \forall t \in \mathscr{T},
\end{aligned} \tag{9}$$

where Q_n is the quota of server n, C_2 represents that each edge server can be selected by at most Q_n devices at the same time, C_3 represents that each device can select only one edge server for computation offloading in an epoch, C_4 represents the edge server computation resource constraint, and C_5 represents the SINR constraint.

We neglect energy consumption because major delay-sensitive EIoT devices can continuously draw energy from the power grid. Nevertheless, the proposed algorithm can also be extended to the scenario where energy consumption is minimized.

3. Empirical Matching-Based Computation Offloading Optimization Algorithm for 5G and Edge Computing-Integrated EIoT

In this section, we introduce the problem transformation. Then, we propose the empirical matching-based large timescale server selection algorithm and Lagrange dual decomposition-based small timescale computation resource allocation for 5G and edge computing-integrated EIoT.

3.1. Problem Transformation. The long-term stochastic joint optimization problem P1 includes two subproblems, i.e., large timescale server selection problem and small timescale computation resource allocation problem. The large timescale server selection problem is first modeled as a many-to-one matching problem, which is decoupled from small timescale computation resource allocation by establishing matching preference list based on empirical performance of delay and reliability. Then, the large timescale server selection problem is solved by pricing-based matching with quota algorithm. Based on the obtained suboptimal result of large timescale server selection strategy, the small timescale computation resource allocation problem is subsequently solved by Lagrange dual decomposition, the result of which is used to update large timescale empirical performance.

3.2. Empirical Matching-Based Large Timescale Server Selection. The large timescale server selection optimization problem is transformed into a many-to-one matching problem, which is defined as a triple $(\mathscr{M}, \mathscr{N}, \mathscr{L})$, where $\mathscr{M}$ and $\mathscr{N}$ represent the sets of matching devices and servers, respectively, and $\mathscr{L}$ represents the preference list of devices.

Theorem 1. *A matching ϕ is a many-to-one correspondence of set $\mathscr{M}$ to set $\mathscr{N}$ based on the preference list $\mathscr{L}$. $\phi(m) = n$ represents the matching of device m and server n, i.e., $x_{m,n}(g) = 1$. Specially, server n can match at most Q_n devices at the same time.*

To decouple the large timescale server selection from the small timescale resource allocation, the preference list is established based on empirical performance of delay and reliability. The empirical matching-based large timescale server selection algorithm is summarized in Algorithm 1, including preference list establishment and the pricing-based iterative matching [27].

3.2.1. Preference List Establishment. In the initialization, each EIoT device first traverses all servers once to obtain the value of $\tilde{\tau}_{m,n}(g)$. At the beginning of the g-th period, the server selection decision is made based on the empirical information up to the current period. Thus, the preference value of device m towards server n in the g-th period is defined as $\theta_{m,n}(g)$, which is expressed as

$$\theta_{m,n}(g) = -\frac{\sum_{i=0}^{g-1} x_{m,n}(i)\tilde{\tau}_{m,n}(i)}{\sum_{i=0}^{g-1} x_{m,n}(i)} - p_n(g), \tag{10}$$

where $p_n(g)$ represents the server price of matching device m with server n, the initial value of which is set as zero and $\tilde{\tau}_{m,n}(g)$ represents the average offloading delay of device m in the g-th epoch, which is expressed as

$$\tilde{\tau}_{m,n}(g) = \frac{1}{T_0} \sum_{t=(g-1)T_0}^{gT_0} \sum_{n=1}^{N} x_{m,n}(t)\tau_{m,n}(t). \quad (11)$$

3.2.2. The Implementation Procedure of Pricing-Based Iterative Matching

3.2.3. Empirical Matching-Based Large Timescale Server Selection Algorithm

(Step 1) When $g = 0$, initialize $x_{m,n}(g) = 0$, $q_n = Q_n$, $p_n(g) = 0$, $\forall m \in \mathscr{M}, \forall n \in \mathscr{N}, \forall g \in \mathscr{G}$ [1]. Define Θ as the set of unmatched devices, and set $\Theta = \mathscr{M}$ at the beginning. Define Γ_n as the set of devices proposed to server n, and set $\Gamma_n = \varnothing$ at the beginning. All devices traverse all servers once to complete the initialization of the preference list $g \le G$.

(Step 2) When $g \le G$, each $m \in \Theta$ calculates its preference value $\theta_{m,n}(g)$ towards n as (10) and sorts $\theta_{m,n}(g)$ in descending order to obtain the preference list $\mathscr{L}_m(g)$.

(Step 3) Each $m \in \mathscr{M}$ proposes its most preferred server in $\mathscr{L}_m(g)$. Each server, e.g., n, adds m into Γ_n if m proposes n. For $n \in \mathscr{N}$, if $|\Gamma_n| \le q_n$, n matches with the devices, e.g., m, from the proposed devices, i.e., $x_{m,n}(t) = 1$. Update $q_n = q_n - |\Gamma_n|$. Update $\Theta = \Theta \setminus m$. Otherwise, n updates the price $p_n(g)$ as (13). Unmatched devices execute steps $2 \sim 3$, until every device m has been matched with a server, i.e., $\phi(m) \ne \varnothing$, or there exists no available server for unmatched device m.

The implementation procedure of pricing-based matching consists of three phases, which are introduced as follows.

Step 1. When $g = 0$, initialize $x_{m,n}(g) = 0$, $q_n = Q$, $p_n(g) = 0$, $\forall m \in \mathscr{M}, \forall n \in \mathscr{N}, \forall g \in \mathscr{G}$, the set of unmatched devices as $\Theta = \mathscr{M}$, and the set of devices which proposed to server n as $\Gamma_n = \varnothing$.

Step 2. Each $m \in \Theta$ calculates the preference value $\theta_{m,n}(g)$ for unmatched device m towards n as (10) and sorts $\theta_{m,n}(g)$ in a descending order to obtain the preference list $\mathscr{L}_m(g)$. We introduce $\succ_m$ to compare the preferences towards different servers. For example, $n \succ_m n'$ means that device m prefers server n than server n', which is expressed as

$$n \succ_m n' \Leftrightarrow \theta_{m,n}(g) > \theta_{m,n'}(g). \quad (12)$$

When a server is selected by multiple devices exceeding to its quota, the server increases the price $p_n(g)$ with increment Δp_n to eliminate matching conflicts, which is expressed as

$$p_n(g) = p_n(g) + \Delta p_n. \quad (13)$$

Then, the preference list $\mathscr{L}_m(g)$ is updated according to new preference value as (10) and (12).

Step 3. Each $m \in \mathscr{M}$ proposes to its most preferred server in $\mathscr{L}_m(g)$. Then, each server, e.g., n, adds m into Γ_n if m proposes to n. n matches with the devices, e.g., m, from the proposed devices, i.e., $x_{m,n}(g) = 1$. Repeat until every device m has been matched with a server.

Therefore, the optimal matching result can be obtained as

$$x_{m,n}(g) = 1 \Leftrightarrow \phi(m) = n. \quad (14)$$

3.2.4. Complexity Analysis. The computation complexity of the proposed server selection algorithm mainly depends on preference value calculation, preference list establishment, and pricing-based iterative matching process, where the complexities of these three steps are, respectively $\mathscr{O}(MN)$, $\mathscr{O}(MN \log(MN))$, and $\mathscr{O}(MN)$. Hence, the complexity of the proposed algorithm is $\mathscr{O}(2MN + MN \log(MN))$.

3.3. Lagrange Dual Decomposition-Based Small Timescale Computation Resource Allocation. Based on the many-to-one matching-based large timescale server selection, the small timescale computation resource allocation is subsequently optimized. For edge server n, the computation resource allocation problem in the t-th slot can be expressed as

$$\begin{aligned} &\text{P2}: \min_{\{f_{m,n}(t)\}} \Psi_n(t) \\ &\text{s.t.} \quad C_3: \sum_{m=1}^{M} f_{m,n}(t) \le f_{n,\max}, \forall n \in \mathscr{N}, \forall t \in \mathscr{T}, \end{aligned} \quad (15)$$

where $\Psi_n(t)$ can be expressed as

$$\Psi_n(t) = \sum_{m=1}^{M} \tau_{m,n}^{C}(t) = \sum_{m=1}^{M} \frac{\varepsilon_m A_m(t)}{f_{m,n}(t)}. \quad (16)$$

Based on Jensen's inequality theorem [28], (16) satisfies $f(\mathbb{E}(\Psi_n(t)) \le \mathbb{E}(f(\Psi_n(t))))$ while $f(\cdot)$ is a convex function. Therefore, P2 is a convex function and can be solved by Lagrange dual decomposition [29]. Define $\zeta_n(t)$ as a Lagrangian subvector corresponding to constraint C_3. The Lagrange equation can be expressed as

$$L(f_{m,n}(t), \zeta_n) = \Psi_n(f_{m,n}(t)) + \zeta_n(t)\left[\sum_{m=1}^{M} f_{m,n}(t) - f_{n,\max}\right]. \quad (17)$$

Based on Lagrange dual decomposition, P2 can be transformed as

$$\max_{\zeta_n(t)>0} \min_{\{f_{m,n}(t)\}} L(f_{m,n}(t), \zeta_n(t)). \quad (18)$$

Based on Karush-Kuhn-Tucker conditions, it can obtain

$$f_{m,n}(t, z+1) = \max\left\{\arg_{f_{m,n}(t)} \nabla_{f_{m,n}}(t) = 0, 0\right\}, \tag{19}$$

where z is the iterative index of Lagrangian multiplier update. Lagrangian multiplier $\zeta_n(t)$ update can be expressed as

$$\zeta_n(t, z+1) = \max\left\{\zeta_n(t, z) + \left[\sum_{m=1}^{M} f_{m,n}(t) - f_{n,\max}\right]\rho_{\zeta_n(t)}(t, z), 0\right\}, \tag{20}$$

where $\rho_{\zeta_n(t)}(t, z)$ is the update step length.

4. Simulation Results

In this section, we introduce the simulation results and analysis. We consider a computation offloading scenario of EIoT with 20 devices and 5 servers. In the case of largescale fading, the channel gain is calculated according to $h_{m,n}(t) = 127 + 30\log(r_{m,n})$ [30], where $r_{m,n}$ is the distance between device m and server n, and distributed within [0.03,0.05] km. According to the actual environment of EIoT, the electromagnetic interference power varies from 28 dBm to 30 dBm. The specific simulation parameters are summarized in Table 2 [31–35]. Two existing algorithms are employed for comparison. The first one is the price matching-based computation offloading (PMCO) algorithm [27], where server-side computation resources are allocated equally. The second one is the greedy matching-based computation offloading (GMCO) algorithm [36], where large timescale server selection is solved based on greedy matching strategy and small timescale computation resource allocation is implemented once in the first time slot of each epoch. In addition, we extend the proposed algorithm to a large-scale computation offloading scenario with 100 devices to verify its effectiveness and robustness. The simulation software is MATLAB 2021. The real execution time of the proposed algorithm is about 49 ms per time slot, which achieves real-time or near-real-time. When adopting dedicated execution module such as DSP, the real execution time can be further reduced to millisecond level or even microsecond level.

4.1. Transmission Delay Performance. Figure 2 shows the box plots of transmission delay. Compared with PMCO and GMCO, the proposed algorithm can reduce the median transmission delay by 13.85% and 19.82% and reduce the transmission delay fluctuation by 65.79% and 68.29%, respectively. The proposed algorithm utilizes empirical matching-based large timescale server selection to reduce the transmission delay, thus performing the best. PMCO outperforms GMCO in terms of transmission delay because PMCO considers the server selection optimization based on price matching, thus effectively reducing the transmission delay.

4.2. Computation Delay Performance. Figure 3 show the box plots of computation delay. Compared with PMCO and GMCO, the median computation delay is decreased by 24.78% and 16.05%, and the computation delay fluctuation is decreased by 46.77% and 31.25%, respectively. The proposed algorithm utilizes Lagrange dual decomposition-based small timescale computation resource allocation to improve computation delay performance, thus performing the best. PMCO performs worse than GMCO in terms of computation delay. The reason is that GMCO considers the computation resource allocation, which leads to a lower computation delay.

Table 2: Simulation parameters.

Parameter	Value	Parameter	Value
M	20	N	5
Q_n	6	G	20
T	200	T_0	10
B_n	0.1 MHz	$\gamma_{\min}$	10 dB
A_m	[0.2,0.5] Mbits	P_{TX}	0.1 W
$f_{n,\max}$	[10, 30] GHz	N_0	−114 dBm
Δp_n	5	ε_m	1000 cycles

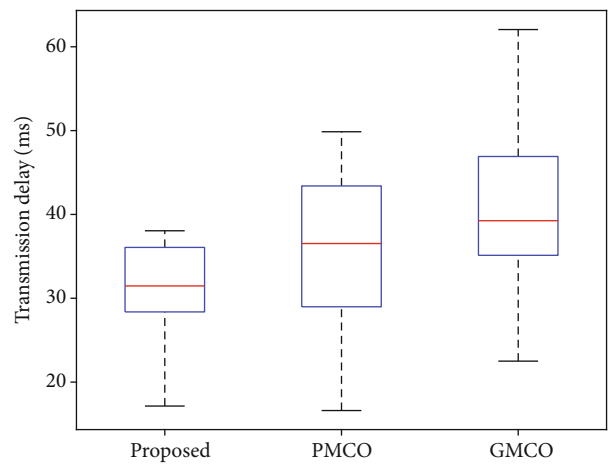


Figure 2: Transmission delay of different algorithms.

4.3. Total Computation Offloading Delay Performance. Figure 4 shows the total computation offloading delay versus time slot. The proposed algorithm outperforms PMCO and GMCO by 14.96% and 19.79%. The reason is that the proposed algorithm can avoid adversarial channels to improve delay performance through large timescale server selection. Then, given the large timescale server selection strategies, it further optimizes the small timescale computation resource allocation in each slot. However, PMCO neglects the optimization of computation resource allocation. The GMCO algorithm based on greedy matching and single-slot computation resource allocation cannot solve matching conflict from the perspective of the whole network and adjust the allocation strategies based on real-time information.

Figure 5 shows the total computation offloading delay versus number of devices. As the number of devices increases from 10 to 100, the total computation offloading

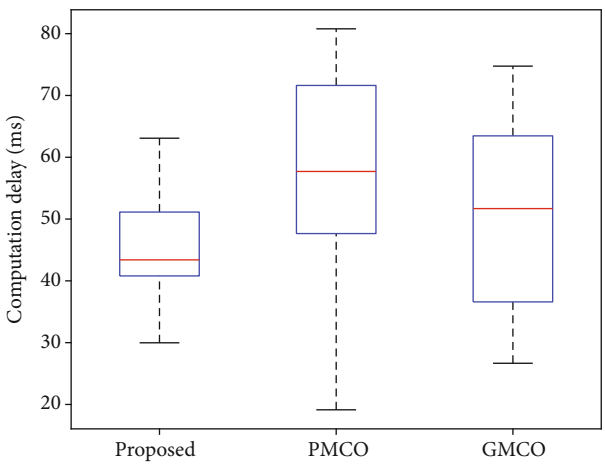


Figure 3: Computation delay of different algorithms.

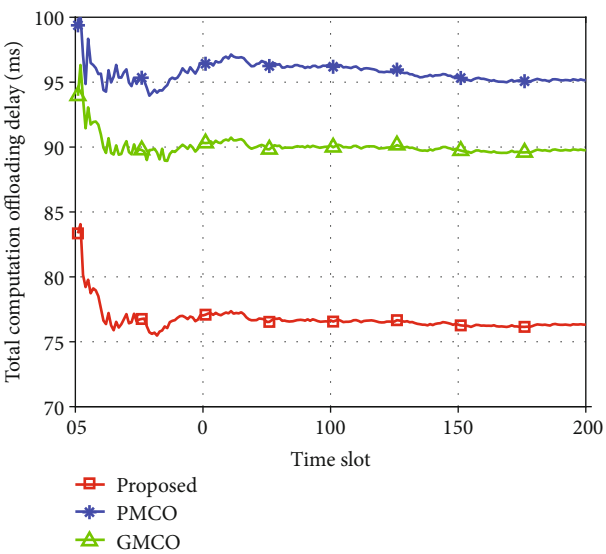


Figure 4: Total computation offloading delay versus time slot.

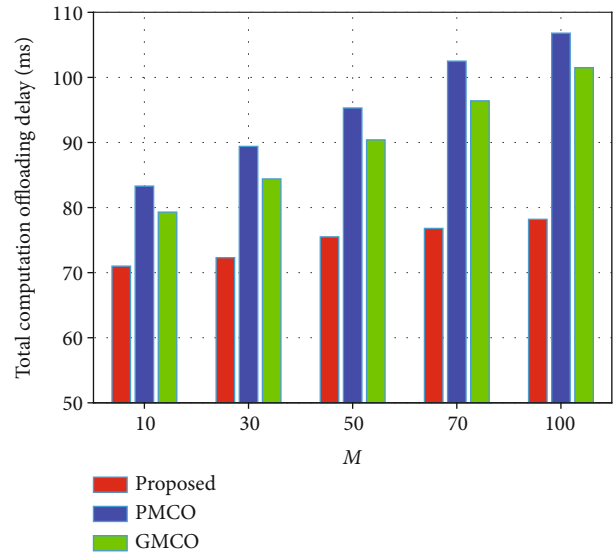


Figure 5: Total computation offloading delay versus number of devices M ($M = 10 \sim 100$).

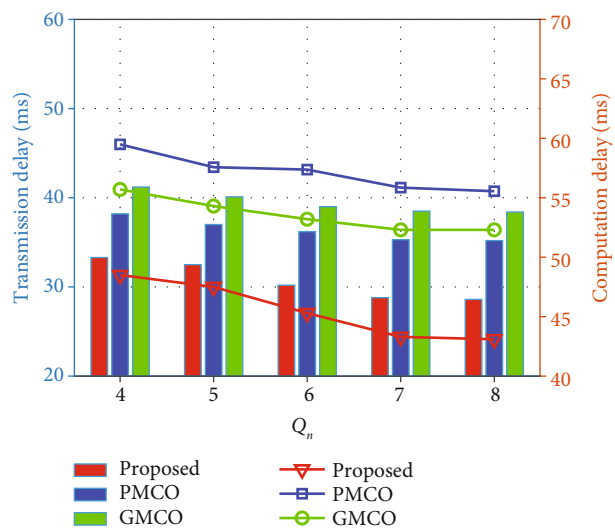


Figure 6: The impact of quota Q_n on transmission delay and computation delay.

delay of PMCO and GMCO increases by 28.21% and 28.99 %, while the total computation offloading delay of the proposed algorithm is slightly increased by 10.14%. The reason why total computation offloading delay increases is that the conflicts caused by multiple devices selecting the same server increases, and the computation resource allocated to each device is poorer. Even in a multi-access scenario with 100 devices, the proposed algorithm can adaptively balance the performance return of all devices through pricing-based iterative matching and adjust the resource allocation strategies based on dynamic data arrival.

Figure 6 shows the impact of the quota Q_n on transmission delay and computation delay. The bars in the figure represent the transmission delay and the curves represent the computation delay. When Q_n increases from 4 to 8, the transmission delay and computation delay of PMCO and GMCO are decreased by 7.85%, 6.62%, 6.79%, and 11.13% , in which those of the proposed algorithm are obviously decreased by 14.11% and 11.13%. The reason for it is that the final delay performance depends on the worst computation offloading situation, where the larger Q_n means the greater offloading possibility to servers with better channels and computation resources for devices. In addition, the proposed algorithm can well coordinate the competition among multiple devices for servers and computation resources based on empirical information and real-time resource allocation.

Table 3 shows the impact of $\gamma_{\min}$ on total computation offloading delay. With the increasing of $\gamma_{\min}$, the total computation offloading delay decreases first and then increases. When $\gamma_{\min} = 12\,\text{dB}$, the total computation offloading delay reaches the minimum value. The reason is that when $\gamma_{\min}$ increases from 8 to 12 dB, stringent reliability constraints impel devices to select servers with better empirical channel

Table 3: The impact of $\gamma_{\min}$ on total computation offloading delay.

$\gamma_{\min}$ (dB)	8	10	12	14	16
Proposed	75.24 ms	74.80 ms	73.25 ms	74.46 ms	75.82 ms
PMCO	95.80 ms	94.23 ms	93.15 ms	95.12 ms	95.60 ms
GMCO	92.73 ms	90.95 ms	89.82 ms	90.23 ms	92.86 ms

states, thus significantly reducing transmission delay at the beginning. However, as $\gamma_{\min}$ continues to increase from 12 to 16 dB, the servers with better channel states will be occupied by more devices resulting in a higher computation delay.

5. Conclusion

In this paper, we investigated a multi-timescale computation offloading optimization problem for 5G and edge computing-integrated EIoT. Specifically, we proposed an empirical matching-based computation offloading optimization algorithm, which includes the joint optimization of large timescale server selection and small timescale computation resource allocation. Compared with the existing PMCO and GMCO algorithms, the proposed algorithm could reduce the total computation offloading delay by 14.96% and 19.79%, respectively. It could also reduce the median transmission delay by 13.85% and 19.82%, the transmission delay fluctuation by 65.79% and 68.29%, the median computation delay by 24.78% and 16.05%, and the computation delay fluctuation by 46.77% and 31.25%. Finally, future research directions are outlined.

5.1. Computation Offloading under Extensive EIoT Scenarios with Incomplete Information. In large-scale EIoT scenarios, considering the prohibitive signaling overheads, global state information such as the channel state between devices and servers, the amount of available computation resources of servers, and the computation resource allocation policy of servers is unknown to devices. Server selection needs to be optimized under incomplete information. A feasible approach is to empower computation offloading for EIoT by learning through interaction with the environment by exploring advanced artificial intelligence algorithms such as deep reinforcement learning and federated learning.

5.2. Insecure Computation Offloading under Malicious Attack. With the extensive access of massive devices, EIoT faces various malicious attacks including distributed denial of services (DDoS), time synchronization attack, sinkhole attack, and tampering attack. These malicious attacks make computation offloading in EIoT insecure, which results in loss and tampering of important information, such as fault data, alarm data, and load fluctuation, endangering the supply-demand balance as well as stable and secure operation of power grid. To solve this problem, advanced security technologies such as blockchain, identity authentication, encryption, and trusted computing can be adopted to ensure the secure data transmission and processing in computing offloading.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Key R&D Program of China (2020YFB0905900) and is funded by the Science and Technology Project of SGCC (State Grid Corporation of China): The Key Technologies for Electric Internet of Things (SGTJDK00DWJS2100223).

References

[1] M. Ullah and J. Park, "Distributed energy trading in smart grid over directed communication network," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3669–3672, 2021.

[2] W. Costa, W. Santos, H. Rocha, M. Segatto, and J. Silva, "Power line communication based smartplug prototype for power consumption monitoring in smart homes," *IEEE Latin America Transactions*, vol. 19, no. 11, pp. 1849–1857, 2021.

[3] X. Li, Y. Tian, G. Ledwich, Y. Mishraand, and C. Zhou, "Minimizing multicast routing delay in multiple multicast trees with shared links for smart grid," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5427–5435, 2019.

[4] P. Thiruvasagam, A. Chakraborty, A. Mathew, and C. Murthy, "Reliable placement of service function chains and virtual monitoring functions with minimal cost in softwarized 5G networks," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1491–1507, 2021.

[5] P. Zhang, Y. Zhang, H. Dong, and H. Jin, "Mobility and dependence-aware QoS monitoring in mobile edge computing," *IEEE Transactions on Cloud Computing*, vol. 9, no. 3, pp. 1143–1157, 2021.

[6] D. Wu, X. Huang, X. Xie, X. Nie, L. Bao, and Z. Qin, "LEDGE: leveraging edge computing for resilient access management of mobile IoT," *IEEE Transactions on Mobile Computing*, vol. 20, no. 3, pp. 1110–1125, 2021.

[7] F. Fang and X. Wu, "A win–win mode: the complementary and coexistence of 5G networks and edge computing," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 3983–4003, 2021.

[8] Y. Chen, Z. Liu, Y. Zhang, Y. Wu, X. Chen, and L. Zhao, "Deep reinforcement learning-based dynamic resource management for mobile edge computing in industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 4925–4934, 2021.

[9] Z. Zhou, K. Ota, M. Dong, and C. Xu, "Energy-efficient matching for resource allocation in D2D enabled cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5256–5268, 2017.

[10] S. K. U. Zaman, A. I. Jehangiri, T. Maqsood et al., "Mobility-aware computational offloading in mobile edge networks: a survey," *Cluster Computing*, vol. 24, no. 4, pp. 2735–2756, 2021.

[11] S. Zhou, W. Jadoon, and J. Shuja, "Machine learning-based offloading strategy for lightweight user mobile edge computing

tasks," *Complexity*, vol. 2021, Article ID 6455617, 11 pages, 2021.

[12] F. Wang, J. Xu, and Z. Ding, "Multi-antenna NOMA for computation offloading in multiuser mobile edge computing systems," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2450–2463, 2019.

[13] P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Cognitive data offloading in mobile edge computing for internet of things," *IEEE Access*, vol. 8, no. 99, pp. 55736–55749, 2020.

[14] Y. Wang, P. Lang, D. Tian et al., "A game-based computation offloading method in vehicular multiaccess edge computing networks," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4987–4996, 2020.

[15] T. Q. Dinh, Q. D. La, T. Q. S. Quek, and H. Shin, "Learning for computation offloading in mobile edge computing," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6353–6367, 2018.

[16] Z. Gao, W. Hao, Z. Han, and S. Yang, "Q-learning-based task offloading and resources optimization for a collaborative computing system," *IEEE Access*, vol. 8, no. 99, pp. 149011–149024, 2020.

[17] Z. Zhou, G. Ma, M. Dong, K. Ota, C. Xu, and Y. Jia, "Iterative energy-efficient stable matching approach for context-aware resource allocation in D2D communications," *IEEE Access*, vol. 4, pp. 6181–6196, 2016.

[18] S. Seng, C. Luo, X. Li, H. Zhang, and H. Ji, "User matching on blockchain for computation offloading in ultra-dense wireless networks," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1167–1177, 2021.

[19] Z. Zhou, H. Liao, X. Zhao, B. Ai, and M. Guizani, "Reliable task offloading for vehicular fog computing under information asymmetry and information uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8322–8335, 2019.

[20] X. Liu, Z. Qin, Y. Gao, and J. McCann, "Resource allocation in wireless powered IoT networks," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4935–4945, 2019.

[21] K. Wang, F. Fang, D. Costa, and Z. Ding, "Sub-channel scheduling, task assignment, and power allocation for OMA-based and NOMA-based MEC systems," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2692–2708, 2021.

[22] Q. Zhang, L. Gui, F. Hou, J. Chen, S. Zhu, and F. Tian, "Dynamic task offloading and resource allocation for mobile-edge computing in dense cloud RAN," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3282–3299, 2020.

[23] T. Liu, Y. Zhang, Y. Zhu, W. Tong, and Y. Yang, "Online computation offloading and resource scheduling in mobile-edge computing," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6649–6664, 2021.

[24] H. Yu, Z. Zhou, Z. Jia, X. Zhao, L. Zhang, and X. Wang, "Multi-timescale multi-dimension resource allocation for NOMA-edge computing-based power IoT with massive connectivity," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1101–1113, 2021.

[25] K. Wang, Y. Zhou, Q. Wu, W. Chen, and Y. Yang, "Task offloading in hybrid intelligent reflecting surface and massive MIMO relay networks," *IEEE Transactions on Wireless Communications*, p. 1, 2021.

[26] Z. Zhou, H. Liao, B. Gu, S. Mumtaz, and J. Rodriguez, "Resource sharing and task offloading in IoT fog computing: A contract-learning approach," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 3, pp. 227–240, 2020.

[27] C. Xu, J. Feng, Z. Zhou, Z. Chang, Z. Han, and S. Mumtaz, "Two-stage matching for energy-efficient resource management in D2D cooperative relay communications," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, Singapore, 2017.

[28] C. Briat, "Convergence and equivalence results for the jensen's inequality— application to time-delay and sampled-data systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1660–1665, 2011.

[29] H. Liao, Z. Zhou, X. Zhao, and Y. Wang, "Learning-based queue-aware task offloading and resource allocation for space–air–ground-integrated power IoT," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5250–5263, 2021.

[30] H. Liao, Z. Zhou, W. Kong et al., "Learning-based intent-aware task offloading for air-ground integrated vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5127–5139, 2021.

[31] M. Tariq, M. Adnan, G. Srivastava, and H. Poor, "Instability detection and prevention in smart grids under asymmetric faults," *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 4510–4520, 2020.

[32] M. Tariq and H. Poor, "Electricity theft detection and localization in grid-tied microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1920–1929, 2018.

[33] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[34] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[35] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[36] S. Zhang, G. Cui, Y. Long, and W. Wang, "Joint computing and communication resource allocation for satellite communication networks with edge computing," *China Communications*, vol. 18, no. 7, pp. 236–252, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9419764, 10 pages
https://doi.org/10.1155/2022/9419764

Research Article

Random Graph-Based M-QAM Classification for MIMO Systems

Mubashar Sarfraz,[1] Sheraz Alam,[1] Sajjad A. Ghauri,[2] Asad Mahmood,[3] M. Nadeem Akram,[2] M. Javvad Ur Rehman,[1] M. Farhan Sohail,[1] and Teweldebrhan Mezgebo Kebedew[4]

[1]*Faculty of Engineering and Computer Science, National University of Modern Languages, Islamabad, Pakistan*
[2]*School of Engineering & Applied Sciences, ISRA University, Islamabad, Pakistan*
[3]*Department of Electrical and Computer Engineering, Comsats University, Islamabad, Wah Campus, Wah Cantt, Pakistan*
[4]*Ethio Telecom, Addis Ababa, Ethiopia*

Correspondence should be addressed to Teweldebrhan Mezgebo Kebedew; tewelde2000@gmail.com

Received 16 February 2022; Revised 6 March 2022; Accepted 21 March 2022; Published 15 April 2022

Academic Editor: Xingwang Li

Automatic modulation classification (AMC) has been identified to perform a key role to realize technologies such as cognitive radio, dynamic spectrum management, and interference identification that are arguably pivotal to practical SG communication networks. Random graphs (RGs) have been used to better understand graph behavior and to tackle combinatorial challenges in general. In this research article, a novel modulation classifier is presented to recognize M-Quadrature Amplitude Modulation (QAM) signals using random graph theory. The proposed method demonstrates improved recognition rates for multiple-input multiple-output (MIMO) and single-input single-output (SISO) systems. The proposed method has the advantage of not requiring channel/signal to noise ratio estimate or timing/frequency offset correction. Undirected RGs are constructed based on features, which are extracted by taking sparse Fourier transform (SFT) of the received signal. This method is based on the graph representation of the SFT of the 2nd, 4th, and 8th power of the received signal. The simulation results are also compared to existing state-of-the-art methodologies, revealing that the suggested methodology is superior.

1. Introduction

Automatic modulation classification (AMC) is utilised by the observation of received data sample, received data samples which are noise-related, and different fading effects to automatically determine the modulation type of the broadcast signal. There are numerous military and civilian applications of AMC such as electronic warfare, signal surveillance and spectrum sensing, and cognitive radio. The modulation classification procedure takes place after preprocessing prior to demodulation of the incoming signal [1–7]. The likelihood-based techniques are based on the signal received probability, while the feature-based methods contain two modules: function extraction and structure of the classifier [8–15].

Random graphs (RGs) have been used to model a multitude of random-like networks, including the unpredictable growth of the Internet's web graph, page ranking, and neural networks. The concept of a random graph was developed by Erdos and Rényi in 1959 as an extension of the probabilistic approach for discovering the existence of certain graph features [16]. They introduce two closely related models $G(n, m)$ and $G(n, P)$, where $G(n, m)$ is the ensemble of graphs having exactly m edges and n vertices with equal probability and $G(n, P)$ is obtained by introducing each of the available chart edges individually with probability of P. The diagram is considered to not be directed, such that no direction is linked to the edges if these pairings are unordered. Moreover, if some randomness is involved in its formation, a graph is said to be random.

QAM signals are used in microwave digital radio for transmission with enhanced spectral efficiency over limited channel bandwidths. It is used widely in many modern domestic applications such as digital cable TV, cable modems, power-line Ethernet, and microwave backhaul systems. In particular, it is used in almost all the 802.11 standards including common a, b, g, and n to latest ac, ah, and ad. On the other hand, one technology has revolutionized wireless communication technologies as a key enabler to provide high data rates in MIMO antenna systems. The combination of M-QAM and MIMO is used in IEEE 802.11 ac and ad standards. This is the reason for choosing M-QAM signals in this research work in combination with different antenna configurations such as SISO, SIMO, MISO, and MIMO for detailed analysis.

1.1. Contribution of the Research. In this paper $G(n, P)$ model is considered for the construction of undirected RGs [17]. The main contribution of our research work is to classify the M-QAM signals using random graphs which to the best of our knowledge is not yet to be done in the literature. Random graph theory is seldom explored for modulation classification, in particular for M-QAM signals that are widely used with MIMO technology. The features are extracted by using sparse Fourier transform (SFT) and RGs for the classification. The most striking feature of the proposed solution is that it neither required channel/SNR estimation and the performance is much better for SISO and MIMO on different channel conditions and in the presence of additive white Gaussian noise.

1.2. Organization of the Article. The rest of the paper has the following structure: Section 2 provides a short overview of related work for AMC. The system model and the major phases of the algorithm suggested are explained in Section 3 and Section 4, respectively. In Section 5, simulation parameters and results are shown. The article is then completed by comparing the suggested method with current state-of-the-art procedures in various future areas in Section 6.

2. Related Work

The AMC literature is largely divided into two techniques: likelihood-based classification (LBC) and feature-based classification (FBC). The existing work on the likelihood-based AMC can be found in [18–23] and feature-based AMC in [24–30].

Convolutional neural networks (CNN) are based on convolution instead of matrix multiplication [31]. In grid constellation matrix- (GCM-) based AMC, a contrastive fully convolutional network (CFCN) is used; this network benefits from a deep learning method and high-dimensional representations from variables via GCMs and a quick training procedure [32].

The constant envelope modulation (CEM) classification is performed using a radial basis function network (RBFN) with a deep learning network and a sparse autoencoder-based neural network (SAEDNN) without any prior knowledge of the CEM or channel [33]. In curriculum learning, two neural networks are used: MentorNet for supervision and StudentNet is trained for the classification [34].

Authors in [35] uses K-nearest neighbour (KNN) with several distance measuring methods, including correlation, Mahalanobis, Euclidean, and Minkowski distance in combination with HOS to identify signals. In [36], a unique feature engineering to improve the efficiency of KNN for NRZ and PAM-4-modulated signals is introduced. The authors of [37] use GP to boost the effectiveness of the KNN classifier by using super features. The Gabor filter network (GFN) was used to extract features and classify digitally modulated signals. The delta rule is used to adaptively modify GFN parameters, which are then tuned with GA using the HMM classifier [38, 39].

Current approaches often adopt a single neural network category, or stack different network categories in series, and seldom properly extract diverse types of information. Softmax is employed in classification to widen the distance between classes when it gets to the output layer. A parallel hybrid network for the AMC problem is suggested in [40].

In [41], the authors present a deeper learning feature fusion system that tries to fuse features from multiple domains of the input signal in order to achieve a more reliable and efficient display of signal modulation kinds. The complementarity of features which may be utilised to remove the impact of the background sounds and the wide dynamic range of (intercepted) signalling is taken into consideration.

For improved feature variety, a hierarchical multifunction fusion approach is presented in [42]. Modulation features are represented by the state signals, and its process of statistical ordering characterises time on a regular basis. Initially, the authors recommended that you use various dimensional convolution filters based on two streams to fully leverage IQ signal characteristics. Parallel to each stream multidimensional convolution filters are used to better characterise individual channels and to improve the interactions between them. A small LSTM network is subsequently used to further exploit sequential data.

In [43], categorization of radar transmitter signal intrapulse modulation refers to the classification of the pulse of each radar emitter to a certain model. An intrapulse modular intrapulse 1-D selective kernel convolutional neural network called 1-D SKCNN is proposed. A basic purpose is to construct a parallel hybrid structure that uses the CNN and Gate Rate Unit (GRU) to draw spatial characteristics and time characteristics separately.

AMC with a feature-based two-lane capsule network based on these concerns is suggested in [44]. As channel characteristics and signal overlaps in MIMO systems lack prior knowledge, standard probability-based and feature-based methods cannot be used directly in these circumstances. Therefore, the time-frequency analysis technique based on the Fourier transform window was used to evaluate the time-frequency characteristics of the time-domain-modulated signals in order to tackle the problem of blind modulation classification in MIMO systems [45].

Chang and Shih [46] proposes a radio transformer that takes instantaneous amplitude, instant phase, and

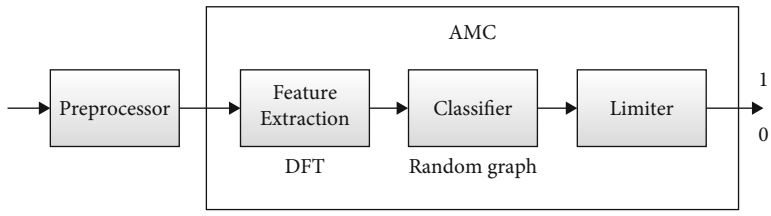


Figure 1: Proposed system model.

instantaneous frequency from received RCB signals and converts them into images by the proposed signal rearrangement to complete the classification of radio modulation using existing ImageNet classification models.

The aims of [47] are to resolve the issue of low identifying precision under single-feature parameters and to decrease the sampling system performance requisites by using a blind modulation classification method based on the compressed sensing with a high-order, cumulative, and cyclic spectrum associated with the decision-tab support vector machine classifier.

3. System Model

The system model and the proposed algorithm for the classification of M-QAM signals are discussed in detail in this section. Figure 1 shows the proposed system model in which the received signal is first preprocessed and discrete Fourier transform (DFT) is applied. After extracting parameters and then selecting distinct features, random graphs are generated. The threshold decides the modulation scheme. The received signal in generalized form is given by

$$r = Hs + \eta, \tag{1}$$

where η is additive white Gaussian noise with zero mean and variance of σ^2, H is a matrix of channel coefficients with order $N_r \times N_t$. s is $N_t \times K$ vector of transmitted symbols, and r is $N_r \times K$ the received signal vector, where N_t and N_r are the number of transmitter and receiver antennas.

4. Proposed Algorithm

The proposed algorithm for M-QAM signal classification based on sparse DFT (SDFT) of the received signal with power p has the following steps.

Step 1. Calculate the $|\mathrm{SDFT}(r^p)|$, where $p = [2, 4, 8]$.

Step 2. To generate the random graphs, the detailed process is as follows:

(1) Normalization of sparse samples is given as

$$\chi_{\mathrm{norm}}(z) = \frac{\chi(\mathrm{z}) - \min\{\chi(\mathrm{z})\}}{\max\{\chi(\mathrm{z}) - \min\{\chi(\mathrm{z})\}\}}, \tag{2}$$

where χ_{norm} is the normalized discrete sample of the received signal and $0 \le \chi_{\mathrm{norm}}(z) \le 1$

(2) Rayleigh cumulative distribution function (RCDF) is employed to obtain the uniform distribution

(3) Uniformly quantized the distribution for predefined quantization levels Q_a

(4) Count the numbers of peaks and set the numbers of nodes n equal to Q_a

(5) Calculate the probability by using equation

$$P = 1 - \left[1 - \frac{2}{Q_a}\right]^{(\kappa-1)}, \tag{3}$$

where κ is the number of observed samples

(6) To construct the adjacency matrix, Bernoulli trials were performed with a biased coin

Step 3. Calculate the maximum numbers of edges using the relationship:

$$N_e = \frac{Qa(Qa-1)}{2}. \tag{4}$$

Calculate the estimated number of edges:

$$\tilde{N}_e = \frac{\text{numbers of one sin adjacency matrix}}{2}. \tag{5}$$

Hence, if $N_e = \tilde{N}_e$, then the graphs are fully connected; otherwise, graphs are disconnected. However, that is a strict decision for graph connectivity. The threshold level for deciding the connectivity of the graph is defined as follows:

$$\lambda = \mathrm{floor}(0.9 \times N_e). \tag{6}$$

Now to check the connectivity of a random graph, compare $\tilde{N}_e$ with λ instead of N_e. The decision function for the connectivity of RGs is defined as follows:

$$G(M, p) = \begin{cases} 1, & \tilde{N}_e \ge \lambda, \\ 0, & ow, \end{cases} \tag{7}$$

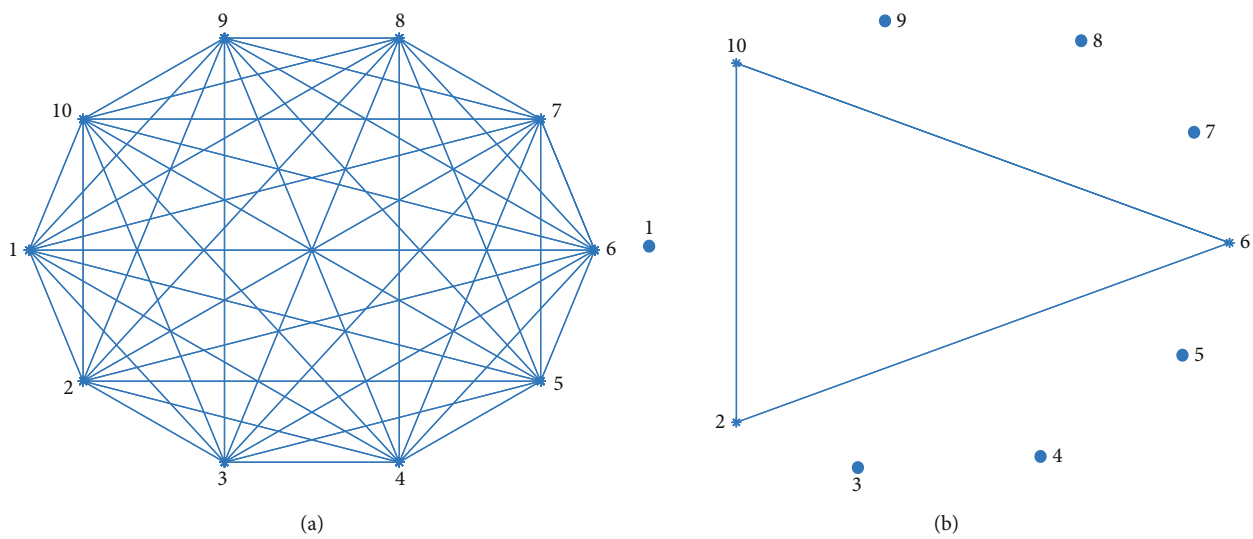


Figure 2: Fully connected and disconnected graph $Q_a = 10$, $N_T = 1$, $N_R = 1$, and Rayleigh channel. (a) 32-QAM, 64-QAM, and 128-QAM and (b) 8-QAM.

```
initialization;
while M-QAM ≠ classified do
    evaluate ⟶ SFT (r) (where r is received signal);
    apply ⟶ Normalization (transformed r);
    evaluate ⟶ Rayleigh CDF (normalized signal);
    for j=1:K do
        for k=1:K-1 do
            connect each ith edge to jth edge;
        end
    end
if graph ==fully connected then
    class 1 is separated;
else
    take next power and continue classification;
end
end
```

Algorithm 1: Algorithm of the proposed problem.

Table 1: Simulation parameters.

Parameter	Values
Numbers of iterations	50
Channel	Pedestrian
Numbers of samples (K)	8000
Quantization levels	10
Channel length	4
Path delays	[0, 110, 190, 410] ns
Average path gains	[0, -0.97, -12.9, -22.8] dB
Maximum Doppler shift	8 Hz
Sample rate	12.5 MHz
Rayleigh CDF parameter	0.25

where $M = 8,32,64,128$ and $p \in 2, 4, 8$. Figure 2(a) depicts the fully connected graph for 32-QAM, 64-QAM, and 128-QAM, and Figure 2(b) shows the disconnected graph for 8-QAM when SDFT is taken with a power $p = 2$ at SNR of 10dB. Similarly, RGs for $p = 4$ are also generated which shows fully connected for 32-QAM while disconnected graph for 64-QAM and 128-QAM. For, $p = 8$, the graph is fully connected for 128-QAM and disconnected for 64-QAM. The pseudocode of the proposed algorithm is shown in Algorithm 1.

5. Simulation and Results

The simulation parameters for the M-QAM signal classification are presented in Table 1. Various types of systems are considered for simulation such as single-input single-output (SISO), single-input multiple-output (SIMO), multiple-input single-output (MISO), and multiple-input multiple-output (MIMO). To validate the results, MATLAB built-in functions and communication system toolbox are used. M-QAM signal classification is also carried out under different fading channel conditions, i.e., Rayleigh, Rician, and Pedestrian channels. Spatial multiplexing is employed for multiple antennas at the transmitter and receiver sides. The figure of merit is the average percentage of correct classification (APCC).

Figure 3 represents the decision tree for the M-QAM classification, where p represents the order of the DFT. The $G = 0$ and 1 represents that the random graphs are disconnected and fully connected, respectively. The higher SNR is considered in Figure 3 for the optimum decision-making. At lower SNRs, the decision-making process is suboptimum, and there are more chances of misclassification. M-QAM signal classification using RGs is achieved into three steps at SNR of 10 dB with 10 quantization levels:

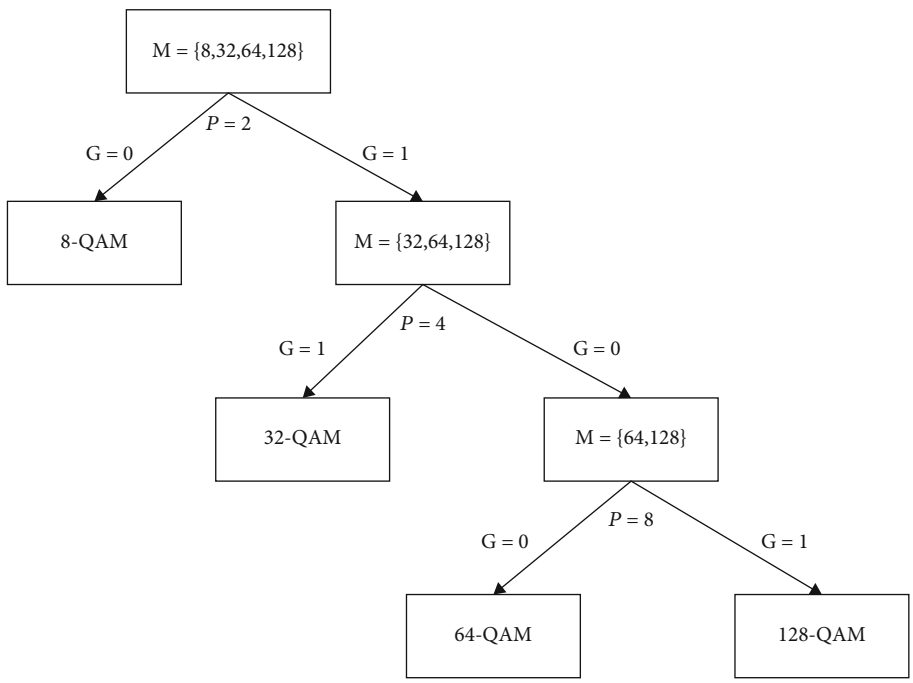


Figure 3: Decision tree for M-QAM signal classification.

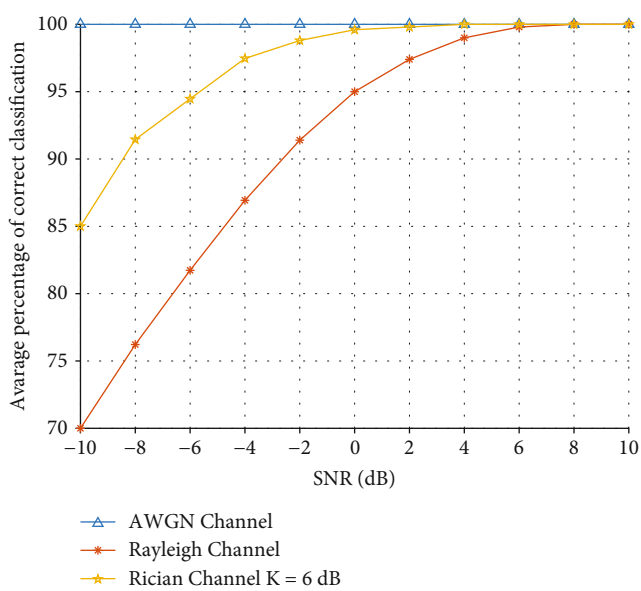


Figure 4: APCC for AWGN, Rayleigh, and Rician channels (64-QAM).

(1) In the first step, if $p = 2$, then there are two possibilities: if a graph is fully connected, then the received signal may be modulated using by using 32-QAM, 64-QAM, and 128-QAM. If a graph is disconnected, then the received signal is modulated using 8-QAM

(2) In 2nd step $p = 4$. If a graph is fully connected, then the received signal is identified to be 32-QAM and the MC process stops here; otherwise, the signals may be modulated using 64-QAM or 128-QAM

(3) In last step, $p = 8$. If a graph is fully connected, then the received signal is modulated using the 128-QAM scheme otherwise 64-QAM

The APCC for M-QAM signals using RGs on the Rayleigh and Rician fading channels with AWGN is simulated in Figure 4. The parameters for the Rayleigh and Rician channels such as path delays and path gains are [0,110] nanosecond and [0,−9.7] dB. K-factor for the Rician channel is selected [6, 6] dB with a maximum Doppler shift of 8 Hz. As shown from Figure 4, APCC for the AWGN channel is 100% because random graph is fully connected for 64-QAM constellations, while for the case of Rayleigh channel model, the APCC is 100% at 8 dB of SNR and for Rician channel the APCC is 100% at 4 dB of SNR. The fully connected random graphs show 100% accuracy at lower SNRs, while disconnected random graphs show lower APCC for lower SNRs.

The proposed algorithm is also validated on various types of systems such as SISO, SIMO, MISO, and MIMO. Therefore simulation scenarios are divided into four cases, and in all cases, the ITU-R Pedestrian fading channel is taken for the classification of M-QAM signals using RGs. For the simplicity of the results, the 8-QAM candidate constellation is selected to check the validity of our proposed cases.

Case 1. In the first case, the number of transmitter and receiver antenna is 1, i.e., SISO system. The APCC for Case 1 is approaching 100% at 6 dB of SNR. Figure 5 shows the APCC is more than 95% at an SNR of 0dB.

Case 2. In the second case, there is only one transmitter antenna and multiple receive antennas, i.e., 2, 3, and 4.

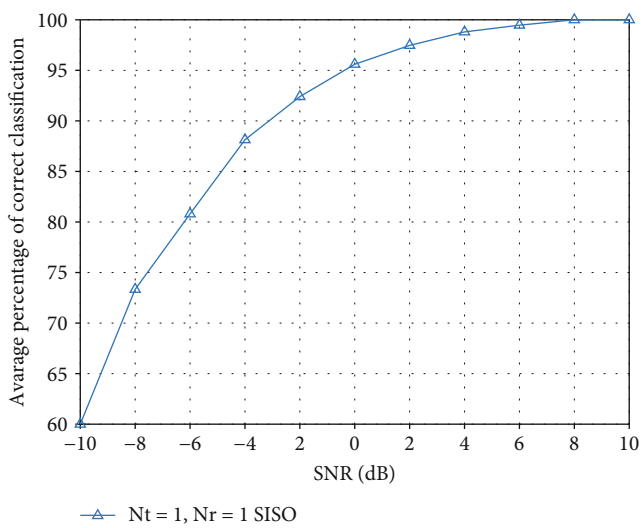


Figure 5: APCC for Case 1 (SISO).

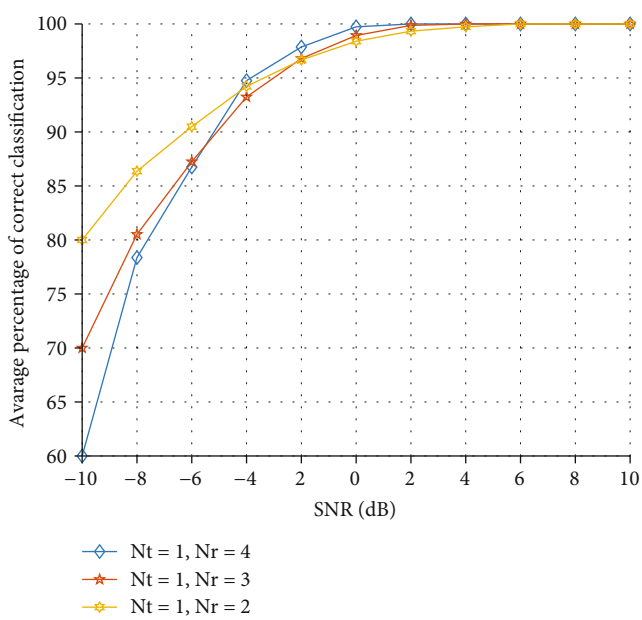


Figure 6: APCC for Case 2 (SIMO).

The APCC is approaching 100% at 6 dB of SNR. At lower SNRs and for $N_r = 2$; APCC is better as compared to $N_r = 3$ and $N_r = 4$. The APCC for Case 2 (SIMO) is given in Figure 6. At higher SNRs, larger number of receive antennas gives better APCC as compared to small number of receive antennas. The performance of proposed algorithm is worst on lower SNRs, i.e., below -4 dB. As the signal is dominant over noise, the APCC for Nr = 4, which is better as compared to the $N_r = 2$ and 3. For example, at SNR = −2 dB, the APCC is 97.5% at $N_r = 4$, while at $N_r = 2$, the APCC is approximately 96%.

Case 3. Multiple transmitter antennas such as 2, 3, and 4 are considered with a single receiver antenna. The APCC is shown in Figure 7, and it is clear from Figure 7 that the accuracy approaches 100% at lower SNRs. The APCC is an

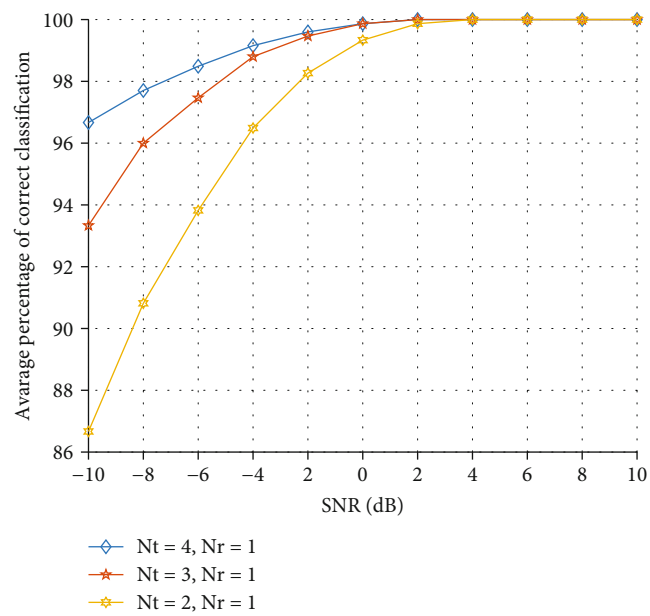


Figure 7: APCC for Case 3 (MISO).

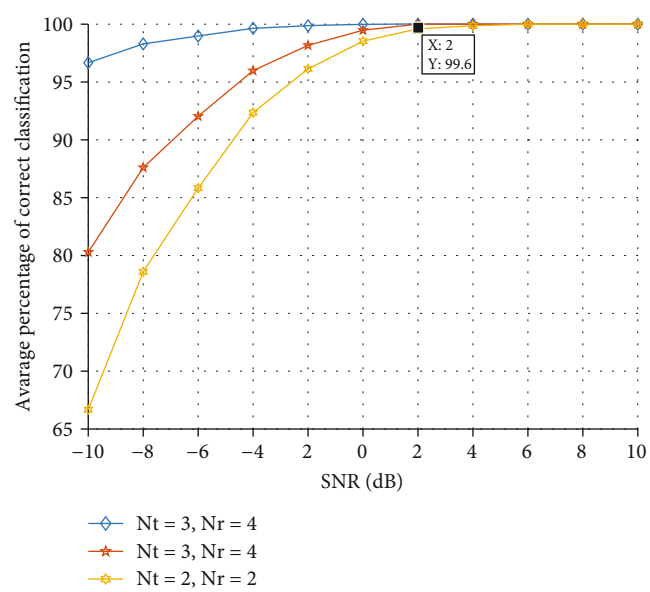


Figure 8: APCC for Case 4 (MIMO).

Table 2: Effect of phase offset on APCC.

SNR (dB)	Phase offset			
	$\pi/12$	$\pi/6$	$\pi/4$	$\pi/3$
-10	66.70%	60%	70%	76.67%
-5	81.78%	77.78%	81%	85.11%
0	89.07%	86.44%	88%	90.63%
5	97.33%	96.3%	96.31%	96.59%
10	100%	100%	100%	100%

Table 3: Comparison with state of art existing techniques.

	SNR (dB)											Complexity analysis
	-10	-8	-6	-4	-2	0	2	4	6	8	10	
[24]	28%	33%	46%	57%	68%	76%	81%	90.5%	96%	96.5%	98.5%	$O(KD^2)$
[31]	80.2%	84.8%	88.6%	95.7%	99%	100%	100%	100%	100%	100%	100%	$O(KND^2)O(N^2)$
[48]	47%	61%	72%	83%	86%	87%	90%	91%	92.5%	93%	94%	$O(KND^2)O(N^2)$
[49]	81%	81.5%	81.5%	83%	86%	91%	93%	95%	96.5%	97%	97.5%	$O(ND^2)$
	Proposed classification scheme											
SISO	60%	70.1%	77.45%	84.34%	89.42%	93.23%	95.9%	97.69%	98.8%	100%	100%	$O(N^2)$
SIMO	80%	86.37%	90.49%	94.22%	96.67%	98.4%	99%	99.75%	99.96%	100%	100%	$O(N^2)$
MISO	93.3%	96%	97.47%	99.6%	99.87%	100%	100%	100%	100%	100%	100%	$O(N^2)$
MIMO	96.67%	98.3%	98.8%	99.64%	99.9%	100%	100%	100%	100%	100%	100%	$O(N^2)$

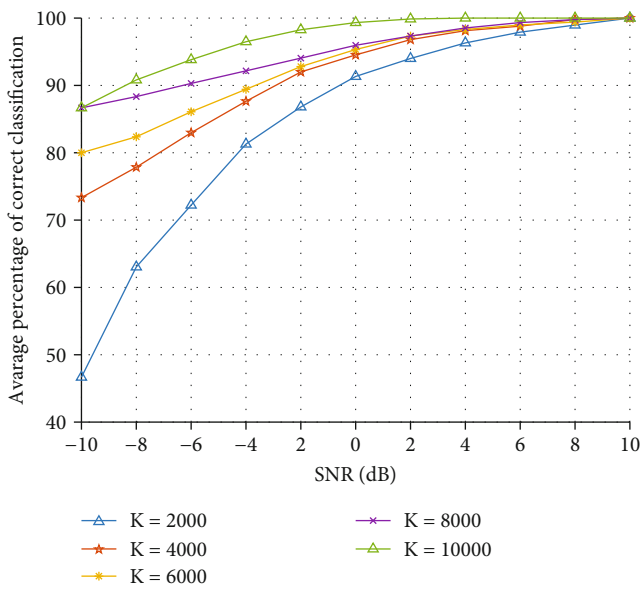


Figure 9: Effect of K on the classification performance 64-QAM signal.

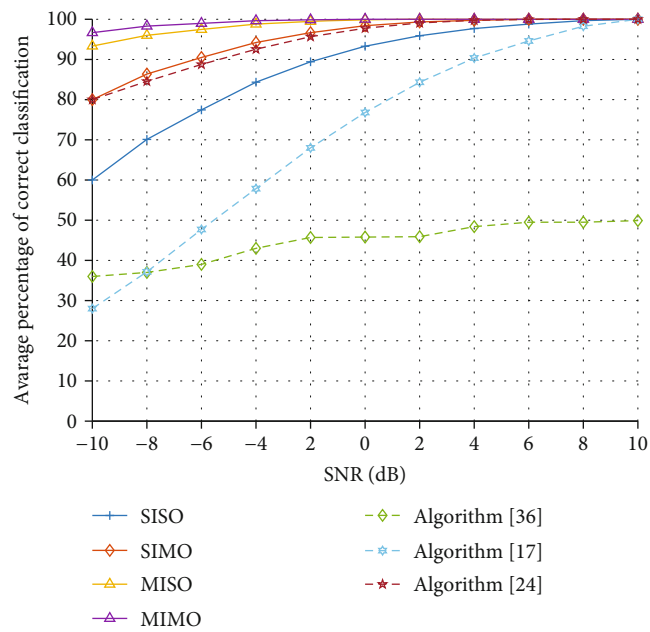


Figure 10: Performance compassion of proposed M-QAM classifier with state of art existing techniques.

approximate approach to 99% at 0 dB of SNR for $N_t = 2$, 3, and 4.

Case 4. In Case 4, multiple transmit and receive antennas are considered, and APCC is shown in Figure 8. At 2 dB of SNR, the APCC is 99.6% for the 2×2 system and 100% for 3×3 and 3×4. From Figure 8, it is clear that the overall APCC of the MIMO system is much better at lower SNRs as compared to SISO, SIMO, and MISO.

To evaluate the effect of phase offset on classifier performance, the APCC is shown in Table 2. Degradation and variation in the classifier performance can be observed at various values of SNR from -10 to 10 dB. From Table 3, it does not mean that bigger offset leads to better classification accuracy. For example, at 5 dB of SNR and $\pi/3$ offset, the APCC is 96.59% and at $\pi/12$ the APCC is 97.33%. At lower SNRs (i.e., -5 dB), the noise is dominant and the behavior of the classifier is random, e.g., at $\pi/4$, the APCC is 81% while at $\pi/12$ the APCC is 81.78%.

APCC is also affected via increasing the number of samples as shown in Figure 9. A various number of samples are taken, i.e., $K = 2000$ to $K = 10000$. Increasing the numbers of samples, the APCC increased and vice versa. In Figure 9, the SISO system with the Pedestrian channel and 10 quantization levels is employed.

The performance of proposed algorithm is also compared with the state of art existing techniques as shown in Figure 10. From Figure 10, it is clear that the proposed algorithm performs much better as compared with the existing techniques for any number of transmitter and receiver antennas.

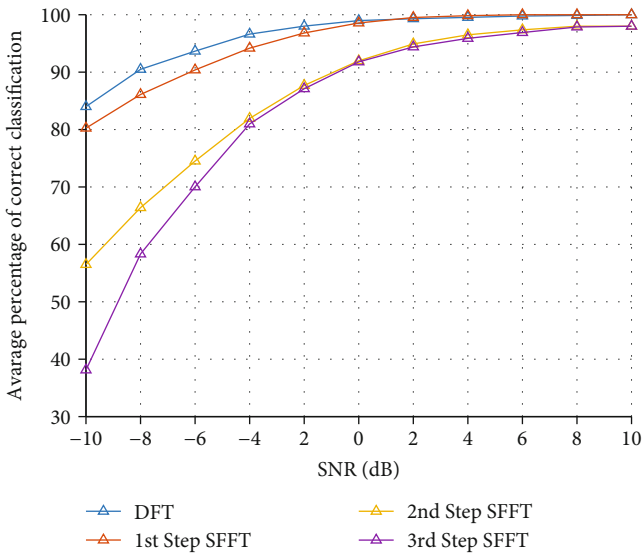


Figure 11: APCC for DFT and multilevel SFFT (M-QAM).

Table 4: Execution time of DFT and SFFT.

Features extraction technique	Time (sec)
DFT	11.18 s
1st step SFFT	10.64 s
2nd step SFFT	10.16 s
3rd step SFFT	9.98 s

The number of samples is 8000 with ITR = 50 and APCC is shown in Figure 11, which shows the approximately same results as for DFT and SFFT. The time complexity of DFT is a little bit more as compared to SFFT of the 3rd level.

The comparison of APCC and their complexity analysis with state of art existing techniques is shown in Table 4. The proposed algorithm performs much better as compared to the existing techniques. The proposed classifier achieves approximately 100% classification accuracy at lower SNRs. Furthermore, authors in [28] considered an AWGN channel model as compare to the proposed simulations the APCC is 100% for lower SNRs as shown in Figure 4. In Table 4, ITU Pedestrian channel model is considered for all the cases.

In this research paper, SFFT is employed to reduce the proposed algorithm time complexity. Table 4 shows the multilevel SFFT with the time in seconds.

Table 4 shows the time complexity comparison of using DFT and sparse FFT in seconds. The feature extraction from the DFT and 3rd step SFFT time complexity is 11.18 seconds and 9.98 seconds, respectively. The time complexity reduces at 3rd step SFFT which will cause the lower classification accuracy. The time saving affects the classification accuracy as shown in Figure 11.

6. Conclusions

The purpose of this research is to improve the classification accuracy of M-QAM signals under low SNR or unsatisfactory channel conditions. In this research, the undirected random graphs are constructed with DFT and sparse transform-based features to classify the M-QAM signals with various system models such as SISO, SIMO, MISO, and MIMO. Sparse Fourier transform is taken for 2nd, 4th, and 8th power of the received signal with the different fading channel scenarios to confirm the validity of the results. The novel classifier proposed in this article has an advantage that it requires neither channel estimates nor the timing/frequency offsets. Additionally, the proposed classifier is also compared with the state-of-the-art existing techniques, and from the simulation results, it is evident that all the considered M-QAM signals are classified accurately even at lower SNRs with less time complexity.

The further classification accuracy may be enhanced using different statistical features and for the application area of massive MIMO.

Data Availability

This paper does not require any dataset whereas the required data is generated uniformly using MATLAB tool.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] O. A. Dobre and F. Hameed, "Likelihood-based algorithms for linear digital modulation classification in fading channels," in *2006 Canadian conference on electrical and computer engineering*, pp. 1347–1350, Ottawa, ON, Canada, May 2006.

[2] A. Ramezani-Kebrya, I.-M. Kim, D. I. Kim, F. Chan, and R. Inkol, "Likelihood-based modulation classification for multiple-antenna receiver," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3816–3829, 2013.

[3] S. A. Ghauri, I. M. Qureshi, T. A. Cheema, and A. N. Malik, "A novel modulation classification approach using Gabor filter network," *The Scientific World Journal*, vol. 2014, 14 pages, 2014.

[4] W. U. Khan, F. Jameel, X. Li, M. Bilal, and T. A. Tsiftsis, "Joint spectrum and energy optimization of NOMA-enabled small-cell networks with QoS guarantee," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8337–8342, 2021.

[5] E. N. Gilbert, "Random graphs," *The Annals of Mathematical Statistics*, vol. 30, no. 4, pp. 1141–1144, 1959.

[6] P. Erdos and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, no. 1, pp. 17–60, 1960.

[7] K. Kim and A. Polydoros, "Digital modulation classification: the BPSK versus QPSK case," in *MILCOM 88, 21st Century Military Communications-What's Possible?*, pp. 431–436, San Diego, CA, USA, Oct 1988.

[8] F. Hameed, O. A. Dobre, and D. C. Popescu, "On the likelihood-based approach to modulation classification," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5884–5892, 2009.

[9] M. R. Bahloul, M. Z. Yusoff, and M. N. M. Saad, "Efficient and low complexity modulation classification algorithm for MIMO systems," *Applied Mathematics & Information Sciences*, vol. 9, no. 1, pp. 58–64, 2015.

[10] M. Gandetto, M. Guainazzo, and C. S. Regazzoni, "Use of timefrequency analysis and neural networks for mode identification in a wireless software-defined radio approach," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, 13 pages, 2004.

[11] S. Zhou, Z. Yin, Z. Wu, Y. Chen, N. Zhao, and Z. Yang, "A robust modulation classification method using convolutional neural networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2019, 15 pages, 2019.

[12] S. Huang, Y. Jiang, Y. Gao, Z. Feng, and P. Zhang, "Automatic modulation classification using contrastive fully convolutional network," *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1044–1047, 2019.

[13] M. H. Shah and X. Dang, "Classification of spectrally efficient constant envelope modulations based on radial basis function network and deep learning," *IEEE Communications Letters*, vol. 23, no. 9, pp. 1529–1533, 2019.

[14] M. Zhang, Z. Yu, H. Wang, H. Qin, W. Zhao, and Y. Liu, "Automatic digital modulation classification based on curriculum learning," *Applied Sciences*, vol. 9, no. 10, p. 2171, 2019.

[15] S. A. Ghauri, "KNN based classification of digital modulated signals," *IIUM Engineering Journal*, vol. 17, no. 2, pp. 71–82, 2016.

[16] M. Bi, J. Yu, X. Miao, L. Li, and W. Hu, "Machine learning classifier based on FE-KNN enabled high-capacity PAM-4 and NRZ transmission with 10-G class optics," *Optics express*, vol. 27, no. 18, pp. 25802–25813, 2019.

[17] A. Hussain, M. Sohail, S. Alam, S. A. Ghauri, and I. M. Qureshi, "Classification of m-qam and m-psk signals using genetic programming," *Neural Computing and Applications*, vol. 31, no. 10, pp. 6141–6149, 2019.

[18] S. A. Ghauri and I. Mansoor Qureshi, "-PAM signals classification using modified Gabor filter network," *Mathematical Problems in Engineering*, vol. 2015, 10 pages, 2015.

[19] S. A. Ghauri, I. M. Qureshi, and A. N. Malik, "A novel approach for automatic modulation classification via hidden Markov models and Gabor features," *Wireless Personal Communications*, vol. 96, no. 3, pp. 4199–4216, 2017.

[20] R. Zhang, Z. Yin, Z. Wu, and S. Zhou, "A novel automatic modulation classification method using attention mechanism and hybrid parallel neural network," *Applied Sciences*, vol. 11, no. 3, article 1327, 2021.

[21] H. Han, Z. Ren, L. Li, and Z. Zhu, "Automatic modulation classification based on deep feature fusion for high noise level and large dynamic input," *Sensors*, vol. 21, no. 6, article 2117, 2021.

[22] N. Wang, Y. Liu, L. Ma, Y. Yang, and H. Wang, "Multidimensional CNN-LSTM network for automatic modulation classification," *Electronics*, vol. 10, no. 14, article 1649, 2021.

[23] S. Yuan, B. Wu, and P. Li, "Intra-pulse modulation classification of radar emitter signals based on a 1-D selective kernel convolutional neural network," *Remote Sensing*, vol. 13, no. 14, 2021.

[24] D. H. Al-Nuaimi, M. F. Akbar, L. B. Salman, I. S. Z. Abidin, and N. A. M. Isa, "AMC2N: automatic modulation classification using feature clustering-based two-lane capsule networks," *Electronics*, vol. 10, no. 1, article 76, 2021.

[25] W. Jiang, X. Wu, Y. Wang, B. Chen, W. Feng, and Y. Jin, "Time–trequency-analysis-based blind modulation classification for multiple-antenna systems," *Sensors*, vol. 21, no. 1, article 231, 2021.

[26] S. Chen, K. Qiu, S. Zheng, Q. Xuan, and X. Yang, "Radio–image transformer: bridging radio modulation classification and imagenet classification," *Electronics*, vol. 9, no. 10, article 1646, 2020.

[27] X. Sun, S. Su, Z. Zuo, X. Guo, and X. Tan, "Modulation classification using compressed sensing and decision tree–support vector machine in cognitive radio system," *Sensors*, vol. 20, no. 5, article 1438, 2020.

[28] P. M. Rodriguez, Z. Fernandez, R. Torrego, A. Lizeaga, M. Mendicute, and I. Val, "Low-complexity cyclostationary-based modulation classifying algorithm," *AEU-International Journal of Electronics and Communications*, vol. 74, pp. 176–182, 2017.

[29] R. Li, C. Song, Y. Song, X. Hao, S. Yang, and X. Song, "Deep geometric convolutional network for automatic modulation classification," *Signal, Image and Video Processing*, vol. 14, no. 6, pp. 1199–1205, 2020.

[30] S. H. Lee, K.-Y. Kim, and Y. Shin, "Effective feature selection method for deep learning-based automatic modulation classification scheme," *Applied Sciences*, vol. 10, no. 2, article 588, 2020.

[31] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Transactions on Communications*, vol. 48, no. 2, pp. 189–193, 2000.

[32] W. U. Khan, M. A. Jamshed, A. Mahmood, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Backscatter-aided NOMA V2X communication under channel estimation errors," 2022, http://arxiv.org/abs/2202.01586.

[33] A. Mahmood, A. Ahmed, M. Naeem, and Y. Hong, "Partial offloading in energy harvested mobile edge computing: A direct search approach," *IEEE Access*, vol. 8, pp. 36757–36763, 2020.

[34] M. Derakhtian, A. Tadaion, and S. Gazor, "Modulation classification of linearly modulated signals in slow flat fading channels," *IET Signal Processing*, vol. 5, no. 5, pp. 443–450, 2011.

[35] W. U. Khan, E. Lagunas, A. Mahmood, S. Chatzinotas, and B. Ottersten, "Integration of backscatter communication with multi-cell NOMA: a spectral efficiency optimization under imperfect SIC," 2021, http://arxiv.org/abs/2109.11509.

[36] A. Mahmood, Y. Hong, M. K. Ehsan, and S. Mumtaz, "Optimal resource allocation and task segmentation in IoT enabled mobile edge cloud," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 13294–13303, 2021.

[37] S. I. H. Shah, S. Alam, S. A. Ghauri, A. Hussain, and F. A. Ansari, "A novel hybrid cuckoo search-extreme learning machine approach for modulation classification," *IEEE Access*, vol. 7, pp. 90525–90537, 2019.

[38] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Communications*, vol. 1, no. 2, pp. 137–156, 2007.

[39] H.-C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 8, pp. 3098–3105, 2008.

[40] W. U. Khan, E. Lagunas, A. Mahmood, S. Chatzinotas, and B. Ottersten, "When RIS meets geo satellite communications: a new optimization framework in 6G," 2022, http://arxiv.org/abs/2202.00497.

[41] W. U. Khan, X. Li, A. Ihsan, M. A. Khan, V. G. Menon, and M. Ahmed, "NOMA-enabled optimization framework for

next-generation small-cell iov networks under imperfect sic decoding," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 1–10, 2021.

[42] P. Sapiano and J. Martin, "Maximum likelihood PSK classifier," *Proceedings of MILCOM'96 IEEE Military Communications Conference*, vol. 3, pp. 1010–1014, 1996.

[43] L. Hong and K. Ho, "Classification of BPSK and QPSK signals with unknown signal level using the Bayes technique," *Proceedings of the 2003 International Symposium on Circuits and Systems*, vol. 4, article 5758, 2003.

[44] A. Abdi, O. A. Dobre, R. Choudhry, Y. Bar-Ness, and W. Su, "Modulation classification in fading channels using antenna arrays," in *IEEE MILCOM 2004. Military Communications Conference*, vol. 1, pp. 211–217, Monterey, CA, USA., Oct 2004.

[45] A. O. A. Salam, R. E. Sheriff, S. R. Al-Araji, K. Mezher, and Q. Nasir, "A unified practical approach to modulation classification in cognitive radio using likelihood-based techniques," in *2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1024–1029, Halifax, NS, Canada, May 2015.

[46] D.-C. Chang and P.-K. Shih, "Cumulants-based modulation classification technique in multipath fading channels," *IET Communications*, vol. 9, no. 6, pp. 828–835, 2015.

[47] S. Kharbech, I. Dayoub, M. Zwingelstein-Colin, and E. P. Simon, "On classifiers for blind feature-based automatic modulation classification over multiple-input–multiple-output channels," *IET Communications*, vol. 10, no. 7, pp. 790–795, 2016.

[48] S. A. Ghauri, I. M. Qureshi, I. Shah, and N. Khan, "Modulation classification using cyclostationary features on fading channels," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 7, no. 24, pp. 5331–5339, 2014.

[49] S. H. Lee, K. Y. Kim, and Y. Shin, "Effective feature selection method for deep learning-based automatic modulation classification scheme using higher-order statistics," *Applied Sciences*, vol. 10, no. 2, article 588, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4763153, 11 pages
https://doi.org/10.1155/2022/4763153

Research Article

A Novel Efficient Data Gathering Algorithm for Disconnected Sensor Networks Based on Mobile Edge Computing

Zeyu Sun,[1,2,3] **Lan Lan**,[1,2] **Cao Zeng**,[1,2] **and Guisheng Liao**[1,3]

[1]*National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China*
[2]*Engineering Research Center of Henan Building Materials Big Data, Luoyang Institute of Science and Technology, Luoyang 471023, China*
[3]*Collaborative Innovation Center of information Sensing, Xidian University, Xi'an 710071, China*

Correspondence should be addressed to Guisheng Liao; gsliao@xidian.edu.cn

Received 21 February 2022; Revised 14 March 2022; Accepted 22 March 2022; Published 14 April 2022

Academic Editor: Xingwang Li

Employing mobile elements is an efficient solution to the performance improvement of wireless sensor networks (WSNs). We propose an efficient data gathering mechanism for disconnected WSNs with rendezvous points (DGM-RPs). The mobile sink traverses the entire network and stops only at the rendezvous points (RPs) while gathers the data from sensors in every disconnected segment. In this paper, mobile sinks perform the task of edge computing and alleviate the load of upper cloud. We measure the shape of disconnected segments, layering them by use of the convex hull, and then design the travelling path of the mobile sink to minimize the travel latency to visit all disconnected segments. At least one RPs will be selected in a segment firstly, and then, on this basis, we consider the distribution density of sensor nodes and the location of the RPs already exist to adding new RPs, which make good use of the margin to reducing the energy consumption and prolong the network lifetime.

1. Introduction

Wireless sensor networks (WSNs) have been playing an increasingly important role in wide-ranging applications, such as military, agriculture, home automation, medical treatment, environmental monitoring, and smart transportation [1–3]. Energy conservation is the most important thing in WSNs, since sensor nodes' capabilities are sustained until their batteries are depleted. In multihop communications due to heavy overload of relaying messages, the nodes which are near the sink have a tendency to die earlier than those which are farther away.

Mobile sink implicitly provides load balancing without extra effort, using mobile sink walks along the WSN to gather data instead of multihop transmission to the sink can dramatically reduce the energy consumption of sensors, which prolongs the lifetime of networks [4, 5]. Sparse and disconnected networks can be better handled with mobile sinks. In WSNs, the delivery latency is defined as the time mobile sink takes to travel around the network to collect data. However, visiting all sensors means a high delay; hence, rendezvous points are selected. Mobile sink is supposed to only visit rendezvous points (RPs) nodes when a delay is restricted. In theoretical research, we require networking a set of disconnected segments within relatively short time duration. Such as sensor nodes deployed in the forest, due to weather or natural disasters and other reasons, some sensor nodes have been damaged, resulting in the network is divided into several segments of nonconnected, and most of the sensor nodes can work properly [6–9]. Redeployment of the network will bring a huge cost; the use of mobile sink with a shorter delay in the collection of data in disconnected segments has become an ideal remedy [10–12]. Network topologies are various; it is more realistic and efficient to design the travelling path of the mobile sink dynamically according to the distribution of sensor nodes.

In this paper, we consider the problem of planning the travelling path of a mobile sink in a disconnected network and reducing the overload of sensor nodes to maximize network lifetimeFirstly, at least one RPs is selected in each

disconnected network to send the data collected by all sensor nodes to the mobile sink. Secondly, we consider distributing sensor nodes location information and add new RPs nodes to reduce the energy consumption of sensor node.

The contribution of this work can be summarized as follows:

(1) According to specific business requirements, we planned the travelling paths of the mobile sink dynamically. Length of paths was compressed maximally when the delay limit was critical, and the number of RPs was increased to reduce energy consumption and extend the lifetime of the network if the delay limit was relatively loose

(2) Instead of using a simplifying model to abstract a disconnected segment into one point, or to use sensor nodes located on the boundary of the segment to represent it, we took full account of the shapes and sizes of segments for their varieties, which more conforms actual applications

(3) With the assist of convex hull scanning, the RPs were placed in corresponding segments of each layer according to the order of outside-to-inside, and shapes of disconnected segments were taken full advantages to reduce the length of paths meanwhile

(4) Under the condition that the mobile SINK node can reach all inter-segment networks and the minimum delay, we consider increasing the number of PRs through the distribution status of sensor nodes and the length of remaining paths to reduce the network energy and prolong the network lifetime

2. Related Work

Recent works have exploited the use of mobile sink in data collection. There are many approaches to the problem of data collection in WSNs with mobile sinks. The paper [13] proposes a static and dynamic conversion scheme to replace faulty nodes. The faulty node is replaced with the adjacent redundant node, and the mobile node locates its best position through simple geometric operations to ensure the connectivity of the network. The paper [14] proposes a path planning strategy of mobile data collection, called the dual approximation of anchor points (DAAP). It is especially designed for disconnected WSNs where sensor nodes are scattered in multiple isolated segments, and it has the least calculation complexity compared with other existing works. The paper [15] proposed prediction model for lost packets based on data reconstruction (PM-LPDR). The lost packet on lossy links can be modeled as a random loss during the transmission, and the matching type of the lost packets can be further predicted. Retransmission is adopted for data recovery when a random packet loss is predicted, while prediction algorithms based on time sequences can be employed for data recovery when a random packet loss cannot be predicted. The paper [16] proposes the optimal deployment strategy of gateway nodes. This strategy can solve the problem of gateway node redundancy, construct fault tolerance with the least number of gateway nodes, and give an approximate solution algorithm. The paper [17] discusses the distributed fault-tolerant clustering algorithm of sensor networks, which is used to construct the k-order dominance set subgraph of wireless sensor networks to complete the optimization of the data transmission link. The paper [18] proves that for any given communication link, constructing a k-connected network topology, there is an approximate solution optimal algorithm, so that the boundary does not exceed $O(n^2)$. On this basis, a greedy distributed algorithm is proposed, and mobile nodes are used to optimize the entire network link. The paper [19] discusses adding appropriate mobile nodes to a wireless sensor network with a two-level cluster structure, so that each node of the network can communicate with at least two gateway nodes, thereby completing a multinode connectivity. The paper [20] is an adaptive clustering algorithm that is executed periodically. The algorithm is divided into two phases in each cycle, namely, the cluster establishment phase and the data stable communication phase. In the cluster establishment phase, adjacent nodes dynamically form clusters and randomly generate cluster heads; in the data stable communication phase, the hydrogen data of the nodes in the cluster is sent to the cluster head, and the cluster head performs data fusion and sends the result to the sink node. The paper [21] proposes a backup survival routing algorithm. The algorithm transmits multiple copies of the same data on multiple independent paths. When the node transmission fails, the original path is repaired or a new path is found to ensure the normal transmission of data; in the case of multipath, the algorithm uses encryption technology to implement redundant routing based on security and diversity and has good fault tolerance. The paper [22] discusses the communication method between the cluster head and the base station and specific methods to improve the efficiency of data transmission. Multipath is used to ensure that the information in the cluster can be reliably transmitted to the base station. By establishing multihop communication between the cluster head and the base station, part of the cluster head energy can be saved, thereby extending the network lifetime.

3. Materials and Methods

As shown in Figure 1, we consider heterogeneous WSNs usually deployed in areas that are dangerous or even unreachable to humans, which consist of many disconnected segments with some static sensor nodes. We have some assumptions before describing deep:

(1) Sensor nodes have a fixed data transmission range

(2) The number, shape, and size of the disconnected segments all three are random

(3) In every disconnected segment, sensor nodes are deployed on a plane randomly, and their positions can be known

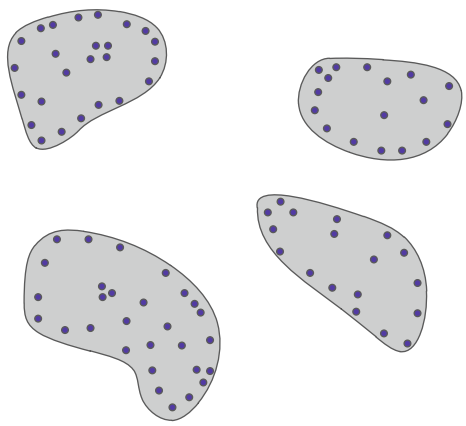


Figure 1: Typical network topology.

```
Input: V_i, the set of sensor nodes in segment i
       d_jm, the distance between node j and node m
       Size(·), get the element number of set
Output: R_i, the auxiliary node of segment i
begin
  for j ←1 to Size (V_i) do
    w ←0
    min ← +∞
    for k ←1 to N and k ≠ i do
      for m ←1 to Size (V_k) do
        w ← w + d_jm
      end for
    end for
    if w < min then
      min = w
      R_i= j
    end if
  end for
end
```

Algorithm 1: Determining the auxiliary node of segment i.

(4) Disconnected segments are isolated; sensor nodes deployed in different segments cannot communicate with each other by multihop transmission

(5) Only one mobile sink can be used for data collection. The mobile sink moves at a fixed speed

(6) Sensor nodes' routing time and the communication time can be negligible, as compared with the mobile sink's travelling time

(7) Each RPs node has a fixed storage capacity to buffer collected data; data can be collected when a mobile sink visits the RPs nodes in a specific time interval

We considered on both aspects of reducing latency and decreasing overall-network energy consumption. Commonly, an application scenario will be accompanied by a specific latency constraint, which brings a limit to length of path of the mobile sink [23–25]. We considered networks in two cases. Some have a small number of segments and a simple topology, and others have a large number of segments and a complex topology. For those simple, we used the center point to select the RP in each segment and plan the mobile sink's travelling strategy by running classical TSP algorithm quickly [26–28]. For those complexes, our approach was divided into two steps, including selection of first RP in every disconnected segment and using remaining available path length to add new RPs. Meanwhile, the first step included using convex hull to stratify the network and organizing the path. The second step included assigning the available path lengths proportionally to each disconnected segment and selecting RPs according to weight, which helped reducing network energy consumption.

3.1. Determining the Center Point. We defined "auxiliary node" for every segment. An auxiliary node is a sensor node that has the least sum of distances to all nodes in other segments. For a node i in segment k, its sum S_i can be determined by formula (1), where j denotes all nodes that are not in segment k. For each segment, we denoted it a number i; and all nodes in this segment formed a set, which was symbolized V_i. Then, we acquire auxiliary node of every segment, after running the algorithm as shown in Algorithm 1. After we found the coordinates of auxiliary nodes, we define the center point of the network with symbol C_0, whose coordinates (x, y) could be calculated by formula (2). Actually, it is the average coordinates of auxiliary nodes of all segments [29, 30]. The set-up of the center point can reduce the influence of shapes of segments on the mobile sink's path length, so we can use the center point to find a short path. If we follow Figure 2 to change the shape and size of one of the segments, the shortest path lengths required to access the four segments in Figure 2(a) and Figure 2(b) are the same.

$$S_i = \sum_j d_{ij} \quad i \in k, j \notin k, \tag{1}$$

$$(x, y) = \left(\frac{1}{N}\sum_{i=1}^{N} x_i, \frac{1}{N}\sum_{i=1}^{N} y_i\right). \tag{2}$$

The node nearest to the center point will be selected as a RP in each segment. For a network with a small number of segments and a simple topology, a travelling path for the mobile sink can be quickly built. The solution is particularly applicable to scenes with harsh latency. As shown in Figure 3(a), the nodes in segments are usually considered as points of uniform quality, and their centers of gravity will be calculated and then be selected as RPs, when people uses the traditional solution, which is inefficient and barely applicable to scenes with harsh latency. As comparison, our approach satisfies the requirement of latency by setting up the center point, which is shown in Figure 3(b). For a network with large numbers of nodes and complex topologies, we perform subsequent steps to adjust the selected RPs in each segment and add new RPs without exceeding the given latency constraint.

3.2. Determining the First RP in Each Disconnected Segment. In the following sections, we will explain how to plan the travelling path for the mobile sink in networks with large

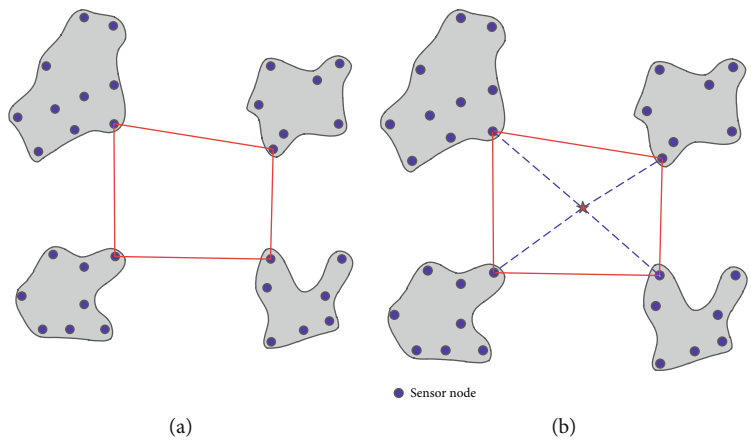


Figure 2: The shortest path of the four segments. (a) The case with similar segments. (b) The case with large difference of segments.

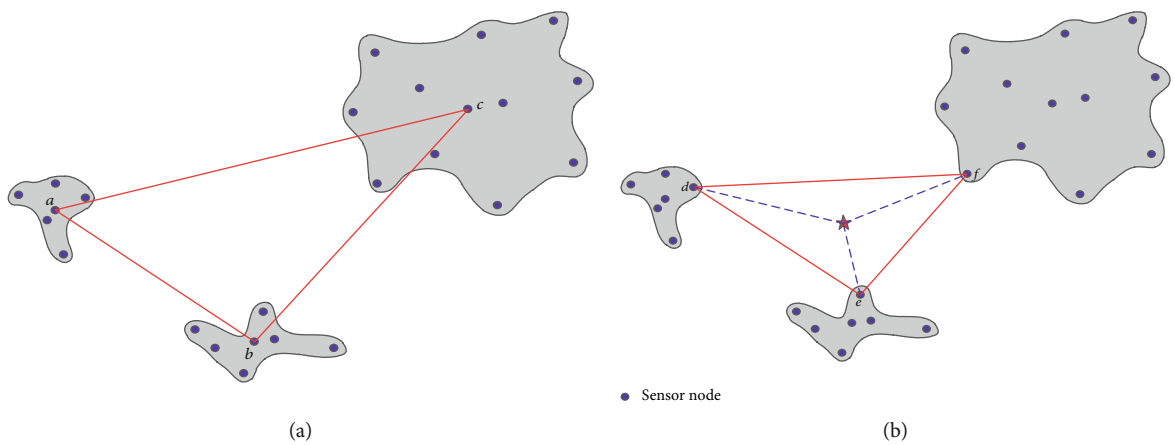


Figure 3: Path scheme for a network with simple topology. (a) Traditional approach. (b) Approach with the center point.

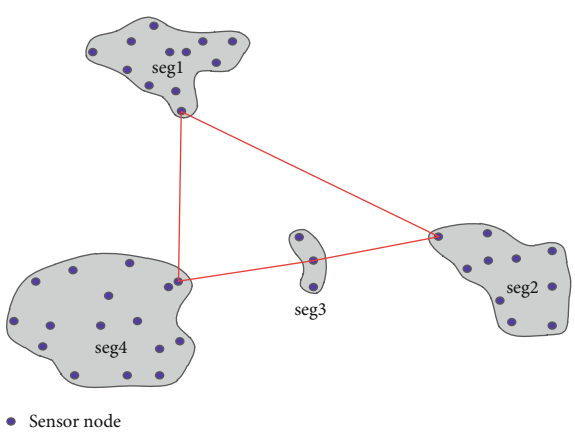


Figure 4: Selecting RPs located near the boundary.

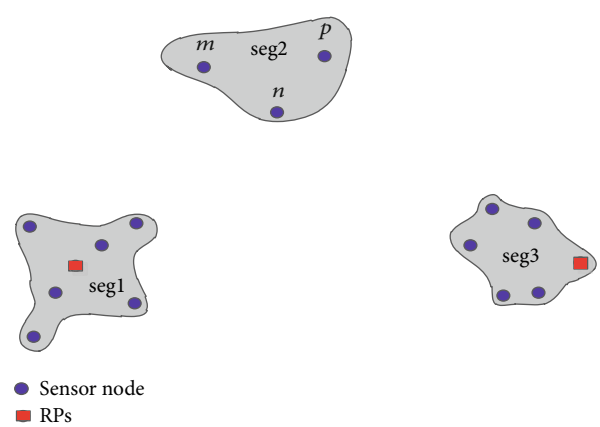


Figure 5: Refer to the known RPs to select RPs.

numbers of nodes and complex topologies, while the shape and size of segments are usually irregular. The first RPs in every segment will be set up primarily as described in Section 3.1. There is one and only one RPs in each segment; after the subsequent steps, the selection of the RPs in some segments will be adjusted. The selection of RPs do not need to be adjusted in some segments which are more possible located near the boundary of the network, and these segments are called as "remote segments." Specifically, whether a segment is a remote segment is determined by the stratification of the entire network, and all segments of the first layer are treated as remote segments. For these remote segments, selecting nodes near the boundary of segments can help reducing length of travelling path of mobile sinks, such as seg1, seg2, and seg4 in Figure 4. On the other hand, the mobile sink is needed to access three segments successively,

```
Input: A = {A_1, A_2, ···, A_i ··· , A_N}, set of representative node
       A', a set represent the convex hull of A
Output: the layer of disconnected segments
begin
  layer ← 1
  while A is not empty
    A' ← find the convex hull of A
    A ← A − A'
    for each i∈A' do
      mark segment i with layer
    end for
    layer++
  end while
end
```

Algorithm 2: Stratifying network.

and positions of RPs in the first segment and the third segment are known. The known two RPs can be used to determine the RPs of the second segment. Obviously, choosing the node that has minimum sum of distances to the known two RPs can reduce the length of mobile sink's path maximally. As shown in Figure 5, if the positions of RPs in seg1 and seg3 are known, node n will be chosen as the RPs in seg2. Our approach combines two main ideas mentioned above; that is, firstly the mobile sink draws a polygon by accessing all remote segments [31, 32]. After that, all other segments except the remote segments refer to the polygon to select the RPs so the entire path is shrunk, which helps reducing the path length. In order to distinguish which segments can be treated as remote segments and form the polygon, which segments can be within the polygon, the center point and the method of convex hull scanning are used to stratify the whole network.

Each segment contains several sensor nodes. To facilitate the stratification of the entire network, a representative node will be selected in each segment to represent the segment. The positional relationship between segment and other segments should be considered in our selection strategy of representative node in a segment, and the center point also can be used as a reference. The first RPs in each segment will be defined as a representative node, and all representative nodes form a collection of points. Given a set of points on a plane, Graham scan can be used to find the corresponding convex hull. More specifically, given a point set, the point was treated as starting point whose x coordinate or the y coordinate is the smallest. The other points are connected to the starting point to form a polar angle, and all points in the set are sorted according to the polar angle. A stack can be used for calculating the convex hull. Each point in the set will be stacked once, and finally, the points on the convex hull will be left on the stack. There are a total of N segments, one of which is represented by the symbol i; the representative node of the segment is represented by the symbol Ai; and the set of points of all representative nodes is denoted by the symbol A, and then, $A = \{A_1, A_2, \cdots, A_i \cdots, A_N\}$. The corresponding convex hull of point set A can be found, which is symbol as A', and A' is also a point set consisting of representative nodes. We retrieved all the elements in A' and mark their corresponding segments as the same layer. Take the difference set of A and A'; repeat the above iterative process; and ultimately, the entire network is divided into several levels. The stratify algorithm is summarized in Algorithm 2.

Through the stratification, the most peripheral segments, which were marked as the first layer of the network, can be found. Accessing all the RPs in the segments of the current layer can form a closed path. Each segment of the next layer will be considered, and the node that has the minimum sum of distances from itself to the two endpoints of the path lines will be adjusted as the RPs, where the path lines include all the lines that make up the closed path. In accordance with the rules described above, we can acquire the position of first RP of each segment. Given the set of RPs, the path design problem is modeled as the TSP, and the mobile sink's travelling path can be acquired by 2-opt algorithm, which is a simple local search algorithm first proposed in [14] for solving the TSP. The main idea behind it is to take a route that crosses over itself and reorder it so that it does not. The advantage of the 2-opt algorithm is easy understanding, and it can help us quickly find the approximate solution in $O(n^2)$ time, where n is the number of points.

After we got the mobile sink's travelling path, we made further optimization to compress the path. Assuming that the RPs is visited in the order of P_i, P_j, and P', which are in segment i, k, and j, respectively. P_i and P_j are connected as a line segment to check the nodes in segment k to see if there is a node P' that is more appropriate than P', which means the distance from P' to line $Pi * Pj$ is closer than P'. The path establishment and optimization algorithm are described in Algorithm 3; after the above stages, a relatively short path is found to ensure that each disconnected segments can be visited.

3.3. RPs Analysis and Realization. The various steps of our approach are explained using a detail example in Figure 6. The network consists of several disconnected segments, as shown in Figure 6(a). The "auxiliary nodes" are shown in

```
Input: V = {V_1, V_2, ⋯, V_N}, set of sensor nodes in segments
       A = {A_1, A_2, ⋯A_i ⋯ , A_N}, set of representative node
       MAX_LAYER, the max layer of disconnected segments
       I, a set represent the visit order of nodes in P
       I_i, the i-th element in the set
       Size (·), get the size of set
       Dis (p, l), calculate the distance from the point p to the line segment l
       Segment(·), find the segment in which the node distributed
       TSP(·), run TSP algorithm to determine the path length
       a ∗ b, line segment whose ends are node a and node b, respectively
Output: P = {P_1, P_2, ⋯, P_N}, set of RPs
procedure PATH_ESTABLISHMENT
begin
  find segments of the first layer, add corresponding A_i to P
  currentLayer ← 2
  while currentLayer < MAX_LAYER
    run TSP algorithm to determine the visit order of nodes in P
    connect adjacent visits to form path lines
    check all segments whose layer are currentLayer+1, add the node to P, the node should has the minimum sum of distances from
itself to the two endpoints of the path lines in its segments
    currentLayer++
  end while
end
end procedure
procedure PATH_OPTIMIZATION
begin
  do
    preLen ← TSP (P)
    optimize ← false
    I← get the visit order of nodes in P
    I←I∪{I_1,I_2}
    Loop: for k =1 to Size (I)-2 do
      for m = 1 to Size(V_Segment(I+1)) do
        if Dis (m, I_k ∗ I_(k+2)) < Dis (P_Segment(I+1), I_k ∗ I_(k+2)) then
          if TSP (P ∪ {m} − { P_Segment(I_+1)}) < pre_Len then
            P ⟵ P ∪ {m} − { P_Segment(I_+1)}
            optimize ← true
            break Loop
          end if
        end if
      end for
    end for
    I ⟵ I − {I_1 − I_2}
  while optimize = true
end
end procedure
```

Algorithm 3: Path organization.

Figure 6(b). Firstly, the center point and the representative node in each segment will be found. The center point is marked as a five-pointed star, and the representative node in each segment is shown in Figure 6(c). All representative nodes form a collection of points, Graham scan is used to find the corresponding convex hull, and one convex hull can split a layer of network. Figure 6(d) shows that the network is divided into three layers, while the outermost layer, the middle layer, and the innermost layer have 4, 3, and 2 segments, respectively. After stratification, the representative node in each segment will be selected as the first RPs, and then, the selection of RPs will be checked and modified layer by layer. For example, as Figure 6(e) shows, connecting all segments of the first layer forms the path lines; segments of layer two are checked to find a new RP, which has the minimum sum of distances from itself to the two endpoints of the path lines. After the above steps, each segment had RPs; the mobile sink's travelling path obtained by TSP algorithm is shown in Figure 6(f), while Figure 6(g) shows the result of our optimization step of the path. Figure 6(h) shows that the detection rectangle representing the distribution boundary of the node can be found, which will be divided

into several cells. The total length of the path cannot exceed a certain value, and the remaining available paths are fully utilized to expand the new RPs.

3.4. Using the Available Path Length to Add the New RPs. After the above steps, a short path for the mobile sink has been planned to visit all disconnected segments, which has only one RP in each segment. If the delay requirement is looser, the number of RPs will increase, and the path will be expanded to reduce the energy consumption of the whole network, which includes assigning available path lengths proportionally to each disconnected segment and weighted selection of RPs.

The numbers of nodes distributed in disconnected segments are different. If a small number of RPs is selected in a segment with many nodes, then some of the nodes need to send the data to the RPs by long-distance multihop transmission, which goes against the principle of extending the network life because of increasing the energy consumption of the network. To set up more RPs in the segments with many nodes is a reasonable program; therefore, the available remaining path length will be allocated by the number of nodes in the segment. For example, the current path length of the mobile sink is L, and the total path length is limited to L_{limit}, and then, the available remaining path length of segment i is determined by

$$L_i = \frac{\text{Size}(V_i)}{\sum_{k=1}^{N} \text{Size}(V_k)} \times (L_{\text{limit}} - L), \tag{3}$$

where function Size $(\cdot)$ can get the number of elements in a set and V_i is the set of sensor nodes distributed in segment i.

As described in the system model, each node knows its own position. Nodes' x coordinates are $X_1, X_2, \cdots, X_N$, and their y coordinates are $Y_1, Y_2, \cdots, Y_N$, respectively. Then, the distribution boundaries of the nodes are within a "detection rectangle," where the bottom left corner of the rectangle is located at $(\min(X_1, X_2, \cdots, X_N), \min(Y_1, Y_2, \cdots, Y_N))$ and the upper right corner of the rectangle is located at $(\max(X_1, X_2, \cdots, X_N), \max(Y_1, Y_2, \cdots, Y_N))$. We divide the rectangle into α^2 cells, with α to represent the granularity of the division. Put it another way, the rectangle is divided into α parts both vertically and horizontally. Next, for each segment, the weight of the nodes will be calculated according to formula (4), and the nodes will be sorted in descending order of weight. In each round, a new RPs will be added in order. If the new RPs does not cause the path length to exceed the limit, it will be retained. The factor β in formula (4) indicates the number of nodes in the cell to which the node belongs and reflects the density of the nodes. On the other hand, it is not appropriate for a large number of RPs to come together; new RPs should have some distance from the existing RPs; therefore, the factor φ is proportional to the number of RPs that already exist in the cell to which the node belongs.

$$W = \beta - \varphi. \tag{4}$$

Finally, we acquire the final RPs set and the corresponding mobile sink's travelling path by the algorithm described in Algorithm 3. The segments with more nodes will set up more RPs. The selection of the RPs strategy includes the consideration of segment size, shape, and distribution of nodes, which is conducive to reducing network energy consumption and extend the network life.

4. Performance Evaluation

The experiment has been implemented in a Java program. Experiments are conducted on a desktop with an Intel i3-4170 processor at 3.7GHz, 8GB memory, and a 64-bit Windows 10 operation system. The simulations are conducted in $1000 \times 1000\,\text{m}^2$ fields wherein random topologies are generated for a varying number of segments. Sensor nodes have a communication range fixed at 30 m, which are deployed in segments randomly. The segment topologies are ever-changing; in order to measure the size of the area and the difference in shape, the standard deviation σ is defined in formula (5), where the factor σ reflects the difference in shape and size between segments. For a segment, its corresponding δ is given by formula (6); the segment is regarded as a quality uniform pattern; and its centroid can be found. S is the area of the segment, and S_{circle} is the area of a circle, where the center of this circle is located in the segment's centroid and has an area equal to that segment. μ is given by formula (7), which represents the average of δ.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\delta_i - \mu)^2}, \tag{5}$$

$$\delta = \frac{S}{S_{\text{circle}}}, \tag{6}$$

$$\mu = \frac{1}{N}\sum_{i=1}^{N}\delta_i. \tag{7}$$

We let each segment to be a circle with a radius of 40 m or 50 m, which means that $S = 800$ or $S = 500$. Figure 7 shows the shortest mobile sink's travelling path length required to connect all segments, where the number of segments is varying from 10 to 60. From the figure, we can see that our approach outperforms DGM-RPS algorithm with about 20% less path length on average, because the DGM-RPS algorithm builds the geometry tree and does not take full account of the shape and size of each segment, and the RPs selection is always based on the product of distance and transmission hop count. In the case of significant differences between segments, DGM-RPS has a lower priority for setting RPs in areas with smaller area and fewer nodes, which may result in using a fairly long path to access all segments. As the standard deviation σ increases, the advantages of our approach will become increasingly apparent; this indicates that our approach is more suitable for the case of large differences between segments.

We changed the shape and size of each segment while maintaining the value of μ constant, so that each topology

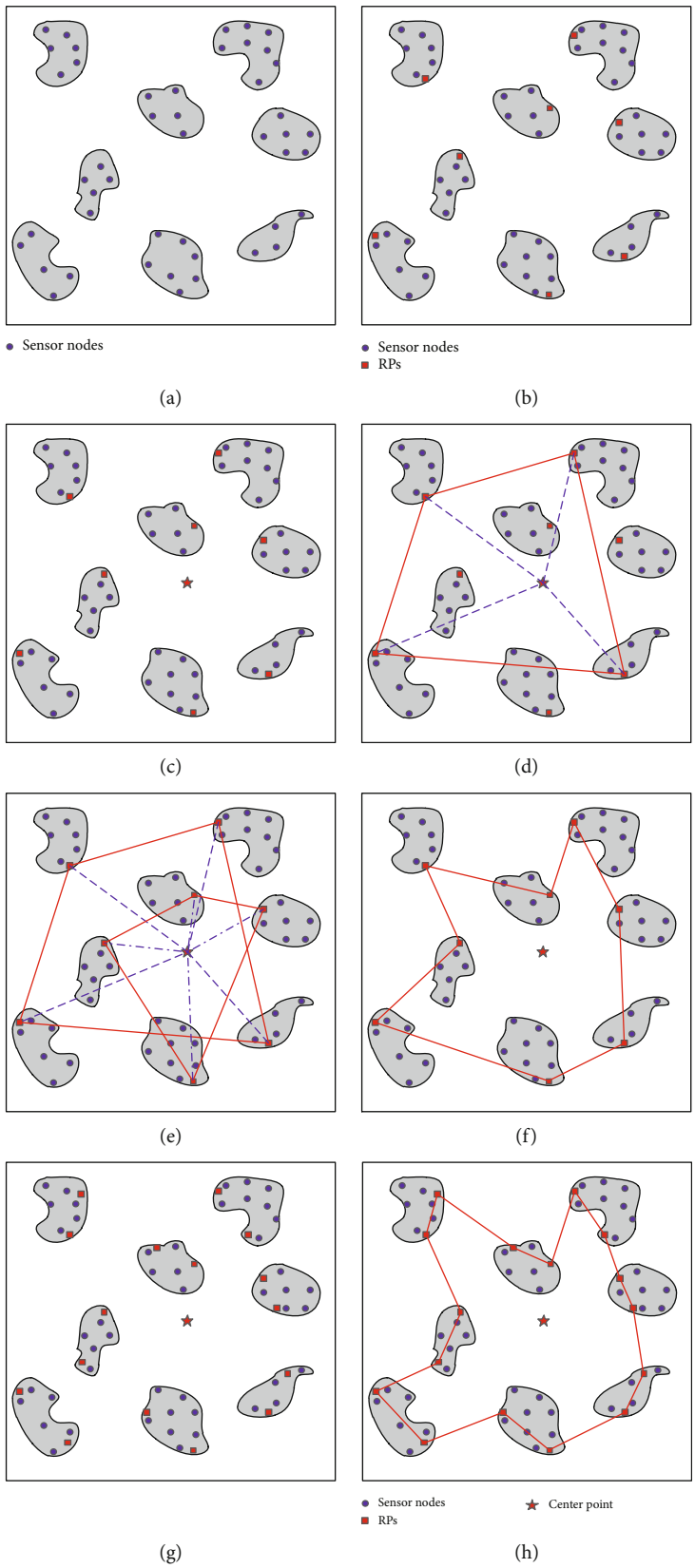
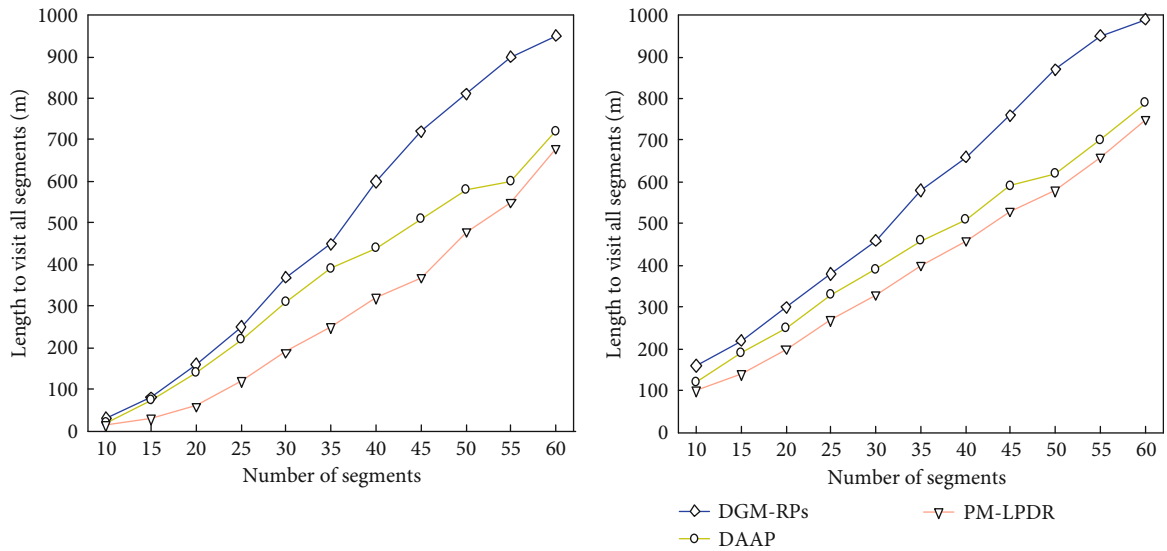


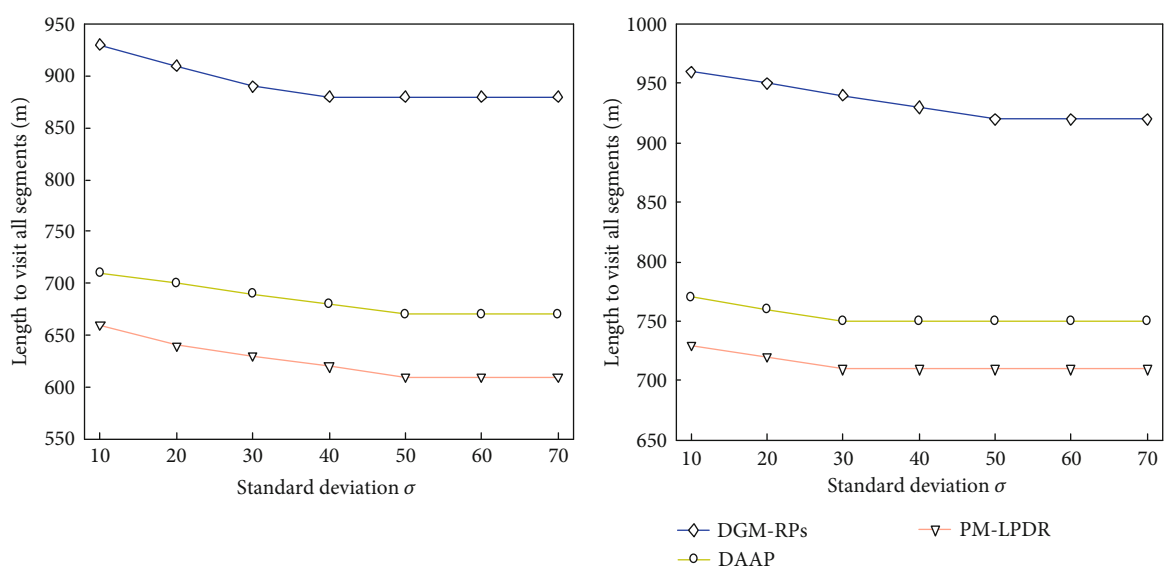
Figure 6: Determines RPs and path: (a) the network topology; (b) the RPs in every segment; (c) the selection of the closest nodes and the center point; (d) scan the convex hull and stratify the network; (e) modify the selection of RPs in the segments except the remote segments; (f) the mobile sink's travelling path obtained by TSP algorithm; (g) the mobile sink's travelling path after optimizing; and (h) divide the whole network into several cells.



(a) $S = 800$ and $\sigma = 20$

(b) $S = 500$ and $\sigma = 20$

Figure 7: Total path length to visit all segments.



(a) Visit 20 segments

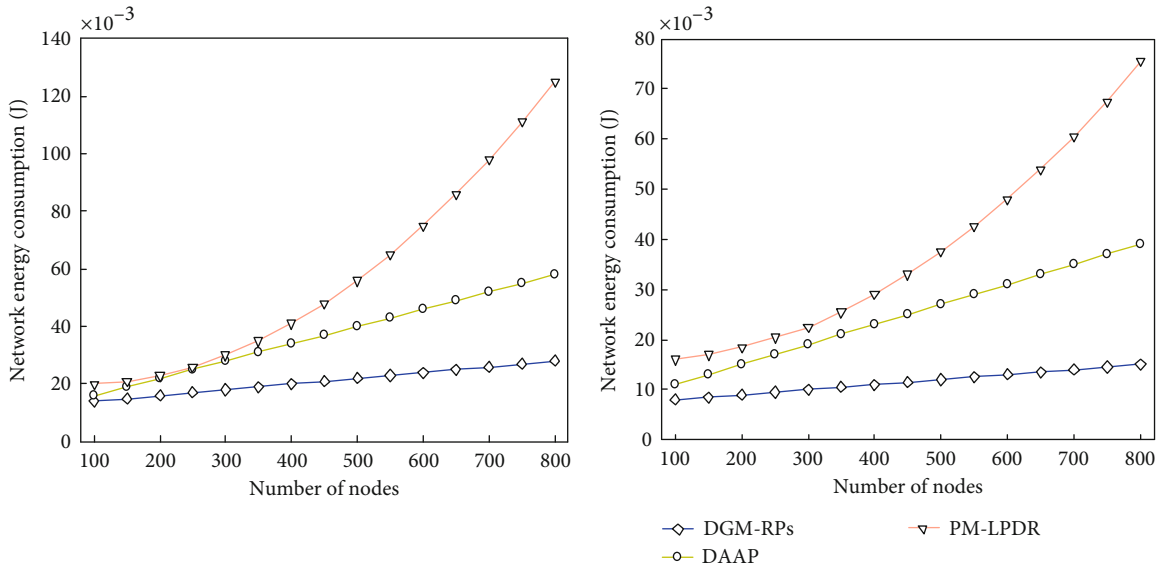
(b) Visit 40 segments

Figure 8: Total path length to visit several segments with $S = 800$.

corresponds to σ. We set the number of segments to 20, which are randomly generated and satisfy the constraints that $S = 800$. Figure 8 shows that when the differences between segments increased, our program performed better. With the increase of σ, the path length required for the DGM-RPS method and PM-LPDR remains essentially unchanged, while the length of visit all segments of our approach is showing a downward trend. Because we fixed μ, a larger value of σ means the area of some segments is relatively large or the shape of the segment is like a long strip, which caused the first RPs to be selected near the boundary of the segment and closer to the center point. Our approach takes full advantage of the shape of the segment, so as to achieve the effect of shortening the path.

The energy consumption of a sensor includes the power for sensing, receiving, and transmission. In this paper, we only take the energy consumption of receiving and transmission into account since the energy consumption for sensing is negligible compared with the other two types of energy consumption. Sensor nodes deployed in different segments cannot communicate with each other by multihop transmission. The energy consumption of a node to transmit k bits data over distance d is shown as

$$E_{Tx}(k, d) = \begin{cases} k \times E_{\text{elec}} + k \times \varepsilon_{fs} \times d^2 & d < d_0 \\ k \times E_{\text{elec}} + k \times \varepsilon_{mp} \times d^4 & d \geq d_0 \end{cases}. \quad (8)$$



(a) $L_{\text{limit}} = 5000$ m (b) $L_{\text{limit}} = 3400$ m

Figure 9: Network energy consumption of three methods with 20 segments and $S = 500$, $\sigma = 20$.

where $E_{\text{elec}} = 50$ nJ/bit, $\varepsilon_{fs} = 10$ pJ/bit/m^2, $d_0 = 87$ m, and $\varepsilon_{mp} = 0.0013$ pJ/bit/m^4.

In the case of total length of the mobile sink's path, travelling length cannot exceed 3400 m or 5000 m, and the detection rectangle is divided into 20×20 cells, while S was set to 500 and σ was set to 20. We consider the network consists of 20 segments, while the number of nodes from 100 to 800 and all nodes generate 1bit data every round. The performances of three methods are shown in Figure 9. Compared to other methods, our solution can save more than 10% of the energy for the network, thanks to the full consideration of the distribution density of nodes, while the design of the attenuation factor to the distance between RPs is too close. Let the total length constraint of the path remains unchanged, while $\mu = 40$ and $\sigma = 20$. Under these circumstances, using DGM-RPS algorithm and our approach to obtain the position of RPs and the corresponding path are shown in Figure 9(a) and Figure 9(b), respectively.

5. Conclusion

In this paper, we propose an efficient data gathering mechanism for disconnected wireless sensor networks with a mobile sink. The mobile sink traverses the entire network and stops only at the rendezvous points (RPs) while gathers the data from sensors in every disconnected segment. We measure the shape of disconnected segments, layering them by use of the convex hull, and then design the travelling path of the mobile sink to minimize the travel latency to visit all disconnected segments. At least one RPs will be selected in a segment firstly, and then, on this basic, we consider the distribution density of sensor nodes and the location of the RPs already exist to adding new RPs, which make good use of the margin to reducing the energy consumption and prolong the lifetime of network. We carry out some simulations, which reflect the advantages of our strategy in terms of path length and network lifetime.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Founding of China under Grant 61771015, 61931016, and 62101402; the National Natural Science Founding of Henan under Grant 202300410286; the Key Science and Technology Projects of Henan Province under Grant 212102210374; the Key Funding Project of Colleges and University of Henan Province under Grant 19A520006 and 20A520027; the Aviation Science Foundation of China (2020000108101); and the China Postdoctoral Science Foundation (2021TQ0261, 2021M702547).

References

[1] T. Haojun, L. Yuxin, L. Anfeng et al., "A novel code data dissemination scheme for internet of things through mobile vehicle of smart cities," *Future Generation Computer Systems*, vol. 94, pp. 351–367, 2019.

[2] L. Yunchang, L. Chunlin, L. Youlong, S. Yanling, and Z. Jing, "Scheduling multimedia services in cloud computing

envieronment," *Enterprise Information Systems.*, vol. 12, no. 2, pp. 218–235, 2018.

[3] Y. Jing, X. Ming, and P. Zhibo, "Distributed for computing based on batched sparse codes for industrial control," *IEEE Transactions on Industrial Informatics.*, vol. 14, no. 10, pp. 4683–4691, 2018.

[4] Z. Yan, "Classified scheduling algorithm of big data under cloud computing," *International Journal of Computers and Applications.*, vol. 41, no. 4, pp. 262–267, 2019.

[5] W. Xiuran and W. Feng, "Massive data balance scheduling in cloud computing environment," *International Journal of Mechatronics and Applied Mechanics*, vol. 2019, no. 5, pp. 100–105, 2019.

[6] Z. Yucheng, H. Wenzhun, Z. Ting, and Z. Tuo, "A novel topology optimization theory and parallel data analysis model based resource scheduling algorithm for cloud computing," *Recent Advances in Electrical and Electronic Engineering*, vol. 11, no. 4, pp. 449–456, 2018.

[7] X. Ruitao and J. Xiaohua, "Data transfer scheduling for maximizing throughput of big-data computing in cloud systems," *IEEE Transactions on Cloud Computing*, vol. 6, no. 1, pp. 87–98, 2018.

[8] X. Siyang, S. Xin, X. Lin, and X. Zhigang, "Energy efficiency maximization for energy harvesting bidirectional cooperative sensor networks with AF mode," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 6, pp. 2686–2708, 2020.

[9] W. Tian, L. Hao, W. Jia, A. Liu, and M. Xie, "MTES: an intelligent trust evaluation scheme in sensor-cloud enabled industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 2054–2062, 2020.

[10] S. Zeyu, R. Tao, N. Xiong, and X. Pan, "CS-PLM: compressive sensing data gathering algorithm based on packet loss matching in sensor networks," *Wireless Communications and Mobile Computing.*, vol. 2018, article 5131949, pp. 1–13, 2018.

[11] W. Liu Tang, B.,. W. Hongyi, and P. Jian, "Low-cost collaborative mobile charging for large-scale wireless sensor networks," *IEEE Transcations on Mobile Computing*, vol. 16, no. 8, pp. 2213–2227, 2017.

[12] H. Ren, X. Zichuan, W. Liang et al., "Efficient algorithms for delay-aware NFV-enabled multicasting in mobile edge clouds with resource sharing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 9, pp. 2050–2066, 2020.

[13] W. Tian, Z. Jiyuan, L. Anfeng, B. M. Zakirul, W. G. Alam, and J. Weijia, "Fog-based computing and storage offloading for data synchronization in IoT," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4272–4282, 2019.

[14] L. Xuxun, Q. Tie, Z. Xiaobo, W. Tian, Y. Lei, and C. Victor, "Latency-aware path planning for disconnected sensor networks with mobile sinks," *IEEE Transactions on Industrial Information*, vol. 16, no. 1, pp. 350–361, 2020.

[15] S. Zeyu, Z. Guozeng, and P. Xiaoyan, "PM-LPDR: a prediction model for lost packets based on data reconstruction on lossy links in sensor networks," *International Journal Computational Science and Engineering*, vol. 19, no. 2, pp. 177–188, 2019.

[16] N. Yalin, W. Haijun, Q. Yujie, and S. Zeyu, "Distributed and morphological operation-based data collection algorithm," *International Journal of Distributed Sensor Networks*, vol. 13, no. 7, 2017.

[17] G. Chan, Y. Bin, J. Xiaohong, I. Hiroshi, and F. Masaru, "Covert communication in relay-assisted IoT systems," *IEEE Internet Things Journal.*, vol. 8, no. 8, pp. 6313–6323, 2021.

[18] B. Salim, Z. Sherali, and M. Abdelhamid, "Fog computing job scheduling optimization based on bees swarm," *Enterprise Information Systems*, vol. 12, no. 4, pp. 373–397, 2018.

[19] L. Xuxun, "Routing protocols based on ant colony optimization in wireless sensor networks: a survey," *IEEE Access*, vol. 5, pp. 26303–26317, 2017.

[20] M. Mansour, B. Ghalem, and N. Abdelkader, "Towards a traceability system based on cloud and fog computing," *Multiagent and Grid Systems*, vol. 13, no. 1, pp. 47–68, 2017.

[21] S. Zeyu, X. Xiaofei, S. Bin, N. Yalin, and S. Hongxiang, "Mobile intelligent computing in internet of things: an optimized data gathering method based on compressive sensing," *IEEE Access*, vol. 7, pp. 66110–66122, 2019.

[22] W. Tian, K. Haoxiong, Z. Xi, W. Kun, S. A. Kumar, and L. Anfeng, "Big data cleaning based on Mobile edge computing in industrial sensor-cloud," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1321–1329, 2020.

[23] W. Tian, L. Yang, W. Guojun, C. Jiannong, B. M. Zakirul, and J. W. Alam, "Sustainable and efficient data collection from WSNs to cloud," *IEEE Transactions on Sustainable Computing*, vol. 4, no. 2, pp. 252–262, 2019.

[24] S. Zeyu, L. Zhiguo, X. Hou Yue, and Y. B. Chen, "MR-DFM: a multi-path routing algorithm based on data fusion mechanism in sensor networks," *Computer Science and Information Systems*, vol. 16, no. 3, pp. 867–890, 2019.

[25] L. Xuxun, Q. Tie, and W. Tian, "Load-balanced data dissemination for wireless sensor networks: a nature-inspired approach," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9256–9265, 2019.

[26] S. Zeyu, L. Longxing, X. Xiaofeng, L. Zhiguo, and N. N. Xiong, "A novel nodes deployment assignment scheme with data association attributed in wireless sensor networks," *Journal of Internet Technology*, vol. 20, no. 2, pp. 509–520, 2019.

[27] S. Zeyu, W. Huihui, L. Baoluo, L. Chuanfeng, P. Xiaoyan, and N. Yalin, "CS-FCDA: a compressed sensing-based on fault-tolerant data aggregation in sensor networks," *Sensors*, vol. 18, no. 11, p. 3749, 2018.

[28] H. Xingyu, P. Zhiwei, W. Jiangtao, and Y. Guisong, "Generic and efficient connectivity determination for IoT applications," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5291–5301, 2020.

[29] S. Zeyu, L. Jun, L. Zhixian et al., "CSR-IM: compress sensing routing-control-method with intelligent migration-mechanism based on sensing cloud-computing," *IEEE Access*, vol. 8, pp. 28437–28449, 2020.

[30] L. Xuxun, W. Tian, J. Weijia, L. Anfeng, and C. Kaikai, "Quick convex hull-based rendezvous planning for delay-harsh mobile data gathering in disjoint sensor networks," *IEEE Transactions on Systems, Man, And Cybernetics: Systems*, vol. 51, no. 6, pp. 3844–3854, 2021.

[31] W. Bin, Z. Xi, and H. Qinfang, "Optimal data routing algorithm for mine WSNs based on maximum lifetime," *IEEE Access*, vol. 8, pp. 131826–131834, 2020.

[32] Y. Guisong, L. Tingting, H. Xingyu, and X. Naixue, "Global and local reliability-based routing protocol for wireless sensor networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3620–3632, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3144509, 13 pages
https://doi.org/10.1155/2022/3144509

Research Article

A 3D Indoor Positioning Method of Wireless Network with Single Base Station in Multipath Environment

Youqing Wang,[1,2] **Kun Zhao**,[1,2] **and Zhengqi Zheng**[1,2]

[1]*Engineering Center of SHMEC for Space Information and GNSS, East China Normal University, Shanghai 200241, China*
[2]*Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China*

Correspondence should be addressed to Kun Zhao; kzhao@ce.ecnu.edu.cn

Received 2 February 2022; Revised 24 February 2022; Accepted 23 March 2022; Published 14 April 2022

Academic Editor: Xingwang Li

The proliferation of indoor location-based services has increased the demand of indoor positioning technology. Severe multipath and coherence effects are the difference between signal propagation indoors and outdoors. Most existing indoor localization methods build their models in 2-dimensional space and try to avoid the influence of multipath. We propose a method to realize 3-dimensional indoor positioning with single base station by using multipath channel. The angles of multipath coherent signals are estimated by MIMO antenna and the delays are estimated by OFDM signal. To avoid the complicated calculation in joint estimation of angles and delays in 3-dimensional space, the angles and delays are estimated separately and matched by the proposed algorithm. The line-of-sight channel is differentiated by time delay, and the reflection paths for non-line-of-sight channels are established with angle information and indoor maps. Finally, combine the angle information of the reflection paths and the line-of-sight path to obtain the target position in 3-dimensional indoor space. We verified the method through simulation in an indoor space of 6m × 8m × 4.5m. The positioning errors are submeter level in 95% cases and less than 0.4m in 60% cases. The proposed method requires only one base station and can be applied in most wireless networks. Compared with existing indoor localization methods, it has lower computational complexity and higher application potential.

1. Introduction

Location information plays an important role in communications, navigation, and other fields. The expanding of indoor space of modern building facilities has derived many new services and applications based on location information, which puts forward new requirements and challenges for indoor positioning technology. Common indoor positioning technologies mainly include Inertial Navigation [1], Infrared [2], Ultrasonic [3], Radio Frequency IDentification [4], Bluetooth [5], WiFi [6], Ultra WideBand [7], Visible Light [8], and Wireless Communication Network [9]. Researchers have developed many indoor positioning methods based on these technologies [10], but most of these technologies cannot balance the positioning accuracy, coverage, and cost. Wireless communication network is one of the signals with the widest indoor coverage and the most equipment. In order to increase the transmission performance, the new technologies such as Orthogonal Frequency Division Multiplexing (OFDM) and Multiple Input Multiple Output (MIMO) are introduced, and the technical specifications also divide the higher frequency band and wider bandwidth to it, such as millimeter wave band in 5G mobile network [11], that also brings new opportunities for indoor positioning. In addition, 3GPP also puts forward the requirement that the positioning error of the User Equipment (UE) in future wireless communication should be less than 3 m for 80% of indoor environment [12]. Therefore, it is of great significance to study for indoor positioning technology based on wireless communication networks.

The biggest difference between signal propagation in indoor and outdoor environment is that there has serious multipath interference in complex indoor environment. Researchers have proposed various solutions to this problem. Garcia et al. proposed a direct localization approach

for Massive MIMO [13], which is based on the framework of a novel compressed sensing. They distinguished the Line-Of-Sight (LOS) and Non-Line-of-Sight (NLOS) signal paths by exploiting the fact that LOS components must originate from a common location, whereas NLOS components have arbitrary AOA. But the model they used is a 2-dimensional (2D) rectangle plane, and four circular array antennas are distributed in four corners. Ma et al. proposed an indoor localization method based on Angle of Arrival (AOA) and Phase Difference Of Arrival (PDOA) using virtual stations for passive ultrahigh frequency radio frequency identification [4]. They choose the two strongest paths according to the received signal strength and use the reflection principle of multipath to establish virtual BS to perform localization. But they did not discuss that the power of LOS path may be lower than NLOS path in some actual indoor positioning environment, for example, the signal can penetrate the obstacle and some energy will be absorbed in LOS path. Zhou et al. proposed an indoor localization method of using the adjacent angle power difference [14]. They use the OMP algorithm to obtain the initial estimation of the direction and adjust it through the power difference of the adjacent angle. And then, the previous location is used as the starting point of the next, so the estimation accuracy and stability of continuous moving target are effectively improved. He and So proposed LTE indoor localization system which combines positioning algorithms based on the Time Difference Of Arrival (TDOA) and fingerprinting [15], they determine a coarse target location through the TDOA algorithm and use it to identify the subarea of monitoring area which includes the best location estimate. Then, the fingerprinting algorithm based on deep neural net will perform the accurate location point by using the corresponding subarea.

In general, positioning technology mainly uses the propagation time, propagation direction, and power attenuation of electromagnetic wave or acoustic wave in space. From the existing research progress, the time-based positioning requires high-precision clock synchronization or large bandwidth, and at least three base stations (BS) are required. The positioning based on signal strength is vulnerable to the environment and needs to update the priori information frequently. The positioning technology based on angle information requires the equipment to have an array antenna, or at least two BS are required. For indoor multipath effect, there are three main processing methods. First way is trying to avoid the influence of multipath signals during positioning [13, 16], the second way is to make full use of multipath signals [4, 17, 18], and the third way is to use multipath signals indirectly according to the characteristics of the environment, like some algorithms which is based on fingerprint information and machine learning [14, 15]. In the future indoor positioning technology, a major trend is to combine these principles to achieve high-precision positioning.

In this paper, we propose a 3-dimensional (3D) indoor positioning method based on multipath information, which makes full use of OFDM technology and MIMO array antenna in wireless communication networks. The main contributions of this paper are as follows.

(1) TDOA and AOA estimation algorithms for indoor multipath signals, the algorithm can avoid the computational complexity of joint estimation

(2) making full use of the multipath signal of indoor complex environment. Combining the TDOA and AOA information with the indoor map to reconstruct the indoor multipath channel model

(3) realizing the indoor positioning in 3D spatial based on single BS, which has lower equipment cost and better application prospect than existing methods

The main works of each section are as follows. We discuss the existing research of indoor positioning and the advantages of our research in the first section. In Section 2, we introduce the indoor space model and signal model of this paper. In Section 3, we describe TDOA and AOA algorithms for multipath coherent signals. In Section 4, we describe the matching methods of AOA and TDOA, and how to obtain the target location. We simulate the proposed method in Section 5 and summarize the full paper in the last section.

2. System Model of 3D Indoor Positioning

2.1. Indoor Environment Model. The 3D indoor environment model of this paper is a cube space, in which the length, width, and height are 8m, 6m, and 4.5m, respectively; here, the m means meter, as shown in Figure 1. Establish a 3D coordinate system with the top center as the origin. The x-axis is perpendicular to the front and rear planes, and the y-axis is perpendicular to the left and right planes. Take the positive direction of the x-axis as the north of indoor environment, the equations of each plane are shown in Table 1.

Suppose UE (x_s, y_s, z_s) is the signal source location in the space, and the BS (x_0, y_0, z_0) is the reference antenna of array antenna, L_p is the path of the multipath channel, where $p = 0, 1, \cdots, P-1$, $R_p(x_p, y_p, z_p)$ is the reflection point of the p-th multipath channel on the wall and the LOS path has no reflection point. Due to the signals after multiple reflections and long-distance propagation have severe fading, only the shorter P paths are considered. In Figure 1, we set the coordinate of UE is UE $(1.5,-2,-3)$,the coordinate of BS is BS $(0, 0, 0)$. Then, we can obtain the reflection points of the three NLOS paths which are $R_1(3,-1.333,-2)$, $R_2(1.125,-1.5,-4.5)$, $R_3(1,-4,-2)$.

Since the array antenna located in the central of the top, we design it as a "+" shape, it can obtain the maximum planar antenna aperture. As shown in Figure 2(a). The whole array has 31 antenna elements, which are divided into two identical Uniform Linear Arrays (ULA) and distributed on the x-axis and y-axis, respectively. Figure 2(b) is the ULA on x-axis. The 9th antenna of two ULA is a common array element and can be regarded as reference array element. As an example, we set the array element spacing of ULA is 0.03 m, which can be applied to any wireless network with frequencies below 5GHz. For higher frequency, the element

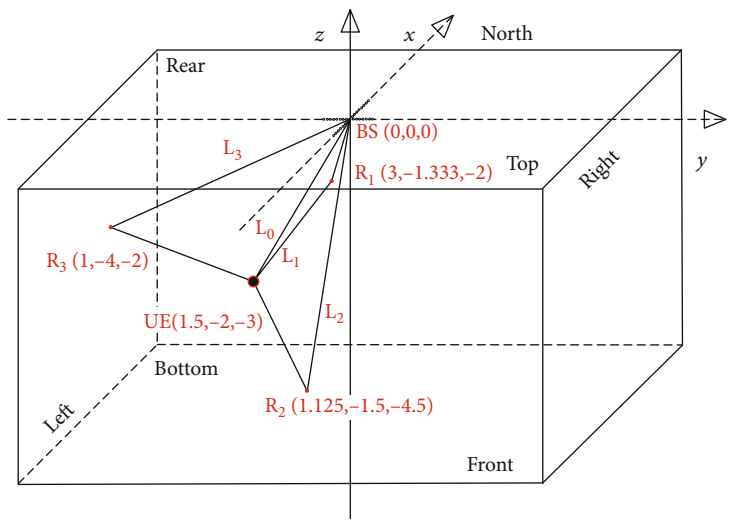


FIGURE 1: 3D indoor environment model, where BS is the location of the base station,UE is the location of the user equipment, L_0, L_1, L_2, L_3 are the channel paths, and R_1, R_2, R_3 are the multipath reflection points.

TABLE 1: Plane equation of indoor environmental model.

Plane	Front	Rear	Left	Right	Top	Bottom
Equation	$x=-3$	$x=3$	$y=-4$	$y=4$	$z=0$	$z=-4.5$

spacing can be adjusted to meet the limit of less than half the signal wavelength.

2.2. Positioning Signal Model. One of the purposes of this paper is to use OFDM pilot subcarrier signals to realize TDOA estimation of multipath signals, so we do not care about the details of information transmission. Another advantage of using pilot subcarrier signals is that it does not occupy the information transmission rate. Consider an OFDM signal with K subcarriers, the transmission symbols of OFDM can be expressed as Equation (1).

$$S(t) = \frac{1}{\sqrt{T}} A_{UE} \sum_{k=0}^{K-1} C_k(t) e^{j2\pi(f_c + k\Delta f)t}, \tag{1}$$

where $t \in [0, T]$,and T is the data duration of OFDM symbols, A_{UE} is the transmission gain of the UE, $C_k(t)$ represents the data on the k-th subcarrier, f_c represents the carrier frequency modulated by the communication system, and Δf is the frequency interval of subcarriers.

When the OFDM signal propagates in indoor environment, the impact response of multipath channel can be expressed as Equation (2).

$$h(t) = \sum_{p=0}^{P-1} a_p(t) \delta(t - \zeta_p), \tag{2}$$

where P is the number of channels, $a_p(t)$ is the amplitude fading of the p-th path, t is the transmission time of the signal, and ζ_p is the delay of multipath channel.

The multipath signals received by a single antenna can be expressed as Equation (3).

$$y(t) = \frac{1}{\sqrt{T}} A_{UE} A_{BS} \sum_{p=0}^{P-1} a_p(t) \sum_{k=0}^{K-1} C_k(t) e^{j2\pi(f_c + k\Delta f)(t - \zeta_p)} + n(t), \tag{3}$$

where A_{BS} is the gain of the receiving array antenna, $n(t)$ is the Gaussian white noise, since the change of channel is very slow compared with the sampling time, $1/\sqrt{T}$, A_{UE}, A_{BS}, $a_p(t)$ can be regarded as constants, and we assume these components are 1. The phase change caused by reflection is very small compared with propagation delay, which has been ignored.

Some time-based methods assume there is high-precision time synchronization between the transmitter and receiver [19]. The delays of these algorithms estimate are ζ_p as shown in Equation (3), but the high cost of atomic clock and the low sampling frequency of communication system make these methods not suitable for practical application. Theoretically, the signal of LOS path will first arrive at the receiver due to the shortest distance. Set L_0-th path is the LOS channel and the signal from L_0 arrive at the receiver at time t_{LOS},where $t_{LOS} = t - \zeta_{LOS}$,and t is the moment when the signal is emitted. The TDOA between NLOS channels and LOS channel is $\tau_p = \zeta_p - \zeta_{LOS}$. Thus, Equation (3) can be changed to Equation 4.

$$y(t) = \sum_{p=0}^{P-1} \sum_{k=0}^{K-1} C_k(t) e^{j2\pi(f_c + k\Delta f)(t_{LOS} - \tau_p)} + n(t). \tag{4}$$

3. AOA and TDOA Calculation

3.1. AOA Algorithm with Multipath Channel. Most existing algorithms obtain the AOA of the signal in 3D space by

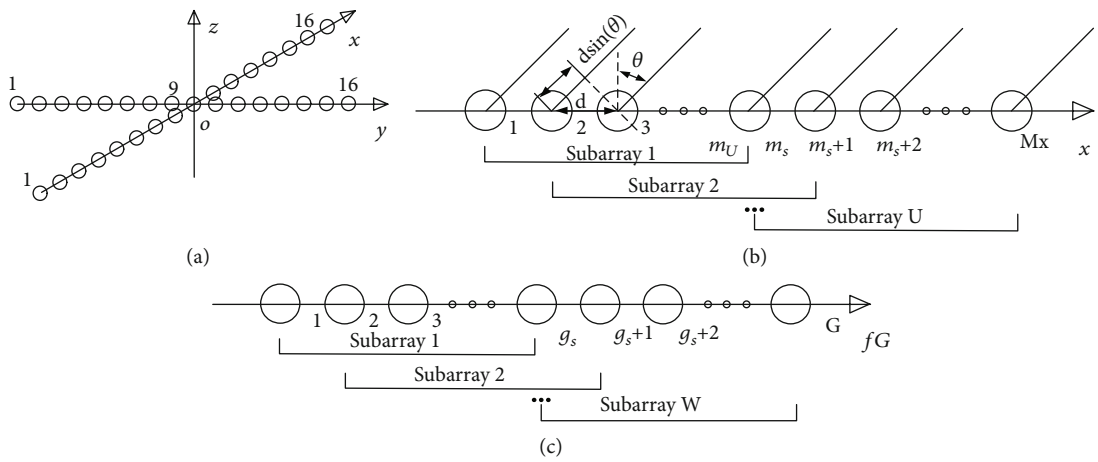


Figure 2: (a) The antenna array model. (b) The model of ULA on x-axis and spatial smoothing algorithm. (c) The model of Uniform Frequency Array and frequency domain smoothing algorithm.

jointly estimating the pitch angles and azimuth angles [20, 21], which makes the calculation very complicated. To simplify the algorithm, we do not estimate the angles jointly.

The spatial smoothing MUltiple SIgnal Classification (MUSIC) algorithm is used to estimate the AOA of multipath signals in this paper. The 1-dimensional MUSIC algorithm has the advantages of superresolution and high precision compared with the traditional beamforming algorithms and has lower computational complexity compared with the compressed sensing algorithms. And the computational complexity of MUSIC algorithm will increase exponentially with the number of dimensions when performing joint estimation, which is one of the shortcomings of the MUSIC algorithm. But we avoid this disadvantage by estimating the azimuth angles on two coordinate axes separately, and then match the two angles to obtain the AOA in 3D space.

Taking the ULA on the x-axis as an example to introduce the principle of spatial smoothing MUSIC algorithm. The structure of ULA is shown in Figure 2(b), there are M_x antennas uniformly distributed in a line with the same spacing d. We set the m_s-th antenna as the reference antenna and the right direction is the positive direction of the x-axis. When the far-field signals $S(t)$ are incident on the array antenna, the signals received by m_x-th antenna can be expressed as Equation (5).

$$y_{m_x}(t) = \sum_{p=0}^{P-1}\sum_{k=0}^{K-1} C_k(t) e^{j2\pi(f_c+k\Delta f)\left(t_{LOS}-\tau_p-\tau_{m_x}\right)} + n(t), \quad (5)$$

where $\tau_{m_x} = (m_x - m_s)dsin\theta/c$ is the delay between the m_x-th antenna and reference antenna, and $m_x = 1, 2, \cdots, M_x$, c is the speed of light. θ is the angle between signal propagation direction and the normal of ULA.

Only one subcarrier frequency is needed to estimate AOA. Thus, the signal matrix of the k-th subcarrier received by the array can be expressed as Equation (6).

$$\boldsymbol{Y}_x = \left[y_1(t) \quad y_2(t) \cdots y_{M_x}(t)\right]', \quad (6)$$

where " $'$ " represents the transpose of matrix.

The steering vector of array antenna is shown in Equation (7).

$$b_x(\theta_p) = \begin{bmatrix} e^{-j2\pi(f_c+k\Delta f)(1-m_s)dsin\theta_p/c} \\ e^{-j2\pi(f_c+k\Delta f)(2-m_s)dsin\theta_p/c} \\ \vdots \\ e^{-j2\pi(f_c+k\Delta f)(M_x-m_s)dsin\theta_p/c} \end{bmatrix}. \quad (7)$$

The multipath signals matrix is shown in Equation (8).

$$S = \begin{bmatrix} C_k(t)e^{j2\pi(f_c+k\Delta f)t_{LOS}} \\ C_k(t)e^{j2\pi(f_c+k\Delta f)(t_{LOS}-\tau_1)} \\ \vdots \\ C_k(t)e^{j2\pi(f_c+k\Delta f)(t_{LOS}-\tau_{P-1})} \end{bmatrix}. \quad (8)$$

The mathematical model of Y_x can be expressed as Equation (9).

$$Y_x = B_x S + N_x, \quad (9)$$

where $\boldsymbol{B}_x = [\boldsymbol{b}_x(\theta_0) \quad \boldsymbol{b}_x(\theta_1) \cdots \boldsymbol{b}_x(\theta_{P-1})]$, and N_x is the noise part.

It can be seen from Equation (8) that the NLOS path signals can be represented by the LOS path signal as $C_k(t_{LOS})$ $e^{j2\pi(f_c+k\Delta f)t_{LOS}}e^{-j2\pi(f_c+k\Delta f)\tau_p}$,and the phase difference between each NLOS and LOS path is fixed. That means they are

coherent signals. We use spatial smoothing to deal with multipath signals. Decompose the array antenna into staggered U subarrays, as shown in Figure 2(b). The number of antennas of each subarray is $m_U = M_x - U + 1$. Taking the first subarray as the reference subarray, the data of the u-th subarray can be expressed as Equation 10.

$$Y_{xu} = \left[y_{xu}(t)\, y_{xu+1}(t) \ \cdots \ y_{xu+m_U-1}(t) \right]' = B_x D_x^{(u-1)} S + N_{xu}, \tag{10}$$

where xu means the u-th subarray of x-axis array antenna, and $u = 1, 2, \cdots, U, \boldsymbol{N}_{xu}$ is the noise part of u-th subarray, the expression matrix $D_x^{(u-1)}$ is shown in Equation (11).

$$D_x^{(u-1)} = \begin{bmatrix} e^{-j2\pi(f_c+k\Delta f)d\sin\theta_0/c} & 0 & \cdots & 0 \\ 0 & e^{-j2\pi(f_c+k\Delta f)d\sin\theta_1/c} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-j2\pi(f_c+k\Delta f)d\sin\theta_{P-1}/c} \end{bmatrix}^{(u-1)}. \tag{11}$$

The covariance matrix of the forward spatial smoothing MUSIC algorithm Q_x^f is obtained as Equation (12).

$$Q_x^f = \frac{1}{U} \sum_{u=1}^{U} B_x D_x^{(u-1)} Q_s \left(D_x^{(u-1)} \right)^H B_x^H + N_x, \tag{12}$$

where $Q_s = \mathrm{E}[SS^H]$,and the $E[\cdots]$ represents the mean value, H represents the conjugate transpose.

If taking the rightmost array as the reference array, the covariance matrix of the backward spatial smoothing MUSIC algorithm Q_x^b as Equation (13) can be obtained by using a similar method.

$$Q_x^b = \frac{1}{U} \sum_{u=1}^{U} B_x J \left(D_x^{(u-1)} Q_s D_x^{(u-1)H} \right)^H J B_x^H + N_x, \tag{13}$$

where J is the $m_U \times m_U$ dimension transformation matrix as shown in Equation (14).

$$J = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & \ddots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}_{m_U \times m_U}. \tag{14}$$

The covariance matrix of forward and backward spatial smoothing algorithm is shown in Equation (15).

$$Q_x^{fb} = \frac{1}{2} \left(Q_x^f + Q_x^b \right). \tag{15}$$

Setting the number of sampling snapshots in practical application is N_S,the maximum likelihood estimation of the covariance matrix is $\widehat{Q}_x^{fb} = 1/N_S \sum_{i=1}^{N_S} Q_{xi}^{fb}$. Eigenvalue decomposition of $\widehat{Q}_x^{fb}$ is shown in Equation 16.

$$\widehat{Q}_x^{fb} = V_{xs} \Sigma_{xs} V_{xs}^H + V_{xn} \Sigma_{xn} V_{xn}^H, \tag{16}$$

where Σ_{xs} is the part of larger eigenvalues. V_{xs} is the eigenvectors corresponding to Σ_{xs}, $V_{xs}\Sigma_{xs}V_{xs}^H$ represents the signal subspace, Σ_{xn} is the part of smaller eigenvalues, V_{xn} is the eigenvectors corresponding to Σ_{xn}, $V_{xn}\Sigma_{xn}V_{xn}^H$ represents the noise subspace. Due to signal subspace which is the same as the subspace formed by the steering vectors of incident signals, the spectral formula of MUSIC algorithm is obtained as Equation 17.

$$P_{x\mathrm{MUSIC}} = \frac{1}{b_x(\theta_p)^H V_{xn} V_{xn}^H b_x(\theta_p)}. \tag{17}$$

Since the noise, $b_x(\theta_p)^H V_{xn}$ will not be equal to 0. By putting all possible values of θ_p into Equation (17), the right angles will make $P_{x\mathrm{MUSIC}}$ get the maximum value.

3.2. TDOA Algorithm for Multipath Channel. We apply the spatial smoothing MUSIC algorithm to the frequency domain and realize the TDOA estimation based on the difference of subcarrier phases. The proposed algorithm can avoid the requirement of time synchronization and high sampling frequency. The phase of the k-th subcarrier from the p-th multipath to the array antenna is set as $\varphi_{k,p}$,where $\varphi_{k,0}$ is the phase of the LOS path. The phase difference between NLOS path and LOS path can be expressed as Equation (18).

$$\Delta\varphi_k = \varphi_{k,p} - \varphi_{k,0} = 2\pi(f_c + k\Delta f)\tau_p. \tag{18}$$

Due to $\tau_p \gg 1/f_c + k\Delta f$,the period number of $(f_c + k\Delta f)$ in τ_p is fuzzy. But we can get the phase difference between the k-th and $(k+i)$-th subcarrier as Equation 19.

$$\Delta\varphi_{k,k+i} = \Delta\varphi_{k+i} - \Delta\varphi_k = 2\pi i \Delta f \tau_p, \tag{19}$$

where $i \in [0, K-1]$.

Equation (19) avoids the fuzzy of period number caused by f_c,and the $i\Delta f$ can be considered the pilot subcarriers of OFDM signal. Pilot subcarriers are discretely inserted into the subcarriers of OFDM. And the pilot is a kind of signals known to both transmitter and receiver. As shown in Figure 3(a), the distance differences of multipath channels in our model are too short than the half wavelength of 7MHz, which makes the phase differences $\Delta\varphi_{k,k+i}$ too small to be distinguished. Figures 3(b)–3(d) show that with $i\Delta f$ increasing, the results of TDOA estimation become better. The simulation results in Figure 3 reflect that only when the frequency interval is properly selected, the TDOA can be estimated. For better TDOA resolution, only partial pilot subcarriers are extracted from the reference antenna to estimate TDOA in our model. Setting the interval $\Delta f_G = i\Delta f$

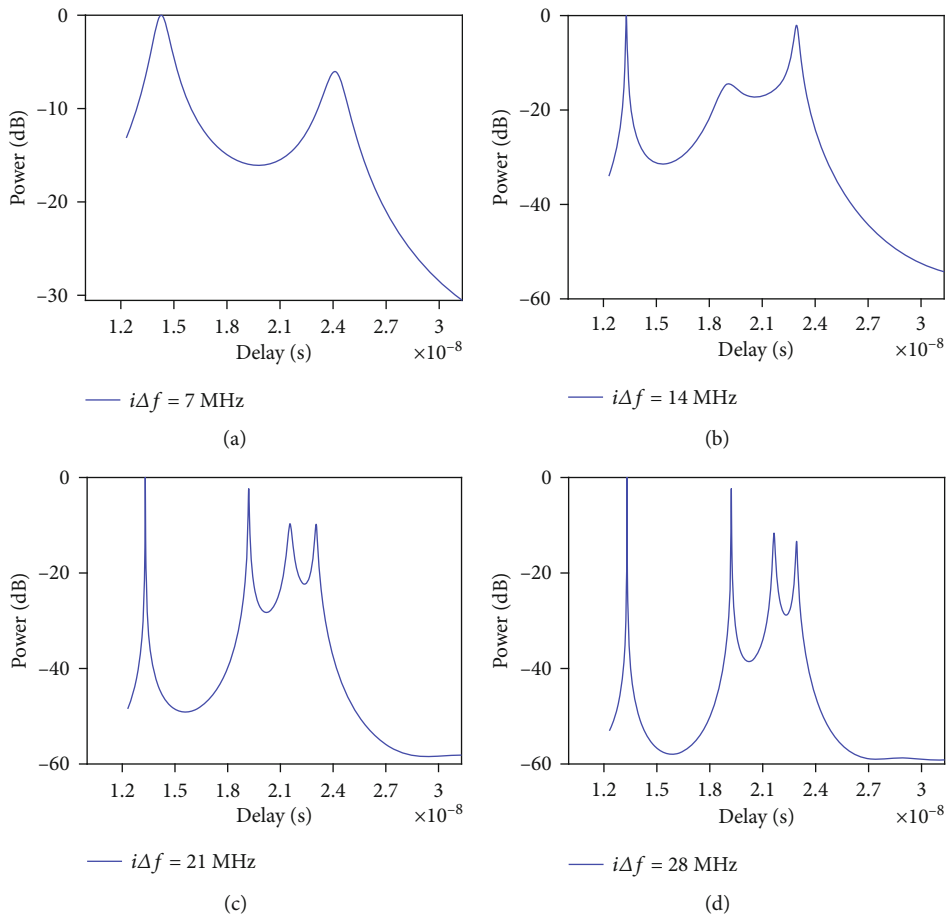


Figure 3: Relationship between frequency interval and delay resolution, where Delay(ns) = (13.31, 22.91, 19.22, 21.66), $SNR = 20$dB. (a) $i\Delta f = 7$MHz. (b) $i\Delta f = 14$MHz. (c) $i\Delta f = 21$MHz. (d) $i\Delta f = 28$MHz.

$= 21MHz$,and suppose there are $G\Delta f_G = 12\Delta f_G$ in the OFDM signal. The g-th pilot subcarriers signal can be expressed as Equation 20.

$$y_g(t) = \sum_{p=0}^{P-1} C_g(t) e^{j2\pi(f_c + g\Delta f_G)(t_{LOS} - \tau_p)} + n(t), \tag{20}$$

where $C_g(\mathrm{t})$ represents the data of pilot signal, some constant terms of this formula are ignored for simplifying the calculation.

Setting the pilot subcarriers be uniformly arranged in the frequency domain as shown in Figure 2(c), and the first frequency on the left side is the reference frequency, the right side is the direction of frequency increase. The pilot signals received by the reference antenna can be expressed as Equation (21).

$$Y_G = [\, y_1(t) \quad y_2(t) \quad \cdots \quad y_G(t) \,]' = LC_G + N_G, \tag{21}$$

where N_G is the noise matrix, C_G is the pilot data received by the reference antenna, and L is the steering matrix of TDOA as shown in Equation (22).

$$L = [\, l(\tau_0) \quad l(\tau_1) \quad \cdots \quad l(\tau_{P-1}) \,], \tag{22}$$

where $l(\tau_p)$ can be expressed as Equation (23).

$$l(\tau_p) = \begin{bmatrix} e^{-j2\pi(f_c + 1\cdot\Delta f_G)(t_{LOS} - \tau_p)} \\ e^{-j2\pi(f_c + 2\cdot\Delta f_G)(t_{LOS} - \tau_p)} \\ \vdots \\ e^{-j2\pi(f_c + G\cdot\Delta f_G)(t_{LOS} - \tau_p)} \end{bmatrix}. \tag{23}$$

Dividing G pilot subcarriers into W interleaved subarrays. Then, the MUSIC algorithm can be applied to the frequency domain using a similar way as AOA estimation.

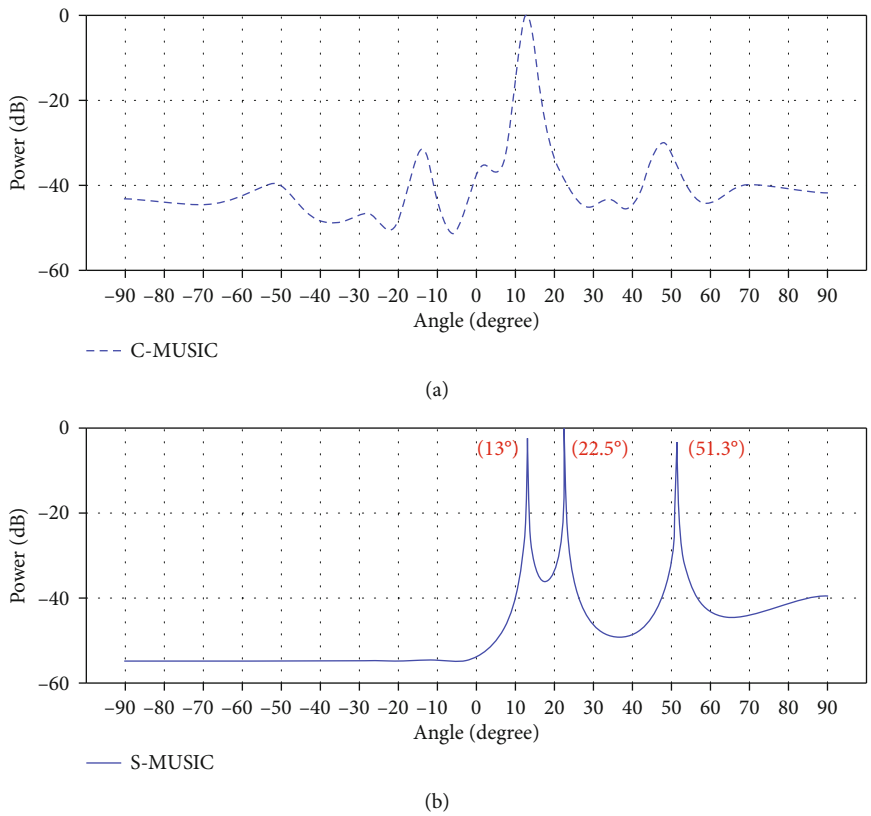


Figure 4: AOA estimation results of two algorithms. (a) AOA estimation results of classical MUSIC algorithm. (b) AOA estimation results of spatial smoothing MUSIC algorithm.

Taking the first subarray as the reference subarray, the signals of the w-th subarray can be expressed as Equation (24).

$$Y_{Gw} = \begin{bmatrix} y_w(t) & y_{w+1}(t) & \cdots & y_{w+g_s-1}(t) \end{bmatrix}' = LD_G^{(w-1)}C_G + N_{Gw}, \tag{24}$$

where $w = 1, 2, \cdots, W$, and $g_s = G - W + 1$ is the number of pilot subcarriers in each subarray, N_{Gw} represents the noise, the matrix $D_G^{(w-1)}$ is shown in Equation (25).

$$D_G^{(w-1)} = \begin{bmatrix} e^{-j2\pi\Delta f_G\tau_0} & 0 & \cdots & 0 \\ 0 & e^{-j2\pi\Delta f_G\tau_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-j2\pi\Delta f_G\tau_{P-1}} \end{bmatrix}^{(w-1)}. \tag{25}$$

4. Positioning Algorithms in Indoor Multipath Environment

4.1. Matching of TDOA and AOA. Although the method for jointly estimating time, azimuth and pitch angles can directly obtain the correspondence between the TDOA and AOA in 3D space, the computational complexity will increase exponentially with the number of estimated parameters. The TDOA and AOA are individually estimated by the proposed method in the previous section, but it needs to be matched before positioning. According to Equation (5), except for the reference antenna, the pilot signals on the m-th ($m \neq m_s$) array element also contain the information of t_{LOS}, τ_p, τ_m, where the t_{LOS} can be regarded as 0. τ_p has been estimated by the phase differences of pilot signals on reference array element. τ_m contains the angle information of the signals and the position information of the array element in the array antenna. Therefore, using the frequency domain smoothing MUSIC algorithm to process the pilot signals on the m-th array element can obtain the delay information of τ_p and τ_m. Based on this principle, we realize the matching of TDOA and AOA. The spectral formula of the m-th array element can be expressed as Equation 26.

$$P_{\text{MUSIC}} = \frac{1}{\boldsymbol{l}(\tau_p, \tau_m)^H \boldsymbol{V}_{Gn} \boldsymbol{V}_{Gn}^H \boldsymbol{l}(\tau_p, \tau_m)}, \tag{26}$$

where V_{Gn} is the noise feature vector, the expression of $l(\tau_p, \tau_m)$ is shown in Equation (27).

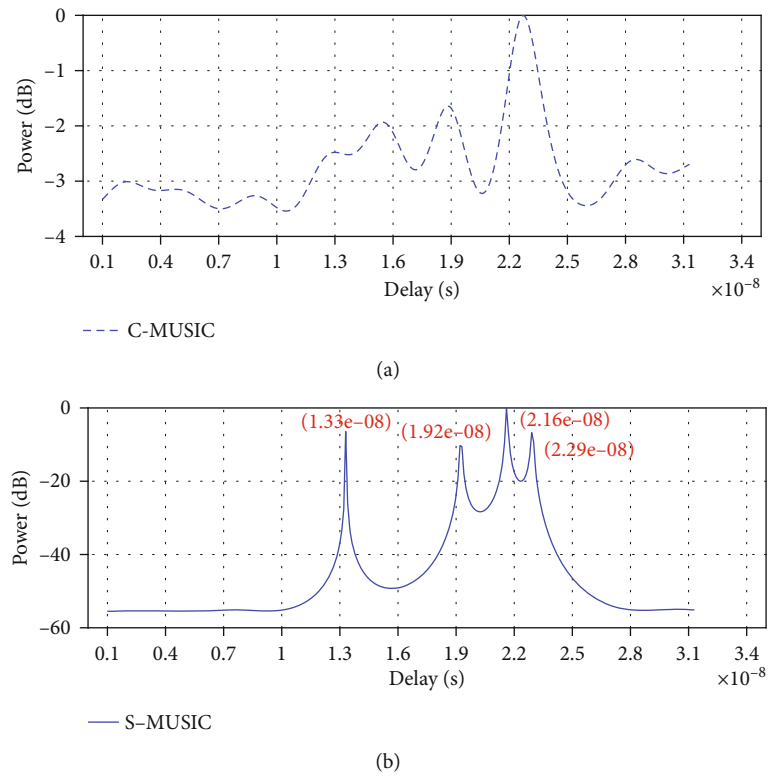


Figure 5: TDOA estimation results of two algorithms. (a) The classical MUSIC algorithm. (b) The frequency domain smoothing MUSIC algorithm.

$$\boldsymbol{l}_m(\tau_p, \tau_m) = \begin{bmatrix} e^{-j2\pi(f_c+1\cdot\Delta f_G)\left(t_{LOS}-\tau_p-\frac{(m-m_s)d\sin\theta_p}{c}\right)} \\ e^{-j2\pi(f_c+2\cdot\Delta f_G)\left(t_{LOS}-\tau_p-\frac{(m-m_s)d\sin\theta_p}{c}\right)} \\ \vdots \\ e^{-j2\pi(f_c+G\cdot\Delta f_G)\left(t_{LOS}-\tau_p-\frac{(m-m_s)d\sin\theta_p}{c}\right)} \end{bmatrix}, \quad (27)$$

where $m = (1, 2, \cdots, m_s - 1, m_s + 1, \cdots, M_x)$.

Combining all the estimated AOA and TDOA in pairs and bringing them into Equation 27. When there is a large solution to Equation (26), it means that AOA and TDOA come from the same multipath signal. The multipath delay spread is larger than the time difference among array elements. To make the matching algorithm more effective, we use the farthest array elements. Similarly, using the same method to match the multipath TDOA with the AOA estimated by the array antenna on y-axis. If the AOA estimated by the array antenna on the x-axis corresponds to the same TDOA as the AOA estimated by the array antenna on the y-axis, it means that they come from the same signal.

4.2. Calculation of Target Position. Selecting the earliest arriving signal as the LOS path signal through TDOA, and the others are NLOS path signals. Then, constructing the reflection paths of NLOS channels combined with indoor map, the location is obtained through the LOS path and the reflection paths. For the 3D spatial coordinate system as shown in Figure 1, the AOA estimated in previous content is the angles between the signal propagation directions and the normal of coordinate axes. Equation (28) can convert (θ_p, μ_p) to $(\alpha_p, \beta_p, \sigma_p)$.

$$\begin{cases} \alpha_p = \dfrac{\pi}{2} - \theta_p \\ \beta_p = \dfrac{\pi}{2} - \mu_p \\ \sigma_p = \pi - \sin^{-1}\sqrt{\left(\cos^2(\alpha_p) + \cos^2(\beta_p)\right.} \end{cases}, \quad (28)$$

where $\alpha_p, \beta_p, \sigma_p$ represent the angle between the p-th multipath signal propagation direction and the positive direction of x-axis, y-axis, and z-axis, respectively, and $\alpha_p, \beta_p \in [0, \pi]$, $\sigma_p \in [(\pi/2), \pi)$ in the model of Figure 1.

The direction vector of the p-th multipath signal received by the array antenna can be expressed as Equation (29).

$$\vec{n}_p = \left(\cos(\alpha_p), \cos(\beta_p), \cos(\sigma_p)\right). \quad (29)$$

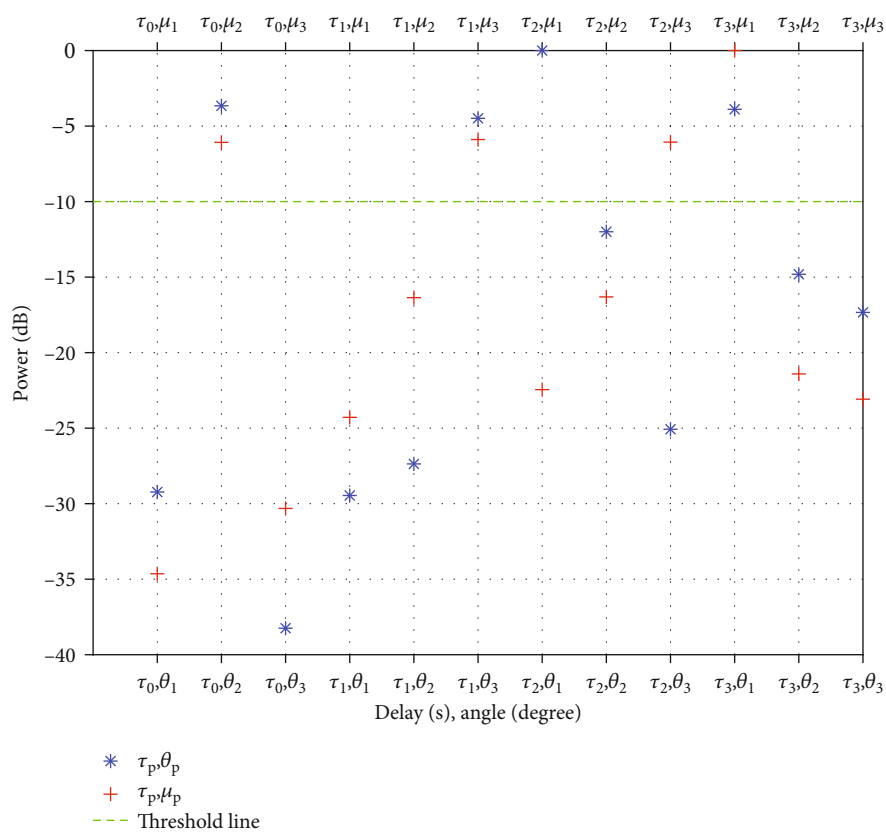


Figure 6: Simulation results of matching algorithms. The point above the dotted line means successfully matched.

And the line equations where $\vec{n}_p$ located can be expressed as Equation (30).

$$SL_p : \begin{cases} x_p = x_0 + \gamma_p \cos\left(\alpha_p\right) \\ y_p = y_0 + \gamma_p \cos\left(\beta_p\right), \\ z_p = z_0 + \gamma_p \cos\left(\sigma_p\right) \end{cases} \tag{30}$$

where (x_0, y_0, z_0) is the coordinate of one known point on the line SL_p, γ_p is the scale factor between the point on the line and the point (x_0, y_0, z_0). For the multipath signals received by the array antenna, (x_0, y_0, z_0) is the coordinate of reference antenna.

The NLOS paths are generated by the reflection of the walls, so the path lines will intersect the plane of the walls. Bring the plane equation parameters shown in Table 1 into Equation (30) to obtain the coordinates of all intersection points. Since the coordinates of the reflection points must meet the conditions of indoor space, that is $x_p \in [-3, 3]$, $y_p \in [-4, 4]$, $z_p \in [-4.5, 0]$,we can exclude those intersections located outside the indoor space. Setting the coordinates of the reflection point of the p-th path in the indoor space is $R_p(x_p, y_p, z_p)$,and the angles between the line where the path of emission signal is located and the positive direction of the 3 axes are $(\alpha_{rp}, \beta_{rp}, \sigma_{rp})$. According to the law of reflection for electromagnetic waves whose incident angle is equal to the reflection angle. If the plane is parallel to the x-axis, the relationship between α_{rp} and α_p is $\alpha_{rp} = -\alpha_p$,and if the plane is perpendicular to the x-axis, the relationship between α_{rp} and α_p is $\alpha_{rp} = \pi - \alpha_p$. The same principle can be applied to β_{rp} and σ_{rp}. Bring $R_p(x_p, y_p, z_p)$ and $(\alpha_{rp}, \beta_{rp}, \sigma_{rp})$ into Equation 30 to obtain the line equations of the NLOS paths before reflection, and we denote these equations of lines as RL_{rp},where $rp = 1, 2, \cdots, P-1$.

Theoretically, the intersection of line SL_0 and lines RL_{rp} is the target location. A system of line equations for the P paths can be constructed as Equation (31).

$$\begin{cases} SL_0 : (x, y, z) = o(x_0, y_0, z_0) + \gamma_0 \vec{n}_0 \\ RL_{rp} : (x, y, z) = R_p\left(x_p, y_p, z_p\right) + \gamma_{rp} \overrightarrow{nr}_{rp} \end{cases}, \tag{31}$$

where $\overrightarrow{nr}_{rp} = (\cos(\alpha_{rp}), \cos(\beta_{rp}), \cos(\sigma_{rp}))$ are the direction vectors of line equations RL_{rp},and $rp = 1, 2, \cdots, P-1$.

Due to the noise and the measurement errors, these lines may not intersect in 3D space. It means that Equation (31) may not have a common solution. In our method, the point in space with the closest distance to SL_0 and all RL_{rp} is

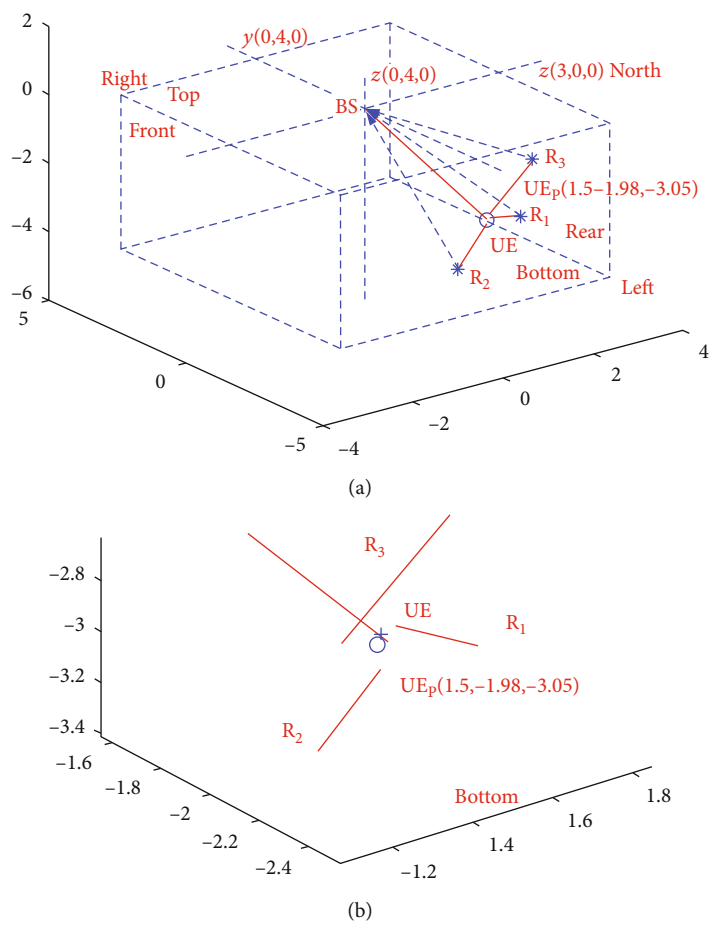


Figure 7: Simulation of positioning algorithm. (a) Location calculation of target. (b) The details of positioning.

regarded as the target location. The parameters in Equation 31 need to be solved can be defined as Equation (32).

$$\boldsymbol{P}_s = \begin{bmatrix} x_{UE} & y_{UE} & z_{UE} & \gamma_0 & \gamma_1 & \cdots & \gamma_{rp} & \cdots & \gamma_{P-1} \end{bmatrix}', \tag{32}$$

where (x_{UE}, y_{UE}, z_{UE}) are the estimated location coordinates of the target.

The direction vector parameters in Equation (31) can be defined as Equation (33).

$$P_n = \begin{bmatrix} I & \left(-\vec{n}_0\right)' & 0 & \cdots & 0 \\ I & 0 & \left(-\vec{nr}_1\right)' & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I & 0 & 0 & \cdots & \left(-\vec{nr}_{P-1}\right)' \end{bmatrix}, \tag{33}$$

where I is a 3×3 dimensional unit vector, 0 is a 3×1 dimensional zero vector.

The coordinate parameters of the reference antenna and reflection points are defined as Equation (34).

$$P_R = \begin{bmatrix} x_0 & y_0 & z_0 & x_1 & y_1 & z_1 & \cdots & x_{P-1} & y_{P-1} & z_{P-1} \end{bmatrix}'. \tag{34}$$

According to equations (32)–(34), Equation (31) can be expressed as the matrix form: $P_n P_s = P_R$. And the P_s can be solved through the method of least squares as Equation (35).

$$P_s = \left(P_n{}'^{P}_{n}\right)^{-1} P_n{}' P_R. \tag{35}$$

Since the multipath signals are transmitted from one point, the lines SL_0 and RL_{rp} will not be parallel to each other in space, so Equation (36) must have a solution and the first three elements of $\boldsymbol{P}_s$ are regarded as the target position coordinates.

5. Simulation and Analysis

In this section, we will build a complete indoor positioning simulation based on the model in Figure 1. Firstly, we analyze the advantages of using smoothing algorithms to process multipath coherent signals. Figures 4(a) and 4(b) show

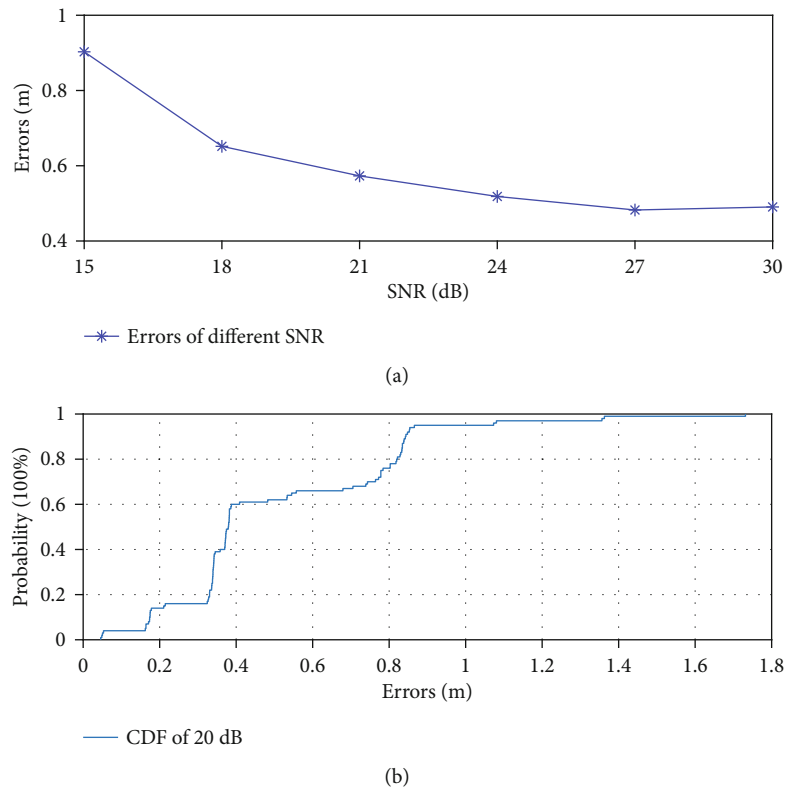


Figure 8: Simulation of positioning error. (a) Influence of SNR on positioning errors, where $N_s = 64$, $T_p = 50$. (b) CDF curve of positioning errors, where $N_s = 64$, $T_p = 100$, $SNR = 20$dB.

the AOA estimation results of classical MUSIC algorithm and spatial smoothing MUSIC algorithm, respectively. It can be seen that the spatial smoothing MUSIC algorithm can estimate the AOA but the classical MUSIC algorithm is not available. An array antenna with M_x elements can estimate up to $2M_x/3$ multipath signals by using the spatial smoothing MUSIC algorithm. That means 16 antennas can estimate 10 multipath signals at the same time which can fully meet the needs of indoor environment. Using the same method can estimate angles $\mu_0, \mu_1, \cdots, \mu_{P-1}$ between the signal propagation directions and the normal of y-axis. The model has 4 multipath signals, but it can be seen from Figure 4(b) , the ULA on the x-axis can only estimate the AOA of 3 paths. That is because the AOA of the two multipath paths is too close to be distinguished by the array antenna. The same happens with the ULA located on the y-axis. But this does not affect the correctness of our algorithm, we will restore the missing angles through the matching algorithm.

Figures 5(a) and 5(b) show the TDOA estimation results of classical MUSIC algorithm and frequency domain smoothing MUSIC algorithm, respectively. It can be seen that the TDOA of coherent signals can be estimated by frequency domain smoothing MUSIC algorithm.

Since the estimated number of AOA is less than the actual number, the matching algorithm needs to use one of the AOA twice to recover the missing AOA information. The proposed matching algorithm makes up for this shortcoming by combining all angles and delays. For better resolution, the first array element which is farthest from the reference array element is selected for matching estimation. Figure 6 shows the simulation results of the proposed matching algorithm, where the AOA of the x-axis and y-axis is matched with the TDOA, respectively. "$*$" corresponds to the axis label at the bottom of the graph indicating the matching result between the TDOA and AOA of the x-axis. "+" corresponds to the axis label at the top of the graph indicating the matching result between the TDOA and AOA of the y-axis. Both the four larger values of "$*$" and "+" above the dotted line in Figure 6 are chosen as the AOA and TDOA for the successful matching. It can be seen that each TDOA corresponds to AOA in the direction of x-axis and y-axis, respectively, where $(\tau_0, \theta_2, \mu_2)$ is the LOS path, $(\tau_1, \theta_3, \mu_3), (\tau_2, \theta_1, \mu_3), (\tau_3, \theta_1, \mu_1)$ are the 3 NLOS paths. The angles $\theta_1 = 12.9^\circ$ and $\mu_3 = -18.8^\circ$ are used twice, respectively. In the actual environment, the corresponding angles of θ_1 are 12.6°, 13.34°, and the corresponding angles of μ_3 are -20.094° and -17.92°. They are very close and considered to be the same value by the array antenna, respectively.

There are two types of AOA and one type of TDOA in our model that need to be estimated. Assuming that the angle range is [0°, 180°] and the resolution is 0.1°, the time delay range is [1, 30]ns and the resolution is 0.1ns. Then,

traditional joint estimation MUSIC algorithm needs to calculate the spectral formula of MUSIC algorithm for $1800 \times 1800 \times 300 = 9.72 \times 10^6$ times in total. We used the 3 times MUSIC algorithm and the twice matching algorithm to obtain these three parameters, respectively, which needs to calculate the spectral formula of MUSIC algorithm for $2 \times 1800 + 300 + 2 \times 12 = 3924$ times in total. Compared with the existing joint estimation algorithm, the proposed algorithm greatly reduces the computational complexity.

After obtaining the AOA and TDOA of the multipath signals in 3D space, we verified the target position calculation algorithm in Section 4.2 through simulations. As shown in Figure 7(a) , $\mathrm{UE}(1.5,-2,-3)$ is the actual target location (represented by "+"). The thicker dotted lines represent the AOA of the multipath signals received by the array antenna. R_1, R_2, R_3 represent the reflection points of NLOS paths, respectively. The solid lines are the path used for positioning. $\mathrm{UE}_P(1.5,-1.93,-3.05)$ is the estimated target coordinate (represented by "○"). It can be seen from Figure 7(b), although the positioning paths are far from the actual location of the target and do not intersect in 3D space, the closest point to all of these positioning paths does not deviate much from the actual location of UE.

Unlike most existing algorithms which treat multipath signals as a kind of interference or require multiple BS to coordinate positioning, the proposed method makes full use of multipath signals and only needs one BS to achieve 3D indoor positioning. In addition, our algorithms do not require precise clock synchronization of transceivers, and it is more accurate to judge the LOS channel by TDOA than by signal strength.

In order to evaluate the influence of noise, we simulate the positioning error under a different Signal-to-Noise Ratio (SNR). The positioning error is defined as Equation (36).

$$d_{\text{error}} = \frac{1}{T_p} \sum_{i=1}^{T_p} \sqrt{(x_i - x_{UE})^2 + (y_i - y_{UE})^2 + (z_i - z_{UE})^2}, \tag{36}$$

where T_p is the number of times for positioning, (x_i, y_i, z_i) is the target coordinates estimated by the proposed method, and (x_{UE}, y_{UE}, z_{UE}) is the actual coordinates of the target.

Figure 8(a) shows the results of the evaluation about the influence of SNR. In the simulation experiments, the number of snapshots is set as 64. To avoid the influence of accidental factors, the positioning error at each SNR is the average value of 50 positioning errors, that is, $\mathrm{TP} = 50$. It can be seen that when the SNR is set to 24dB,the improvement of positioning errors will be very small. That is because when the SNR is improved, the errors caused by SNR will be very small, but the errors caused by resolution of the array antenna will become the main factor restricting the positioning accuracy. That is in the process of AOA estimation, the two angles with very small difference are regarded as the same AOA.

In order to evaluate the positioning errors of the proposed method, Figure 8(b) shows the Cumulative Distribution Function (CDF) curve of errors, where the simulation conditions are set as $N_s = 64$, $\mathrm{SNR} = 24\mathrm{dB}$, $T_p = 100$. It can be seen that in 95% of the cases, the positioning accuracy is within submeter level, and within 0.4m in 60% of the cases. Note that the CDF curve is approximately stepped, the reason for this is the amplification effect of the reflection paths on the errors. For example, when the AOA estimated by the ULA at one of coordinate axes has errors, that will cause this multipath signal to have AOA errors on the other coordinate axes, too. And these errors will be increased again on the NLOS paths after reflection, and the longer the multipath paths, the more obvious the increase of these errors; finally, all these errors constitute the positioning errors.

6. Conclusions

The purpose of this research is to realize the 3D positioning with a single BS in complex indoor space. Signal propagation in indoor environment with serious multipath effects and those multipath signals are coherent, which make indoor positioning very difficult. Combining MIMO technology and OFDM technology in wireless communication system, we use the array antenna to estimate the AOA of the multipath signals and apply the spatial smoothing algorithm to the frequency domain to realize the TDOA estimation of multiple coherent signals. Then, we proposed a matching method of TODA and AOA, which reduced the computational complexity. After obtaining the TODA and AOA of the multipath signals, the LOS channel and NLOS channels are distinguished by TODA, and then, the AOA are combined with the indoor map to establish the reflection paths. The target location in 3D indoor space is obtained by combining the line equation where the reflection paths and the LOS path are located in 3D coordinate system. Finally, the proposed algorithm is verified by simulations. The results show the correctness of our method, and the localization accuracy is kept in the submeter level in 95% of the cases, and within 0.4m in 60% of the cases.

Most of existing indoor positioning methods are based on 2D models, the algorithms require at least two or more BS for collaborative positioning, some methods try to avoid the effects of multipath signals. However, from the perspective of cost and application, single BS positioning is one of the most potential solutions for indoor positioning in the future. The method proposed in this paper not only utilizes multipath signals to realize the positioning of a single BS but also extends the indoor positioning algorithm to 3D space. That makes our positioning method have greater application prospects.

Due to the resolution of ULA is not fine enough, the two AOA with very small difference are regarded as the same one, which will negatively affect the positioning results. This problem also occurs in other AOA estimation algorithms. How to avoid the influence of these errors remains to be further studied.

Data Availability

The simulation data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no competing interests.

References

[1] Z. Wang, X. Li, X. Zhang, Y. Bai, and C. Zheng, "An attitude estimation method based on monocular vision and inertial sensor fusion for indoor navigation," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 27051–27061, 2021.

[2] E. Aparicio-Esteve, Á. Hernández, and J. Ureña, "Design, calibration, and evaluation of a long-range 3-D infrared positioning system based on encoding techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, article 7005113, pp. 1–13, 2021.

[3] R. Carotenuto, M. Merenda, D. Iero, and F. G. Della Corte, "An indoor ultrasonic system for autonomous 3-D positioning," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 7, pp. 2507–2518, 2019.

[4] Y. Ma, B. Wang, S. Pei, Y. Zhang, S. Zhang, and J. Yu, "An indoor localization method based on AOA and PDOA using virtual stations in multipath and NLOS environments for passive UHF RFID," *IEEE Access*, vol. 6, pp. 31772–31782, 2018.

[5] R. Faragher and R. Harle, "Location fingerprinting with Bluetooth low energy beacons," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2418–2428, 2015.

[6] C. Yang and H. Shao, "WiFi-based indoor positioning," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 150–157, 2015.

[7] D. Feng, C. Wang, C. He, Y. Zhuang, and X. -G. Xia, "Kalman-filter-based integration of IMU and UWB for high-accuracy indoor positioning and navigation," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3133–3146, 2020.

[8] Y. Zhuang, L. Hua, L. Qi et al., "A survey of positioning systems using visible LED lights," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1963–1988, 2018.

[9] A. Blanco, N. Ludant, P. J. Mateo, Z. Shi, Y. Wang, and J. Widmer, "Performance evaluation of single base station ToA-AoA localization in an LTE testbed," in *2019 IEEE 30th annual international symposium on personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, Istanbul, Turkey, 2019.

[10] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[11] A. Zaidi, F. Athley, J. Medbo, U. Gustavsson, G. Durisi, and X. Chen, *5G Physical Layer Principles, Models and Technology Components*, Elsevier, 2019.

[12] "Directory Listing /ftp/Specs/2019-03/Rel-16/38_series, 3GPP," 2019, https://www.3gpp.org/ftp/Specs/2021-03/Rel-16/38_series.

[13] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2475–2487, 2017.

[14] X. Zhou, L. Chen, J. Yan, and R. Chen, "Accurate DOA estimation with adjacent angle power difference for indoor localization," *IEEE Access*, vol. 8, pp. 44702–44713, 2020.

[15] J. He and H. C. So, "A hybrid TDOA-fingerprinting-based localization system for LTE network," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13653–13665, 2020.

[16] X. Yang, "NLOS mitigation for UWB localization based on sparse pseudo-input Gaussian process," *IEEE Sensors Journal*, vol. 18, no. 10, pp. 4311–4316, 2018.

[17] R. Mendrzik, H. Wymeersch, G. Bauch, and Z. Abu-Shaban, "Harnessing NLOS components for position and orientation estimation in 5G millimeter wave MIMO," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 93–107, 2019.

[18] M. Si, Y. Wang, C. K. Seow, H. Cao, H. Liu, and L. Huang, "An adaptive weighted Wi-Fi FTM-based positioning method in an NLOS environment," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 472–480, 2022.

[19] F. Wen and C. Liang, "Fine-grained indoor localization using single access point with multiple antennas," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1538–1544, 2015.

[20] R. Feng, Y. Liu, J. Huang, J. Sun, and C. -X. Wang, "Comparison of music, unitary ESPRIT, and SAGE algorithms for estimating 3D angles in wireless channels," in *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–6, Qingdao, China, 2017.

[21] M. Sánchez-Fernández, V. Jamali, J. Llorca, and A. M. Tulino, "Gridless multidimensional angle-of-arrival estimation for arbitrary 3D antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4748–4764, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2638461, 13 pages
https://doi.org/10.1155/2022/2638461

Research Article

Joint Decoding Order and Power Allocation Design for a NOMA-Based Overlay Cognitive Integrated Satellite-Terrestrial Relay Network

Rui Liu,[1] **Kefeng Guo**,[1,2] **Kang An**,[3] **Shibing Zhu**,[1] **and Haifeng Shuai**[1]

[1]*School of Space Information, Space Engineering University, Beijing 101407, China*
[2]*College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*
[3]*Sixty-Third Research Institute, National University of Defense Technology, Nanjing 210007, China*

Correspondence should be addressed to Kefeng Guo; guokefeng.cool@163.com

Received 27 February 2022; Revised 15 March 2022; Accepted 22 March 2022; Published 1 April 2022

Academic Editor: Han Wang

Due to the shortage of spectrum resources, a nonorthogonal multiple access- (NOMA-) based overlay cognitive integrated satellite-terrestrial relay network is established in this paper, where cognitive radio (CR) and NOMA scheme are utilized to enhance spectral efficiency. In this system, a secondary terrestrial transmitter assists the transmission of primary signal and transmits its own signal simultaneously. To maximize the achievable rate of a secondary network on the premise of ensuring the quality of service of primary network, the decoding order and power allocation are jointly optimized. Particularly, we first derive the optimal power allocation factor for a given decoding order. After searching all different decoding orders, the globally optimal joint design of the decoding order and power allocation is obtained. To gain deep insights, the exact and asymptotic expressions for outage probability of both primary and secondary networks are derived. Numerical results demonstrate the accuracy of our derivation and indicate the impacts of key parameters on the proposed system.

1. Introduction

Although the standard of the next-generation wireless network has not been unified, the introduction of satellite communication (SatCom) into the existing wireless communication network has become the consensus of experts and engineers [1, 2], which has the capabilities of all-weather operation, long transmission distance, and seamless coverage. In addition, the integrated satellite-terrestrial relay networks (ISTRNs) are becoming the promising architecture of the next-generation communication network and the Internet of things (IoT), which can offer comprehensive and convenient services [3, 4]. With the rapid development of ISTRN, the establishment of large-scale networks and the exponential growth of the number for terminals will cause serious shortage of spectrum resources. To solve the above problem, the application of cognitive radio (CR) and nonorthogonal multiple access (NOMA) into ISTRN have been regarded as effective approaches to improve spectrum utilization and deal with large-scale users [5, 6]. In general, CR enables unauthorized users to reasonably utilize the authorized spectrum [7, 8]. On the other hand, the NOMA technique allows different users to multiplex the same time/code/frequency [9, 10].

1.1. Related Works. Over recent years, the exploration of ISTRN has been investigated in various research areas. Based on fractional programming, a power allocation optimization scheme for downlink land mobile satellite (LMS) was proposed in [11], and the optimal average energy efficiency (EE) was derived. The authors of [12] investigated the performance of the LMS system, which considered the inherent hardware impairments (HIs) and jamming signal. In [13], the authors discussed the physical layer security of ISTRN, in which the secrecy outage probability (SOP) and average security capacity were derived under multiple illegal eavesdroppers and a threshold-based user scheduling scheme. A variety of methods to enhance the physical layer security

and reliability of ISTRN were discussed in [14], and a kind of effective achievable rate and a new beamforming (BF) scheme were developed. The authors of [15] carried out joint BF and optimization for reconfigurable intelligent surface-(RIS-) aided ISTRN.

To deal with the problem of spectrum shortage caused by increasingly scarce spectrum resources, many existing works have considered the combination of CR and ISTRN. The authors of [16] investigated the cognitive ISTRN with multiple secondary networks, and the outage probability (OP) of the primary network was minimized by the proposed secondary network selection scheme. In [17], the authors proposed a secure BF scheme to deal with the situation of multiple eavesdroppers in rate-splitting multiple access-based cognitive ISTRN. In the case of full channel state information (CSI) and statistical CSI, the authors proposed two BF schemes to maximize the instantaneous rate of terrestrial users in cognitive ISTRN by using hybrid zero forcing and partial zero forcing methods in [18]. The authors of [19] utilized terrestrial base stations and cooperative terminals to enhance the secrecy performance of the cognitive ISTRN system. On this basis, a joint artificial noise produce and cooperative interference BF scheme was proposed.

For the sake of solving the service request of ever-increasing large-scale users, NOMA is considered a significant scheme in ISTRN [20]. Joint influence of channel estimation error and HIs on the security performance of cognitive integrated satellite-terrestrial relay networks (CISTRN) was investigated in [21]. The authors of [22] investigated the security performance of ISTRN with the NOMA scheme, where the HIs and multiple relays were considered and the SOP of the system was derived. In [23], the performance of ISTRN which considered cache and NOMA was studied, the OP and hit probability of the considered system were derived based on stochastic geometry. The performance of NOMA-based ISTRN with content delivery was discussed in [24], where the transmission delay can be reduced.

1.2. Motivations. The system performance of the NOMA-based system can be enhanced by designing the decoding order or power allocation scheme [25–27]. The authors of [25] minimized the overall OP under the cases of different decoding orders. In [26], to maximize the achievable rate of the secondary network with energy harvesting, the power allocation and time portion were optimized. In [27], the fair throughput for the NOMA-based system and orthogonal multiple access- (OMA-) based system was maximized. However, the above works cannot be simply utilized to optimize the performance of a priority-based NOMA system. In the traditional NOMA system, power is always allocated based on channel conditions. However, the priority-based NOMA system should guarantee the quality of service (QoS) of users with high priority. In fact, the QoS of high-priority users can be guaranteed and the system performance can be improved by optimizing the power allocation. On this basis, the optimization of the decoding order can make the system achieve optimal performance. A few existing works have investigated the joint optimization of the decoding order and power allocation [28–30], but the QoS of high-priority users was not strictly considered. Moreover, the authors of [31] considered the QoS of high-priority users, but the method was carried out in terrestrial networks, and there is no similar work in ISTRN.

Motivated by the above discussion, a NOMA-based overlay CISTRN is considered. As far as we know, no prior work has discussed the NOMA-based overlay CISTRN with joint optimization of the decoding order and power allocation.

1.3. Our Contributions. The main contributions of this paper can be summarized as follows:

(1) Firstly, an overlay CISTRN with the NOMA scheme is established, in which the secondary network assists the transmission of the primary network in exchange for the opportunity to utilize the authorized spectrum of the primary network to transmit its own signals simultaneously. Besides, the satellite signal cannot reach the primary user and secondary user due to some severe shadowing

(2) Secondly, to improve the achievable performance of the secondary network, the decoding order and power allocation are optimized with the aim of maximizing the achievable rate of the secondary network under the predefined constraints of the primary satellite network. Specifically, the closed-form expressions for the optimal decoding order and power allocation factors are derived, which is of great significance for the study of NOMA-based CISTRN

(3) Finally, basing on the joint optimal design of the decoding order and power allocation, the exact OPs of both the primary network and secondary network are obtained. To get further insights, the asymptotic expressions and diversity order (DO) of the two networks are also derived. In addition, the numerical results validate our theoretical analysis. On this foundation, the impacts of many key parameters on the system performance are revealed

1.4. Paper Organization. The structure of the rest of this paper is as follows. Section 2 provides the system model and problem formulation. The joint decoding order and power allocation design are given in Section 3. The performance of the proposed system is discussed in Section 5. Section 4 gives the numerical results and reveals the impacts of key parameters on system performance. Finally, the full text summary is given in Section 6.

Notations. $CN(a, b)$ denotes the Gaussian distribution with mean a and variance b. $E(\cdot)$ is the expectation operator, and $\Gamma(\cdot)$ is the gamma function (Eq. 8.339.1 in [32]). The probability density function (PDF) and the cumulative distribution function (CDF) of variable z are, respectively, represented by $f_z(.)$ and $F_z(.)$. $(.)_n$ is the Pochhammer symbol

(p. xliii in [32]). $K_v(\cdot)$ stands for the second kind modified Bessel function (Eq. 8.432.6 in [32]).

2. System Model and Problem Formulation

As shown in Figure 1, a NOMA-based overlay CISTRN is considered in our paper, which consists of a primary satellite source (S) and a secondary terrestrial transmitter (R) with their corresponding receivers (U and D). It is assumed that there is no direct satellite (DS) link between S and U due to severe shadowing effects, such as obstacles. Therefore, the secondary network is asked for assisting the transmission from S to U as a relay [6]. As a reward, R utilizes the authorized spectrum of S to communicate with D. The NOMA scheme is adopted in R to transmit the expected signals of U and D simultaneously, and the two-user pair is considered in our paper to balance the complexity and effectiveness. Each node works in half-duplex mode and applies an omni-directional antenna to reduce the complexity of the system [16], and the decode-and-forward (DF) protocol is employed by R. Specifically, the channel coefficients of S-R link, R-U link, and R-D link are denoted by h_{sr}, h_{ru}, and h_{rd}, respectively. Moreover, the satellite-terrestrial channel experiences Shadowed-Rician (SR) fading, and the terrestrial channels undergo Nakagami-m fading. Without loss of generality, it is assumed that the channel coefficients are variable but remain constant in a time slot (furthermore, the perfect hardware impairments and decoding errors are assumed in our proposed system to simplify the analysis. Besides, the perfect CSI can be obtained by feedback and training from users). In addition, all receivers sustain additive white Gaussian noise (AWGN) with $CN(0, \sigma^2)$.

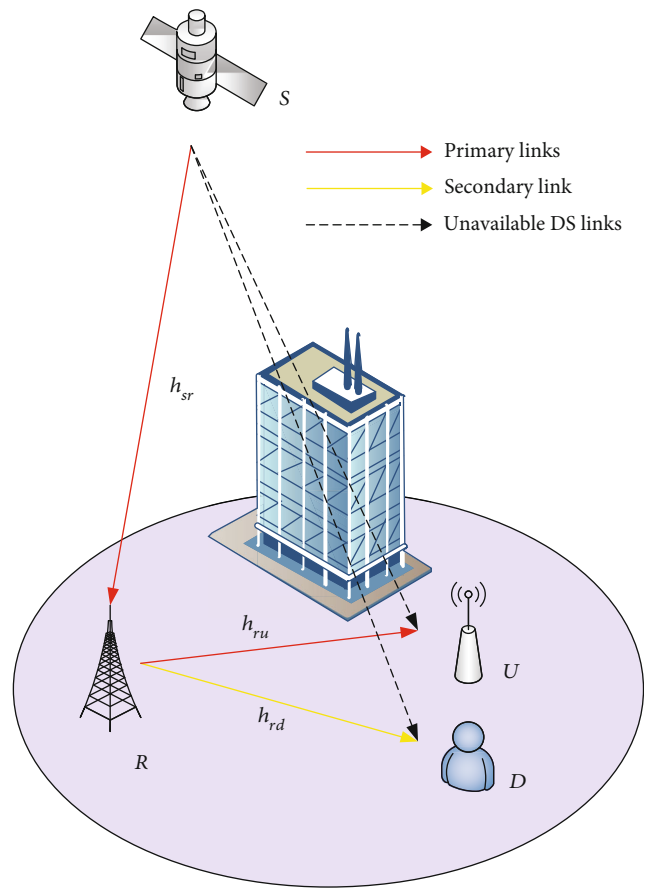


Figure 1: System model.

2.1. Problem Formulation. There are two phases in the entire transmission process, which is shown in Figure 2.

During the first phase, S transmits its signal to R, and the signal received by R is represented by

$$y_{sr} = \sqrt{P_s} h_{sr} x_s + n_r, \tag{1}$$

where P_s is the transmit power of S, x_s denotes the satellite message with $E(|x_s|^2) = 1$, and n_r is AWGN. From (1), the achievable rate of x_s at R can be expressed as

$$R_r = \frac{1}{2} \log_2(1 + \gamma_{sr}), \tag{2}$$

where $\gamma_{sr} = \bar{\gamma}_s |h_{sr}|^2$ and $\bar{\gamma}_s = P_s/\sigma^2$ is the signal-to-noise ratio (SNR) of the satellite. The target rate of the primary network is denoted as R_{th}^u. To decode the primary signal successfully, the following constraint must be satisfied, namely, $R_r \geq R_{th}^u$.

In the second phase, R integrated primary and its own signal with the superposition coding technique (SCT), and the signal received by U and D can be expressed by

$$y_{ra} = h_{ra}\left(\sqrt{P_r \beta} x_s + \sqrt{P_r(1-\beta)} x_r\right) + n_a, \tag{3}$$

where $a = \{r, d\}$, P_r is the transmit power of R, x_r denotes the expected signal of D with $E(|x_r|^2) = 1$, β is the power allocation factor with $0 < \beta \leq 1$, and n_a represents AWGN.

At the receiving ends, there are two decoding orders. The first is that U decodes the desired signal directly, and D performs successive interference cancellation (SIC). The second is that D detects its expected signal, while U adopts SIC. Besides, the achievable rates of x_s at U and x_r at D under the conditions of two decoding orders can be given as follows, respectively.

2.1.1. The First Decoding Order. In this order, U decodes x_s while x_r is regarded as the serial interference. Then, the achievable rate of x_s at U can be expressed as

$$R_u^1 = \frac{1}{2} \log_2\left(1 + \frac{\beta \gamma_{ru}}{(1-\beta)\gamma_{ru} + 1}\right), \tag{4}$$

where $\gamma_{ru} = \bar{\gamma}_r |h_{ru}|^2$, $\bar{\gamma}_r = P_r/\sigma^2$ is the SNR of the secondary transmitter, and the tag "1" indicates the first decoding order.

Recalling to the properties of the NOMA scheme, SIC is adopted at D; thus, x_s is decoded by D with x_r as the cochannel interference (CCI); then, the decoded message is removed from y_{rd}. Finally, D decodes x_r from the remaining signal. Hence, the achievable rates of x_s and x_r at D can be,

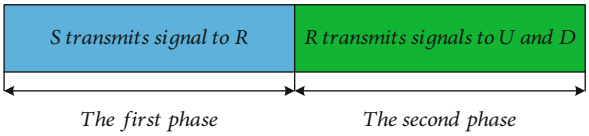


Figure 2: The slotted structure of the system.

Table 1: Simulation parameters.

Parameter name	Parameter value
Satellite	GEO
Carrier frequency	2 GHz
3 dB angle	0.4°
Maximal satellite gain	48 dB
Maximal relay gain	4 dB
Carrier bandwidth	15 MHz
Noise temperature	300°
Variance of AWGN	1
Frequent heavy shadowing (FHS)	(1,0.063,0.0007)
Average shadowing (AS)	(5,0251,0.279)
Infrequent light shadowing (ILS)	(10,0.158,1.29)

respectively, represented as

$$R_{u\longrightarrow d}^{1}=\frac{1}{2}\log_2\left(1+\frac{\beta\gamma_{rd}}{(1-\beta)\gamma_{rd}+1}\right),\quad R_{d}^{1}=\frac{1}{2}\log_2(1+(1-\beta)\gamma_{rd}),\tag{5}$$

where $\gamma_{rd}=\bar{\gamma}_r|h_{rd}|^2$.

2.1.2. The Second Decoding Order. In the second decoding order, SIC is applied at U, and x_s is treated as the serial interference at both U and D firstly. Therefore, the achievable rates of x_r at U and D are expressed as, respectively,

$$R_{d\longrightarrow u}^{2}=\frac{1}{2}\log_2\left(1+\frac{(1-\beta)\gamma_{ru}}{\beta\gamma_{ru}+1}\right),\tag{6}$$

$$R_{d}^{2}=\frac{1}{2}\log_2\left(1+\frac{(1-\beta)\gamma_{rd}}{\beta\gamma_{rd}+1}\right),\tag{7}$$

where the tag "2" indicates the second decoding order.

If x_r can be decoded by U successfully, the achievable rate of x_s at U can be represented as

$$R_{u}^{2}=\frac{1}{2}\log_2(1+\beta\gamma_{ru}).\tag{8}$$

2.2. Channel Model. In this section, the statistical characteristics of SR fading and Nakagami-m fading are provided. Firstly, the PDF of h_{sr} is given by [33]

$$f_{|h_{sr}|^2}(x)=\alpha_{sr}e^{-\beta_{sr}x}{}_1F_1(m_{sr};1;\delta_{sr}x),\tag{9}$$

where $\alpha_{sr}\Delta=1/2b_{sr}(2b_{sr}m_{sr}/(2b_{sr}m_{sr}+\Omega_{sr}))^{m_{sr}}$, $\beta_{sr}\Delta=1/2b_{sr}$, and $\delta_{sr}\Delta=\Omega_{sr}/2b_{sr}(2b_{sr}m_{sr}+\Omega_{sr})$. $m_{sr}>0$ denotes the Nakagami-m parameter, and $2b_{sr}$ and Ω_{sr} are the average power of the multipath component and that of the line-of-sight (LOS) component, respectively.

When m_{sr} is an integer, with the help of Eq. 07.20.03.009.01 and Eq. 07.02.03.0014.01 in [34], we can get

$${}_1F_1(m_{sr};1;\delta_{sr}x)=e^{-\delta_{sr}x}\sum_{n=0}^{m_{sr}-1}\frac{(-\delta_{sr})^n(1-m_{sr})}{(n!)^2}x^n.\tag{10}$$

By taking (10) into (9), it can be rewritten as

$$f_{|h_{sr}|^2}(x)=\alpha_{sr}e^{-(\beta_{sr}-\delta_{sr})x}\sum_{n=0}^{m_{sr}-1}\boxtimes\frac{(-\delta_{sr})^n(1-m_{sr})_n}{(n!)^2}x^n.\tag{11}$$

Owing to $\gamma_{sr}=\bar{\gamma}_s|h_{sr}|^2$, the PDF of γ_{sr} can expressed as

$$f_{\gamma_{sr}}(x)=\alpha_{sr}\sum_{n=0}^{m_{sr}-1}\boxtimes\zeta(n)x^ne^{-\Delta_{sr}x},\tag{12}$$

where $\zeta(n)=(-\delta_{sr})^n(1-m_{sr})_n/(n!)^2(\bar{\gamma}_s)^{n+1}$ and $\Delta_{sr}=(\beta_{sr}-\delta_{sr})/\bar{\gamma}_s$.

With the utilization of Eq. 3.351.2 in [32], we can get CDF of γ_{sr} as

$$F_{\gamma_{sr}}(x)=1-\alpha_{sr}\sum_{n=0}^{m_{sr}-1}\boxtimes\sum_{t=0}^{n}\boxtimes\frac{n!\zeta(n)}{t!\Delta_{sr}^{n-t+1}}x^te^{-\Delta_{sr}x},\tag{13}$$

In addition, the PDF and CDF of γ_{ra} are expressed as, respectively [35],

$$f_{\gamma_{ra}}(x)=\frac{1}{\Gamma(m_{ra})}\Lambda_{ra}^{m_{ra}}x^{m_{ra}-1}e^{-\Xi_{ra}x},\tag{14}$$

$$F_{\gamma_{ra}}(x)=1-e^{-\Lambda_{ra}x}\sum_{p=0}^{m_{ra}-1}\boxtimes\frac{1}{p!}\Lambda_{ra}^{p}x^p,\tag{15}$$

where $\Lambda_{ra}=m_{ra}/\Omega_{ra}\bar{\gamma}_r$, m_{ra} is the fading severity parameter, and $\Omega_{ra}=(d_{ra})^{-\omega}$ denotes the average fading power, in which d_{ra} is the distance from R to U or D and ω denotes the path loss exponent.

3. Joint Decoding Order and Power Allocation Design

In this section, by jointly designing the decoding order and power allocation factor, we maximize the achievable rate of the secondary network under the predefined constraints of the primary network. Firstly, we find the optimal power allocation factor under the two decoding orders. Then, the optimal allocation factors are utilized to obtain the corresponding achievable rates. By comparing the calculated achievable rates under different decoding orders, the desired

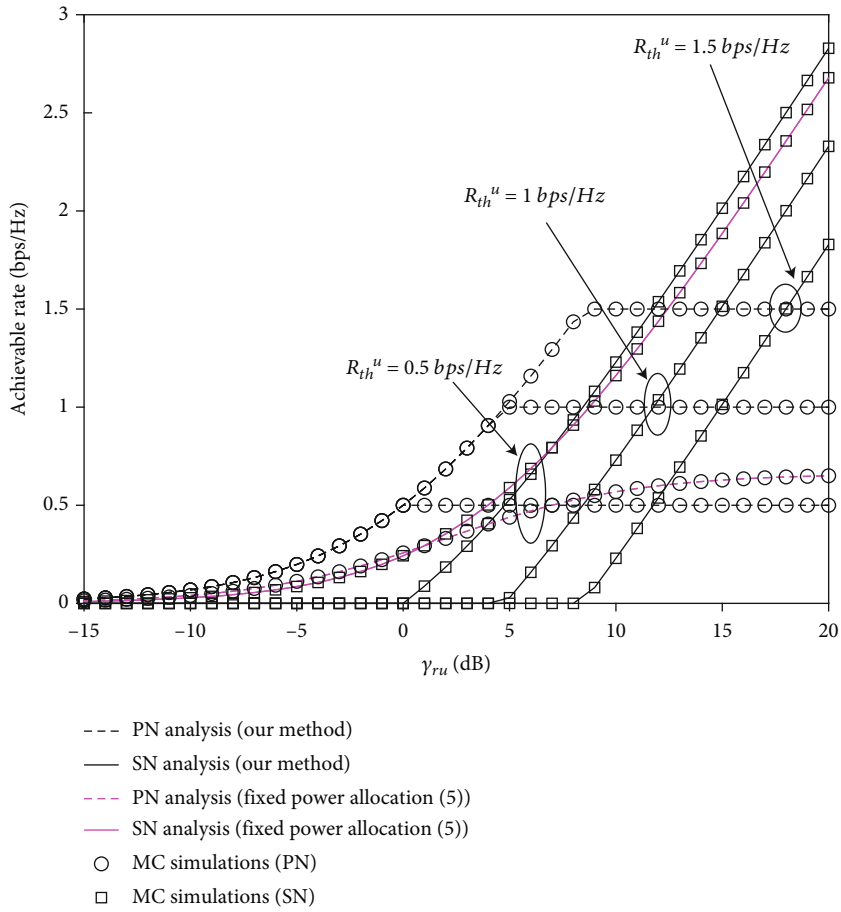


Figure 3: The achievable rates versus different target rates of the primary system.

decoding order corresponding to the optimal achievable rate can be obtained.

3.1. The First Decoding Order. Under the first decoding order, the optimization problem can be formulated as follows.

Problem 1.

$$\max_{\beta} R_d^1 = \frac{1}{2}\log_2(1+(1-\beta)\gamma_{rd}), \tag{16}$$

$$\text{s.t.} \min\left\{R_u^1, R_{u\longrightarrow d}^1\right\} \geq R_{th}^u, \tag{17}$$

$$0 < \beta \leq 1, \tag{18}$$

where (17) guarantees that the primary signal can be decoded correctly and SIC can be performed successfully. Besides, (17) can be reexpressed as

$$\begin{aligned} \beta &\geq \frac{k_u(1+\gamma_{ru})}{\gamma_{ru}(1+k_u)}, \\ \beta &\geq \frac{k_u(1+\gamma_{rd})}{\gamma_{rd}(1+k_u)}, \end{aligned} \tag{19}$$

where $k_u = 2^{2R_{th}^u} - 1$.

From (18), we can obtain $\tilde{\gamma} \geq k_u$ with $\tilde{\gamma} = \min(\gamma_{ru}, \gamma_{rd})$. It is worth noting that if the constraints of (17) and (18) cannot be satisfied, all power is allocated to transmit the message of the primary network.

It can be observed that R_d^1 is a decreasing function of β; thus, the optimal power allocation factor can be expressed as

$$\beta^* = \begin{cases} \dfrac{k_u(1+\tilde{\gamma})}{\tilde{\gamma}(1+k_u)}, & \tilde{\gamma} \geq k_u, \\ 1, & \text{else}. \end{cases} \tag{20}$$

3.2. The Second Decoding Order. Under the second decoding order, the optimization problem can be formulated as follows.

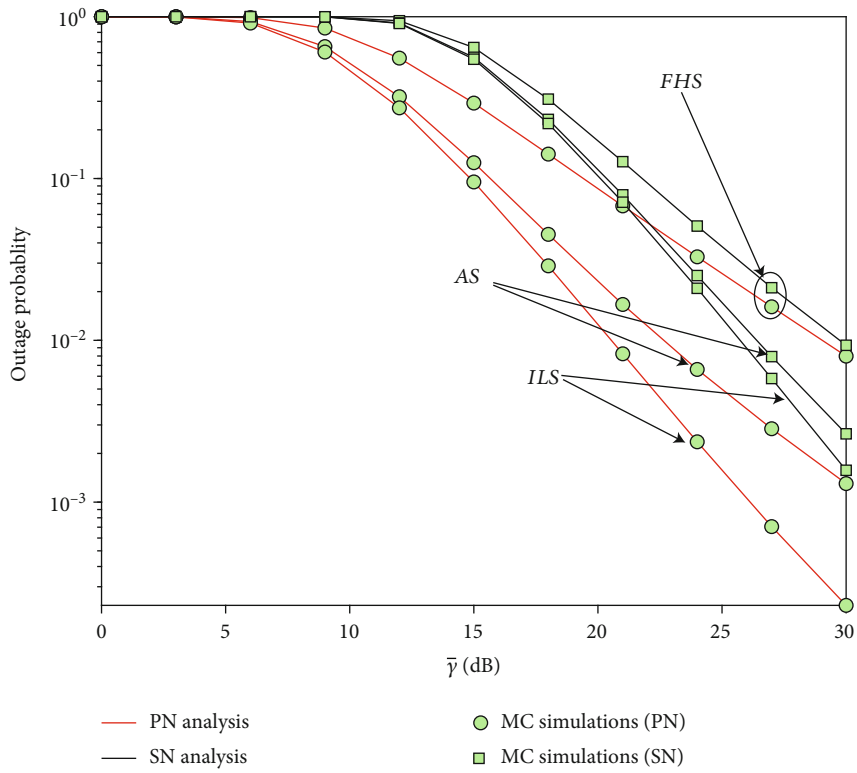


Figure 4: OP versus different SR fading parameters.

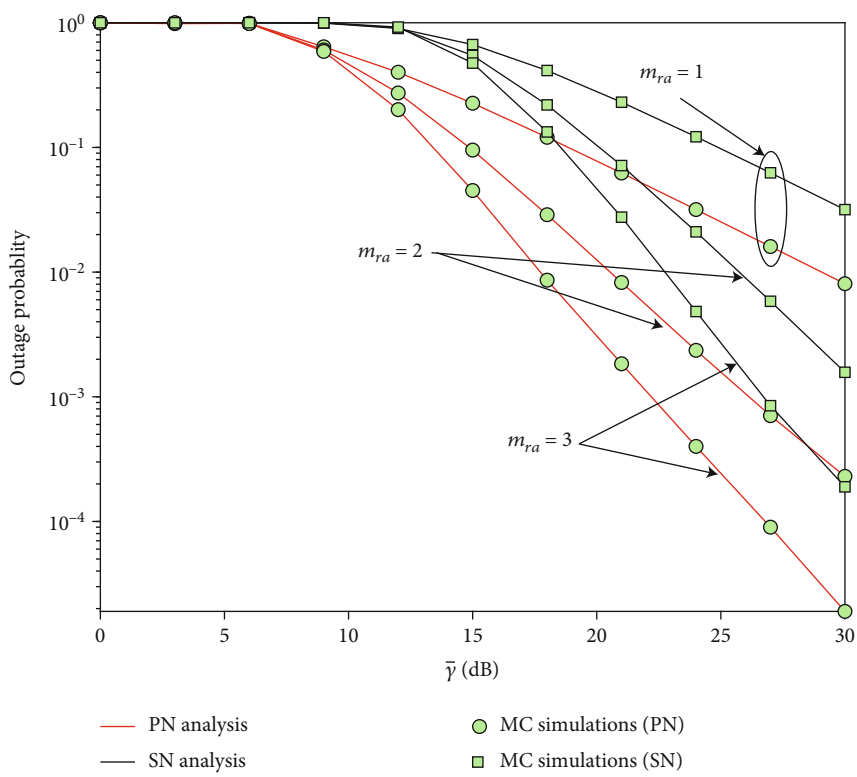


Figure 5: OP versus different Nakagami-m fading parameters.

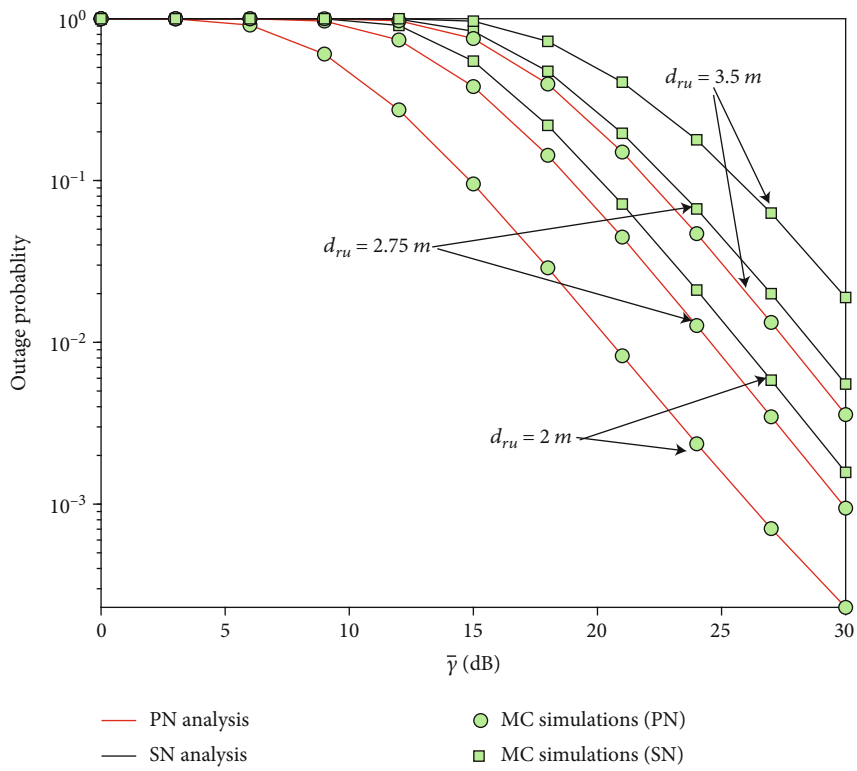


Figure 6: OP versus different distances from R to U.

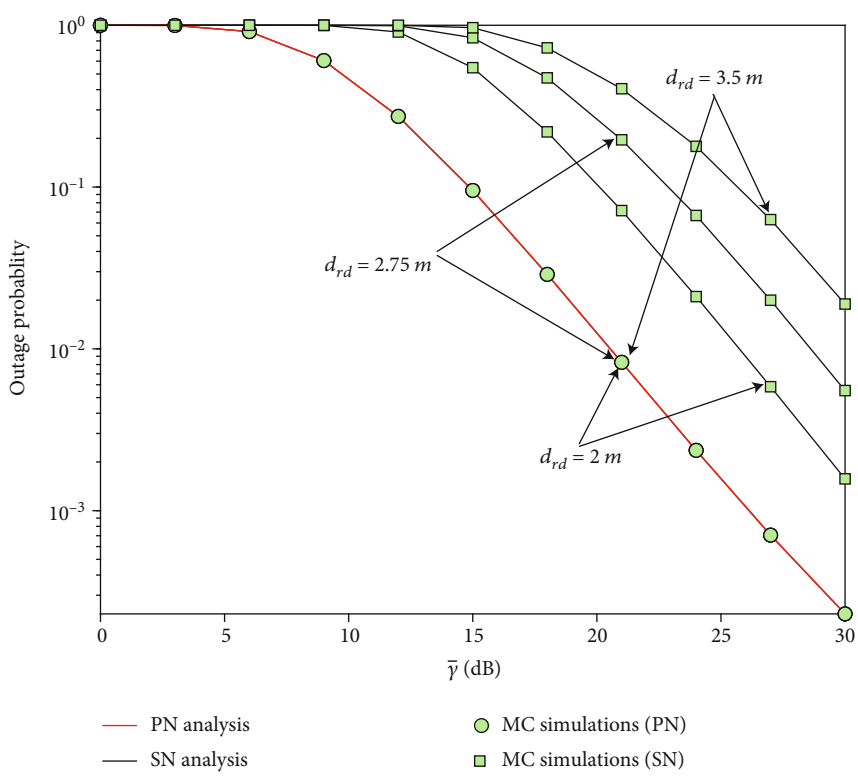


Figure 7: OP versus different distances from R to D.

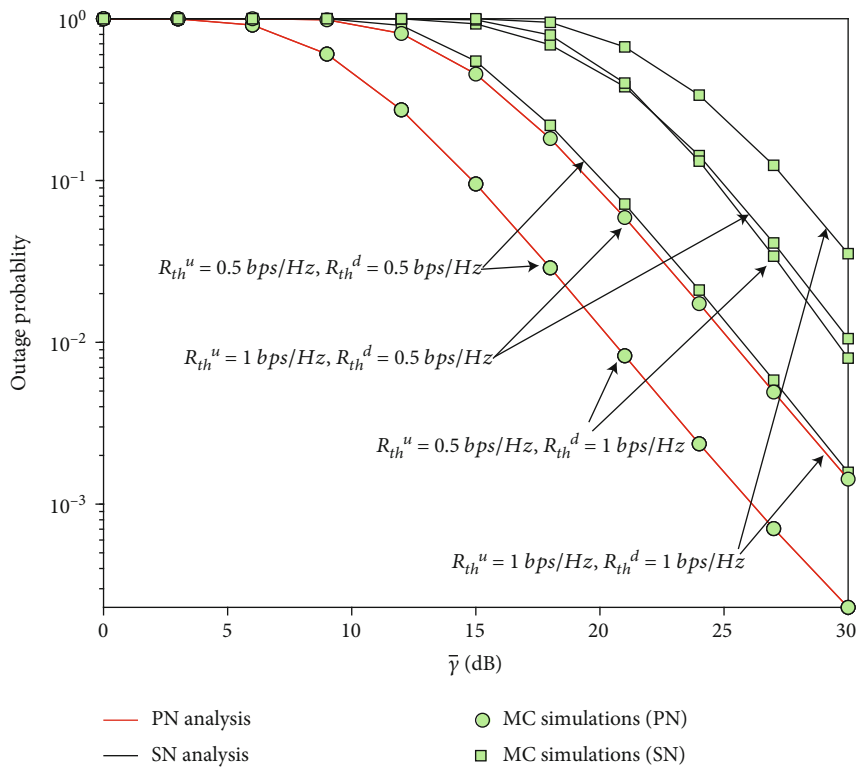


Figure 8: OP versus different target rates.

Problem 2.

$$\max_{\beta} R_d^2 = \frac{1}{2}\log_2\left(1 + \frac{(1-\beta)\gamma_{rd}}{\beta\gamma_{rd}+1}\right), \tag{21}$$

$$\text{s.t.} R_{d\longrightarrow u}^2 \geq R_{th}^d, \tag{22}$$

$$R_u^2 \geq R_{th}^u, \tag{23}$$

$$0 < \beta \leq 1, \tag{24}$$

where R_{th}^d is the target rate of the secondary network, (22) and (23) ensure that the primary network can decode the expected signal successfully, and we can get

$$\frac{k_u}{\gamma_{ru}} \leq \beta \leq \frac{\gamma_{ru} - k_d}{\gamma_{ru}(1+k_d)}, \tag{25}$$

where $k_d = 2^{2R_{th}^d} - 1$. Moreover, to ensure the feasible region of this problem, not an empty set, $(\gamma_{ru} - k_d)/\gamma_{ru}(1+k_d) \geq k_u/\gamma_{ru}$ should be guaranteed. Hence, we can get $\gamma_{ru} \geq k = k_u + k_d + k_u k_d$.

From (21), we can obtain that R_d^2 is a strict decreasing function of β. Thus, the optimal power allocation factor is expressed as

$$\beta^{\dagger} = \frac{k_u}{\gamma_{ru}}, \quad \gamma_{ru} \geq k. \tag{26}$$

Only when $\gamma_{ru} \geq k$ can the system work under the second order. Thus, we just need to compare the optimal achievable rates under the first and second order with $\gamma_{ru} \geq k$. After comparison, we find that when $\gamma_{ru} > \gamma_{rd}$, $R_d^2(\beta^{\dagger}) > R_d^1(\beta^*)$. Therefore, the joint optimal design of the decoding order and power allocation is given by

$$(\lambda, \beta) = \begin{cases} \left(2, \beta^{\dagger}\right), & \gamma_{ru} > \max\left(\gamma_{rd}, k\right), \\ \left(1, \beta^*\right), & \text{else}, \end{cases} \tag{27}$$

where λ represents the decoding order.

According to the above joint optimization, we can get the achievable rates of x_s and x_r at U and D as, respectively,

$$R_u = \begin{cases} R_{th}^u, & \gamma_{ru} > k_u, \\ \frac{1}{2}\log_2(1+\gamma_{ru}), & \text{else}, \end{cases} \tag{28}$$

$$R_d = \begin{cases} \frac{1}{2}\log_2(1+\Xi_1), & \gamma_{ru} > \max\left(\gamma_{rd}, k\right), \\ \frac{1}{2}\log_2(1+\Xi_2), & \text{else}, \end{cases} \tag{29}$$

where $\Xi_1 = (\gamma_{ru} - k_u)\gamma_{rd}/(k_u\gamma_{rd} + \gamma_{ru})$ and $\Xi_2 = (\tilde{\gamma} - k_u)\gamma_{rd}/\tilde{\gamma}(1+k_u)$.

Remark. The above derivations show that the user with low priority should be decoded firstly when the channel gain of

the high-priority user is larger than that of the low-priority user, as well as the specific value. In great majority of existing works on the NOMA system, the high-priority users should be decoded firstly, which is diverse with our observation.

4. Performance Analysis

In this section, the exact and asymptotic expressions of OP for both the primary network and secondary network are obtained based on the derivations of achievable rates, where OP is defined as the probability of the achievable rate lower than the target rate. Besides, we derive the DOs to provide more insights.

4.1. Primary Network. The primary network will be interrupted if the achievable rate of x_s at S or U is lower than R_{th}^u; thus, the OP can be expressed as

$$P_{\text{out}}^p = 1 - \Pr\{R_r \geq R_{th}^u\} \Pr\{R_u \geq R_{th}^u\} = 1 - \left[1 - F_{\gamma_{sr}}(k_u)\right]\left[1 - F_{\gamma_{ru}}(k_u)\right]. \tag{30}$$

Then, by taking (13) and (15) into (30), we can get

$$P_{\text{out}}^p = 1 - \alpha_{sr} \sum_{n=0}^{m_{sr}-1} \boxtimes \sum_{t=0}^{n} \boxtimes \frac{n!\zeta(n)}{t!\Delta_{sr}^{n-t+1}} \sum_{p=0}^{m_{ru}-1} \boxtimes \frac{1}{p!} \Lambda_{ru}^p k_u^{p+t} e^{-(\Delta_{sr}+\Lambda_{ru})k_u}. \tag{31}$$

To derive the asymptotic OP, we first give the asymptotic CDF of γ_{sr} and γ_{ra} in the high SNR regime as [35]

$$F_{\gamma_{sr}}(x) = \frac{\alpha_{sr}}{\bar{\gamma}_s} x, \tag{32}$$

$$F_{\gamma_{ra}}(x) = \frac{1}{\Gamma(m_{ra}+1)} \Lambda_{ra}^{m_{ra}} x^{m_{ra}}. \tag{33}$$

Finally, we substitute (32) and (33) into (30); the asymptotic expression for OP of the primary system can be expressed as

$$P_{\text{out}}^{p,\infty} \simeq \begin{cases} \dfrac{\alpha_{sr}}{\bar{\gamma}_s} k_u, & m_{ru} > 1, \\ \dfrac{\alpha_{sr}}{\bar{\gamma}_s} k_u + \dfrac{\Lambda_{ru}^{m_{ru}} k_u^{m_{ru}}}{\Gamma(m_{ru}+1)}, & m_{ru} = 1. \end{cases} \tag{34}$$

According to (34), the DO of the primary network can be expressed as $\text{DO}_p = \min(1, m_{ru})$.

4.2. Secondary Network. In the secondary network, OP means the probability of achievable rates of x_s at R lower than R_{th}^u or that of x_r at D lower than R_{th}^d, which can be given by

$$P_{\text{out}}^s = 1 - \underbrace{\Pr\{R_r \geq R_{th}^u\}}_{T_1} \underbrace{\Pr\{R_d \geq R_{th}^u\}}_{T_2}. \tag{35}$$

We can easily get

$$T_1 = \alpha_{sr} \sum_{n=0}^{m_{sr}-1} \boxtimes \sum_{t=0}^{n} \boxtimes \frac{n!\zeta(n)}{t!\Delta_{sr}^{n-t+1}} k_u^{\,t} e^{-\Delta_{sr}k_u}. \tag{36}$$

From (29), T_2 can be reexpressed as

$$\begin{aligned} T_2 &= \Pr\{\gamma_{ru} \geq \max\{\gamma_{rd}, k\}, \Xi_1 \geq k_d\} \\ &\quad + \Pr\{\gamma_{ru} \leq \max\{\gamma_{rd}, k\}, \Xi_2 \geq k_d\} \\ &= \Pr\left\{\gamma_{ru} \geq k, \gamma_{ru} \geq \gamma_{rd} \geq \frac{\gamma_{ru}k_d}{\gamma_{ru} - k + k_d}\right\} \\ &\quad + \Pr\left\{\gamma_{rd} \geq k, \gamma_{rd} \geq \gamma_{ru} \geq \frac{\gamma_{rd}k_d}{\gamma_{rd} - k + k_u}\right\}. \end{aligned} \tag{37}$$

To simplify the calculation, some mathematical processing is conducted for (37) as follows [31].

$$T_2 = \Pr\left\{\gamma_{ru} \geq k, \gamma_{ru} \geq \gamma_{rd} \geq \frac{\gamma_{ru}k_d}{\gamma_{ru} - k + k_d}\right\} \tag{38a}$$

$$+\Pr\left\{\gamma_{ru} \geq k, \gamma_{rd} \geq \gamma_{ru} \geq \frac{\gamma_{ru}k_d}{\gamma_{ru} - k + k_d}\right\} \tag{38b}$$

$$+\Pr\left\{\gamma_{rd} \geq k, \gamma_{rd} \geq \gamma_{ru} \geq \frac{\gamma_{rd}k_d}{\gamma_{rd} - k + k_u}\right\} \tag{38c}$$

$$+\Pr\left\{\gamma_{rd} \geq k, \gamma_{ru} \geq \gamma_{rd} \geq \frac{\gamma_{rd}k_d}{\gamma_{rd} - k + k_u}\right\} \tag{38d}$$

$$-\Pr\left\{\gamma_{ru} \geq k, \gamma_{rd} \geq \gamma_{ru} \geq \frac{\gamma_{ru}k_d}{\gamma_{ru} - k + k_d}\right\} \tag{38e}$$

$$-\Pr\left\{\gamma_{rd} \geq k, \gamma_{ru} \geq \gamma_{rd} \geq \frac{\gamma_{rd}k_d}{\gamma_{rd} - k + k_u}\right\}. \tag{38f}$$

By merging (38a) and (38b), (38c) and (38d), and (38e) and (38f), we can get

$$T_2 = \underbrace{\Pr\left\{\gamma_{ru} \geq k, \gamma_{rd} \geq \frac{\gamma_{ru}k_d}{\gamma_{ru} - k + k_d}\right\}}_{Q_1} + \underbrace{\Pr\left\{\gamma_{rd} \geq k, \gamma_{ru} \geq \frac{\gamma_{rd}k_d}{\gamma_{rd} - k + k_u}\right\}}_{Q_2} - \underbrace{\Pr\{\gamma_{ru} \geq k, \gamma_{rd} \geq k\}}_{Q_3}. \tag{39}$$

Similar to the derivation of P_{out}^{p}, with the utilization of (15), we can obtain

$$Q_3 = \left[1 - F_{\gamma_{rd}}(k)\right]\left[1 - F_{\gamma_{ru}}(k)\right] = e^{-(\Lambda_{rd}+\Lambda_{ru})k} \sum_{p_1=0}^{m_{rd}-1} \sum_{p_2=0}^{m_{ru}-1} \frac{1}{p_1!}\frac{1}{p_2!} \Lambda_{rd}^{p_1}\Lambda_{ru}^{p_2} k^{p_1+p_2}. \tag{40}$$

Then, the derivation of Q_1 is carried on, which is given by

$$Q_1 = \int_k^{+\infty} \left[1 - F_{\gamma_{rd}}\left(\frac{xk_d}{x-k+k_d}\right)\right] f_{\gamma_{ru}}(x)dx. \tag{41}$$

Let $x - k + k_d = t$; (46) can be rewritten as

$$Q_1 = \underbrace{\int_0^{+\infty} \left[1 - F_{\gamma_{rd}}\left(\frac{(t+k-k_d)k_d}{t}\right)\right] f_{\gamma_{ru}}(t+k-k_d)dt}_{Q_{11}} - \underbrace{\int_0^{k_d} \left[1 - F_{\gamma_{rd}}\left(\frac{(t+k-k_d)k_d}{t}\right)\right] f_{\gamma_{ru}}(t+k-k_d)dt}_{Q_{12}}. \tag{42}$$

By substituting (14) and (15) into Q_{11}, with the help of Eq. 1.111 and Eq. 3.471.9 in [32], the expression of Q_{11} can be derived as

$$\begin{aligned} Q_{11} = {} & \frac{2\Lambda_{ru}^{m_{ru}}}{\Gamma(m_{ru})} e^{-[(k-k_d)\Lambda_{ru}+\Lambda_{rd}k_d]} \sum_{p_1=0}^{m_{rd}-1} \frac{1}{p_1!}(\Lambda_{rd}k_d)^{p_1} \sum_{p_2=0}^{m_{ru}-1+p_1} \binom{m_{ru}-1+p_1}{p_2} \\ & \times (k-k_d)^{m_{ru}-1+p_1-p_2} \left(\frac{\Lambda_{rd}k_d(k-k_d)}{\Lambda_{ru}}\right)^{(p_2-p_1+1)/2} K_{p_2-p_1+1}\left(2\sqrt{\Lambda_{ru}\Lambda_{rd}k_d(k-k_d)}\right). \end{aligned} \tag{43}$$

Let $t = ((u+1)/2)k_d$; Q_{12} can be rewritten as

$$Q_{12} = \frac{k_d}{2}\int_{-1}^{1} \left[1 - F_{\gamma_{rd}}\left(\frac{(u+1)k_d + 2k - 2k_d}{u+1}\right)\right] f_{\gamma_{ru}}\left(\frac{(u+1)k_d + 2k - 2k_d}{2}\right)du. \tag{44}$$

By utilizing the Gaussian-Chebyshev quadrature [33], we can get

$$\begin{aligned} Q_{12} \approx {} & \frac{\pi k_d \Lambda_{ru}^{m_{ru}}}{I\Gamma(m_{ru})} e^{-[(k-k_d)\Lambda_{ru}+\Lambda_{rd}k_d]} \sum_{p_1=0}^{m_{rd}-1} \frac{1}{p_1!}(\Lambda_{rd}k_d)^{p_1} \sum_{p_2=0}^{m_{ru}-1+p_1} \binom{m_{ru}-1+p_1}{p_2} \\ & \times (k-k_d)^{m_{ru}-1+p_1-p_2} \sum_{i_1=1}^{I} \sqrt{1-u_{i_1}^2}\, e^{-\left[\left((2k-2k_d)\Lambda_{rd}/\left(u_{i_1}+1\right)\right)+\left(\Lambda_{ru}k_d\left(u_{i_1}+1\right)/2\right)\right]} \left(\frac{u_{i_1}+1}{2}\right)^{p_2-p_1} k_d^{p_2}, \end{aligned} \tag{45}$$

where $u_i = \cos\left[(2i-1)\pi/2I\right]$ and I is the Gaussian-Chebyshev approximation.

By taking similar derivations of Q_{11} and Q_{12}, Q_{21} and Q_{22} are given by, respectively,

$$\begin{aligned} Q_{21} = {} & \frac{2\Lambda_{rd}^{m_{rd}}}{\Gamma(m_{rd})} e^{-[(k-k_u)\Lambda_{rd}+\Lambda_{ru}k_u]} \sum_{p_1=0}^{m_{ru}-1} \frac{1}{p_1!}(\Lambda_{ru}k_u)^{p_1} \sum_{p_2=0}^{m_{rd}-1+p_1} \binom{m_{rd}-1+p_1}{p_2} \\ & \times (k-k_u)^{m_{rd}-1+p_1-p_2} \left(\frac{\Lambda_{ru}k_u(k-k_u)}{\Lambda_{rd}}\right)^{(p_2-p_1+1)/2} K_{p_2-p_1+1}\left(2\sqrt{\Lambda_{ru}\Lambda_{rd}k_u(k-k_u)}\right). \end{aligned} \tag{46}$$

$$\begin{aligned} Q_{22} \approx {} & \frac{\pi k_u \Lambda_{rd}^{m_{rd}}}{I\Gamma(m_{rd})} e^{-[(k-k_u)\Lambda_{rd}+\Lambda_{ru}k_u]} \sum_{p_1=0}^{m_{ru}-1} \frac{1}{p_1!}(\Lambda_{ru}k_u)^{p_1} \sum_{p_2=0}^{m_{rd}-1+p_1} \binom{m_{rd}-1+p_1}{p_2} \\ & \times (k-k_u)^{m_{rd}-1+p_1-p_2} \sum_{i_2=1}^{I} \sqrt{1-u_{i_2}^2}\, e^{-\left[\left((2k-2k_u)\Lambda_{ru}/\left(u_{i_2}+1\right)\right)+\left(\Lambda_{rd}k_u\left(u_{i_2}+1\right)/2\right)\right]} \left(\frac{u_{i_2}+1}{2}\right)^{p_2-p_1} k_u^{p_2}. \end{aligned} \tag{47}$$

Therefore, the expression of OP for the secondary network is given by

$$
\begin{aligned}
P_{\text{out}}^{s} = 1 - \Bigg[& \alpha_{sr} \sum_{n=0}^{m_{sr}-1} \sum_{t=0}^{n} \frac{n!\zeta(n)}{t!\Delta_{sr}^{n-t+1}} k_u^{\,t} e^{-\Delta_{sr} k_u} \Bigg] \\
& \Bigg\{ e^{-[(k-k_d)\Lambda_{ru}+\Lambda_{rd}k_d]} \sum_{p_1=0}^{m_{rd}-1} \frac{1}{p_1!} (\Lambda_{rd}k_d)^{p_1} \times \sum_{p_2=0}^{m_{ru}-1+p_1} \binom{m_{ru}-1+p_1}{p_2} (k-k_d)^{m_{ru}-1+p_1-p_2} \\
& \Bigg[\frac{2\Lambda_{ru}^{m_{ru}}}{\Gamma(m_{ru})} \left(\frac{\Lambda_{rd}k_d(k-k_d)}{\Lambda_{ru}} \right)^{(p_2-p_1+1)/2} \times K_{p_2-p_1+1}\left(2\sqrt{\Lambda_{ru}\Lambda_{rd}k_d(k-k_d)}\right) \\
& - \frac{\pi k_d \Lambda_{ru}^{m_{ru}}}{I\Gamma(m_{ru})} \sum_{i_1=1}^{I} \sqrt{1-u_{i_1}^2}\, e^{-\left[\left((2k-2k_d)\Lambda_{rd}/\left(u_{i_1}+1\right)\right)+\left(\Lambda_{ru}k_d\left(u_{i_1}+1\right)/2\right)\right]} \times \left(\frac{u_{i_1}+1}{2}\right)^{p_2-p_1} k_d^{p_2} \Bigg] \\
& + e^{-[(k-k_u)\Lambda_{rd}+\Lambda_{ru}k_u]} \sum_{p_1=0}^{m_{ru}-1} \frac{1}{p_1!} (\Lambda_{ru}k_u)^{p_1} \sum_{p_2=0}^{m_{rd}-1+p_1} \binom{m_{rd}-1+p_1}{p_2} \times (k-k_u)^{m_{rd}-1+p_1-p_2} \Bigg[\frac{2\Lambda_{rd}^{m_{rd}}}{\Gamma(m_{rd})} \left(\frac{\Lambda_{ru}k_u(k-k_u)}{\Lambda_{rd}} \right)^{(p_2-p_1+1)/2} \\
& K_{p_2-p_1+1}\left(2\sqrt{\Lambda_{ru}\Lambda_{rd}k_u(k-k_u)}\right) - \frac{\pi k_u \Lambda_{rd}^{m_{rd}}}{I\Gamma(m_{rd})} \sum_{i_2=1}^{I} \sqrt{1-u_{i_2}^2}\, e^{-\left[\left((2k-2k_u)\Lambda_{ru}/\left(u_{i_2}+1\right)\right)+\left(\Lambda_{rd}k_u\left(u_{i_2}+1\right)/2\right)\right]} \times \left(\frac{u_{i_2}+1}{2}\right)^{p_2-p_1} k_u^{p_2} \Bigg] \\
& - e^{-(\Lambda_{rd}+\Lambda_{ru})k} \sum_{p_1=0}^{m_{rd}-1} \sum_{p_2=0}^{m_{ru}-1} \frac{1}{p_1!}\frac{1}{p_2!} \Lambda_{rd}^{p_1}\Lambda_{ru}^{p_2} k^{p_1+p_2} \Bigg\}.
\end{aligned}
\tag{48}
$$

Similar to the asymptotic OP of the primary network, we can easily obtain the asymptotic OP of the secondary network as

$$
P_{\text{out}}^{s,\infty} = \begin{cases}
\dfrac{\alpha_{sr}}{\bar{\gamma}_s} k_u, & m_{ru} > 1, m_{rd} > 1, \\
\dfrac{\alpha_{sr}}{\bar{\gamma}_s} k_u + \dfrac{\Lambda_{ru}^{m_{ru}} k^{m_{ru}}}{\Gamma(m_{ru}+1)}, & m_{ru} = 1, m_{rd} > 1, \\
\dfrac{\alpha_{sr}}{\bar{\gamma}_s} k_u + \dfrac{\Lambda_{rd}^{m_{rd}} k^{m_{rd}}}{\Gamma(m_{rd}+1)}, & m_{rd} = 1, m_{ru} > 1, \\
\dfrac{\alpha_{sr}}{\bar{\gamma}_s} k_u + \dfrac{\Lambda_{ru}^{m_{ru}} k^{m_{ru}}}{\Gamma(m_{ru}+1)} + \dfrac{\Lambda_{rd}^{m_{rd}} k^{m_{rd}}}{\Gamma(m_{rd}+1)}, & m_{ru} = m_{rd} = 1.
\end{cases}
\tag{49}
$$

In addition, the DO of the secondary network is $\text{DO}_s = \min(1, m_{ru}, m_{rd})$.

5. Numerical Results

In this section, numerical results are provided to prove the correctness of our theoretical derivation. The simulation tool is MATLAB 2019a. Besides, the impacts of key parameters are revealed. Particularly, we set $\bar{\gamma}_{sr} = \bar{\gamma}_{ru} = \bar{\gamma}_{rd} = \bar{\gamma}$, $\omega = 3$, and $I = 20$ [3]. Moreover, the system and channel parameters of SR fading are provided in Table 1 [12, 36].

Figure 3 illustrates that the achievable rates versus different target rates of the primary system with setting $m_{ru} = m_{rd} = 2$, $d_{ru} = d_{rd} = 2m$, and $R_{th}^{d} = 0.5$ bps/Hz. It can be observed that our theoretical analysis is consistent with the Monte Carlo (MC) simulations, which demonstrates the accuracy of the analytical results. We can find that the achievable rate of the primary network increases with γ_{ru} and reaches a fixed value, namely, R_{th}^{u}. Noting that when R_u achieves R_{th}^{u}, R_d begins to increase. This is due to the fact that if the primary network cannot correctly decode its expected signal, a secondary transmitter allocates all power to the primary network. Besides, R_d decreases with stricter target rates. In a high SNR regime, it can be clearly found that the achievable rate of the secondary network with our method is higher than that with the method in [6], which proves that our method is superior to the fixed power allocation method. Moreover, in a low SNR regime, the achievable rate of the secondary network is lower than that in [6]; it is because our method should ensure the QoS of the primary network.

Figure 4 depicts that OP versus different SR fading parameters with $m_{ru} = m_{rd} = 2$, $d_{ru} = d_{rd} = 2m$, and $R_{th}^{u} = R_{th}^{d} = 0.5$ bps/Hz. It can be observed that the primary network outperforms the secondary network. This is because

we give higher priority to the primary network in the optimal design for ensuring its QoS. Moreover, the curves in Figure 3 display that lighter SR channel fading can lead to better system performance.

Figure 5 shows that OP versus different Nakagami-m fading parameters with setting $d_{ru} = d_{rd} = 2m$ and $R_{th}^{u} = R_{th}^{d} = 0.5$ bps/Hz. We can also observe that the performance of the primary network is superior to that of the secondary network. Besides, the larger Nakagami-m fading parameters mean lighter channel fading, which can improve the system performance.

Figure 6 plots the OP versus different distances from R to U with $m_{ru} = m_{rd} = 2$ and $R_{th}^{u} = R_{th}^{d} = 0.5$ bps/Hz, AS. It can be clearly seen that whether it is the primary network or the secondary network, the increase in the distance between the secondary transmitter R and the primary user U improves the OP of the system. It is not difficult to understand that longer distance degrades the communication quality.

However, in Figure 7, which illustrates the OP versus different distances from R to D with $m_{ru} = m_{rd} = 2$ and $R_{th}^{u} = R_{th}^{d} = 0.5$ bps/Hz, AS, the distance between the secondary transmitter and the secondary user only affects the OP of the secondary network. It indicates that the OP of the primary network is not related to d_{rd}, while that of the secondary network will rise with the increase in both d_{ru} and d_{rd}.

Figure 8 depicts the OP versus different target rates with setting $m_{ru} = m_{rd} = 2$ and $d_{ru} = d_{rd} = 2m$, AS. We can find that the larger target rate of the primary network will increase the OP of both the primary network and the secondary network. It can be explained by the fact that a larger target rate means higher communication requirements, which is more difficult to achieve. Besides, the stricter target rate of the secondary network only reduces the system performance of the secondary network. Like d_{rd}, R_{th}^{d} has no effect on the performance of the primary network.

6. Conclusions

In this paper, we established an overlay CISTRN with the NOMA scheme, where the secondary network multiplexed spectrum resources by accessing the spectrum of the primary network. In exchange, the secondary network helped the information transmission of the primary network as a relay. To improve the system performance of the secondary network, we jointly optimized the decoding order and power allocation, where the QoS of the primary network has been ensured. In addition, we derived the exact OP of both the primary network and secondary network. To get further insights, the asymptotic OPs and DOs were provided. From the theoretical results, it can be observed that when the channel coefficient of the high-priority user is higher than that of the low-priority user as well as a fixed value, the system will decode the message of the low-priority user first, which will make the system have better performance. This conclusion is diverse from great majority of existing works on a NOMA-based system. Finally, our theoretical results were validated by simulation results, and the impacts of key system parameters on optimal design as well as system performance were revealed.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the National Science Foundation of China under Grants 61901502 and 62001517, the National Postdoctoral Program for Innovative Talents under Grant BX20200101, and the Research Project of National University of Defense Technology (NUDT) under Grants ZK18-02-11 and 18-QNCXJ-029 and in part by the Research Project of Space Engineering University under Grants 2020XXAQ01 and 2019XXAQ05 and in part by the Science and Technology Innovation Cultivation Fund of Space Engineering University. Thanks are due to all the authors for their contribution to this paper.

References

[1] Z. Lin, M. Lin, T. de Cola, J. B. Wang, W. P. Zhu, and J. Cheng, "Supporting IoT with rate-splitting multiple access in satellite and aerial-integrated networks," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11123–11134, 2021.

[2] K. Guo, K. An, B. Zhang et al., "On the performance of the uplink satellite multiterrestrial relay networks with hardware impairments and interference," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2297–2308, 2019.

[3] R. Liu, K. Guo, K. An, S. Zhu, and S. Shuai, "NOMA-based integrated satellite-terrestrial relay networks under spectrum sharing environment," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1266–1270, 2021.

[4] V. Singh and P. K. Upadhyay, "Exploiting cache-free/cache-aided TWR-NOMA in cognitive hybrid satellite-terrestrial networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 1778–1793, 2022.

[5] R. Liu, K. Guo, K. An, and S. Zhu, "NOMA-based overlay cognitive integrated satellite-terrestrial relay networks with secondary network selection," *IEEE Transactions on Vehicular Technology.*, vol. 71, no. 2, pp. 2187–2192, 2022.

[6] X. Zhang, D. Guo, K. An et al., "Performance analysis of NOMA-based cooperative spectrum sharing in hybrid satellite-terrestrial networks," *IEEE Access*, vol. 7, pp. 172321–172329, 2019.

[7] X. Li, Y. Zheng, W. U. Khan et al., "Physical layer security of cognitive ambient backscatter communications for green Internet-of-things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[8] Z. Lin, M. Lin, W. P. Zhu, J. B. Wang, and J. Cheng, "Robust secure beamforming for wireless powered cognitive satellite-terrestrial networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 567–580, 2021.

[9] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[10] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[11] K. An, T. Liang, X. Yan, Y. Li, and X. Qiao, "Power allocation in land mobile satellite systems: an energy-efficient perspective," *IEEE Communications Letters*, vol. 22, no. 7, pp. 1374–1377, 2018.

[12] K. Guo, M. Lin, B. Zhang, W. Zhu, J. Wang, and T. A. Tsiftsis, "On the performance of LMS communication with hardware impairments and interference," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1490–1505, 2019.

[13] K. Guo, K. An, B. Zhang et al., "Physical layer security for multiuser satellite communication systems with threshold-based scheduling scheme," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5129–5141, 2020.

[14] M. Lin, Q. Huang, T. de Cola et al., "Integrated 5G-satellite networks: a perspective on physical layer reliability and security," *IEEE Wireless Communications*, vol. 27, no. 6, pp. 152–159, 2020.

[15] Z. Lin, H. Niu, K. An et al., "Refracting RIS aided hybrid satellite-terrestrial relay networks: joint beamforming design and optimization," *IEEE Transactions on Aerospace and Electronic Systems*, 2022.

[16] P. K. Sharma, P. K. Upadhyay, D. B. da Costa, P. S. Bithas, and A. G. Kanatas, "Performance analysis of overlay spectrum sharing in hybrid satellite-terrestrial systems with secondary network selection," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6586–6601, 2017.

[17] Z. Lin, M. Lin, B. Champagne, W. P. Zhu, and N. al-Dhahir, "Secure and energy efficient transmission for RSMA-based cognitive satellite-terrestrial networks," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 251–255, 2021.

[18] K. An, M. Lin, J. Ouyang, and W. P. Zhu, "Secure transmission in cognitive satellite terrestrial networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 3025–3037, 2016.

[19] Z. Lin, M. Lin, B. Champagne, W. P. Zhu, and N. Al-Dhahir, "Secure beamforming for cognitive satellite terrestrial networks with unknown eavesdroppers," *IEEE Systems Journal*, vol. 15, no. 2, pp. 2186–2189, 2021.

[20] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[21] K. Guo, C. Dong, and K. An, "NOMA-based cognitive satellite terrestrial relay network: secrecy performance under channel estimation errors and hardware impairments," *IEEE Internet of Things Journal*, 2022.

[22] K. Guo, K. An, F. Zhou, T. Tsiftsis, G. Zheng, and S. Chatzainotas, "On the secrecy performance of NOMA-based integrated satellite multiple-terrestrial relay networks with hardware impairments," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3661–3676, 2021.

[23] X. Zhang, B. Zhang, K. An, G. Zheng, S. Chatzinotas, and D. Guo, "Stochastic geometry-based analysis of cache-enabled hybrid satellite-aerial-terrestrial networks with non-orthogonal multiple access," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 1272–1287, 2021.

[24] X. Zhang, B. Zhang, K. An et al., "On the performance of hybrid satellite-terrestrial content delivery networks with non-orthogonal multiple access," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 454–458, 2021.

[25] V.-L. Dao, L.-N. Hoang, S. Girs, and E. Uhlemann, "Outage performance of pairwise NOMA allowing a dynamic decoding order and optimal pairs of power levels," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1886–1906, 2020.

[26] F. Li, H. Jiang, R. Fan, and P. Tan, "Cognitive non-orthogonal multiple access with energy harvesting: an optimal resource allocation approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 7080–7095, 2019.

[27] G. Li, D. Mishra, and H. Jiang, "Resource allocation in power-beacon-assisted IoT networks with nonorthogonal multiple access," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14385–14398, 2021.

[28] Y. Yu, H. Chen, Y. Li, Z. Ding, L. Song, and B. Vucetic, "Antenna selection for MIMO non-orthogonal multiple access systems," *IEEE Transactions on vehicular technology*, vol. 67, no. 4, pp. 3158–3171, 2018.

[29] Y. Li, Y. Li, X. Chu, Y. Ye, and H. Zhang, "Performance analysis of relay selection in cooperative NOMA networks," *IEEE Communications Letters*, vol. 23, no. 4, pp. 760–763, 2019.

[30] J. Zhang, L. Zhu, Z. Xiao, X. Cao, D. O. Wu, and X.-G. Xia, "Optimal and sub-optimal uplink NOMA: joint user grouping, decoding order, and power control," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 254–257, 2020.

[31] G. Li, D. Mishra, and H. Jiang, "Channel-aware power allocation and decoding order in overlay cognitive NOMA networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6511–6524, 2020.

[32] I. S. Gradshteyn and I. M. Ryzhik, *Table of Intergals, Series, and Products*, Academic Press, Amsterdam, The Netherlands, 7th edition, 2007.

[33] N. I. Miridakis, D. D. Vergados, and A. Michalas, "Dual-hop communication over a satellite relay and shadowed Rician channels," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4031–4040, 2015.

[34] *The Wolfram Function Site*http://functions.wolfram.com.

[35] X. Yan, H. Xiao, C. Wang, and K. An, "Outage performance of NOMA-based hybrid satellite-terrestrial relay networks," *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 538–541, 2018.

[36] X. Li, M. Huang, Y. Liu, V. G. Menon, A. Paul, and Z. Ding, "I/Q imbalance aware nonlinear wireless-powered relaying of B5G networks: security and reliability analysis," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 4, pp. 2995–3008, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7424854, 9 pages
https://doi.org/10.1155/2022/7424854

Research Article

UCB-Based Route and Power Selection Optimization for SDN-Enabled Industrial IoT in Smart Grid

Xiaoyue Li, Chaoqun Zhou, Zilong Liang, Qiang Yu, Xiankai Chen, and Zhiyuan He

State Grid Qingdao Power Supply Company, Qingdao 266000, China

Correspondence should be addressed to Xiaoyue Li; lxy-qdpower@outlook.com

Received 28 November 2021; Revised 18 December 2021; Accepted 24 December 2021; Published 20 March 2022

Academic Editor: Xingwang Li

As an essential building block for smart grid, the industrial internet of things (IIoT) plays a significant role in providing powerful sensing capability and ubiquitous connectivity for differentiated power services. The rapid development of smart grid imposes higher data monitoring and transmission requirements in terms of delay and energy efficiency. However, due to the severe electromagnetic interference (EMI) caused by massive electrical equipment, the transmission performance of IIoT becomes inferior. The traditional single-hop transmission mode evolves towards a multihop cooperation mode to satisfy differentiated quality of service (QoS) requirements. In this paper, we propose an upper confidence bound- (UCB-) based joint route and power selection optimization algorithm to support multihop cooperation mode evolution, which adopts a software-defined networking- (SDN-) enabled IIoT network framework to simplify network configuration and management. Compared with existing local-side-information-based route selection (LSI-RS) and random route selection (RRS) algorithms, simulation results demonstrate that the proposed algorithm has superior performances in total delay, energy efficiency, and utility.

1. Introduction

The industrial internet of things (IIoT) is an essential building block for smart grid, which has powerful sensing capability and ubiquitous connectivity. With the development of smart grid, a large number of IIoT devices need to be deployed to collect information such as voltage, current, power, temperature, and humidity and transmit the information back for real-time analysis. IIoT has strict requirements on transmission delay, energy efficiency, and network coverage [1]. However, electrical equipment in smart grid emits electromagnetic interference (EMI), which affects the transmission performance of IIoT. Therefore, the traditional single-hop transmission mode needs to evolve towards a multihop cooperation mode to satisfy the quality of service (QoS) requirements [2]. IIoT utilizes massive devices laid in different routes to form a mesh network for multihop transmission. In multihop transmission, dynamic route selection can avoid worse routes with long distance and low quality and enable IIoT to reduce transmission delay and enhance energy efficiency.

Routing selection needs to be optimized according to the dynamic network environment. However, the traditional network architecture with tight coupling between control and data planes cannot adapt to complex IIoT application scenarios. Software-defined networking (SDN) provides a solution by separating the control plane from the data plane [3]. SDN can manage and control IIoT networks through a standard and open programmable interface, which supports more efficient and flexible route selection solutions [4]. However, the research on route selection optimization for SDN-enabled IIoT in smart grid still faces many challenges, which are summarized below.

First, considering the highly time-varying channel states and complex EMI, the global state information (GSI) is incomplete [5]. Traditional GSI-based route selection optimization cannot be applied. Second, IIoT devices based on battery have strict requirements on energy efficiency. Improving energy efficiency via dynamic power selection not only makes route selection more complicated but also possibly leads to larger transmission delay. Therefore, how to meet differentiated

QoS requirements through joint route and power selection optimization is also a challenge. Finally, electric equipment such as inverters and insulation switches emits EMI [6, 7], which greatly reduces QoS performance and brings severe challenges for joint route and power selection optimization.

Route selection of IoT has always been a research hotspot. In [8], Desuo et al. employed an improved Dijkstra algorithm to find the shortest path between two consecutive points for IoT networks. In [9], He et al. proposed an energy-aware route selection algorithm for simultaneous information and power transfer to decrease energy consumption. However, these works do not consider SDN architecture and only consider a single QoS metric. In [10], Saha et al. proposed a traffic-aware QoS route selection scheme by exploiting the flow-based nature of SDN and obtained the optimal route based on Yen's K-shortest path algorithm. In [11], Li et al. proposed an SDN-enabled IoT adaptive transmission architecture for different delay flow situations. However, these works assume that perfect GSI is available, which is not applicable for smart grid with incomplete information.

Upper confidence bound (UCB) as a reinforcement learning algorithm has emerged as a powerful solution to address problems without perfect GSI [12]. In [13], Sun et al. designed an energy-aware mobility management (EMM) scheme based on UCB to optimize energy consumption. In [14], Maghsudi and Stanczak proposed two joint power and channel selection strategies based on UCB to maximize energy efficiency. However, these works only consider energy consumption optimization, which ignore delay and other QoS requirements. In [15], Zhao et al. proposed a delay minimization algorithm based on UCB, but neglected the joint optimization of energy consumption and delay. In [16], Bae et al. proposed a downlink network routing algorithm based on UCB to jointly optimize throughput and delay, but ignored the influence of complex EMI and service priority. Moreover, all the abovementioned works do not consider the impact of complex EMI and service priority of smart grid on the joint optimization of route and power selection.

To address the abovementioned challenges, we propose a UCB-based joint route and power selection optimization algorithm. Firstly, considering the influence of EMI, we construct an SDN-based multihop IIoT framework and formulate the joint route and power selection optimization problem. The objective is to maximize the overall network utility function under the threshold constraints of signal-to-interference-plus-noise ratio (SINR) and energy efficiency. Second, we model the joint optimization problem as a multiarmed bandit (MAB) problem, where the options of route and power are combined to form an arm. Finally, we utilize UCB to learn the optimal route and power combination based on local and historical information. The main contributions of this work are summarized as follows:

(i) We propose an SDN-enabled multihop IIoT framework for smart grid, which greatly simplifies network management through separating control and data planes. In addition, the control plane also supports the configuration of intelligent route and power selection algorithms.

(ii) The route and power options are combined to form a set of arms in MAB. The proposed algorithm dynamically learns the optimal combination by interacting with the environment.

(iii) Through dynamically adjusting the values of weight parameters, the proposed algorithm can satisfy differentiated QoS requirements of smart grid by adjusting the tradeoff between delay, energy efficiency, and service priority.

The rest of this paper is organized as follows: Section 2 describes the system model and problem formulation. The proposed joint route and power selection algorithm is introduced in Section 3. Section 4 provides simulation results. Finally, the conclusion is provided in Section 5.

2. System Model

In this section, the system model and the problem formulation are introduced.

2.1. Network Model of SDN-Enabled IIoT for Smart Grid. The SDN-enabled IIoT for smart grid is shown in Figure 1, which consists of two planes, i.e., the data plane and the control plane [17]. The data plane mainly contains IIoT devices which provide data forwarding services. The control plane mainly contains the SDN controller, which locates in the gateway. The SDN controller can obtain IIoT network topology, learn the optimal route and power selection strategy, and send the strategy to the IIoT source device (SD) [18].

The SDN-enabled IIoT network topology is represented by a directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{L})$ [19], where $\mathscr{V}$ denotes I IIoT devices. The set is defined as $\mathscr{V} = \{v_1, \ldots, v_i, \ldots, v_I\}$. v_1 and v_I are the SD and destination device (DD). v_i, $i = 2, 3, \ldots, I-1$, is the relay device. $\mathscr{L}$ denotes physical links, and the set is defined as $\mathscr{L} = \{L(v_i, v_j) | v_i \in \mathscr{V}, v_j \in \mathscr{V}_{N(v_i)}\}$, where $\mathscr{V}_{N(v_i)}$ is the set of devices connected with v_i. There exist M routes between v_1 and v_I, and the set is represented as $\mathscr{F} = \{f_1, \ldots, f_m, \ldots, f_M\}$. Each route consists of K devices, which are SD v_1, DD v_I, and $K-2$ relay devices. The set of devices in f_m is denoted as $\mathscr{A}_m = \{d_m^1, \ldots, d_m^k, \ldots, d_m^k\}$ in the order from SD to DD, where $\mathscr{A}_m \subseteq \mathscr{V}$, $L(d_m^k, d_m^{k+1}) \in \mathscr{L}, 1 \le k \le K$.

In this paper, the set of T time slots is represented as $\mathscr{T} = \{1, \ldots, t, \ldots, T\}$. The slot length depends on the transmission delay from SD to DD [20]. At the beginning of the t-th slot, v_1 generates a data packet of size U_t, $U_{\min} \le U_t \le U_{\max}$, which needs to be transmitted to v_I. Each data packet can only be transmitted in one route [21]. The transmission is unsuccessful if the delay exceeds τ.

2.2. Delay Model. We assume that the data packets are transmitted by wireless channels. We denote P_h as the transmission power, which contains H levels. The set of transmission power levels is given by

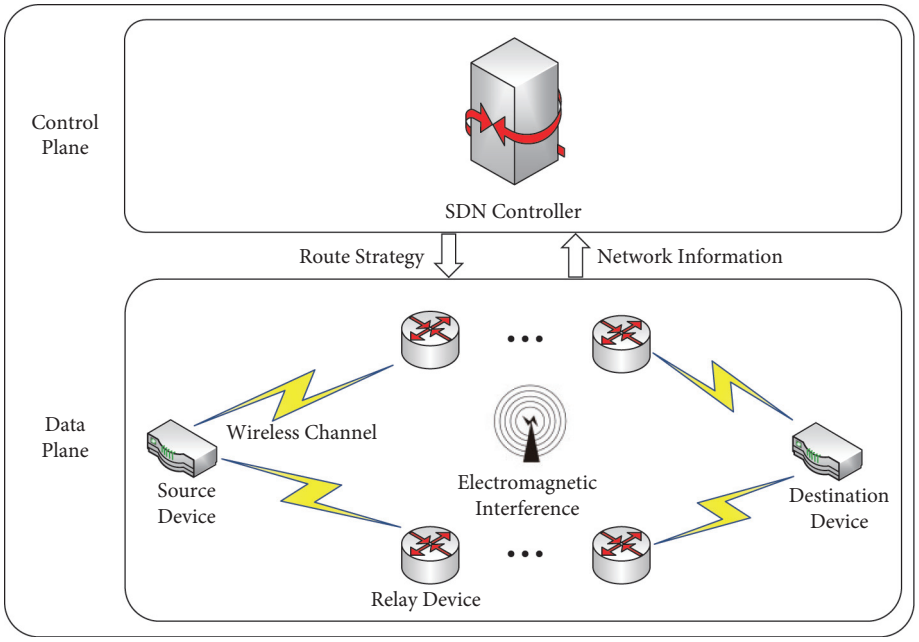


Figure 1: SDN-enabled IIoT framework for smart grid.

$$\mathscr{P} = \left\{ P_{\min}, \ldots, P_{\min} + \frac{(h-1)(P_{\max} - P_{\min})}{H-1}, \ldots, P_{\max} \right\}. \tag{1}$$

The achievable transmission rate from d_m^k to d_m^{k+1} is given by

$$R_t^{d_m^k, d_m^{k+1}}(h) = B(m)\log_2\left(1 + \mathrm{SINR}^{d_m^k, d_m^{k+1}}(h)\right), \tag{2}$$

where $B(m)$ is the transmission bandwidth of route f_m. $\mathrm{SINR}^{d_m^k, d_m^{k+1}}(h)$ is the SINR [22] between d_m^k and d_m^{k+1} and is given by

$$\mathrm{SINR}^{d_m^k, d_m^{k+1}}(h) = \frac{P_h g_t^{d_m^k, d_m^{k+1}}}{\sigma_0 + \lambda_t^{d_m^k, d_m^{k+1}}}, \tag{3}$$

where $g_t^{d_m^k, d_m^{k+1}}$ is the channel gain. σ_0 is the noise power. $\lambda_t^{d_m^k, d_m^{k+1}}$ is the EMI power.

We denote the power selection variable as $y_t(h) \in \{0, 1\}$. $y_t(h) = 1$ represents that v_1 selects P_h; otherwise, $y_t(h) = 0$. The transmission delay from d_m^k to d_m^{k+1} and the total forwarding delay on the route f_m are given by

$$\begin{aligned} D_t^{d_m^k, d_m^{k+1}}(h) &= \frac{U_t}{R_t^{d_m^k, d_m^{k+1}}(h)}, \\ D_t(m, h) &= \sum_{k=1}^{K-1} D_t^{d_m^k, d_m^{k+1}}(h). \end{aligned} \tag{4}$$

We denote the route selection variable $x_t(m) \in \{0, 1\}$. $x_t(m) = 1$ represents that v_1 selects f_m; otherwise, $x_t(m) = 0$ [23]. The total forwarding delay is given by

$$\tau_t = \sum_{m=1}^{M} \sum_{h=1}^{H} y_t(h) x_t(m) D_t(m, h). \tag{5}$$

2.3. Energy Efficiency Model. The energy consumption for data packet transmission from d_m^k to d_m^{k+1} and the total energy consumption on route f_m are given by

$$\begin{aligned} E_t^{d_m^k, d_m^{k+1}}(h) &= D_t^{d_m^k, d_m^{k+1}}(h) P_h, \\ E_t(m, h) &= \sum_{k=1}^{K-1} E_t^{d_m^k, d_m^{k+1}}(h). \end{aligned} \tag{6}$$

We define $\gamma_t(m, h)$ as the energy efficiency of data packet transmission on route f_m with power $P_{\min} + ((h-1)(P_{\max} - P_{\min})/(H-1))$ in the t-th time slot, which is given by

$$\gamma_t(m, h) = \frac{U_t}{B(m) E_t(m, h)}. \tag{7}$$

Therefore, the total energy efficiency is given by

$$\alpha_t = \sum_{m=1}^{M} \sum_{h=1}^{H} x_t(m) y_t(h) \gamma_t(m, h). \tag{8}$$

2.4. Problem Formulation. Since the data packets have different QoS requirements, the service priority needs to be taken into account. We use η_t to represent the priorities of different data packets. We define the overall network utility function related to the total forwarding delay, service priority, and total energy efficiency as

$$\Phi = \frac{1}{T} \sum_{t=1}^{T} \left(\alpha_t + V \frac{\eta_t}{\tau_t} \right), \tag{9}$$

where V is the weight used to balance the order of magnitude.

Therefore, the objective is to maximize Φ by optimizing the route and power selection strategies. The optimization problem is formulated as

$$\begin{aligned}
&\mathbf{P1}\colon \max_{\{x_t(m), y_t(h)\}} \Phi,\\
&\text{s.t. } C_1\colon \sum_{m=1}^{M} x_t(m) = 1, \quad \forall t \in \mathcal{T},\\
&C_2\colon \sum_{h=1}^{H} y_t(h) = 1, \quad \forall t \in \mathcal{T},\\
&C_3\colon P_h \in \mathcal{P}, \quad h = 1, \ldots, H,\\
&C_4\colon \mathrm{SINR}^{d_m^k, d_m^{k+1}}(h) \geq \mathrm{SINR}_{\min}, \quad \forall d_m^k, d_m^{k+1} \in \mathcal{A}_m,\\
&C_5\colon \alpha_t \geq \alpha_{\min}, \quad \forall t \in \mathcal{T},
\end{aligned} \tag{10}$$

where $\mathrm{SINR}_{\min}$ and $\alpha_{\min}$ represent the thresholds of SINR and energy efficiency, respectively. C_1 is the route selection constraint; i.e., each data packet can only select one route. C_2 is the power selection constraint; i.e., each data packet can only select one power level. C_3 is the transmission power constraint. C_4 is the SINR constraint. C_5 is the energy efficiency constraint.

3. UCB-Based Route and Power Selection Optimization for SDN-Enabled Industrial IoT in Smart Grid

It is impractical to obtain the perfect GSI due to the dynamic network topology and complex EMI, and IIoT devices should optimize route and power selection based on the local-side information. MAB is an effective solution to solve decision-making problems with incomplete information [24]. In each slot, the decision maker pulls an arm. Then, the pulled arm generates a reward. The decision maker's goal is to maximize the cumulative reward.

We transform **P1** into an MAB problem. The decision maker, arm, and reward are modeled as follows:

(i) Decision maker: the decision maker generates the decision. In this paper, we define the SDN controller as the decision maker.

(ii) Arm: we define $\mathcal{C} = \{c_{1,1}, \ldots, c_{m,h}, \ldots, c_{M,H}\}$ as the set of arms which satisfy $|\mathcal{C}| = M \times H$, where $|\mathcal{C}|$ represents the number of elements in $\mathcal{C}$. The arm $c_{m,h}$ represents the route f_m and power P_h.

(iii) Reward: we define a reward function $\theta_t(c_{m,h})$ to represent the reward obtained by selecting $c_{m,h}$, which is given by

$$\theta_t(c_{m,h}) = \begin{cases} \gamma_t(m,h) + V\dfrac{\eta_t}{D_t(m,h)}, & \text{satisfy } C_3 \text{ and } C_4,\\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

If $\mathrm{SINR}^{d_m^k, d_m^{k+1}}(h) \geq \mathrm{SINR}_{\min}$ and $\alpha_t \geq \alpha_{\min}$, the reward is $\gamma_t(m,h) + V(\eta_t/D_t(m,h))$. Otherwise, the reward is zero.

We propose a UCB-based joint route and power selection algorithm for SDN-enabled IIoT in smart grid to address the MAB problem. UCB is a low-complexity learning-based algorithm to balance exploitation and exploration [25]. The proposed algorithm enables the SDN controller to take action based on local state information such as delay. Afterwards, the obtained reward and updated state information is perceived by the SDN controller for the next selection. The implementation of the proposed algorithm is shown in Figure 2.

The proposed algorithm consists of three phases, which is summarized in Algorithm 1.

(i) Phase I: $x_t(m)$, $y_t(h)$, $\theta_t(c_{m,h})$, and $\hat{n}_t(c_{m,h})$ are initialized as zero. When $t \leq |\mathcal{C}|$, the controller sequentially selects each arm and obtains the initial value.

(ii) Phase II: based on (12), the preference of the SD towards arm $c_{m,h}$ in the t-th slot is given by

$$\tilde{\theta}_t(c_{m,h}) = \bar{\theta}_{t-1}(c_{m,h}) + \omega\sqrt{\frac{\ln t}{\hat{n}_{t-1}(c_{m,h})}}, \tag{12}$$

where $\bar{\theta}_{t-1}(c_{m,h})$ is the average reward of $c_{m,h}$ up to the $(t-1)$-th time slot. $\hat{n}_{t-1}(c_{m,h})$ is the number of times to select $c_{m,h}$. ω is the weight of exploration. The second term allows the controller to explore arms with selections to improve estimation and to focus on the exploitation when arms have been estimated enough. After obtaining $\tilde{\theta}_t(c_{m,h})$, the selected arm is given by

$$c_{m^*,h^*} = \mathrm{argmax}_{\mathcal{C}}\{\tilde{\theta}_t(c_{m,h})\}. \tag{13}$$

c_{m^*,h^*} represents SD selects f_{m^*} and P_{h^*}, which is given by

$$c_{m^*,h^*} \Rightarrow \{x_t(m^*) = 1, y_t(h^*) = 1\}. \tag{14}$$

(iii) Phase III: the controller observes delay and energy efficiency performances as well as service priority. Then, $\theta_t(c_{m^*,h^*})$ is updated as (11). Accordingly, $\bar{\theta}_t(c_{m,h})$ and $\hat{n}_t(c_{m,h})$ are updated as

$$\bar{\theta}_t(c_{m,h}) = \frac{\bar{\theta}_{t-1}(c_{m,h})\hat{n}_{t-1}(c_{m,h}) + \theta_t(c_{m,h})x_t(m)y_t(h)}{\hat{n}_{t-1}(c_{m,h}) + x_t(m)y_t(h)}, \tag{15}$$

$$\hat{n}_t(c_{m,h}) = \hat{n}_{t-1}(c_{m,h}) + x_t(m)y_t(h). \tag{16}$$

Finally, the algorithm terminates until $t > T$.

4. Simulation Results

In this section, we firstly introduce the simulation parameter setting. Then, the simulation analysis is described.

4.1. Simulation Parameter Setting. In this section, we evaluate the proposed algorithm through simulations. The

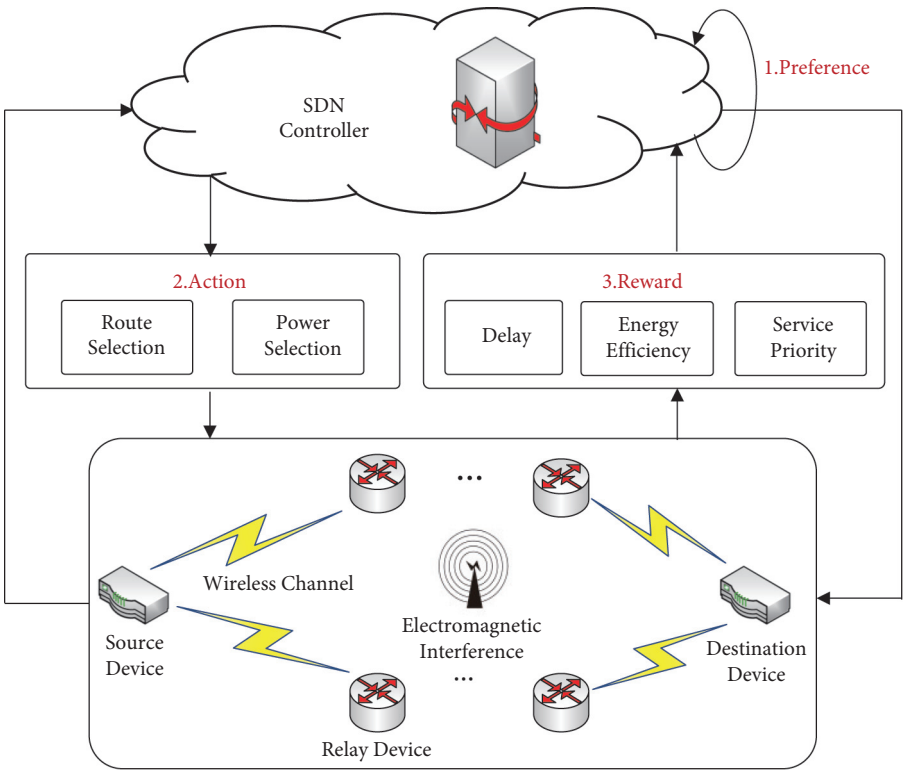


Figure 2: The implementation of the proposed algorithm.

(1) **Input**: ω, $\{\eta_t\}$, $\{D_t(m,h)\}$, $\{\gamma_t(m,h)\}$.
(2) **Phase 1**:
(3) Set $x_t(m) = 0$, $y_t(h) = 0$, $\theta_t(c_{m,h}) = 0$, and $\hat{n}_t(c_{m,h}) = 0$, $\forall t \in \mathcal{T}, \forall f_m \in \mathcal{F}, \forall P_h \in \mathcal{P}$.
(4) **for** $t = 1$ to $|\mathcal{C}|$ **do**
(5) Select arms sequentially, and obtain the initial values.
(6) **end for**
(7) **for** $t = |\mathcal{C}| + 1$ to T **do**
(8) **Phase 2**:
(9) Calculate the preference of the SD towards arm $c_{m,h}$ as (12).
(10) Select c_{m^*,h^*} based on (13).
(11) **Phase 3**:
(12) Observe delay and energy efficiency performances.
(13) Calculate $\theta_t(c_{m^*,h^*})$ based on (11).
(14) Update $\bar{\theta}_t(c_{m,h})$ and $\hat{n}_t(c_{m,h})$ based on (15) and (16).
(15) **end for**

Algorithm 1: UCB-based route and power selection optimization for SDN-enabled industrial IoT in smart grid.

considered IIoT route topology is shown in Figure 3, which includes 9 IIoT devices and 6 routes. v_1 and v_9 are the SD and DD, respectively. The distances of adjacent devices on each route are shown in Table 1. In the case of large-scale fading, the channel gain is calculated according to $g_{i,j}(t) = 127 + 30\log(r_{i,j})$ [26], where $r_{i,j}$ is the distance between v_i and v_j. The EMI varies from 28 dBm to 30 dBm. The service priority η_t is set as $[0.1, 0.2, 0.3, 0.4, 0.5]$. The setting of simulation parameters is summarized in Table 2 [27, 28]. We consider two existing algorithms for comparison. The first one is the UCB-based route selection algorithm named UCB-RS [29]. The other one is the shortest route selection algorithm named SRS [30]. Both UCB-RS and SRS neglect the optimization of power selection.

4.2. Simulation Analysis. Figure 4 shows the average utility versus time slot. Compared with UCB-RS and SRS, the simulation result demonstrates that the proposed algorithm improves the performance of utility by27.28% and 37.31%, respectively. The reason is that the proposed algorithm jointly optimizes the route and power selection. In contrast,

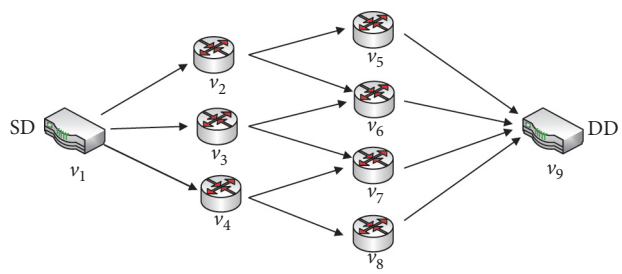


Figure 3: IIoT route topology.

Table 1: Distance between devices.

Path	Distance (m)
$v_1 \longrightarrow v_2$	50
$v_1 \longrightarrow v_4$	70
$v_2 \longrightarrow v_6$	50
$v_3 \longrightarrow v_7$	50
$v_4 \longrightarrow v_8$	70
$v_6 \longrightarrow v_9$	30
$v_8 \longrightarrow v_9$	70
$v_1 \longrightarrow v_3$	30
$v_2 \longrightarrow v_5$	50
$v_3 \longrightarrow v_6$	30
$v_4 \longrightarrow v_7$	50
$v_5 \longrightarrow v_9$	70
$v_7 \longrightarrow v_9$	70

Table 2: Simulation parameter settings.

Simulation parameters	Value
The number of time slots T	1000
The number of IIoT devices M	9
Transmission power P_h (W)	$[0.1, 0.2, 0.3, 0.4, 0.5]$
Packet size U_t (Mbits)	$[1, 2]$
The number of transmission power levels H	5
Transmission bandwidth B (MHz)	$[2, 3]$
Noise power σ_0 (dBm)	-114
EMI λ_t (dBm)	$[28, 30]$
Service priority η_t	$[0.1, 0.2, 0.3, 0.4, 0.5]$
Exploration weight ω	2
The weight V	12

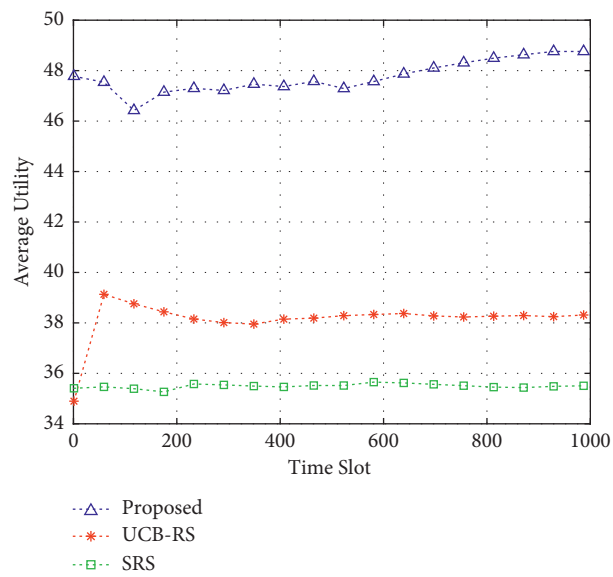


Figure 4: **Q2-2:** average utility versus time slot.

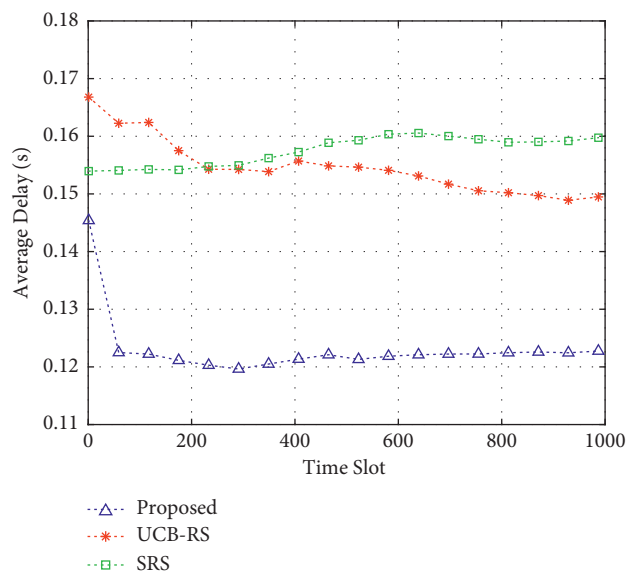


Figure 5: Average delay versus time slot.

UCB-RS neglects the power selection. SRS always selects the shortest route, which cannot overcome the adverse impact caused by the dynamic change of channel state, thereby performing the worst.

Figure 5 shows the average delay versus time slot. The simulation result shows that the proposed algorithm outperforms UCB-RS and SRS by 17.93% and 23.17% in delay performance, respectively. Both UCB-RS and SRS do not take the optimization of power selection into consideration, which result in worse delay performance.

Figure 6 shows the average energy efficiency versus time slot. Compared with UCB-RS and SRS, the proposed algorithm improves the performance of energy efficiency by 28.61% and 51.15%, respectively. The proposed algorithm can select suitable power to optimize energy efficiency.

Figure 7 shows the ratio of optimal route selection versus time slot. SRS performs the worst. The reason is that the proposed algorithm and UCB-RS can dynamically adjust the route selection strategy. However, SRS always selects the shortest route fixedly and cannot get rid of the adverse impact of EMI.

Figures 8–10 show the average energy efficiency, average delay, and average utility versus $\mathrm{SINR}_{\min}$. With the increase of $\mathrm{SINR}_{\min}$, the energy efficiency and delay of the proposed algorithm decrease, while the utility increases

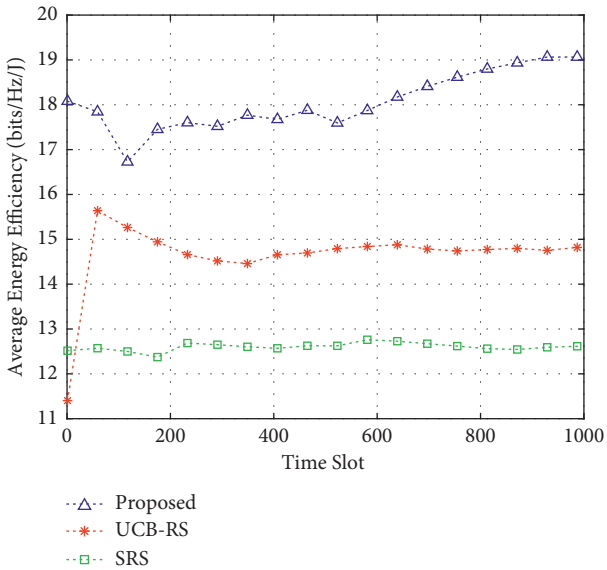


Figure 6: Average energy efficiency versus time slot.

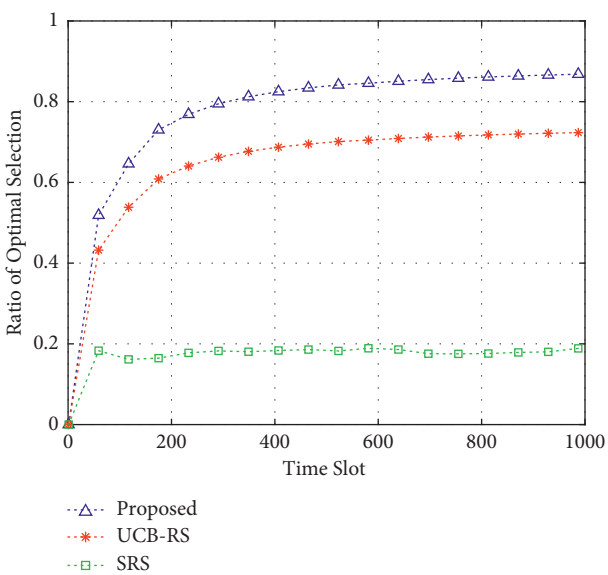


Figure 7: Ratio of optimal selection versus time slot.

first and then decreases. When $\mathrm{SINR}_{\min} = 14\,\mathrm{dB}$, the utility reaches the maximum value. The performance of UCB-RS fluctuates. The reason is that the proposed algorithm can learn to optimize power selection to meet more stringent SINR constraint. UCB-RS neglects power selection, which makes it difficult to adapt to different SINR constraints.

Figure 11 shows the impact of V on the delay and energy efficiency. As V increases, the proposed algorithm lays more emphasis on energy efficiency rather than delay. The proposed algorithm can dynamically balance the tradeoff between energy efficiency and delay. Moreover, the simulation results provide a reference for the setting of the weight V.

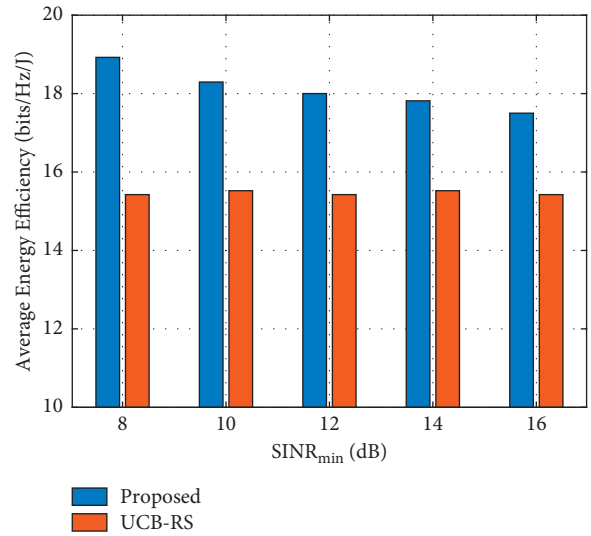


Figure 8: Average energy efficiency versus $\mathrm{SINR}_{\min}$.

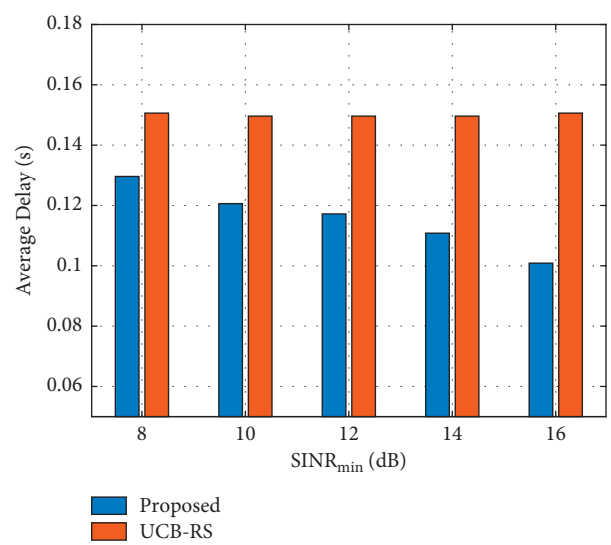


Figure 9: Average delay versus $\mathrm{SINR}_{\min}$.

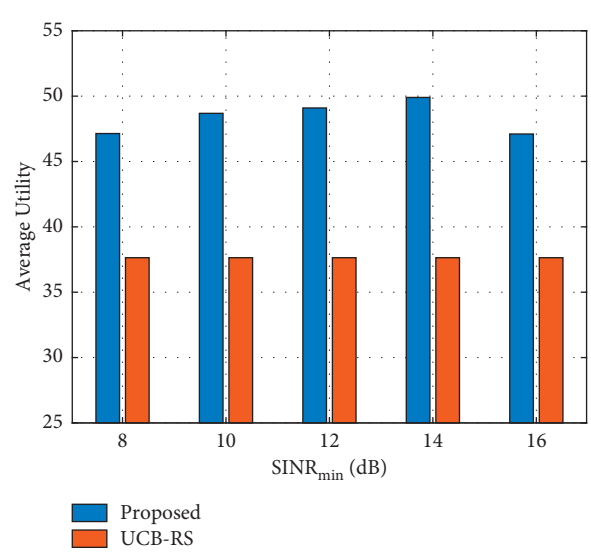


Figure 10: Average utility versus $\mathrm{SINR}_{\min}$.

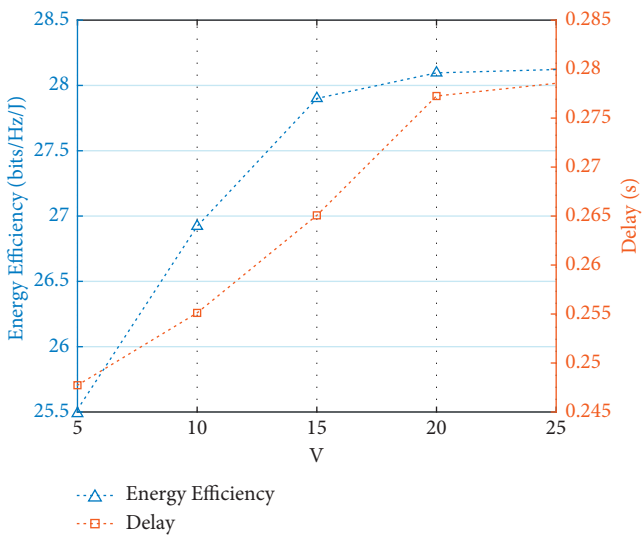


FIGURE 11: The impact of V on the delay and energy efficiency.

5. Conclusions

In this paper, we proposed an UCB-based joint route and power selection optimization algorithm for SDN-enabled IIoT. The proposed algorithm can effectively optimize route and power selection strategies based only on local information and historical observations. It can provide a low-complexity route and power selection strategy while maximizing the overall network utility. Simulation results show that the proposed algorithm has superior performances in delay, energy efficiency, and utility. Compared with existing LSI-RS and RRS algorithms, the proposed algorithm reduces the delay by 17.93% and 23.17%, improves the utility by 27.28% and 37.31%, and improves the energy efficiency by 28.61% and 51.15%. In the future, we will use deep reinforcement learning to optimize the multidimensional resource allocation in SDN-enabled IIoT.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Science and Technology Project of the State Grid Shandong Power Supply Company under grant no. 5206021900VV.

References

[1] L. You, Y. Huang, D. Zhang, Z. Chang, W. Wang, and X. Gao, "Energy efficiency optimization for multi-cell massive MIMO: centralized and distributed power allocation algorithms," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5228–5242, 2021.

[2] G. Chen, J. P. Coon, and S. E. Tajbakhsh, "Secure routing for multihop Ad Hoc networks with inhomogeneous eavesdropper clusters," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10660–10670, 2018.

[3] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[4] Z. Zhou, J. Gong, Y. He, and S. E. Tajbakhsh, "Software defined machine-to-machine communication for smart energy management," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 52–60, 2017.

[5] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[6] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA system: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[7] B. Wang, Y. Sun, Q. Cao, S. Li, and Z. Sun, "Bandwidth slicing for socially-aware D2D caching in SDN-enabled networks," *IEEE Access*, vol. 6, no. 99, pp. 50910–50926, 2018.

[8] L. Desuo, M. Bessani, R. Fanucchi, T. Gross, and C. D. Maciel, "A multi-objective swarm intelligence approach for field crews patrol optimization in power distribution systems restoration," *IEEE Latin America Transactions*, vol. 17, no. 2, pp. 338–346, 2019.

[9] S. He, K. Xie, W. Chen, D. Zhang, and J. Wen, "Energy-aware routing for SWIPT in multi-hop energy-constrained wireless network," *IEEE Access*, vol. 6, no. 99, pp. 17996–18008, 2018.

[10] N. Saha, S. Bera, and S. Misra, "Sway: traffic-aware QoS routing in software-defined IoT," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 1, pp. 390–401, 2021.

[11] X. Li, D. Li, J. Wan, C. Liu, and M. Imran, "Adaptive transmission optimization in SDN-based industrial internet of things with edge computing," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1351–1360, 2018.

[12] M. Tariq, M. Ali, F. Naeem, and H. V. Poor, "Vulnerability assessment of 6G-enabled smart grid cyber–physical systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5468–5475, 2021.

[13] Y. Sun, S. Zhou, and J. Xu, "EMM: energy-aware mobility management for mobile edge computing in ultra dense networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2637–2646, 2017.

[14] S. Maghsudi and S. Stanczak, "Joint channel selection and power control in infrastructureless wireless networks: a multi-player multi-armed bandit framework," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4565–4578, 2014.

[15] Y. Zhao, J. Lee, and W. Chen, "Q-greedy UCB: a new exploration policy to learn resource-efficient scheduling," *China Communications*, vol. 18, no. 6, pp. 12–23, 2021.

[16] J. Bae, J. Lee, and S. Chon, "Learning to schedule network resources throughput and delay optimally using Q+-learning," *IEEE/ACM Transactions on Networking*, vol. 29, no. 2, pp. 750–763, 2021.

[17] A. Singh, G. Aujla, S. Garg, S. Kaddoum, and G. Singh, "Deep-learning-based SDN model for internet of things: an incremental tensor train approach," *IEEE Internet of Things Journal*, vol. 6, no. 7, pp. 6302–6311, 2020.

[18] F. Naeem, M. Tariq, and H. V. Poor, "SDN-enabled energy-efficient routing optimization framework for industrial

internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5660–5667, 2021.

[19] X. Lai, X. Ji, X. Zhou, and L. Chen, "Energy efficient link-delay aware routing in wireless sensor networks," *IEEE Sensors Journal*, vol. 18, no. 2, pp. 837–848, 2018.

[20] Z. Zhou, Y. Guo, Y. He, X. Zhao, and W. M. Bazzi, "Access control and resource allocation for M2M communications in industrial automation," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 3093–3103, 2019.

[21] Y. Li and Y. Liang, "Compressed sensing in multi-hop large-scale wireless sensor networks based on routing topology tomography," *IEEE Access*, vol. 6, no. 99, pp. 27637–27650, 2018.

[22] S. Shi, S. Pang, Y. Li, F. Wang, H. Gacanin, and D. Zhang, "Buffer-aided relaying network with hybrid BNC for the internet of things: protocol and performance analysis," *IEEE Access*, vol. 8, no. 99, pp. 19646–19656, 2020.

[23] Y. Wang, R. Zhao, Y. Huang et al., "Asymptotic performance analysis of massive MIMO relay systems with multi-pair devices over correlated fading channels," *IEEE Access*, vol. 7, no. 99, pp. 27565–27578, 2019.

[24] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, "A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in internet of vehicle," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4550–4559, 2021.

[25] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, USA, 2018.

[26] H. Yu, Z. Zhou, Z. Jia, X. Zhao, L. Zhang, and X. Wang, "Multi-timescale multi-dimension resource allocation for NOMA-edge computing-based power IoT with massive connectivity," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1101–1113, 2021.

[27] Z. Wang, Z. Jia, and H. Liao, "Energy-aware and URLLC-aware task offloading for internet of health things," in *Proceedings of the GLOBECOM 2020-2020 IEEE Global Communications Conference*, pp. 1–6, Taipei, Taiwan, February 2021.

[28] Z. Jia, H. Liao, and Z. Zhou, "Multi-dimension resource allocation for NOMA-edge computing-based 6G power IoT," in *Proceedings of the 2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Montreal, QC, Canada, June 2021.

[29] X. He, H. Jiang, Y. Song, C. He, and H. Xiao, "Routing selection with reinforcement learning for energy harvesting multi-hop CRN," *IEEE Access*, vol. 7, no. 99, pp. 54435–54448, 2019.

[30] H. Liao, Z. Zhou, W. Kong et al., "Learning-based intent-aware task offloading for air-ground integrated vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5127–5139, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6795205, 10 pages
https://doi.org/10.1155/2022/6795205

Research Article

Matching Learning-Based Relay Selection for Substation Power Internet of Things

Wei Wang,[1,2] **Ruiqiuyu Wang**,[1,2] **Hao Zhang**,[1,2] **Zhenyu Zhou**,[1,2] **and Yanhua He**[3]

[1]*State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing, China*
[2]*National Mobile Communications Research Laboratory, Southeast University, Nanjing, Jiangsu, China*
[3]*State Grid Zhejiang Electric Power Company Ningbo Bureau, Hangzhou, Zhejiang, China*

Correspondence should be addressed to Hao Zhang; haozhang@ncepu.edu.cn

Received 30 November 2021; Revised 12 January 2022; Accepted 27 January 2022; Published 21 February 2022

Academic Editor: Xingwang Li

Wireless sensor network (WSN) can effectively solve the problems of weak coverage, blind coverage, and low survivability of smart substation communication networks by deploying multiple relay nodes and adopting multihop transmission. However, there are still some challenges in the traditional relay selection strategy of WSN in substation, including incomplete information and the selection conflicts among multisource nodes. In this paper, we propose a matching learning-based relay selection mechanism for WSN-based substation power Internet of things (SPIoT). Firstly, considering the electromagnetic interference caused by the operation of high-voltage equipment, a multihop transmission model of SPIoT is built. Furthermore, based on the upper confidence bound (UCB) algorithm and matching theory, a matching learning-based relay selection (MLRS) algorithm is proposed to minimize the energy consumption of SPIoT devices. Simulation results demonstrate that MLRS outperforms existing algorithms in terms of energy consumption and optimal selection probability.

1. Introduction

Substation has great significance to ensure long-distance power transmission and stable power supply [1]. To achieve 24-hour uninterrupted monitoring of the substation, a large number of substation Internet of things (SPIoT) devices are deployed to collect various types of information including temperature, humidity, and smoke [2, 3]. However, traditional fiber-optical communication cannot meet the on-demand coverage and data transmission requirements of SPIoT due to poor scalability and high construction cost [4, 5]. Wireless sensor network (WSN) has the advantages of flexible networking and low deployment cost. WSN adopts multihop transmission to realize coverage enhancement and increase fault tolerance for on-demand monitoring of substations [6, 7]. It is complementary to fiber-optical communication and acts as an effective enabler for SPIoT.

In multihop transmission of WSN-enabled SPIoT, the source node selects relays for data transmission to shorten the transmission distance and enhance coverage under strong electromagnetic interference [8, 9]. To fully utilize the spectrum and energy resources, relay selection needs to be optimized dynamically according to the network state and service requirements. However, relay selection optimization in SPIoT still faces several critical challenges as below [10].

First, the relay selection strategies are coupled across different devices when multiple devices compete for the same relay. Each device faces an adversarial relay selection problem in which its strategy is affected by the strategies of other devices. Second, the global state information (GSI), including channel state information (CSI) and electromagnetic interference [11], is uncertain due to the limitations of network sensing capability and signaling overhead. The devices are forced to optimize relay selection under incomplete state information. Finally, SPIoT devices impose strict requirements on energy consumption due to limited

battery capacity. It is necessary to take the optimization of energy consumption into consideration, thereby improving the sustainability of SPIoT networks.

There exist some works that have addressed relay selection problems in IoT. In [12], Muller and Speidel proposed optimal relay selection strategies aiming at either maximizing the mean mutual information or minimizing the outage probability. In [13], Mousavi et al. proposed a relay subset selection method for two-hop WSN. However, these works ignore the optimization of energy consumption, and it is infeasible for SPIoT devices with limited battery. In [14], an analysis model for relay selection under the constraint of energy consumption was developed. In [15], Bakhsh et al. proposed a low-energy distributed relay selection algorithm to achieve reliable communication. However, the above studies have neglected the decision coupling among multiple devices.

Matching theory provides a method to solve combination problem and is widely used in relay selection optimization. In [16, 17], Wang et al. and Baidas et al. proposed a relay selection method based on matching theory, but the establishment of matching preference list requires complete GSI. In substations with dynamic network environment and complex electromagnetic interference, the preference list cannot be constructed without complete GSI, thereby making the traditional matching theory-based relay selection approaches unsuitable.

Reinforcement learning (RL) provides a powerful tool to deal with multistage decision-making problems under incomplete information [18, 19]. In [20], Su et al. proposed a deep RL-based relay selection scheme to achieve lower outage probability and higher utility. In [21], Liang et al. modeled the resource allocation problem in vehicular networks as a semi-Markov decision process and utilized DL algorithms to solve the problem. However, when dealing with problems with high-dimensionality spaces, RL invokes the dimensionality curse and has inferior performances in optimality and convergence [22, 23]. The relay selection strategies obtained by RL are unstable, and the influence of electromagnetic interference is ignored.

Motivated by the aforementioned challenges, we propose a matching learning-based relay selection (MLRS) algorithm to minimize the energy consumption of SPIoT devices. First, considering the electromagnetic interference of substations, a two-hop transmission model for SPIoT with electromagnetic interference is established. Second, we transform the relay selection problem into a one-to-one matching problem between multisource nodes and multirelay nodes. Based on the upper confidence bound (UCB) algorithm, the SPIoT gateway learns to build the preference lists of source nodes based on the number of selections and empirical performances. Third, the matching conflicts between multisource nodes and multirelay nodes are resolved based on matching with rising price. Finally, we compare MLRS with existing relay selection algorithms to validate its performance. The major contributions are presented as follows:

(i) **Learning-based matching preference construction without precise GSI:** MLRS utilizes UCB to learn to construct preference lists based on historical observations of relay node selection times and empirical energy consumption performances. MLRS enables the implementation of iterative matching without precise GSI and achieves an effective compromise between exploration and exploitation.

(ii) **Stable matching between source and relay nodes:** MLRS utilizes matching theory to avoid selection conflicts between multiple sources and relay nodes, which achieve a stable matching based on the learned preference lists. In MLRS, the conflicts among multiple nodes are resolved by iteratively raising matching prices, and each source node is allocated with the most suitable relay node that ranks highest in its updated preference lists.

(iii) **Low energy consumption:** MLRS can dynamically learn the relay selection preference, i.e., the historical energy consumption, thus effectively reducing transmission energy consumption based on the optimal relay selection strategy. Extensive simulations are carried out to demonstrate the low energy consumption performances of MLRS compared with existing relay selection algorithms.

The reminder of the work is organized as follows. Section 2 introduces the SPIoT model. Section 3 presents MLRS. Section 4 presents simulation results. Section 5 concludes this article.

2. System Model

A relay selection model of SPIoT network considering electromagnetic interference in complex substation environment is shown in Figure 1. The SPIoT network consists of two parts, i.e., SPIoT devices and gateway. The gateway provides decision-making and computing services for SPIoT devices. The SPIoT network is a many-to-one two-hop transmission network that includes I source nodes, J relay nodes, and one destination node. Each source node collects various types of information including temperature, humidity, smoke, and gas and transmits the collected data to a relay node. The relay node receives the collected data from the source node and forwards the data to the destination node. Specifically, decode-and-forward relay node is adopted to avoid the amplification of noise power. The destination node is the substation interval measurement and control cabinet, which serves as the receiving end of monitoring data. Denote the sets of source nodes and relay nodes as $\mathcal{S} = \{s_1, \ldots, s_i, \ldots, s_I\}$ and $\mathcal{R} = \{r_1, \ldots, r_j, \ldots, r_J\}$, respectively. Denote the destination node as u_0. The total optimization period is divided to T time slots, the set of which is indexed as $\mathcal{T} = \{1, \ldots, t, \ldots, T\}$. At the beginning of the tth slot, the source node s_i has to transmit $A_i(t)$ amount of data to the destination node. The tth slot ends when all the data have arrived at the destination node. Considering the impact of the dynamic CSI and electromagnetic interference on the transmission delay, the slot duration is unequal. In each slot,

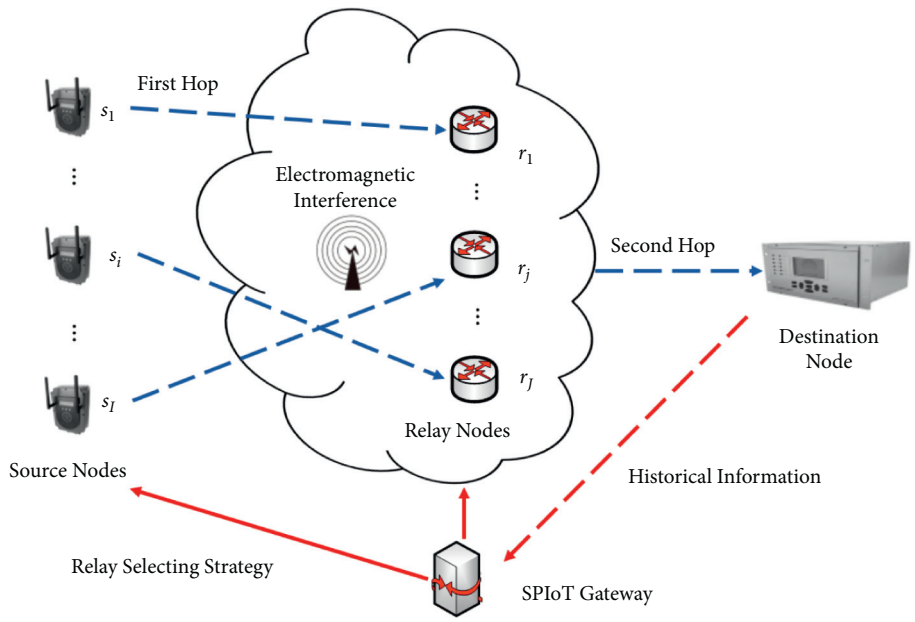


Figure 1: Two-hop transmission model for SPIoT with electromagnetic interference.

the gateway learns the optimal relay selection strategy ϕ for each source node based on the historical performance and sends the relay selection strategy to source nodes. The transmission delay and energy consumption models are introduced as follows.

2.1. Transmission Delay Model. In the tth slot, the source node s_i selects a relay node, e.g., r_j, to transmit $A_i(t)$ amount of data to the destination node u_0 based on the relay selection strategy from the gateway. Let $a_{i,j,t} = 1$ represent that s_i selects r_j in the tth slot; otherwise, $a_{i,j,t} = 0$. The transmission process is divided into two hops. The first hop is from the source node to the relay node, and the second hop is from the relay node to the destination node.

The signal-to-noise ratio of the first-hop transmission from s_i to r_j is given by [24, 25]

$$\gamma_{i,j}(t) = \frac{P_{TX,i}G_{i,j}(t)}{\sigma_0 + \varepsilon_{i,j}(t)}, \tag{1}$$

where $P_{TX,i}$ represents the transmission power of s_i, $G_{i,j}(t)$ represents the channel gain between s_i and r_j, σ_0 represents the noise power, and $\varepsilon_{i,j}(t)$ represents the electromagnetic interference power during the first-hop transmission. According to Shannon's formula [26], the first-hop transmission rate is given by

$$R_{i,j}(t) = B\log\left(1 + \gamma_{i,j}(t)\right) = B\log\left(1 + \frac{P_{TX,i}G_{i,j}(t)}{\sigma_0 + \varepsilon_{i,j}(t)}\right), \tag{2}$$

where B represents the bandwidth. Therefore, the first-hop transmission delay from s_i to r_j is given by

$$\tau_{i,j,t} = \frac{A_i(t)}{R_{i,j}(t)} = \frac{A_i(t)}{B\log\left(1 + \left(\left(P_{TX,i}G_{i,j}(t)\right)/\left(\sigma_0 + \varepsilon_{i,j}(t)\right)\right)\right)}. \tag{3}$$

In the tth slot, the signal-to-noise ratio of the second-hop transmission from r_j to the destination node u_0 is given by

$$\gamma_{j,0}(t) = \frac{P_{TX,j}G_{j,0}(t)}{\sigma_0 + \varepsilon_{j,0}(t)}, \tag{4}$$

where $P_{TX,j}$ represents the transmission power of r_j, $G_{j,0}(t)$ represents the channel gain between r_j and u_0, and $\varepsilon_{j,0}(t)$ represents the electromagnetic interference power during the second-hop transmission. The transmission rate from r_j to u_0 in the second hop is given by

$$R_{j,0}(t) = B\log\left(1 + \gamma_{j,0}(t)\right) = B\log\left(1 + \frac{P_{TX,j}G_{j,0}(t)}{\sigma_0 + \varepsilon_{j,0}(t)}\right). \tag{5}$$

Therefore, in the tth slot, the second-hop transmission delay from r_j to u_0 is given by

$$\tau_{j,0,t} = \frac{A_i(t)}{R_{j,0}(t)} = \frac{A_i(t)}{B\log\left(1 + \left(\left(P_{TX,j}G_{j,0}(t)\right)/\left(\sigma_0 + \varepsilon_{j,0}(t)\right)\right)\right)}. \tag{6}$$

Therefore, the total transmission delay from s_i to u_0 through r_j in the tth slot, i.e., the length of slot t, is the sum of the first-hop transmission delay and the second-hop transmission delay, which is given by

$$\tau_t = \tau_{i,j,t} + \tau_{j,0,t} = A_i(t)\left(\frac{1}{B\log\left(1+\left(\left(P_{TX,i}G_{i,j}(t)\right)/\left(\sigma_0+\varepsilon_{i,j}(t)\right)\right)\right)}+\frac{1}{B\log\left(1+\left(\left(P_{TX,j}G_{j,0}(t)\right)/\left(\sigma_0+\varepsilon_{j,0}(t)\right)\right)\right)}\right). \tag{7}$$

2.2. Energy Consumption Model. The energy consumption of the two-hop data transmission includes transmission energy consumption and reception energy consumption. The transmission energy consumption includes data transmission energy consumption and transmitting circuit energy consumption. The reception energy consumption is receiving circuit energy consumption. The first-hop data transmission energy consumption from s_i to r_j is given by [27]

$$E^1_{i,j,t} = P_{TX,i}\tau_{i,j,t}. \tag{8}$$

The second-hop data transmission energy consumption from r_j to u_0 is given by

$$E^2_{i,j,t} = P_{TX,j}\tau_{j,0,t}. \tag{9}$$

The transmission circuit energy consumption of s_i is given by

$$E^{tra}_{i,t} = A_i(t)E_{tra}, \tag{10}$$

where E_{tra} is energy consumption coefficient of transmission circuit.

The transmission circuit and reception circuit energy consumption of r_j is given by

$$E^{trec}_{i,t} = A_i(t)E_{trec}. \tag{11}$$

where E_{trec} is energy consumption coefficient of transmission circuit and reception circuit.

The reception circuit energy consumption of u_0 is given by

$$E^{rec}_{i,t} = A_i(t)E_{rec}. \tag{12}$$

where E_{rec} is energy consumption coefficient of reception circuit.

The total energy consumption is the sum of transmission energy consumption and reception energy consumption, which is given by

$$E_{i,j,t} = E^{tra}_{i,t} + E^1_{i,j,t} + E^{trec}_{i,t} + E^2_{i,j,t} + E^{rec}_{i,t}, \tag{13}$$

where the first and second terms represent the energy consumption of the first-hop transmission from s_i to r_j, the third term represents the forwarding energy consumption of r_j, and the fourth and fifth terms represent the energy consumption of the second-hop transmission from r_j to u_0.

2.3. Problem Formulation. We address the relay selection problem for SPIoT under dynamic CSI and electromagnetic interference. The relay selection problem is formulated as

$$\begin{aligned}
&\mathbf{P1}\colon \min_{\{a_{i,j,t}\}} \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J} a_{i,j,t}E_{i,j,t},\\
&\text{s.t. } C_1\colon a_{i,j,t} \in \{0,1\}, \quad \forall s_i \in \mathcal{S}, \forall r_j \in \mathcal{R}, \forall t \in \mathcal{T},\\
&C_2\colon \sum_{r_j \in \mathcal{R}} a_{i,j,t} \le 1, \quad \forall s_i \in \mathcal{S}, \forall t \in \mathcal{T},\\
&C_3\colon \sum_{s_i \in \mathcal{S}} a_{i,j,t} \le 1, \quad \forall r_j \in \mathcal{R}, \forall t \in \mathcal{T},
\end{aligned} \tag{14}$$

where C_1 defines the value of the selected indicator variable and C_2 and C_3 ensure that there is no selection conflict in each time slot. In other words, a source node selects at most one relay node for data forwarding in each time slot, and a relay node is selected by one source node at most.

3. Matching Learning-Based Relay Selection for SPIoT

In this section, the implementation process of the proposed MLRS algorithm for two-hop data transmission in SPIoT is elaborated.

3.1. Problem Transformation. **P1** is intractable because the relay selection strategies of all source nodes are coupled. Based on matching theory [28, 29], the selection problem between multisource node and multirelay node can be transformed into a one-to-one matching problem between I source nodes and J relay nodes. To solve the problem of incomplete global information in preference lists' construction of matching theory, MLRS utilizes UCB to enable the implementation of iterative matching without precise GSI and utilizes matching theory to avoid selection conflicts, which achieves a stable matching based on the learned preference lists. Since the relay node is unconditionally selected by the source node for data forwarding, the unilateral matching theory is adopted [30]. The definition of matching is given below.

Definition 1. (matching). A matching ϕ is a one-to-one correspondence of the set $\mathcal{S} \cup \mathcal{R}$ to itself. $\phi(s_i) = r_j$ means $a_{i,j,t} = 1$, i.e., s_i and r_j are matched. Otherwise, $\phi(s_i) = s_i$ means s_i does not match with any r_j.

However, the traditional matching theory needs to construct a preference list for each source node based on global information, which is not suitable to real-world SPIoT scenario, due to the dynamic changes of electromagnetic interference and the uncertain global information. Reinforcement learning can solve decision-making problems with incomplete information through continuous interaction with the environment, thereby enabling a high-precision and low-complexity preference list construction. It has the advantages of fast convergence and strong adaptability.

3.2. The Proposed MLRS Algorithm. Based on the multi-armed bandit (MAB) framework in reinforcement learning [31, 32], the source nodes and the relay nodes are modeled as players and arms, respectively. In the tth slot, s_i selects r_j to transmit data, and the performance of r_j can only be observed afterwards. The UCB algorithm is an effective method to solve the MAB problem. It is employed by MLRS to construct the preference list. MLRS is mainly divided into two steps: UCB-based preference list construction and iterative matching with rising price.

3.2.1. UCB-Based Preference List Construction. Denote the historical energy consumption of $\phi(s_i) = r_j$ up to slot t as $\overline{E}_{i,j,t-1}$, i.e., the empirical estimation of the two-hop transmission energy consumption $E_{i,j,t-1}$. Denote the total number of times that s_i selects r_j up to slot t as $k_{i,j,t-1}$, which is given by

$$k_{i,j,t-1} = \sum_{m=1}^{m=t-1} a_{i,j,m}. \tag{15}$$

Based on the UCB algorithm and the optimization objective **P**1, the matching preference value of $s_i \in \mathcal{S}$ towards $r_j \in \mathcal{R}$ is constructed to minimize the total energy consumption of the two-hop transmission, which is given by

$$\widetilde{Q}_{i,j,t} = \frac{1}{\overline{E}_{i,j,t-1}} + \sqrt{\frac{\omega \ln t}{k_{i,j,t-1}}} - H_j, \tag{16}$$

where the first item represents the empirical performance of r_j up to slot t. The matching preference value of s_i towards r_j decreases as the historical energy consumption increases. The second item is the confidence bound, which represents the estimation uncertainty. It decreases as $k_{i,j,t-1}$ increases, indicating that the estimated performance gradually approaches the actual expected value. $\omega > 0$ represents the weight for exploration. H_j represents the matching cost of r_j. The initial value of H_j is set as zero.

Denote the preference list of source nodes $s_i \in \mathcal{S}$ toward all the relay nodes $r_j \in \mathcal{R}$ as $\mathcal{F}_i$. Sort all $\widetilde{Q}_{i,j,t}$, $\forall r_j \in \mathcal{R}$, in the descending order to construct the preference list $\mathcal{F}_i$ of the source nodes.

3.2.2. Iterative Matching with Rising Price. The SPIoT gateway makes matching decisions based on the following steps. The implementation procedure is provided in Algorithm 1, which includes three steps.

Step 1: initialization.

Set $\phi(s_i) = \varnothing$, $\forall s_i \in \mathcal{S}$, $\Omega = \varnothing$, $H_j = 0$, and $\forall r_j \in \mathcal{R}$, where Ω is denoted as the set of relay nodes that have received more than one matching request.

Step 2: iterative matching.

Repeat

$\exists \phi(s_i) = \varnothing$, each $s_i \in \mathcal{S}$ makes a request to its most preferred r_j according to $\mathcal{F}_i$.

For any $r_j \in \mathcal{R}$, it is matched with the source node which sends the initial matching request. If r_j has been requested by more than one source node, add r_j into Ω.

If $\Omega \neq \varnothing$ • Each $r_j \in \Omega$ raises its price as

$$H_j = H_j + \Delta H_j, \tag{17}$$

where ΔH_j represents the rising step.

Each s_i update $\mathcal{F}_i$ based on H_j and renew its request.

Repeat price rising process until r_j is not requested by more than one source node. Remove r_j from Ω.

For $s_i \in \mathcal{S}$ that $\phi(s_i) = s_i$, the data transmission is suspended and waits for the next iteration.

Until $\phi(s_i) \neq \varnothing$.

Finally, the gateway sends the result ϕ to source nodes. Each source node $s_i \in \mathcal{S}$ selects $\phi(s_i)$ and transmits data.

Step 3: learning.

Observe the transmission performance and get the energy consumption $E_{i,j,t}$. Update $\overline{E}_{i,j,t}$ and $k_{i,j,t}$ as

$$\overline{E}_{i,j,t} = \frac{\overline{E}_{i,j,t-1} k_{i,j,t-1} + E_{i,j,t} a_{i,j,t}}{k_{i,j,t-1} + a_{i,j,t}}, \tag{18}$$

$$k_{i,j,t} = k_{i,j,t-1} + a_{i,j,t}. \tag{19}$$

3.3. Performance Analysis

Theorem 1 (stability). *For any $s_i \in \mathcal{S}$, there is no situation, where s_i prefers r_j than $\phi(s_i)$ and ϕ is stable.*

The detailed proof is in [33, 34].

Theorem 2 (convergence). *Due to the stability derived in Theorem 1, MLRS is convergent.*

Proof. Based on proof by contradiction, we assume MLRS is not convergent. Hence, there exist s_i preferring r_j' when $\phi(s_i) = r_j$ and $\phi(s_i') = r_j'$. Due to the higher preference value of s_i towards r_j', r_j' should be requested by s_i before r_j is requested by s_i. However, r_j' is matched with s_i', i.e., r_j' refused s_i, which is in contradiction with the assumption. Therefore, MLRS is convergent.

Complexity of MLRS: the computational complexity of MLRS consists of three parts. In the first step, the computational complexity is $\mathcal{O}(J)$. In the second step, we assume that MLRS takes ζ iterations to resolve matching conflicts. When the number of source nodes is greater than or equal to the number of relay nodes, i.e., $I \geq J$, the computational complexity is $\mathcal{O}(J) + \mathcal{O}(J \log J) + \mathcal{O}(I\zeta)$, while the computational complexity of the enumeration method is $\mathcal{O}(J^I I!)$. In the third step, the computational complexity is $\mathcal{O}(J)$. Therefore, the overall computational complexity of MLRS is $3\mathcal{O}(J) + \mathcal{O}(J \log J) + \mathcal{O}(I\zeta)$. □

Input: ω, T, s, $\mathscr{R}$.
Output: ϕ.
In the first $I \times J$ time slots, $\forall s_i \in \mathscr{S}$ traverse to select $\forall r_j \in \mathscr{R}$ once to observe the transmission performance.
while $t \leq T$ **do**
  All source nodes $s_i \in \mathscr{S}$ construct a preference list $\mathscr{F}_i$ as (16).
  Step 1: Initialization
  Set $\phi(s_i) = \varnothing$, $\forall s_i \in \mathscr{S}$, $\Omega = \varnothing$, $H_j = 0$, $\forall r_j \in \mathscr{R}$.
  Step 2: Iterative Matching
  while $\exists \phi(s_i) = \varnothing$ **do**
    for $s_i \in \mathscr{S}$ **do**
      $\forall s_i \in \mathscr{S}$ makes a request to its most preferred r_j according to $\mathscr{F}_i$.
    end for
    Add the relay nodes selected by more than one source node into Ω
    if $\Omega \neq \varnothing$ **then**
      for $r_j \in \Omega$ **do**
        r_j raises its price as (17).
        $\forall s_i$ selecting r_j, update $\mathscr{F}_i$ and renew its request.
        Repeat price rising process until r_j is not requested by more than one source node. Remove r_j from Ω.
      end for
    end if
  end while
  The gateway sends the result ϕ to source nodes. $\forall s_i \in \mathscr{S}$ selects $\phi(s_i)$ and transmits data.
  Step 3: Learning
  Observe transmission performance and get energy consumption $E_{i,j,t}$.
  Update $\overline{E}_{i,j,t}$ and $k_{i,j,t}$ as (18) and (19).
end while

Algorithm 1: MLRS.

4. Simulation Results

In this section, we validate the performance of MLRS by simulations. We consider a scenario with 10 source nodes, 15 relay nodes, and 1 destination node. The number of time slots is 500. The transmission power $P_{TX,i}$ and $P_{TX,j}$ are set as $[0.1, 1]$ W. The channel gains from the source node to the relay node and from the relay node to the destination node all satisfy the normal distribution $\mathscr{N}(127 + 30\log(l), 1)$ [35], where l is the distance. The simulation parameters are summarized in Table 1 [36–38]. We consider two state-of-the-art algorithms for comparison. The first one is the traditional UCB algorithm [39], where each source node calculates its preference value for each relay node based on historical performance and sends a data transmission request to the favorite relay node. Considering the case of matching conflicts, the source nodes with conflicts' preferences towards a relay node will be randomly matched with the remaining unmatched relay nodes. The second one is the energy efficiency performance selection (EEPS) algorithm [40], where each source node sends a transmission request to the relay node with the best historical performance, and the rest process is the same as the traditional UCB algorithm.

The α-stable distribution is employed to describe the electromagnetic interference. The characteristic function of the electromagnetic interference power from s_i to r_j is given by

$$\log \varepsilon_{i,j}(t) = \begin{cases} -\sigma^{\alpha}|t|^{\alpha}\left\{1 - j\beta \operatorname{sign}(t)\tan\left(\frac{\pi\alpha}{2}\right)\right\} + j\mu t, & \alpha \neq 1, \\ -\sigma|t|\left\{1 + j\beta \operatorname{sign}(t)\frac{2}{\pi}\log|t|\right\} + j\mu t, & \alpha = 1, \end{cases} \tag{20}$$

where $\alpha \in (0, 2]$ is the characteristic exponent, $\beta \in [-1, 1]$ is the skew parameter, $\beta = 0$ indicates a symmetric α-stable distribution, $\mu \in R$ is the location parameter, and σ is the scale parameter [41].

Figure 2 shows the average energy consumption versus time slot. When $t = 500$, compared with UCB and EEPS, the average energy consumption of MLRS is reduced by 17.49% and 24.22%, respectively. The reason is that the random selection mechanism in UCB and EEPS leads to unmatched source nodes and suspended data transmission when multiple source nodes compete for the same relay node. Moreover, UCB and EEPS randomly assign unmatched source nodes to relay nodes. As a result, these source nodes may be matched with some relay nodes with poor historical performance, leading to higher energy consumption. MLRS can effectively resolve matching conflicts between multisource nodes and multirelay nodes based on dynamically learned relay selection preference.

Figure 3 shows the optimal relay node selection probability versus time slot. In the initial stage for all the algorithms, the optimal relay node selection probability is low. The probability increases gradually with the number of

Table 1: Simulation parameters.

Parameters	Value
The number of source nodes	10
The number of relay nodes	15
The number of time slots	500
The transmission power, $P_{TX,i}$ and $P_{TX,j}$	$[0.1, 1]$ W
The channel gains from s_i to r_j and from r_j to u_0	$\mathcal{N}(127 + 30\log(l), 1)$
The amount of task data at the source node, $A_i(t)$	$[30, 40]$ kbit
Channel bandwidth, B	0.1 MHz
The noise power, n_0	−114 dBm
The energy consumption coefficient, E_{tra} and E_{rec}	50 nJ/bit
The energy consumption coefficient, E_{trec}	100 nJ/bit
The weight of exploration, ω	0.3
The characteristic exponent, α	1.4
The skew parameter, β	−0.791
The location parameter, μ	$-4 * 10^{-14}$
The scale parameter, σ	10^{-15}

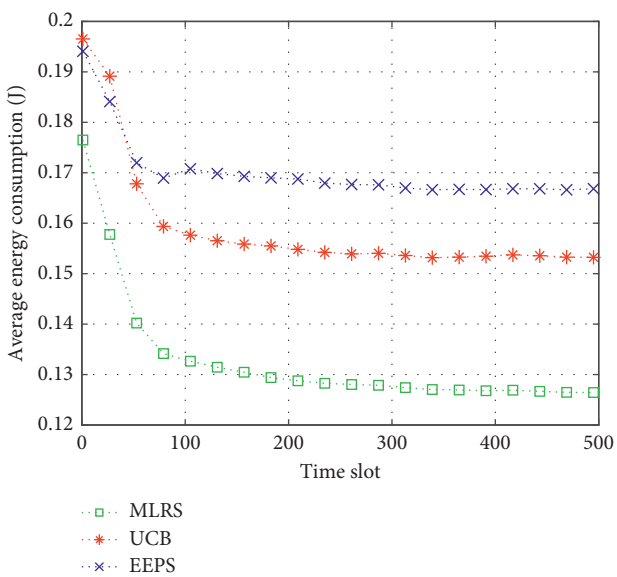


Figure 2: The average energy consumption versus time slot.

selections and finally converges as to 68%, 54%, and 34% for MLRS, UCB, and EEPS, respectively. MLRS always outperforms UCB and EEPS. The reason is that MLRS adopts a tendency exploration scheme instead of random selection, which can achieve an effective compromise between exploration and exploitation.

Figure 4 shows the optimal relay node selection probability versus the number of relay nodes. The optimal relay node selection probability of MLRS converges around 67% and outperforms UCB and EEPS by 18.08% and 18.20%, respectively. When the number of relay nodes increases, i.e., the network topology becomes more complex, MLRS can better adapt to complex network topology with more relay nodes and learn the most appropriate relay selection strategy by combining with learning-based matching preference lists' construction and matching-based conflict resolution.

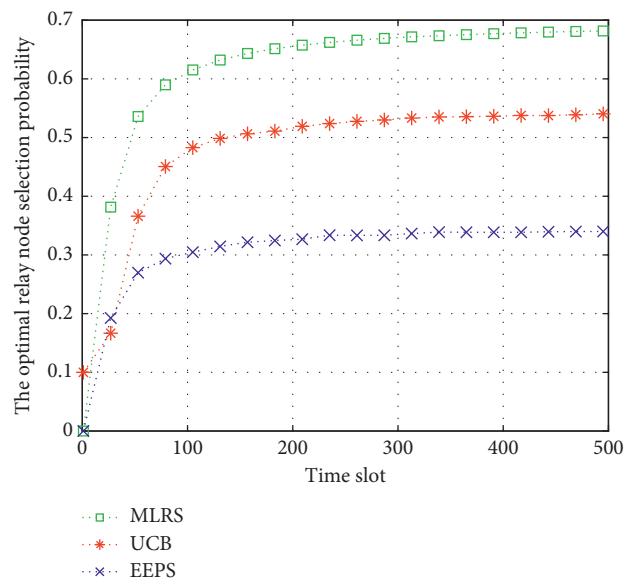


Figure 3: The optimal relay node selection probability versus time slot.

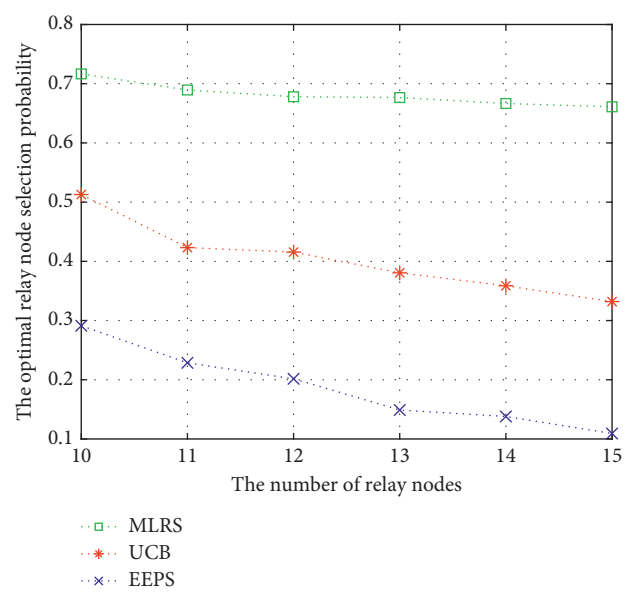


Figure 4: The optimal relay node selection probability versus the number of relay nodes.

However, with the increase of the number of relay nodes, the problem of matching conflict becomes prominent in UCB and EEPS, which leads to sharp performance decrease.

Figure 5 shows the average energy consumption versus ω. As ω increases, the performance of MLRS improves. When $\omega = 0.3$, the algorithm performance is optimal. A too small ω leads to biased preference towards exploitation and inability to explore potential better options, while a too large ω leads to inability to exploit existing optimal option.

Figure 6 shows the average energy consumption versus the electromagnetic interference intensity. We divide the electromagnetic interference into five levels by changing the

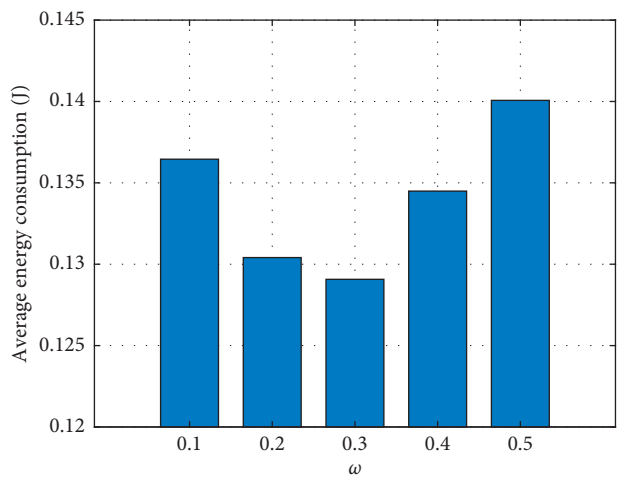


Figure 5: The average energy consumption versus the exploration weight ω in MLRS.

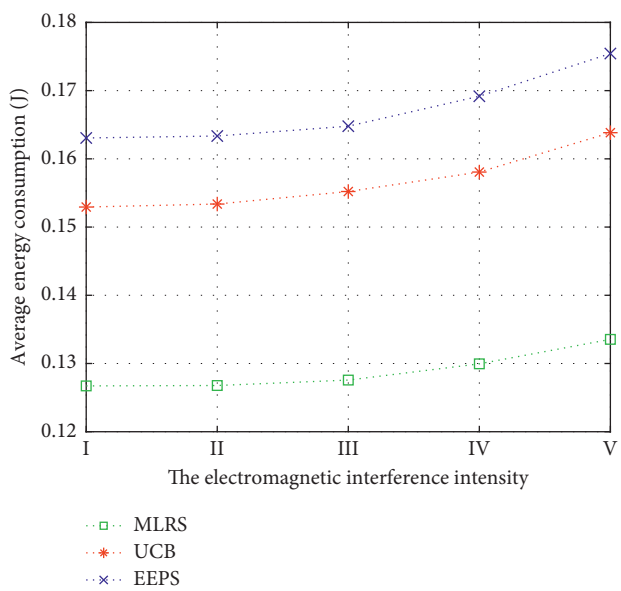


Figure 6: The average energy consumption versus the electromagnetic interference intensity.

Table 2: Electromagnetic interference levels.

Level	I	II	III	IV	V
μ	$4 * 10^{-16}$ W	$4 * 10^{-15}$ W	$4 * 10^{-14}$ W	$4 * 10^{-13}$ W	$4 * 10^{-12}$ W

location parameter μ of the α-stable distribution, which are summarized in Table 2. As the electromagnetic interference intensity level increases, compared with UCB and EEPS, MLRS always has the lowest average energy consumption. The reason is that MLRS can learn the optimal relay selection strategy and resolve the matching conflict regardless of the electromagnetic interference level, which verifies that MLRS shows adaptability to various wireless environment.

5. Conclusion

In this paper, a novel relay selection algorithm named MLRS was proposed for SPIoT. MLRS can minimize the energy consumption of SPIoT devices under complex electromagnetic interference through relay selection optimization without GSI. Simulation results indicate that, compared with UCB and EEPS, MLRS reduces the energy consumption by

17.49% and 24.22%, respectively. In the future, we will focus on the joint optimization of multiple quality of service (QoS) performance metrics including delay and throughput, considering the differentiated QoS requirements of SPIoT.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, under Grant LAPS202125, and supported by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant no. 2021D12.

References

[1] L. Zhao, I. Matsuo, Y. Zhou, and W. J. Lee, "Design of an industrial IoT-based monitoring system for power substations," *IEEE Transactions on Industry Applications*, vol. 55, no. 4, pp. 5666–5674, 2019.

[2] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[3] M. Tariq, M. Ali, F. Naeem, and H. V. Poor, "Vulnerability assessment of 6G-enabled smart grid cyber-physical systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5468–5475, 2021.

[4] A. Amari, O. A. Dobre, R. Venkatesan, O. S. S. Kumar, P. Ciblat, and Y. Jaouën, "A survey on fiber nonlinearity compensation for 400 gb/s and beyond optical communication systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 3097–3113, 2017.

[5] L. Zhao, X. Li, and B. Gu, "Vehicular communications: standardization and open issues," *IEEE Commun. Standards Mag.*vol. 2, no. 4, pp. 74–80, 2020.

[6] J. Cheng, P. Yang, K. Navaie, Q. Ni, and H. Yang, "A low-latency interference coordinated routing for wireless multi-hop networks," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8679–8690, 2021.

[7] G. Sheng, H. Wang, F.-Q. Wen, and X. Wang, "Fast angle estimation and sensor self-calibration in bistatic MIMO radar with gain-phase errors and spatially colored noise," *IEEE Access*, vol. 8, pp. 123701–123710, 2020.

[8] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[9] Z. Zhou, H. Liao, X. Zhao, B. Ai, and M. Guizani, "Reliable task offloading for vehicular fog computing under information asymmetry and information uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8322–8335, 2019.

[10] C. Wang, X. Gaimu, C. Li, H. Zou, and W. Wang, "Smart mobile crowdsensing with urban vehicles: a deep reinforcement learning perspective," *IEEE Access*, vol. 7, pp. 37334–37341, 2019.

[11] Z. Tang and T. Jian, "Research on influences factors of electromagnetic interferences for partial discharge detection in substations," in *Proceedings of the IEEE Electrical Insultion Conference (EIC)*, pp. 42–45, Baltimore, MD, USA, August 2017.

[12] A. Muller and J. Speidel, "Relay selection in dual-hop transmission systems: selection strategies and performance results," in *Proceedings of the IEEE International Conference on Communication (ICC)*, pp. 4998–5003, Beijing, China, May 2008.

[13] S. H. Mousavi, J. Haghighat, and W. Hamouda, "A relay subset selection scheme for wireless sensor networks based on channel state information," in *Proceedings of the IEEE International Conference on Communication (ICC)*, pp. 1–6, Kuala Lumpur, Malaysia, May 2016.

[14] M. Adil, R. Khan, J. Ali, B.-H. Roh, Q. T. H. Ta, and M. A. Almaiah, "An energy proficient load balancing routing scheme for wireless sensor networks to maximize their lifespan in an operational environment," *IEEE Access*, vol. 8, pp. 163209–163224, 2020.

[15] S. T. Bakhsh, "Energy-efficient distributed relay selection in wireless sensor network for internet of things," in *Proceedings of the 13th International Wireless Communication on Mobile Computers Conference (IWCMC)*, pp. 1802–1807, Valencia, Spain, July 2017.

[16] B. Wang, Y. Sun, H. M. Nguyen, and T. Q. Duong, "A novel socially stable matching model for secure relay selection in D2D communications," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 162–165, 2020.

[17] M. W. Baidas, M. M. Afghah, and F. Afghah, "A matching-theoretic approach to distributed swipt in ad-hoc wireless networks," in *Proceedings of the International Symposium on Network, Computer Communication (ISNCC)*, pp. 1–6, Istanbul, Turkey, June 2019.

[18] H. Liao, Z. Zhou, X. Zhao et al., "Learning-based context-aware resource allocation for edge-computing-empowered industrial IoT," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4260–4277, 2020.

[19] L. Zhao, K. Yang, and Z. Tan, "A novel cost optimization strategy for SDN-enabled UAV-assisted vehicular computation offloading," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3664–3674, 2020.

[20] Y. Su, X. Lu, Y. Zhao, L. Huang, and X. Du, "Cooperative communications with relay selection based on deep reinforcement learning in wireless sensor networks," *IEEE Sensors Journal*, vol. 19, no. 20, pp. 9561–9569, 2019.

[21] H. Liang, X. Zhang, X. Hong et al., "Reinforcement learning enabled dynamic resource allocation in the internet of vehicles," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 4957–4967, 2021.

[22] Z. Zhou, J. Feng, B. Gu et al., "When mobile crowd sensing meets UAV: energy-efficient task assignment and route planning," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5526–5538, 2018.

[23] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," in *Proceedings of the IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, September 2021.

[24] J. Wang, B. Li, M. Liu, and J. Li, "SNR estimation of time-frequency overlapped signals for underlay cognitive radio,"

IEEE Communications Letters, vol. 19, no. 11, pp. 1925–1928, 2015.

[25] M. A. Raza and A. Hussain, "Maximum likelihood SNR estimation of hyper cubic signals over Gaussian channel," *IEEE Communications Letters*, vol. 20, no. 1, pp. 45–48, 2016.

[26] S.-H. Lee and S.-Y. Chung, "Capacity scaling of wireless ad hoc networks: Shannon meets Maxwell," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1702–1715, 2012.

[27] C. Wang, D. Deng, L. Xu, W. Wang, and F. Gao, "Joint interference alignment and power control for dense networks via deep reinforcement learning," *IEEE Wireless Communications Letters*, vol. 10, no. 5, pp. 966–970, 2021.

[28] Z. Zhou, C. Zhang, C. Xu, F. Xiong, Y. Zhang, and T. Umer, "Energy-efficient industrial internet of UAVs for power line inspection in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2705–2714, 2018.

[29] S. Jeong, T.-H. Lin, and M. M. Tentzeris, "A real-time range-adaptive impedance matching utilizing a machine learning strategy based on neural networks for wireless power transfer systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 12, pp. 5340–5347, 2019.

[30] Y. Yuan, T. Yang, Y. Hu, H. Feng, and B. Hu, "Two-timescale resource allocation for cooperative D2D communication: a matching game approach," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 543–557, 2021.

[31] N. Papandreou and T. Antonakopoulos, "Resource allocation management for indoor power-line communications systems," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 893–903, 2007.

[32] L. Zhao, W. Zhao, A. Hawbani et al., "Novel online sequential learning-based adaptive routing for edge software-defined vehicular networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 2991–3004, 2021.

[33] Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Social big-data-based content dissemination in internet of vehicles," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 768–777, 2018.

[34] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, "A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4550–4559, 2021.

[35] Y. Sun, S. Zhou, and J. Xu, "EMM: energy-aware mobility management for mobile edge computing in ultra dense networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2637–2646, 2017.

[36] H. Liao, Z. Zhou, W. Kong et al., "Learning-based intent-aware task offloading for air-ground integrated vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5127–5139, 2021.

[37] J. Zhang, J. Tang, and F. Wang, "Cooperative relay selection for load balancing with mobility in hierarchical WSNs: a multi-armed bandit approach," *IEEE Access*, vol. 8, pp. 18110–18122, 2021.

[38] J. Luo, J. Hu, D. Wu, and R. Li, "Opportunistic routing algorithm for relay node selection in wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 1, pp. 112–121, 2015.

[39] M. Endo, T. Ohtsuki, T. Fujii, and O. Takyu, "Secure channel selection using multi-armed bandit algorithm in cognitive radio network," in *Proceedings of the IEEE 85th Vehicle Technology Conference (VTC Spring)*, pp. 1–5, Sydney, NSW, Australia, November 2017.

[40] B. Klaiqi, X. Chu, and J. Zhang, "Energy-Efficient and low signaling overhead cooperative relaying with proactive relay subset selection," *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 1001–1015, 2016.

[41] Z. Zhou, X. Yang, and C. Xu, "Performance evaluation of multi-antenna based m2m communications for substation monitoring," in *Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 97–102, Jeju, Korea (South), October 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8176261, 11 pages
https://doi.org/10.1155/2022/8176261

Research Article

Secrecy Analysis of NOMA-Based Integrated Satellite-Multiple UAV Networks under Nonideal Hardware and Colluding Scheme

Yuanyuan Wang,[1] **Feng Zhou**,[1] **Rugang Wang**,[1] **Xueling Wang**,[2] **and Kefeng Guo**[3]

[1]*School of Information Technology, Yancheng Institute of Technology, Yancheng 224051, China*
[2]*Space E-Star Communication Technology Co. Ltd., Nanjing 210007, China*
[3]*Space Information School, Space Engineering University, Beijing 101417, China*

Correspondence should be addressed to Kefeng Guo; guokefeng.cool@163.com

Received 25 December 2021; Revised 9 January 2022; Accepted 15 January 2022; Published 3 February 2022

Academic Editor: Xingwang Li

This paper studies the secrecy outage probability (SOP) of nonorthogonal multiple access (NOMA)-based integrated satellite-multiple unmanned aerial vehicle networks (ISMUAVNs) under nonideal hardware and colluding scheme. Colluding scheme is defined as that the eavesdroppers cooperate with each other to overhear the legitimate users' signals. On this foundation, we obtain the closed-form expression for the SOP for the considered NOMA-based ISMUAVNs along with the partial unmanned aerial vehicle (UAV) selection scheme. To gain more things in the high signal-to-noise ratio (SNR) regime, we derive the asymptotic expression for the SOP, which gives fast ways to evaluate the benefit of the NOMA scheme and the impact of nonideal hardware impairments on the SOP. Finally, Monte Carlo simulation results are provided to show the correctness of the analytical results. In addition, with the help of representative results, some important and interesting findings are obtained.

1. Introduction

The integration of satellite communication (SatCom) networks with the terrestrial mobile networks is one of the key elements for the fifth-generation (5G) and beyond 5G networks due to enhanced capability and extended coverage [1]. However, the coverage region may be restricted by the obstacles and shadowing between the terrestrial users and satellite, which leads to the unavailability of the line-of-sight (LOS) links [2]. Under this situation, the terrestrial relay is utilized to overcome this challenge such as fading, shadowing, and path loss [3], which leads to the formation of the integrated satellite-unmanned aerial vehicle networks (ISUAVNs). Particularly, the main idea of ISUAVNs is to utilize terrestrial relays/unmanned aerial vehicles (UAVs) to forward the satellite signal, which has been a good framework for the SatCom, they have great effects on the fixed satellite service and high-speed broadband mobile satellite service, and SiriusXM is the real system which relies on the foundation of this architecture [4].

As the wide coverage of the satellite beam, there are often many wireless communication nodes in one satellite beam, which result in the suitable selection scheme for the wireless communication nodes [1]. In [5], the outage probability (OP) was analyzed with the opportunistic wireless node scheduling scheme for the integrated satellite-terrestrial relay networks (ISTRNs). In [6], the OP was investigated for the ISTRNs under the consideration of cognitive technology and multiple terrestrial transmission nodes.

With the breakthrough of UAV technology, UAV communication has been widely utilized in many temporary large-scale activities and natural disasters [7]. Total power constraint-beamforming and per-antenna power constraint-beamforming schemes were proposed to optimize the energy efficiency of integrated satellite-UAV terrestrial networks in [7]. The authors of [8] investigated the UAV system with mobile edge computing, where the energy expenditure of users was minimized under assignment delay demands and resource limitations. In [9], the game theory was applied to optimize the resources' allocation of the UAV

communication network for more rewards, and a multiagent reinforcement learning framework was proposed to find the optimal policy. The potential game, mean-field game, Stackelberg game, graphical game, and coalition game were utilized to solve the resource management problem of large-scale UAVs in [10]. The authors of [11] designed a decaying deep Q-network-based algorithm to minimize the energy consumption, in which the trajectory of the UAV, phase shifts of the reconfigurable intelligent surfaces, power allocation strategy, and dynamic decoding sequence were considered.

To alleviate the spectrum shortage and meet the requirements of improved spectral efficiency of ISTRNs, nonorthogonal multiple access (NOMA) has been viewed as the new paradigm of multiple access techniques [12–14], which transmits signals simultaneously in the same time/frequency resource block. On this foundation, successive interference cancellation (SIC) is used at the receivers to distinguish the users' message by different power levels [15]. In NOMA systems, the message for the weaker was first decoded by the stronger users in which case, the stronger can be regarded as the relay to help the user with poor channel environment to enhance the system performance [16, 17]. In [18], the energy efficiency, ergodic capacity, and outage probability (OP) of NOMA-based downlink SatCom were investigated. In [19], the cooperative transmission method was analyzed for the NOMA-based SatCom. In [15], the authors derived the closed-form expression of OP in the presence of secondary networks.

As presented before, the satellite beam has a wide coverage, which results in the security problem in reality for the ISTRNs [20]; physical layer security (PLS) is considered as a new method to present the security problem of the wireless transmission networks. The major idea of PLS is to keep the secrecy transmission rate in the case of no interception under malicious nodes. In [21], the authors listed enough problems in the SatCom and investigated the secrecy performance. In [22], the authors investigated the SOP and ASC with the maximum user scheduling scheme for the SatCom. In [23], the authors analyzed the impact of imperfect channel stat information (CSI) for SatCom with multiple legitimate users and multiple eavesdroppers. In [24, 25], the ASC and SOP for the ISTRNs with the opportunistic relaying scheme were analyzed, respectively.

It is mentioned that, in real wireless communication systems, the transmission nodes are often not ideal owing to several reasons [26, 27]. In [28], the authors summarized all the reasons and proposed a general nonideal hardware model to research the impact of nonideal hardware on the wireless communication networks. In [29], the authors investigated the performance for the amplify-and-forward multiple relay networks in the presence of nonideal hardware and available direct transmission link. In [30], the OP was analyzed with the nonideal hardware and cognitive technology existing in the ISTRNs.

Until now, the impact of nonideal hardware on the NOMA-based SatCom was only investigated in several former papers, such as [31–34]. In [31], the authors analyzed the SOP for the considered SatCom with nonideal hardware and channel estimation errors (CEEs). In [32], the impact of nonideal hardware was investigated for the ISTRNs. In [33], Guo analyzed the effect of nonideal hardware on the ISTRNs with multiple terrestrial relays under the partial wireless node selection scheme.

In this paper, we investigate the secrecy NOMA-based integrated satellite-multiple unmanned aerial vehicle networks (ISMUAVNs) in the presence of multiple eavesdroppers and HIs by utilizing the partial UAV selection scheme. In detail, the main contributions of this paper are shown in the following:

(i) Firstly, we establish a NOMA-based secrecy ISMUAVN with nonideal hardware. The partial UAV selection scheme with low implementation complexity is applied in the considered system.

(ii) Secondly, the closed-form expression for SOP of the considered network in the presence of nonideal hardware and the proposed UAV selection scheme is derived.

(iii) Finally, to get insight of the SOP at high signal-to-noise ratios (SNRs), the asymptotic expression for SOP of the considered secrecy ISMUAVNs is also derived.

The rest of this paper is given in the following. Section 2 introduces the system model and formulates the secrecy problem of the considered secrecy ISMUAVNs. Section 3 provides the detailed analysis for the SOP. In Section 3, the secrecy diversity order and secrecy coding gain are analyzed along with the asymptotic analysis. In Section 4, numerical Monte Carlo (MC) results are obtained. In Section 5, the summarization is presented.

1.1. Notations. $|\cdot|$ denotes the absolute value of a complex scalar, $\exp(\cdot)$ is the exponential function, $E[\cdot]$ represents the expectation operator, and $\mathcal{CN}(\mathbf{a}, b)$ is the complex Gaussian distribution of a random vector $\mathbf{a}$ and covariance matrix b.

2. System Model

As plotted in Figure 1, in this paper, we investigate NOMA-based secrecy ISMUAVNs, which consist of a satellite (S), namely, Alice, multiple UAVs ($R_p, p \in \{1, \ldots, M\}$), two destinations (namely, D_1 and D_2), and N eavesdroppers. Decode-and-forward (DF) protocol is utilized at R. Two time slots are used for the total transmission. The direct transmission link is not considered between S and D_i $i \in \{1, 2\}$ 29, for some reasons (Owing to some practical reasons, such as heavy fading, shadowing, or obstacle blockage, no direct transmission link is assumed in this paper.). In general, all transmission nodes in the considered NOMA-based ISMUAVNs are equipped with one antenna, respectively (In this paper, only two users are considered; however, the derived results can also be suitable for multiple users, which will be investigated in the near future. In this paper, the single antenna is assumed; however, the derived results will also be suitable for the case with multiple antennas.).

II. SYSTEM MODEL

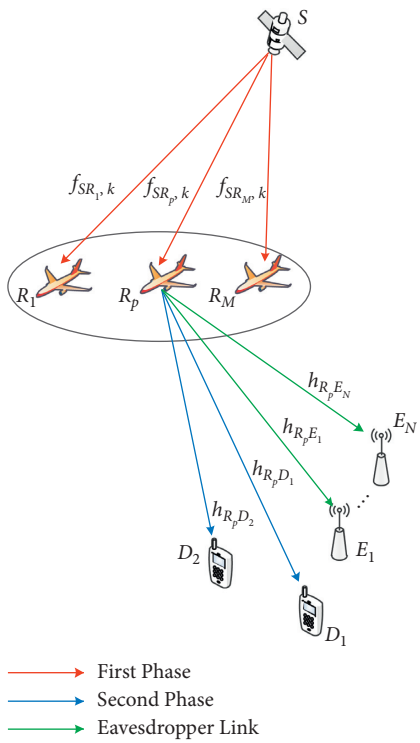


Figure 1: Illustration of the system model.

For the first time slot, the signal $s(t)$ is forwarded by S with $E[|s(t)|^2] = 1$ to pth R. The received signal at pth R is derived as

$$y_{R_p}(t) = f_{SR_p,k}\sqrt{P_S}\left[s(t) + \eta_S(t)\right] + n_{R_p}(t), \tag{1}$$

where $f_{SR_p,k}$ represents the channel coefficient modeled as shadowed-Rician (SR) fading (In this paper, the perfect CSI is assumed, and our results will be as the branch for the case with imperfect CSI.) and P_S denotes the transmit power from S. When applying the NOMA scheme, the signal is shown as $s(t) = a_{1p}x_1(t) + a_{2p}x_2(t)$, where $x_1(t)$ and $x_2(t)$ present the target transmitted signal to D_1 and D_2, respectively. Besides, it is assumed that $a_{1p} > a_{2p}$ for the reason that the channel D_1 suffers heavy fading than that of D_2. On this foundation, much power should be allocated to D_1. a_{1p} and a_{2p} represent the transmission power allocation (PA) factors for D_1 and D_2, respectively. In addition, $a_{1p}^2 + a_{2p}^2 = 1$. $\eta_S(t)$ is the distortion noise due to nonideal hardware with $\eta_S(t) \sim C\mathcal{N}(0, a_{1p}^2k_{1p}^2 + a_{2p}^2k_{2p}^2)$. k_{1p} and k_{2p} denote the nonideal levels for nonideal hardware which are defined in [3]. $n_{R_p}(t)$ represents the additive white Gaussian noise (AWGN) at pth R distributed as $n_{R_p}(t) \sim C\mathcal{N}(0, \delta_{R_p}^2)$.

When the signal is received by R, SIC works as follows: $x_1(t)$ is firstly decoded and then deleted from the obtained signal [12]. Thereafter, $x_2(t)$ is decoded from the left signal (In this paper, perfect SIC is assumed. Besides, different users are distinguished by the channel quality. We assume that the nearby user has a better channel quality than that of the distant users which has been assumed in the literature studies.).

Hence, the signal-to-interference-plus-noise-and-distortion ratio (SINDR) of detecting $x_1(t)$ at pth R is given by

$$\gamma_{R_p,1} = \frac{\gamma_{1p}a_{1p}^2}{\gamma_{1p}\left(a_{1p}^2k_{1p}^2 + a_{2p}^2 + a_{2p}^2k_{2p}^2\right) + 1}, \tag{2}$$

where $\gamma_{1p} = |f_{SR_p,k}|^2P_S/\delta_{R_p}^2$.

Moreover, we can obtain the signal-to-noise-and-distortion ratio (SNDR) of detecting $x_2(t)$ at pth R as

$$\gamma_{R_p,2} = \frac{\gamma_{1p}a_{2p}^2}{\gamma_{1p}\left(a_{1p}^2k_{1p}^2 + a_{2p}^2k_{2p}^2\right) + 1}. \tag{3}$$

As announced before, the partial UAV selection scheme [35] is adopted for the considered system model, and the SINDR of $x_1(t)$ and $x_2(t)$ during the first transmission link is, respectively, shown as

$$\gamma_{R,1} = \max_{p\in\{1,\ldots,M\}}\left(\gamma_{R_p,1}\right), \tag{4}$$

$$\gamma_{R,2} = \max_{p\in\{1,\ldots,M\}}\left(\gamma_{R_p,2}\right). \tag{5}$$

For the second time slot, the selected pth R forwards received signal $s(t)$ to D_1 and D_2, respectively; thus, the signal received at D_i can be represented as

$$y_{D_{pi}}(t) = h_{R_pD_i}\sqrt{P_{R_p}}\left(\xi_{1p}x_1(t) + \xi_{2p}x_2(t) + \eta_p(t)\right) + n_{D_i}(t), \tag{6}$$

where $h_{R_pD_i}$ denotes the channel coefficient between pth R and D_i which undergoes Rayleigh fading, P_{R_p} represents the transmit power for pth R, $x_i(t)$ denotes the expected signal, ξ_{jp} $j \in \{1,2\}$ which suits $\xi_{1p} > \xi_{2p}$, and $\xi_{1p}^2 + \xi_{2p}^2 = 1$ represents the PA factors at pth R, respectively (We apply the same assumption which means the channel quality of user D_1 is worse than that of user D_2.). $n_{D_i}(t)$ denotes the AWGN at D_i distributed as $n_{D_i}(t) \sim \mathcal{CN}(0, \delta_{D_i}^2)$. $\eta_p(t)$ with $\eta_p(t) \sim \mathcal{CN}(0, \xi_{1p}^2k_{p1}^2 + \xi_{2p}^2k_{p2}^2)$ represents the distortion noise which is caused by nonideal hardware. ξ_{1p} along with ξ_{2p} denotes the nonideal levels of nonideal hardware [35]. Therefore, D_2 performs SIC to decode $x_1(t)$ (while $x_2(t)$ is treated as the cochannel noise).

From (6), the SINDR of $x_1(t)$ sent by pth R at D_2 is represented as

$$\gamma_{D_{p1},D_2} = \frac{\gamma_{2p}\xi_{1p}^2}{\gamma_{2p}\left(\xi_{1p}^2k_{1p}^2 + \xi_{2p}^2 + \xi_{2p}^2k_{2p}^2\right) + 1}, \tag{7}$$

where $\gamma_{2p} = |h_{R_pD_2}|^2P_{R_p}/\delta_{D_2}^2$.

γ_{D_{p1},D_2} must be satisfied as $\gamma_{D_{p1},D_2} \geq \gamma_{th}$ due to the same condition like (2). D_2 deleted the decoded information from the received signal by observing its own signal; hence, the final SNDR for signal $x_2(t)$ forwarded by pth R is given by

$$\gamma_{D_{p2},D_2} = \frac{\gamma_{2p}\xi_{2p}^2}{\gamma_{2p}\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right) + 1}. \tag{8}$$

Then, $x_1(t)$ is decoded by D_1, while $x_2(t)$ is regarded as the serial interference; the SINDR received at D_1 from pth R can be obtained as

$$\gamma_{D_{p1},D_1} = \frac{\gamma_{3p}\xi_{1p}^2}{\gamma_{3p}\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 + \xi_{2p}^2 k_{2p}^2\right) + 1}, \tag{9}$$

where $\gamma_{3p} = |h_{R_p D_1}|^2 P_{R_p}/\delta_{D_1}^2$.

At the same time, due to the fact that the signal sent to relays can be received by all eavesdroppers, the heard signal of the jth eavesdropper can be represented as

$$y_{E_{pj}}(t) = h_{R_p E_j}\sqrt{P_{R_p}}\left(\xi_{1p}x_1(t) + \xi_{2p}x_2(t) + \eta_p(t)\right) + n_{E_j}(t), \tag{10}$$

where $h_{R_p E_j}$ represents the channel coefficient between pth R and the jth Eve which is modeled as Rayleigh fading. $n_{E_j}(t)$ with $n_{E_j}(t) \sim \mathcal{CN}(0, \delta_{E_j}^2)$ denotes the AWGN at the jth Eve.

Through the whole paper, we assume that parallel interference cancellation (PIC) is available at each eavesdropper [36, 37]. Hence, the received SNR of ith D's expected messages at the jth Eve can be shown as

$$\lambda_{D_{pi}E_j} = \frac{\xi_{ip}^2 \gamma_{D_{pi}E_j}}{\gamma_{D_{pi}E_j}\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right) + 1}, \tag{11}$$

where $\gamma_{D_{pi}E_j} = |h_{R_p E_j}|^2 P_{R_p}/\delta_{E_j}^2$.

Due to the cooperation of all eavesdroppers, we can get the SNR of all Eves' link as

$$\gamma_{D_{pi}E} = \sum_{j=1}^{N} \lambda_{D_{pi}E_j}. \tag{12}$$

The authors in [1] introduced the definition of the secrecy capacity; with the help of (7), (8), (9), and (12), we can obtain the secrecy capacity for the ith signal as

$$C_{S_{pi}\rho} = \left[C_{B_{pi}\rho} - C_{D_{pi}E}\right]^+, \left(\rho \in \{D_1, D_2\}\right), \tag{13}$$

where $[x]^+ \triangleq \max[x, 0]$, $C_{B_{pi}\rho} = \log_2(1 + \gamma_{D_{pi},\rho})$, and $C_{D_{pi}E} = \log_2(1 + \gamma_{D_{pi}E})$ ($CBpi\rho$ means the secrecy capacity at the ρ user to detect the ith signal from pth R.).

3. Performance Analysis

3.1. The Channel Model

3.1.1. Satellite Channel Model. As mentioned before, the satellite-UAV links are considered to be modeled as SR fading, so the probability density function (PDF) of γ_{1p} can be shown as

$$f_{\gamma_{1p}}(x) = \frac{\alpha_{1p}}{\overline{\gamma}_{1p}} e^{-\beta_1/\overline{\gamma}_1} {}_1F_1\left(m_{1p}; 1; \frac{\delta_{1p}}{\overline{\gamma}_{1p}} x\right), \; x > 0, \tag{14}$$

where ${}_1F_1(a; b; x)$ denotes the confluent hypergeometric function [38]. $\overline{\gamma}_{1p}$ represents the average SNR between Alice and pth R, $\alpha_{1p} = (2b_{1p}m_{1p}/2b_{1p}m_{1p} + \Omega_{1p})^{m_{1p}}/2b_{1p}$, $\beta_{1p} = 1/2b_{1p}$, and $\delta_{1p} = \Omega_{1p}/2b_{1p}(2b_{1p}m_{1p} + \Omega_{1p})$ with $m_{1p} \geq 0$, $2b_{1p}$, and Ω_{1p} being the fading severity parameter ranging from 0 to ∞, the average power of the multipath component, and the average power of the LOS component, respectively. In this paper, we assume that m_{1p} is an integer, and the PDF and cumulative distribution function (CDF) of γ_{1p} are, respectively, written as

$$f_{\gamma_{1p}}(x) = \alpha_{1p} \sum_{k_1=0}^{m_{1p}-1} \frac{\left(1 - m_{1p}\right)_{k_1}\left(-\delta_{1p}\right)^{k_1}}{(k_1!)^2 \overline{\gamma}_{1p}^{k_1+1}} x^{k_1} \exp\left(-\Delta_{1p}x\right),$$

$$F_{\gamma_{1p}}(x) = 1 - \alpha_{1p} \sum_{k_1=0}^{m_{1p}-1} \sum_{t=0}^{k_1} \frac{\left(1 - m_{1p}\right)_{k_1}\left(-\delta_{1p}\right)^{k_1}}{k_1!\left(\overline{\gamma}_{1p}\right)^{k_1+1} t! \Delta_{1p}^{k_1-t+1}} x^t e^{-\Delta_{1p}x}, \tag{15}$$

where $\Delta_{1p} = \beta_{1p} - \delta_{1p}/\overline{\gamma}_{1p}$ and $(\cdot)_k$ denotes the Pochhammer symbol [38].

3.1.2. Terrestrial Channel Model. In the former section, we assume the channel model between R and intended users or eavesdroppers which undergoes independent and identically distributed (i.i.d) Rayleigh fading. Thus, the CDF and PDF of γ_U $(U \in \{2p, 3p, D_i E_j\})$ are, respectively, shown as

$$F_{\gamma_U} = 1 - e^{-x/\overline{\gamma}_U} \tag{16}$$

and

$$f_{\gamma_U} = \frac{1}{\overline{\gamma}_U} e^{-x/\overline{\gamma}_U}, \tag{17}$$

where $\overline{\gamma}_U$ is the average channel gain.

By utilizing (11) and (12) along with [5], the CDF and PDF of $\gamma_{D_{pi}E}$ are, respectively, derived as

$$F_{\gamma_{D_{pi}E}}(x) = 1 - \sum_{v=0}^{N-1} \frac{1}{v} \left[\frac{x}{\overline{\gamma}_{D_{pi}E}\xi_{ip}^2 - x\overline{\gamma}_{D_{pi}E}\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right)} \right]^v e^{-x/\overline{\gamma}_{D_{pi}E}\left[\xi_{ip}^2 - x\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right)\right]},$$

$$f_{\gamma_{D_{pi}E}}(x) = \frac{\xi_{1p}^2 \overline{\gamma}_{D_{pi}E}^{-N}}{\left[\xi_{ip}^2 - x\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right)\right]^2 (N-1)!} \left(\frac{x}{\overline{\gamma}_{D_{pi}E}\xi_{ip}^2 - x\overline{\gamma}_{D_{pi}E}\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right)} \right)^{N-1} \times \exp\left(-\frac{x}{\overline{\gamma}_{D_{pi}E}\left[\xi_{ip}^2 - x\left(\xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2\right)\right]} \right). \tag{18}$$

3.2. Secrecy Outage Probability. SOP refers to the probability of reaching the nonnegative target confidentiality rate. In this section, the SOP of the NOMA-based ISMUAVN describes the confidentiality behavior, and it can be expressed as

$$P_{\text{out}}(\gamma_0) = P_1(\gamma_0) + P_{2D_2}(\gamma_0) - P_1(\gamma_0)P_{2D_2}(\gamma_0), \tag{19}$$

where

$$P_1(\gamma_0) = P_{1D_1}(\gamma_0) + P_{1D_2}(\gamma_0) - P_{1D_1}(\gamma_0)P_{1D_2}(\gamma_0) \tag{20}$$

and

$$P_{2D_2}(\gamma_0) = \mathrm{P}_{12D_2}(\gamma_0) + \mathrm{P}_{22D_2}(\gamma_0) - \mathrm{P}_{12D_2}(\gamma_0)\mathrm{P}_{22D_2}(\gamma_0), \tag{21}$$

where

$$P_{1\rho}(\gamma_0) = \mathrm{P}_{11\rho}(\gamma_0) + \mathrm{P}_{21\rho}(\gamma_0) - \mathrm{P}_{11\rho}(\gamma_0)\mathrm{P}_{21\rho}(\gamma_0). \tag{22}$$

By using (4) and (5), $P_{1i\rho}$ and $P_{2i\rho}$ are, respectively, presented as

$$\begin{aligned} P_{1i\rho}(\gamma_0) &= \left[\Pr\left(\gamma_{R_p,i} \le \gamma_0\right)\right]^M, \quad i \in \{1,2\}, \\ P_{2i\rho}(\gamma_0) &= \Pr\left(C_{S_{pi\rho}} \le C_0\right) = \int_0^\infty \int_0^{\gamma_0 + y(1+\gamma_0)} f_{\gamma_{D_{pi,\rho}}}(x) f_{\gamma_{D_{pi}E}}(y)\,\mathrm{d}x\mathrm{d}y. \end{aligned} \tag{23}$$

Also, $C_0 = \log_2(1+\gamma_0)$.

In the following, to make the analysis easier, $\overline{\gamma}_{D_{p1}E} = \overline{\gamma}_{D_{p2}E} = \overline{\gamma}_E$ is assumed, and the detailed analysis of SOP for D_1 and D_2 is, respectively, given in the following theorems.

Theorem 1. *At D_1, the closed-form expression for P_{1D_1} by detecting $x_1(t)$ is given by*

$$P_{1D_1}(\gamma_0) = P_{11D_1}(\gamma_0) + P_{21D_1}(\gamma_0) - P_{11D_1}(\gamma_0)P_{21D_1}(\gamma_0), \tag{24}$$

where

$$P_{11D_1}(\gamma_0) = \left[1 - \sum_{k_1=0}^{m_{1p}-1} \sum_{t=0}^{k_1} \frac{\alpha_{1p}\left(1-m_{1p}\right)_{k_1}\left(-\delta_{1p}\right)^{k_1}}{k_1!\left(\overline{\gamma}_{1p}\right)^{k_1+1} t!\Delta_{1p}^{k_1-t+1}} \times e^{-\Delta_{1p}\gamma_0/a_{1p}^2 - \gamma_0 B} \left(\frac{\gamma_0}{a_{1p}^2 - \gamma_0 B} \right)^t \right]^M \tag{25}$$

and

$$P_{21D_1}(\gamma_0) = 1 - \frac{H_1}{2\overline{\gamma}_E^{2N-1}(N-1)!} \sum_{i=1}^{K_1} w_i H_1(z_i). \tag{26}$$

In (25), $B = a_{1p}^2 k_{1p}^2 + a_{2p}^2 + a_{2p}^2 k_{2p}^2$. In (26),

$$H_1(z) = \frac{\xi_{1p}^2 z^{N-1}}{\left(\xi_{1p}^2 - zC\right)^{N+1}} \exp\left(-\frac{z}{\xi_{1p}^2 \overline{\gamma}_E \left(\xi_{1p}^2 - zC\right)} \right) \exp\left\{ -\frac{\gamma_0 + z(1+\gamma_0)}{\overline{\gamma}_{3p}\xi_{1p}^2 - [\gamma_0 + z(1+\gamma_0)]\overline{\gamma}_{3p}A} \right\}, \tag{27}$$

where K_1 denotes the number of terms, $H_1 = \min(\xi_{1p}^2/C, (\xi_{1p}^2 - A\gamma_0)/(A + A\gamma_0))$, $A = \xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 + \xi_{2p}^2 k_{2p}^2$, $C = \xi_{1p}^2 k_{1p}^2 + \xi_{2p}^2 k_{2p}^2$, $z_i = t_i + 1$ presents the ith zero of Legendre polynomials, and w_i denotes the Gaussian weight, which is given in Table (25.4) of [39].

With the similar method, at D_2, the exact expression of P_{1D_2} to decode $x_1(t)$ can be expressed as

$$P_{1D_2}(\gamma_0) = P_{11D_2}(\gamma_0) + P_{21D_2}(\gamma_0) - P_{11D_2}(\gamma_0)P_{21D_2}(\gamma_0), \tag{28}$$

where $P_{11D_2}(\gamma_0) = P_{11D_1}(\gamma_0)$,

$$P_{21D_2}(\gamma_0) = 1 - \frac{H_2}{2\overline{\gamma}_E^{2N-1}(N-1)!}\sum_{i=1}^{K_2} w_i H_2(z_i), \tag{29}$$

and

$$H_2(z) = \frac{\xi_{1p}^2 z^{N-1}}{\left(\xi_{1p}^2 - zC\right)^{N+1}}\exp\left(-\frac{z}{\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - zC\right)}\right)\exp\left\{-\frac{\gamma_0 + z(1+\gamma_0)}{\overline{\gamma}_{2p}\xi_{1p}^2 - [\gamma_0 + z(1+\gamma_0)]\overline{\gamma}_{2p}A}\right\}, \tag{30}$$

where $H_2 = \min(\xi_{1p}^2/C, (\xi_{1p}^2 - A\gamma_0)/(A + \gamma_0 A))$.

Proof. See Appendix A. □

Theorem 2. *$x_2(t)$ is only detected by D_2; therefore, the exact expression of P_{2D_2} to decode $x_1(t)$ can be represented as*

$$P_{2D_2}(\gamma_0) = P_{12D_2}(\gamma_0) + P_{22D_2}(\gamma_0) - P_{12D_2}(\gamma_0)P_{22D_2}(\gamma_0), \tag{31}$$

where

$$P_{12D_2}(\gamma_0) = \left[1 - \sum_{k_1=0}^{m_{1p}-1}\sum_{t=0}^{k_1}\frac{\alpha_{1p}\left(1-m_{1p}\right)_{k_1}\left(-\delta_{1p}\right)^{k_1}}{k_1!\left(\overline{\gamma}_{1p}\right)^{k_1+1}t!\Delta_{1p}^{k_1-t+1}}\times e^{-\Delta_{1p}\gamma_0/a_{2p}^2-\gamma_0 C}\left(\frac{\gamma_0}{a_{2p}^2 - \gamma_0 C}\right)^t\right]^M \tag{32}$$

and

$$P_{22D_2}(\gamma_0) = 1 - \frac{H_3}{2\overline{\gamma}_E^{2N-1}(N-1)!}\sum_{i=1}^{K_3} w_i H_3(z_i), \tag{33}$$

where

$$H_3(z) = \frac{\xi_{2p}^2 z^{N-1}}{\left(\xi_{2p}^2 - zC\right)^{N+1}}\exp\left(-\frac{z}{\xi_{2p}^2\overline{\gamma}_E\left(\xi_{2p}^2 - zC\right)} - \frac{\gamma_0 + z(1+\gamma_0)}{\overline{\gamma}_{2p}\xi_{2p}^2 - [\gamma_0 + z(1+\gamma_0)]\overline{\gamma}_{2p}C}\right). \tag{34}$$

Also, $H_3 = \min(\xi_{2p}^2/C, (\xi_{2p}^2 - C\gamma_0)/(C + \gamma_0 C))$.

Proof. See Appendix B. □

3.3. Asymptotic SOP. In the behind description, to derive deeper insights of the impact of the significant system parameters of the considered NOMA-based ISMUAVN in the high SNR regime, we show the following analysis. According to (16) and (15), when $\overline{\gamma}_{1p} \longrightarrow \infty$, we can derive

$$F_{\gamma_{1p}}(x) \approx \frac{\alpha_{1p}}{\overline{\gamma}_{1p}}x + o(x), \tag{35}$$

where $o(x)$ denotes the infinitesimal of x in the higher order. Similarly, when $\overline{\gamma}_U \longrightarrow \infty$, (17) can be rewritten as

$$F_{\gamma_U}(x) \approx \frac{x}{\overline{\gamma}_U} + o(x). \tag{36}$$

Proposition 1. *At D_1, the asymptotic SOP for decoding $x_1(t)$ of our proposed system is represented as*

$$P_{1D_1}^{\infty}(\gamma_0) = 1 + \left[\frac{\alpha_{1p}}{\overline{\gamma}_{1p}}\left(\frac{\gamma_0}{a_{1p}^2 - \gamma_0 B}\right)\right]^M - \sum_{i=1}^{K_1}\frac{z_i^{N-1}H_1 w_i \xi_{1p}^2}{2\overline{\gamma}_E^{2N-1}(N-1)!} \times \frac{e^{-z_i/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - z_i C\right)}}{\left(\xi_{1p}^2 - z_i C\right)^{N+1}}\left(1 - \frac{\gamma_0 + z_i(1+\gamma_0)}{\overline{\gamma}_{3p}\xi_{1p}^2 - [\gamma_0 + z_i(1+\gamma_0)]\overline{\gamma}_{3p}A}\right). \tag{37}$$

At D_2, the asymptotic SOP for decoding $x_1(t)$ of our proposed system is obtained as

$$P_{1D_2}^{\infty}(\gamma_0) = 1 + \left[\frac{\alpha_{1p}}{\overline{\gamma}_{1p}}\left(\frac{\gamma_0}{a_{1p}^2 - \gamma_0 B}\right)\right]^M - \sum_{i=1}^{K_2} \frac{z_i^{N-1} H_2 w_i \xi_{1p}^2}{2\overline{\gamma}_E^{2N-1}(N-1)!} \times \frac{e^{-z_i/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - z_i C\right)}}{\left(\xi_{1p}^2 - z_i C\right)^{N+1}}\left(1 - \frac{\gamma_0 + z_i(1+\gamma_0)}{\overline{\gamma}_{2p}\xi_{1p}^2 - [\gamma_0 + z_i(1+\gamma_0)]\overline{\gamma}_{2p}A}\right). \tag{38}$$

$x_2(t)$ is only detected by D_2; therefore, the asymptotic expression of SOP for decoding $x_2(t)$ for our proposed system is shown as

$$P_{2D_2}^{\infty}(\gamma_0) = 1 + \left[\frac{\alpha_{1p}}{\overline{\gamma}_{1p}}\left(\frac{\gamma_0}{a_{2p}^2 - \gamma_0 C}\right)\right]^M - \sum_{i=1}^{K_3} \frac{z_i^{N-1} H_3 w_i \xi_{2p}^2}{2\overline{\gamma}_E^{2N-1}(N-1)!} \times \frac{e^{-z_i/\xi_{2p}^2\overline{\gamma}_E\left(\xi_{2p}^2 - z_i C\right)}}{\left(\xi_{2p}^2 - z_i C\right)^{N+1}}\left(1 - \frac{\gamma_0 + z_i(1+\gamma_0)}{\overline{\gamma}_{2p}\xi_{2p}^2 - [\gamma_0 + z_i(1+\gamma_0)]\overline{\gamma}_{2p}C}\right). \tag{39}$$

Proof. At first, by replacing (15) and (16) with (35) and (36) and then with the same method of Appendixes A and B, the exact equations for the asymptotic analysis (namely, (37), (38), and (40)) are derived.

In this sequel, we obtain the diversity order along with the coding gain. By letting $\overline{\gamma}_{1p} = \overline{\gamma}_{2p} = \overline{\gamma}_{3p} = \overline{\gamma}$, (22) can be expressed as

$$P_{\text{out}}^{\infty}(\gamma_0) \approx G_a\left(\frac{1}{\overline{\gamma}}\right)^{G_d}, \tag{40}$$

where $G_d = \min(M, 1)$ denotes the secrecy diversity order, and the secrecy coding gain of the considered NOMA-based ISMUAVN includes two cases, i.e., when $M \le 1$,

$$G_a = \left\{ \frac{2\alpha_{1p}\gamma_0}{a_{1p}^2 - \gamma_0 B} + \frac{\alpha_{1p}\gamma_0}{a_{2p}^2 - \gamma_0 C} + 2\sum_{i=1}^{K_1} \frac{z_i^{N-1} H_1 w_i \xi_{1p}^2 e^{-z_i/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - z_i C\right)}[\gamma_0 + z_i(1+\gamma_0)]}{2\overline{\gamma}_E^{2N-1}(N-1)!\left(\xi_{1p}^2 - z_i C\right)^{N+1}\left\{\xi_{1p}^2 - [\gamma_0 + z_i(1+\gamma_0)]A\right\}} \right. \quad + \sum_{i=1}^{K_3} \frac{z_i^{N-1} H_3 w_i \xi_{2p}^2 e^{-z_i/\xi_{2p}^2\overline{\gamma}_E\left(\xi_{2p}^2 - z_i C\right)}[\gamma_0 + z_i(1+\gamma_0)]}{2\overline{\gamma}_E^{2N-1}(N-1)!\left(\xi_{2p}^2 - z_i C\right)^{N-1}\left\{\xi_{2p}^2 - [\gamma_0 + z_i(1+\gamma_0)]C\right\}}. \tag{41}$$

When $M > 1$,

$$G_a = \left\{ 2\sum_{i=1}^{K_1} \frac{z_i^{N-1} H_1 w_i \xi_{1p}^2 e^{-z_i/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - z_i C\right)}[\gamma_0 + z_i(1+\gamma_0)]}{2\overline{\gamma}_E^{2N-1}(N-1)!\left(\xi_{1p}^2 - z_i C\right)^{N+1}\left\{\xi_{1p}^2 - [\gamma_0 + z_i(1+\gamma_0)]A\right\}} + \sum_{i=1}^{K_3} \frac{z_i^{N-1} H_3 w_i \xi_{2p}^2 e^{-z_i/\xi_{2p}^2\overline{\gamma}_E\left(\xi_{2p}^2 - z_i C\right)}[\gamma_0 + z_i(1+\gamma_0)]}{2\overline{\gamma}_E^{2N-1}(N-1)!\left(\xi_{2p}^2 - z_i C\right)^{N-1}\left\{\xi_{2p}^2 - [\gamma_0 + z_i(1+\gamma_0)]C\right\}}. \tag{42}$$

□

4. Numerical Results

The correctness of the performance analysis of our considered system is verified by Monte Carlo simulations. In general, we assume $\delta_R^2 = \delta_{D_1}^2 = \delta_{D_2}^2 = \delta_{E_j}^2 = 1$, $K = M = 8$, and $\overline{\gamma}_{1p} = \overline{\gamma}_{2p} = \overline{\gamma}_{3p} = \overline{\gamma}$. The channel parameters are given in Table 1 [5]. Different power coefficient combinations are assumed as follows: scenario 1: $a_1^2 = \xi_1^2 = 0.9$ and $a_2^2 = \xi_2^2 = 0.1$; scenario 2: $a_1^2 = \xi_1^2 = 0.8$ and $a_2^2 = \xi_2^2 = 0.2$; scenario 3: $a_1^2 = \xi_1^2 = 0.7$ and $a_2^2 = \xi_2^2 = 0.3$ (In this paper, the fixed power allocation factors are considered; however, the results derived can also be utilized for our future work when power allocation is studied.).

As proved in [40], a satellite channel model is expected to be general and applicable for a wide range of elevation angles, under which the satellite can be observed. In this regard, the most common approach to evaluate the impact of elevation angle on channel statistical parameters is based on the transformation from empirical expression. According to Abdi et al. [40], the maximum elevation angle is around 80°, while the minimum elevation angle is considered usually around 20° in order to tackle the geographical terrain affects. Hence, this is particularly useful when we apply a set of data parameters with moderate variation to a model with specific shadowing and infrequent light shadowing conditions (i.e., frequent heavy shadowing, average shadowing, and infrequent light shadowing). Please note that, for specific shadowing conditions, different parameters have been employed to cover a range of elevation angles in many existing works [5].

Figure 2 illustrates the SOP versus $\overline{\gamma}$ and k by setting $\overline{\gamma}_E = 1$ dB, $\gamma_0 = 0$ dB, $N = 2$, scenario 1, and $M = 3$ under different shadowing fading. We can find that when $\overline{\gamma}$ grows to be a fixed value, the SOP tends to be constant. Besides, it can be found that the SOP descends with the increase of the nonideal hardware level. In addition, it should be noted that the SOP under different channel conditions is consistent for high SNR regimes, which means that channel conditions will seriously affect the SOP.

Figure 3 plots the SOP versus $\overline{\gamma}$ for different N and k with $\gamma_0 = 0$ dB, $\overline{\gamma}_E = 1$ dB, $M = 3$, and scenario 1 under the AS scenario. It can be found that, with N being larger, the SOP will be larger because more eavesdroppers are utilized to

Table 1: Channel parameters.

Shadowing	m_{1p}	b_{1p}	Ω_{1p}
Frequent heavy shadowing (FHS)	1	0.063	0.0007
Average shadowing (AS)	5	0.251	0.279
Infrequent light shadowing (ILS)	10	0.158	1.29

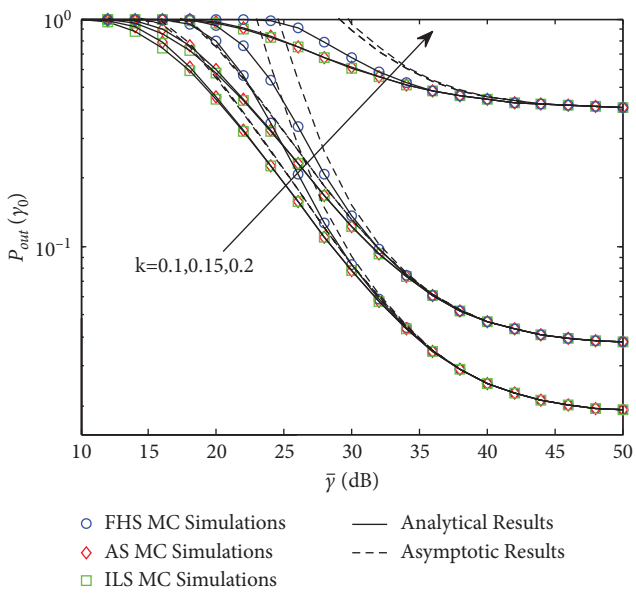


Figure 2: The SOP versus $\bar{\gamma}$ and k by setting $\bar{\gamma}_E = 1$ dB, $\gamma_0 = 0$ dB, $N = 2$, scenario 1, and $M = 3$ under different shadowing fading.

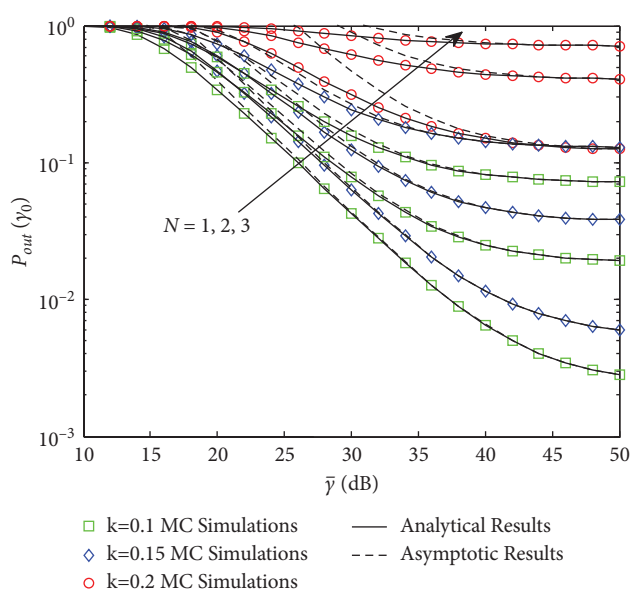


Figure 3: The SOP versus $\bar{\gamma}$ for different N and k with $\gamma_0 = 0$ dB, $\bar{\gamma}_E = 1$ dB, $M = 3$, and scenario 1 under the AS scenario.

steal the legitimate signals. Similar to Figure 2, the SOP will be a fixed value for big enough $\bar{\gamma}$. This is due to the fact that

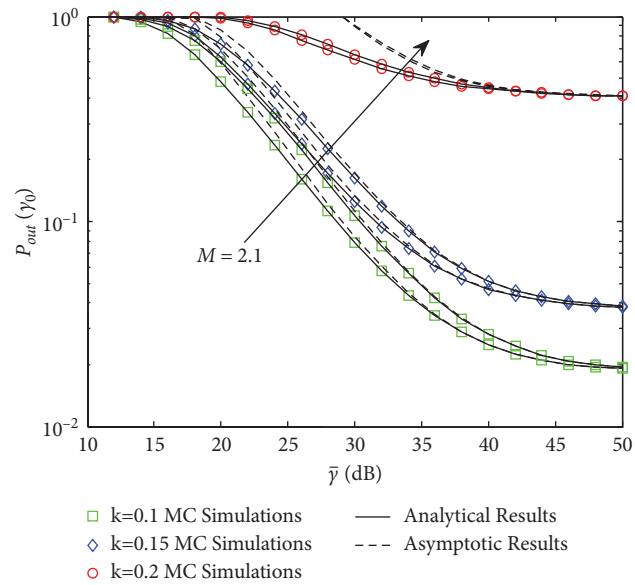


Figure 4: The SOP versus $\bar{\gamma}$ for different M and k with $\bar{\gamma}_E = 1$ dB, $\gamma_0 = 0$ dB, $N = 2$, and scenario 1 under the AS scenario.

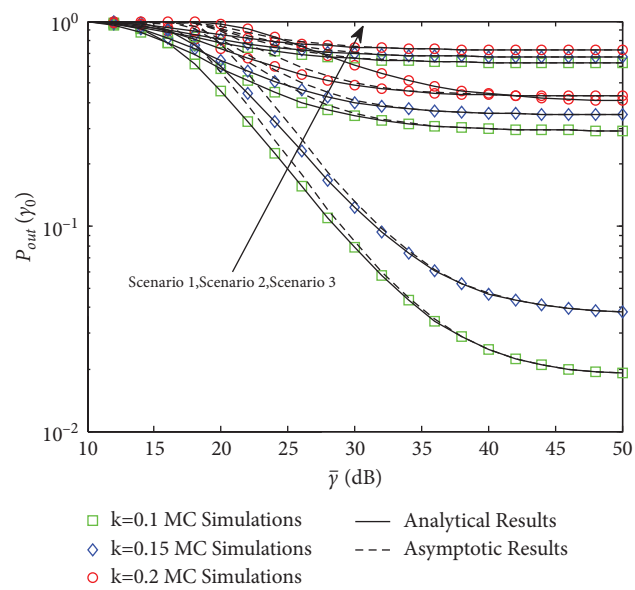


Figure 5: The SOP versus $\bar{\gamma}$ for different scenarios and k with $\bar{\gamma}_E = 1$ dB, $\gamma_0 = 0$ dB, $N = 2$, and $M = 3$ under the AS scenario.

an upper bound of the SINR is existed to the legitimate user, which is related to (2), (3), (7), (8), and (9).

Figure 4 examines the SOP versus $\bar{\gamma}$ for different M and k with $\bar{\gamma}_E = 1$ dB, $\gamma_0 = 0$ dB, $N = 2$, and scenario 1 under the AS scenario. We may obtain that when M grows larger, the SOP will be lower. Figure 5 illustrates the SOP versus $\bar{\gamma}$ for different scenarios and k with $\bar{\gamma}_E = 1$ dB, $\gamma_0 = 0$ dB, $N = 2$, and $M = 3$ under the AS scenario. We can obtain that, with the gap of the power allocation factors becoming larger, the SOP will be lower; this is because the communication quality of the worse user is ensured. At last, we find that the secrecy

performance will deteriorate with the nonideal hardware level.

5. Conclusions

In this paper, we studied the secrecy outage probability of NOMA-based ISMUAVNs in which nonideal hardware is considered and the partial UAV selection scheme is adopted. The exact closed-form expression of SOP for our proposed system was derived. Moreover, in the high SNR regime, we derived the asymptotic SOP. To get more insights, the secrecy coding gain and secrecy diversity order were obtained. The impacts of NOMA technology, nonideal hardware, and the partial UAV selection scheme on the performance of the considered system were revealed by theoretical analysis.

Abbreviations

AS: Average shadowing
ASC: Average secrecy capacity
AWGN: Additive white Gaussian noise
CDF: Cumulative distribution function
CSI: Channel state information
DF: Decode and forward
5G: Fifth generation
FHS: Frequency heavy shadowing
ILS: Infrequent light shadowing
ISMUAVNs: Integrated satellite-multiple unmanned aerial vehicle networks
ISTRNs: Integrated satellite-terrestrial relay networks
LOS: Line of sight
MC: Monte Carlo
NOMA: Nonorthogonal multiple access
OP: Outage probability
PDF: Probability density function
PIC: Parallel interference cancellation
PLS: Physical layer security
SatCom: Satellite communication
SIC: Successive interference cancellation
SINDR: Signal-to-interference-plus-noise-and-distribution ratio
SNDR: Signal-to-noise-and-distortion ratio
SNR: Signal-to-noise ratio
SNRs: Signal-to-noise ratios
SOP: Secrecy outage probability
SR: Shadowed-Rician
UAV: Unmanned aerial vehicle.

Appendix

A. Proof of Theorem 1

By utilizing (22), to get P_{11D_1}, we should first derive $P_{11D_1}(\gamma_0)$ and $P_{21D_1}(\gamma_0)$. In the following, they are derived.

By utilizing (25) and (2), (25) is rewritten as

$$P_{11D_1}(\gamma_0) = \Pr\left[\max_{p\in\{1,\ldots,M\}}\left(\gamma_{R_p,1}\right) \le \gamma_0\right] = \left[F_{\gamma_{1p}}\left(\frac{\gamma_0}{a_{1p}^2 - \gamma_0 B}\right)\right]^M. \tag{A.1}$$

From (23) and after some mathematical derivations, it can be seen that

$$\mathrm{P}_{21D_1}(\gamma_0) = \int_0^{H_1} F_{\gamma_{D_{p1},D_1}}(\gamma_0 + y(1+\gamma_0)) f_{\gamma_{D_{p1}E}}(y)\,\mathrm{d}y + \int_{H_1}^{\infty} f_{\gamma_{D_{p1}E}}(y)\,\mathrm{d}y. \tag{A.2}$$

In (A.2), obtaining the CDF of γ_{D_{p1},D_1} and PDF of $\gamma_{D_{p1}E}$ is the earlier thing needed to be done.

By utilizing (9) and (16), the CDF of γ_{D_{p1},D_1} can be written as

$$F_{\gamma_{D_{p1},D_1}}(x) = 1 - e^{-\gamma_0 + y(1+\gamma_0)/\overline{\gamma}_{3p}\xi_{1p}^2 - [\gamma_0 + y(1+\gamma_0)]\overline{\gamma}_{3p}A}. \tag{A.3}$$

By inserting (A.3) and (18) into (A.2), we can obtain

$$\begin{aligned} P_{21D_1}(\gamma_0) &= \int_0^{\infty} \frac{\xi_{1p}^2 e^{-y/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - yC\right)}\left\{1 - e^{-\gamma_0 + y(1+\gamma_0)/\overline{\gamma}_{3p}\xi_{1p}^2 - [\gamma_0 + y(1+\gamma_0)]\overline{\gamma}_{3p}A}\right\}}{\overline{\gamma}_E^N y^{1-N}(N-1)!\left(\xi_{1p}^2 - yC\right)\left(\overline{\gamma}_E\xi_{1p}^2 - y\overline{\gamma}_E C\right)^{N-1}}\,\mathrm{d}y \\ &= 1 - \frac{\overline{\gamma}_E^{-N}\xi_{1p}^2}{(N-1)!}\underbrace{\left[\begin{array}{c}\displaystyle\int_0^{H_1} \frac{e^{-\gamma_0 + y(1+\gamma_0)/\overline{\gamma}_{3p}\xi_{1p}^2 - [\gamma_0 + y(1+\gamma_0)]\overline{\gamma}_{3p}A}\left(\overline{\gamma}_E\xi_{1p}^2 - y\overline{\gamma}_E C\right)^{1-N}}{e^{y/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - yC\right)} y^{1-N}\left(\xi_{1p}^2 - yC\right)}\,\mathrm{d}y \\ \displaystyle + \int_{H_1}^{\infty} \frac{e^{-y/\xi_{1p}^2\overline{\gamma}_E\left(\xi_{1p}^2 - yC\right)} y^{N-1}}{\left(\xi_{1p}^2 - yC\right)\left(\overline{\gamma}_E\xi_{1p}^2 - y\overline{\gamma}_E C\right)^{N-1}}\,\mathrm{d}y \end{array}\right]}_{H_1(y)}. \end{aligned} \tag{A.4}$$

However, trying the authors' best, the exact expression of $H_1(\cdot)$ is difficult to be obtained. Thus, by utilizing the Gaussian–Chebyshev quadrature [39], $H_1(\cdot)$ is rewritten as (27). By inserting (27) into (26) and after the simple process, we can finally derive the exact expression for $P_{21D_1}(\gamma_0)$. Utilizing the similar methods, the exact

expression for $P_{21D_2}(\gamma_0)$ can also be obtained. The proof is completed.

B. Proof of Theorem 2

With the help of (8) and (16), the CDF of γ_{D_{p2},D_2} can be expressed as

$$F_{\gamma_{D_{p2},D_2}}(x) = 1 - e^{-\gamma_0 + z(1+\gamma_0)/\overline{\gamma}_{2p}\xi_{2p}^2 - [\gamma_0 + z(1+\gamma_0)]\overline{\gamma}_{2p}C}. \quad \text{(B.1)}$$

Then, utilizing (3) and (15), the exact expression of $P_{12D_2}(\gamma_0)$ will be derived. Then, by inserting (B.1) and (18) into (23), we can obtain the exact expression of $P_{22D_2}(\gamma_0)$.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the Natural Science Foundation of China (Grant nos. 61673108 and 62001517), in part by the Colleges and Universities Natural Science Foundation in Jiangsu Province (Grant no. 19KJA110002), and in part by the Industry-University-Research Cooperation Project of Jiangsu Province (nos. BY2020335 and BY2020358). Thanks go to all the authors for their contribution to this paper.

References

[1] K. Guo, K. An, B. Zhang et al., "Physical layer security for multiuser satellite communication systems with threshold-based scheduling scheme," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5129–5141, 2020.

[2] K. Guo, K. An, F. Zhou, T. A. Tsiftsis, G. Zheng, and S. Chatzinotas, "On the secrecy performance of NOMA-based integrated satellite multiple-terrestrial relay networks with hardware impairments," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3661–3676, 2021.

[3] K. Guo, M. Lin, B. Zhang, W.-P. Zhu, J.-B. Wang, and T. A. Tsiftsis, "On the performance of LMS communication with hardware impairments and interference," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1490–1505, 2019.

[4] K. Guo, M. Lin, B. Zhang et al., "Performance analysis of hybrid satellite-terrestrial cooperative networks with relay selection," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9053–9067, 2020.

[5] P. K. Sharma, P. K. Upadhyay, D. B. da Costa, P. S. Bithas, and A. G. Kanatas, "Performance analysis of overlay spectrum sharing in hybrid satellite-terrestrial systems with secondary network selection," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6586–6601, 2017.

[6] K. Guo, K. An, B. Zhang, Y. Huang, and D. Guo, "On the performance of cognitive satellite-terrestrial relay networks with channel estimation error and hardware impairments," *Sensors*, vol. 18, pp. 1–19, 2018.

[7] Q. Huang, M. Lin, J.-B. Wang, T. A. Tsiftsis, and J. Wang, "Energy efficient beamforming schemes for satellite-aerial-terrestrial networks," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3863–3875, 2020.

[8] Y. Qu, H. Dai, H. Wang et al., "Service provisioning for UAV-enabled mobile edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 11, pp. 3287–3305, 2021.

[9] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2020.

[10] J. Chen, P. Chen, Q. Wu, Y. Xu, N. Qi, and T. Fang, "A game-theoretic perspective on resource management for large-scale UAV communication networks," *China Communications*, vol. 18, no. 1, pp. 70–87, 2021.

[11] X. Liu, Y. Liu, and Y. Chen, "Machine learning empowered trajectory and passive beamforming design in UAV-RIS wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 2042–2055, 2021.

[12] S. M. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource allocation for downlink NOMA systems: key techniques and open issues," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 40–47, 2018.

[13] X. Li, "Cognitive amBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[14] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[15] R. Liu, K. Guo, K. An, S. Zhu, and H. Shuai, "NOMA-based integrated satellite-terrestrial relay networks under spectrum sharing environment," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1266–1270, 2021.

[16] H. Lei, Z. Yang, K.-H. Park et al., "Secrecy outage analysis for cooperative NOMA systems with relay selection schemes," *IEEE Transactions on Communications*, vol. 67, no. 9, pp. 6282–6298, 2019.

[17] F. Zhou, Y. Wu, Y.-C. Liang, Z. Li, Y. Wang, and K.-K. Wong, "State of the art, taxonomy, and open issues on cognitive radio networks with NOMA," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 100–108, 2018.

[18] X. Yan, H. Xiao, C.-X. Wang, K. An, A. T. Chronopoulos, and G. Zheng, "Performance analysis of NOMA-based land mobile satellite networks," *IEEE Access*, vol. 6, pp. 31327–31339, 2018.

[19] R. Wan, L. Zhu, T. Li, and L. Bai, "A NOMA-PSO based cooperative transmission method in satellite communication systems," in *Proceedings of the 2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, vol. 2017, pp. 1–6, Nanjing, China, October 2017.

[20] K. Guo, K. An, B. Zhang, Y. Huang, and D. Guo, "Physical layer security for hybrid satellite terrestrial relay networks with joint relay selection and user scheduling," *IEEE Access*, vol. 6, pp. 55815–55827, 2018.

[21] B. Li, Z. Fei, C. Zhou, and Y. Zhang, "Physical-layer security in space information networks: a survey," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 33–52, 2020.

[22] K. Guo, M. Lin, B. Zhang, J. Ouyang, and W.-P. Zhu, "Secrecy performance of satellite wiretap channels with multi-user opportunistic scheduling," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 1054–1057, 2018.

[23] K. Guo, K. An, Y. Huang, and B. Zhang, "Physical layer security of multiuser satellite communication systems with channel estimation error and multiple eavesdroppers," *IEEE Access*, vol. 7, pp. 96253–96262, 2018.

[24] V. Bankey and P. K. Upadhyay, "Ergodic secrecy capacity analysis of multiuser hybrid satellite-terrestrial relay networks with multiple eavesdroppers," in *Proceedings of the 2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Shanghai, China, May 2019.

[25] K. Guo, K. An, and X. Tang, "Secrecy performance for integrated satellite terrestrial relay systems with opportunistic scheme," in *Proceedings of the 2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Shanghai, China, May 2019.

[26] X. Li, M. Huang, Y. Liu, V. G. Menon, A. Paul, and Z. Ding, "I/Q imbalance aware nonlinear wireless-powered relaying of B5G networks: security and reliability analysis," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 4, pp. 2995–3008, 2021.

[27] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[28] E. Bjornson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying: performance limits with hardware impairments," *IEEE Transactions on Communications*, vol. 61, no. 11, pp. 4512–4525, 2013.

[29] P. K. Sharma and P. K. Upadhyay, "Cognitive relaying with transceiver hardware impairments under interference constraints," *IEEE Communications Letters*, vol. 20, no. 4, pp. 820–823, 2016.

[30] K. Guo, K. An, B. Zhang, Y. Huang, and G. Zheng, "Outage analysis of cognitive hybrid satellite-terrestrial networks with hardware impairments and multi-primary users," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 816–819, 2018.

[31] F. Zhou, R. Wang, and K. Guo, "Secrecy outage analysis of satellite communication networks with hardware impairments and channel estimation errors," *Electronics Letters*, vol. 56, no. 20, pp. 1059–1062, 2020.

[32] X. Tang, K. An, K. Guo, Y. Huang, and S. A. Wang, "Outage analysis of non-orthogonal multiple access-based integrated satellite-terrestrial relay networks with hardware impairments," *IEEE Access*, vol. 7, pp. 141258–141267, 2019.

[33] K. Guo, "NOMA-based integrated satellite-terrestrial multi-relay networks with hardware impairments and partial relay selections scheme," in *Proceedings of the 2019 IEEE 19th International Conference on Communication Technology (ICCT)*, pp. 1–6, Xi'an, China, October 2019.

[34] F. Zhou, R. Wang, and J. Bian, "Performance analysis of non-orthogonal multiple access based-satellite communication networks with hardware impairments and channel estimations," *Electronics Letters*, vol. 56, no. 1, pp. 52–55, 2020.

[35] K. Guo, K. An, B. Zhang et al., "On the performance of the uplink satellite multiterrestrial relay networks with hardware impairments and interference," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2297–2308, 2019.

[36] Z. Qin, Y. Liu, Z. Ding, Y. Gao, and M. Elkashlan, "Physical layer security for 5G non-orthogonal multiple access in large-scale networks," in *Proceedings of the 2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kuala Lumpur, Malaysia, May 2016.

[37] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and L. Hanzo, "Enhancing the physical layer security of non-orthogonal multiple access in large-scale networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1656–1672, 2017.

[38] I. S. Gradshteyn, I. M. Ryzhik, A. Jeffrey, and D. Zwillinger, *Table of Integrals, Series, and Products*, Elsevier/Academic Press, Amsterdam, The Netherlands, 7th edition, 2007.

[39] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, Springfield, IL, USA, 9th edition, 1972.

[40] A. Abdi, W. C. Lau, M. Alouini, and M. Kaveh, "A new simple model for land mobile satellite channels: first- and second-order statistics," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 519–528, 2003.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4351518, 5 pages
https://doi.org/10.1155/2022/4351518

Review Article

UAV-Based Physical-Layer Intelligent Technologies for 5G-Enabled Internet of Things: A Survey

Changyu Wang,[1] Weili Yu,[1] Jinrong Lu,[2,3] Fusheng Zhu,[4] Lihua Fan,[5] and Shengping Li[6]

[1]*Air Force Aviation University, Changchun, Jilin, China*
[2]*School of Computer Science, Guangzhou University, Guangzhou, China*
[3]*Guangdong Provincial Key Laboratory of Building Energy Efficiency and Application Technologies, Guangzhou University, Guangzhou, China*
[4]*Guangdong New Generation Communication and Network Innovative Institute (GDCNi), Guangzhou, China*
[5]*School of Electronic Information, Shunde Polytechnic, Foshan, Guangdong, China*
[6]*College of Engineering, Shantou University, Shantou, Guangdong, China*

Correspondence should be addressed to Fusheng Zhu; zhufusheng@gdcni.cn

Received 8 December 2021; Revised 4 January 2022; Accepted 10 January 2022; Published 28 January 2022

Academic Editor: Xingwang Li

In recent years, the utilization and application of unmanned aerial vehicle (UAV) have attracted much attention, both from academy and industry. UAVs have been widely used in many practical communication scenarios, due to its high flexibility, high mobility, and low cost. Therefore, this paper addresses the key technologies of UAV communication and reviews the current research status, from various aspects including UAV communication transmission, UAV formation control and networking, UAV resource allocation, and intelligent communication from artificial intelligence algorithms. Then, artificial intelligence is introduced into multiple aspects of UAV communication, including channel transmission, control and networking, and resource scheduling, to organically integrate artificial intelligence into UAV communication, which can help reduce the complexity of communication algorithms and improve system spectrum efficiency. This paper will help improve the efficiency of UAV communication and promote the development of UAV-related industries.

1. Research Background

Unmanned aerial vehicle (UAV) is a new and evolving technology that has gained much attention over the past few decades due to its flexibility, high speed mobility, and low cost [1–3]. Initially, UAVs were mainly used in military operations to perform critical missions and reduce pilot damage. But as the cost of manufacturing UAVs have continued to decrease and the size of the equipment has been miniaturized, numerous other applications for UAVs have emerged [4–6]. As shown in Figure 1, typical UAV applications include emergency search and rescue, climate monitoring, traffic control, forest fire detection, cargo logistics, aerial photography, and communications.

Motivated by the development and deployment of the fifth-generation (5G) communication [7–9], UAV communication has been successfully applied in the academy and industry, due to the advantages of flexible deployment, wide coverage, and no regional restrictions. Many countries around the world have carried out research on UAV communication. In 2016, the Central Committee of the Communist Party of China, the State Council and the Central Military Commission issued the Opinions on the Integrated Development of Economic Construction and National Defence Construction, pointing out that UAV communication is an important part of the integration of military and civilian construction of air, space, and sky networks. In 2017, the State Council issued the Development Plan for a New Generation of Artificial Intelligence, which called for the construction of an innovative theoretical system for the autonomous collaborative control and optimal decision of intelligent

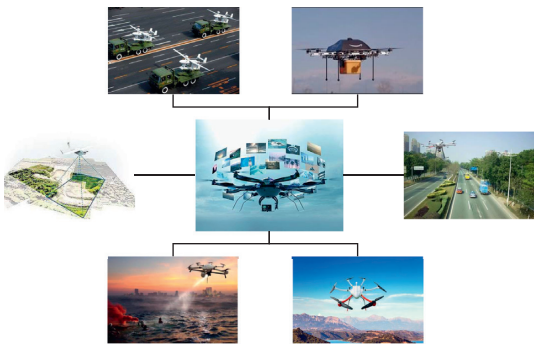


Figure 1: Typical application diagram of UAV.

unmanned systems. In March 2020, in order to implement the overall plan of the State Council's New Generation Artificial Intelligence Development Plan, the Ministry of Science and Technology further launched the implementation of the New Generation Artificial Intelligence Major Science and Technology Project Implementation Plan.

In order to achieve efficient UAV swarm communication coverage, smooth communication links need to be maintained at all times between swarm UAVs and UAVs or between UAVs and service terminals. The future UAV communication network needs to support a user experience rate of gigabits per second, fast and flexible UAV swarm networking, low latency control signaling, and 100 megabit system bandwidth. Considering the demand for high-capacity and high-rate communications for swarm communications, the technological focus of UAV communications is beginning to shift to the high frequency band such as millimeter wave (30–300 GHz). Millimeter waves can provide larger bandwidth resources, enabling user rates up to gigabits per second. At the same time, the tiny millimeter wave wavelength allows hundreds of antennas to be mounted in a small space or integrated into a chip, giving UAVs ultra-high-resolution beam assignment capabilities. Therefore, with the help of millimeter wave and multiantenna technology [10, 11], UAV swarm communication system is expected to realize broadband wireless transmission through collaboration and networking and improve network capacity and link reliability through multi-UAV formation, rapid deployment, and collaboration. In addition to the aforementioned techniques, some other techniques, such as the reconfigurable intelligent surfaces [12–17] and multiple access [18–20], can be integrated to the UAV communication systems, in order to further enhance the system performance.

However, UAV swarm communication based on millimeter-wave and multiple antennas also faces great challenges. The sparse characteristics of millimeter-wave propagation in high frequency bands, the application of multiple antennas, the high-speed mobility of UAVs, the dynamic change of formations, and the limited multidimensional resources of swarms cause that the key technologies of swarm communication need to be reexamined. UAV swarm communication breaks the application scope of traditional mobile communication systems. The unified technical framework which supports the extremely different and complicated applications of UAV swarm faces unprecedented challenges.

In recent years, artificial intelligence techniques [21–23] have continued to evolve with the rapid development of intelligent mobile computing and data storage technologies [24–26]. Artificial intelligence can find its typical applications for many scenarios including statistics, inference, fitting, optimization, and clustering of data. The rapid development of artificial intelligence techniques offers potential possibilities to address the challenges of swarm communication and to surpass design concepts and performance of the traditional mobile communication. This paper will design a swarm intelligent communication transmission strategy to break the traditional communication model and introduce artificial intelligence into multiple levels of UAV swarm communication, such as broadband transmission, formation control and networking, and multidimensional resource allocation, to substantially improve the performance of the swarm communication system, realize the organic integration of artificial intelligence and UAV swarm communication system, and promote the development and maturity of UAV swarm communication.

2. Research Progress and Challenges of UAV Swarm Communication Transmission

With the rapid development of automatic control and artificial intelligence technology, UAV communication is widely used in various fields such as aerial photography, logistics, and entertainment, which has attracted much attention from researches and engineers at home and abroad. For swarm communication, the design and performance evaluation of the system require the establishment of UAV communication channel models. In general, the channel model can be divided into two categories: deterministic channel models and random channel model. Unlike the deterministic model, the stochastic model has high accuracy and low complexity, and it can be more easily extended to various wireless communication application scenarios. In this direction, a three-dimensional geometric stochastic model could be proposed to describe the communication channel, which assumed that mobile transmitter and receiver moved in a flat environment and did not consider the effect of ground reflection on the channel modeling. A 3D ellipsoidal channel model could be designed for high-altitude platforms, which did not consider the effect of platform elevation angle on the channel modeling. The technique of multiple-input multiple-output could be integrated into the high-altitude platform system and proposed a three-dimensional geometric single-hop reflection reference model. At the same time, different from traditional terrestrial cellular mobile communication networks, millimeter wave wavelengths are shorter and have severe path loss, and millimeter wave-based UAV broadband communication links are highly susceptible to atmospheric environment, terrain, and various obstacles. The beamforming technique could be applied to improve the signal quality at the UAV

receiver. Moreover, single- and multilayer beamforming methods could be utilized for large-scale antenna arrays. Furthermore, a simulated beam assignment method could be provided for a single RF link to multiple antenna channels. However, the existing works lack a comprehensive consideration of the number and cost of RF links and the performance of analog and digital beamforming algorithms. It is hence urgent to study the intelligent communication strategies to support UAV swarms in complied application scenarios, including channel modeling, channel information acquisition, and hybrid beam assignment and tracking.

3. Research Progress and Challenges of UAV Swarm Communication Formation Control and Networking

Formation control is the key technology of UAV swarm cooperative communication and networking. In the direction, a formation strategy was proposed based on a star network structure to ensure UAV network routing performance. A multihop reactive AODV routing protocol was devised to adapt to the change in dynamic wireless links. However, the protocol search time was long when the link was unstable, which could not meet the high dynamic communication requirements of UAVs. A UAV formation strategy could be designed by aiming at energy and time optimization. The use of UAVs could be used as the network structure of aerial base stations, where we could analyze the relationship between UAV height and ground coverage and obtain the UAV height under the best ground coverage of UAV. The relationship between the number of drone hovering times and the coverage probability of the ground terminal could be obtained through theoretical analysis. A spatial three-dimensional Poisson point process model could be utilized to obtain the communication interruption probability under different UAV densities and navigation altitudes of the base station deployment. Moreover, a wireless network relaying model for UAVs could be deployed to solve the problem of blind area coverage overlap. Furthermore, the multilayer core network of UAV based on millimeter-wave could be proposed. According to above literature review, we can find that there is lack of research on drone swarm control and collaborative networking in the existing works. Moreover, the UAV motion model based on the fixed route cannot fully characterize the dynamic and random change of the UAV path, and the existing routing method is limited by the communication range of the star network node, which cannot meet the requirement of large and dynamic UAV swarm formation activities. Therefore, it is urgent to study the UAV swarm communication control and networking technology suitable for UAV swarm communication characteristics and dynamic topology change.

4. Research Progress and Challenges of UAV Swarm Communication Resource Allocation

Due to the limited resources available for drone swarms in air navigation, the key to improving the communication performance in drone swarms is the allocation of various resources of drone swarms timely and efficiently. In the direction, the opportunities and challenges of millimeter-wave massive MIMO could be reviewed in 5G hotspot coverage, dense networking, and heterogeneous networking. The lower bound of the spectrum efficiency of massive MIMO systems could be analyzed, and the energy efficiency of distributed massive MIMO systems could be optimized through antenna selection, user clustering, and power control. A power consumption model could be established to optimize the number of base station antennas and improve the system energy efficiency. A joint carrier selection and resource allocation method could be devised based on a greedy algorithm to improve system throughput. Aiming at the problem of the limited supply of UAV platforms, an energy consumption optimization algorithm could be used for multi-UAV cooperative assistance ground communication. Aiming at different service quality requirements, random user scheduling strategies could be optimized based on the service and lacking analysis and developed multiple differentiated service types for future drone swarm communication. Considering the high mobility, dynamic topology, energy limitation, and communication link performance of drone swarm communication, it is urgent to study the resource allocation method that takes into account the multicategory business requirement and resource limitation of drones, which can help obtain the optimal system performance under limited resources, solve the bottleneck problem of drone swarm scheduling, and achieve reliable information transmission.

5. Research Progress and Challenges of Intelligent Communication

AI is a new technique that studies and exploits theories, methods, technologies, and application systems for simulating, extending, and expanding human intelligence. Intelligent communication methods jointly designed by AI and communications are expected to significantly improve the performance of communication systems [27, 28]. In this area, the authors in [29–31] proposed to utilize the federated learning to wireless communications, which opens a new window for the design of UAV-based communication systems. Moreover, some works based on the reinforcement learning in mobile edge computing networks [32, 33] could be applied to the UAV communication systems, in order to achieve a fine scheduling and trajectory design of multiple UAVs, which utilized Q-learning to improve the scheduling performance of the system. Furthermore, deep reinforcement learning and federation learning could be combined with edge computing in a joint design to improve the performance of mobile communication systems. A channel prediction method for large-scale MIMO communication systems could be proposed based on deep learning, which effectively reduced the complexity of downlink channel information acquisition and training feedback overhead. A machine learning-based hybrid precoding method for UAV millimeter-wave large-scale MIMO communication systems was proposed to improve the robustness of UAV communication transmission. However, the above existing works combining AI and communication mainly consider traditionally fixed base station scenarios and

lack consideration of mobility, formation topology, and resource constraint of swarm UAV communication systems. Hence, it is urgent to study the use of advanced AI techniques to solve UAV swarm communication challenges.

It is clear from the above wide literature review that the existing works lack the study that integrates intelligent and efficient UAV swarm communication, formation control and collaborative networking, and resource allocation, and it is urgent to study communication, control, and networking methods that are applicable to the characteristics of swarms and dynamic topological change. Therefore, this paper introduces AI into multiple levels of UAV communication, including channel transmission, control and networking, and resource scheduling, to integrate AI with UAV swarm communication, thereby reducing the complexity of swarm communication algorithms and improving system spectrum efficiency. To address the challenges of swarm network navigation trajectory, attitude, formation topology, and large-scale dynamic change in space and time through efficient UAV swarm intelligent collaborative control and networking, we take into account the system spectrum efficiency, energy efficiency, user transmission rate, quality of service, and fairness and other requirements. We, in future, use AI to optimally schedule UAV swarm antenna, position, motion, attitude, carrier, spectrum, power, beam, storage, computation, and other resources, in order to achieve an efficient, robust, and intelligent resource allocation for the UAV swarm communication.

6. Conclusions

With its flexibility, high maneuverability, and low cost, UAV communication has promoted the application of UAV communication in emergency search and rescue, climate monitoring, traffic control, forest fire detection, cargo logistics, aerial photography, and other fields. Based on the key technologies of UAV communication, this paper summarized the current research status of UAV swarm communication transmission, UAV swarm communication formation control and networking, UAV swarm communication resource allocation, and intelligent communication and so on. Furthermore, AI was introduced into multiple aspects of UAV communication, including channel transmission, control and networking, and resource scheduling. In this work, AI was integrated into the UAV swarm communication to reduce the complexity of swarm communication algorithms and improve the system spectrum efficiency. The research in this paper will improve the efficiency of UAV communication and promote the development of UAV-related industries.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key-Area Research and Development Program of Guangdong Province, China (no. 2018B010124001), National Natural Science Foundation of China (no. 61573233), Team Project of the University of Guangdong Province (no. 2015KCXTD018), Key Project of Natural Science Foundation of Guangdong Province (no. 2015A030311017), research program of Guangzhou University (no. YJ2021003), and International Science and Technology Cooperation Projects of Guangdong Province (no. 2020A0505100060).

References

[1] Z. Na, J. Wang, C. Liu, M. Guan, and Z. Gao, "Join trajectory optimization and communication design for uav-enabled ofdm networks," *Ad Hoc Networks*, vol. 98, Article ID 102031, 2020.

[2] J. Liu, F. Gao, W. Jia, and W. Zhang, "Gridless compressed sensing based channel estimation for UAV wideband communications with beam squint," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10265–10277, 2021.

[3] L. Chen, "Intelligent ubiquitous computing for future UAV-enabled MEC network systems," *Cluster Computing*, vol. 2021, no. 25, pp. 1–10, 2021.

[4] Y. Liu, J. Shi, C. Liu, and Z. Gao, "UAV-supported clustered NOMA for 6G-enabled internet of things: trajectory planning and resource allocation," *IEEE Internet of Things Journal*, vol. 8, no. 20, 2020.

[5] D. Deng, X. Li, V. Menon, M. J. Piran, H. Chen, and M. A. Jan, "Learning-based joint UAV trajectory and power allocation optimization for secure IoT networks," *Digital Communications and Networks*, vol. 54, pp. 2352–2361, 2021.

[6] J. Liu, J. Jiang, and F. Gao, "Efficient deployment with geometric analysis for mmwave UAV communications," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1115–1119, 2020.

[7] S. Ni, "Multiband cooperation for 5g hetnets: a promising network paradigm," *IEEE Vehicular Technology Magazine*, vol. 14, no. 4, pp. 85–93, 2019.

[8] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Di Renzo, "Safeguarding 5g wireless communication networks using physical layer security," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 20–27, 2015.

[9] X. Sun, Q. Li, and X. Ma, "Edge caching and computation management for real-time internet of vehicles: an online and distributed approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2183–2197, 2021.

[10] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.

[11] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, 2017.

[12] X. Hu, C. Zhong, Y. Zhu, X. Chen, and Z. Zhang, "Programmable metasurface-based multicast systems: design and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1763–1776, 2020.

[13] G. Li, H. Liu, G. Huang, X. Li, B. Raj, and F. Kara, "Effective capacity analysis of reconfigurable intelligent surfaces aided

NOMA network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, p. 198, 2021.

[14] Q. Tao, J. Wang, and C. Zhong, "Performance analysis of intelligent reflecting surface aided communication systems," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2464–2468, 2020.

[15] X. Hu, C. Zhong, Y. Zhang, X. Chen, and Z. Zhang, "Location information aided multiple intelligent reflecting surface systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7948–7962, 2020.

[16] X. Hu, J. Wang, and C. Zhong, "Statistical CSI based design for intelligent reflecting surface assisted MISO systems," *Science China Information Sciences*, vol. 63, no. 12, Article ID 222303, 2020.

[17] J. Zhang, Y. Zhang, C. Zhong, and Z. Zhang, "Robust design for intelligent reflecting surfaces assisted MISO systems," *IEEE Communications Letters*, vol. 24, no. 10, pp. 2353–2357, 2020.

[18] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[19] Y. Zheng, W. U. Khan, M. Zeng, D. Li, G. K. Ragesh, and L. Li, "Physical layer security of cognitive ambient backscatter communications for green internet-of-things," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1066–1076, 2021.

[20] Y. Zheng, M. D. Alshehri, L. Hai, V. Balasubramanian, M. Zeng, and G. Nie, "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 53, no. 445, pp. 232–244, 2021.

[21] S. Tang, "Dilated convolution based CSI feedback compression for massive MIMO systems," *IEEE Transactions on Vehicular Technology*, 2021.

[22] Z. Zhao, "System optimization of federated learning networks with a constrained latency," *IEEE Transactions on Vehicular Technology*, vol. 71, 2021.

[23] L. He and K. He, "Towards optimally efficient search with deep learning for large-scale MIMO systems," *IEEE Transactions on Communications*, pp. 1–12, 2022.

[24] X. Lai, "Outdated access point selection for mobile edge computing with cochannel interference," *IEEE Transactions on Vehicular Technology*, 2021.

[25] J. Lu, "Analytical offloading design for mobile edge computing based smart internet of vehicle," *EURASIP Journal on Advances in Signal Processing*, vol. 99, pp. 1–10, 2022.

[26] X. Lai, "Secure mobile edge computing networks in the presence of multiple eavesdroppers," *IEEE Transactions on Communications*, vol. 70, no. 1, 2021.

[27] L. Fan and X. Lei, "Learning-Based MIMO detection with dynamic spatial modulation," *Science China Information Sciences*, vol. 99, pp. 1–10, 2022.

[28] K. He and Y. Deng, "Efficient memory-bounded optimal detection for GSM-MIMO systems," *IEEE Transactions on Communications*, vol. 99, pp. 1–12, 2022.

[29] Y. Guo, S. Lai, H. H. Yang, and T. Q. S. Quek, "Distributed machine learning for multiuser mobile edge computing systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 99, 2021.

[30] R. Zhao and M. Tang, "Profit maximization in cache-aided intelligent computing networks," *Physical Communication*, vol. 99, pp. 1–10, 2022.

[31] S. Tang and L. Chen, "Computational intelligence and deep learning for next-generation edge-enabled industrial IoT," *IEEE Transactions on Network Science and Engineering*, vol. 99, pp. 1–12, 2022.

[32] L. Zhang, "Dqn based mobile edge computing for smart internet of vehicle," *EURASIP Journal on Advances in Signal Processing*, vol. 99, pp. 1–10, 2022.

[33] J. Chen and F. ZHou, "Physical-layer security on mobile edge computing for emerging cyber physical systems," *IEEE Transactions on Vehicular Technology*, vol. 2022, no. 8, pp. 115–126, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6518066, 12 pages
https://doi.org/10.1155/2022/6518066

Research Article

Channel Measurement and Noise Estimation in FBMC/OQAM-Based IoT Networks

Jun Sun,[1,2] Xiaomin Mu,[1] and Dejin Kong[3]

[1]*School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China*
[2]*School of Electronic and Information, Zhongyuan University of Technology, Zhengzhou 450007, China*
[3]*State Key Laboratory of New Textile Materials and Advanced Processing Technologies, The School of Electronic and Electrical Engineering, Wuhan Textile University, Wuhan 430200, China*

Correspondence should be addressed to Dejin Kong; djkou@wtu.edu.cn

Received 8 November 2021; Revised 14 December 2021; Accepted 17 December 2021; Published 10 January 2022

Academic Editor: Xingwang Li

Channel measurement plays an important role in the emerging 5G-enabled Internet of Things (IoT) networks, which reflects the channel quality and link reliability. In this paper, we address the channel measurement for link reliability evaluation in filter-bank multicarrier with offset quadrature amplitude modulation- (FBMC/OQAM-) based IoT network, which is considered as a promising technique for future wireless communications. However, resulting from the imaginary interference and the noise correlation among subcarriers in FBMC/OQAM, the existing frequency correlation method cannot be directly applied in the FBMC/OQAM-based IoT network. In this study, the concept of the block repetition is applied in FBMC/OQAM. It is demonstrated that the noises among subcarriers are independent by the block repetition and linear combination, instead of correlated. On this basis, the classical frequency correlation method can be applied to achieve the channel measurement. Then, we also propose an advanced frequency correlation method to improve the accuracy of the channel measurement, by assuming channel frequency responses to be quasi-invariant for several successive subcarriers. Simulations are conducted to validate the proposed schemes.

1. Introduction

The evaluation of the link reliability is required in the emerging 5G-enabled Internet of Things (IoT) networks, and based on the evaluation, the network performance can be enhanced by the power control or cell selection [1]. Therefore, the channel measurement plays a very important role in 5G-enabled IoT networks since it reflects the channel quality and link reliability. Generally speaking, the evaluation of the link reliability can be performed by measuring the received power of predefined reference signals [2]. In case of poor link, the base station can conduct the link switch, reducing the probability of network outage. In this study, we focus on the channel measurement for link reliability evaluation in filter-bank multicarrier with offset quadrature amplitude modulation- (FBMC/OQAM-) based IoT networks.

As one promising physical-layer technique for IoT networks, FBMC/OQAM exhibits the advantages of asynchronous transmissions owing to the low sidelobe of the spectrum [3, 4] and high spectral efficiency due to the avoidance of the cyclic prefix [5], compared with the classical orthogonal frequency division multiplexing (OFDM). However, due to the real-valued orthogonality, there is imaginary-valued interference among symbols in FBMC/OQAM, called as intrinsic imaginary interference. Resulting from intrinsic imaginary interference [6, 7], many existing algorithms for OFDM cannot be directly used in FBMC/OQAM, such as channel measurement. Since the long-term evolution (LTE) era, reference signal received power (RSRP) has been the key one of channel measurements by utilizing reference signals [8]. Currently, there have existed several algorithms for channel measurement and noise estimation. In [9], a discrete Fourier transform- (DFT-) based time-domain algorithm was presented by utilizing the property that energy of the time-domain channel is only concentrated on the taps corresponding to propagation paths. In

[10, 11], the authors proposed a frequency correlation algorithm by computing the correlation of the estimated frequency channel responses at adjacent subcarriers. Due to the low complexity and satisfactory performance, the classical frequency correlation method has been widely adopted for the channel measurement and noise estimation in various 3GPP proposals [10, 11]. However, these existing algorithms were proposed for OFDM-based networks. Due to the imaginary interference, channel measurement and noise estimation are more complicated in FBMC/OQAM systems. To the authors' knowledge, channel measurement has not been addressed for FBMC/OQAM-based IoT networks.

In this paper, we address channel measurement and estimation based on the FBMC/OQAM structure with block repetition. Firstly, we derive the noise distribution mathematically, which indicates that the noises at the receiver subcarriers are independent. On this basis, the classical frequency correlation method can be applied to achieve the channel measurement. Then, we also propose an advanced frequency correlation method to improve the accuracy of the channel measurement in FBMC/OQAM-based wireless networks, by assuming channel frequency responses to be quasi-invariant for several successive subcarriers. Numerical simulations have been conducted to validate the proposed schemes. The innovations are summarized as follows. The corresponding text is shown in the following:

(i) By using the block repetition, the noises at the receiver subcarriers are demonstrated to not be correlated, which is a key difference from the existing concept on noise distribution in FBMC/OQAM systems

(ii) Based on the mentioned-above features, the classical frequency correlation method can be employed in FBMC/OQAM to achieve the channel measurement

(iii) We propose an improved frequency correlation method to improve the accuracy of channel measurement."

2. Classical Frequency Correlation Method

The classical frequency correlation method has been proposed for the channel measurement in OFDM-based networks, which has been widely adopted in various 3GPP proposals [10, 11]. In this section, we firstly present the classical frequency correlation method and then specify the drawbacks when it is directly used for the channel measurement in FBMC/OQAM-based networks.

2.1. Channel Measurement in OFDM. For classical OFDM systems, the received signal at the receiver can be obtained [12]

$$y_m = H_m x_m + \eta_m, \tag{1}$$

where x_m is the transmitted symbol at the m-th subcarrier and H_m is the channel frequency response. The noise η_m is usually supposed to satisfy the Gaussian distribution with zero mean and variance σ^2 and be independent for different subcarriers in the classical OFDM systems.

Based on Equation (1), the frequency-domain channel estimation is obtained as

$$\widehat{H}_m = \frac{y_m}{x_m} = H_m + \frac{\eta_m}{x_m}. \tag{2}$$

Suppose that the transmitted symbol x_m has a unit power. The noise term η_m/x_m still satisfies the Gaussian distribution with variance σ^2 and is independent for different subcarriers, which can be denoted as ξ_m for simplicity.

Supposing M subcarriers used for the estimation and the channel frequency responses are quasi-invariant for adjacent subcarriers, then the channel measurement, i.e., RSRP, can be obtained by the following frequency correlation method [10, 11],

$$\mathrm{RSRP}_o = \left| \frac{1}{M-1} \sum_{m=0}^{M-2} \widehat{H}_m \widehat{H}_{m+1}^* \right|, \tag{3}$$

where $\widehat{H}_{m+1}^*$ is the conjugate of $\widehat{H}_{m+1}$. When $H_m = H_{m+1}$ and M goes to infinity, Equation (3) can be rewritten as

$$\begin{aligned} \mathrm{RSRP}_o &= \lim_{M\longrightarrow\infty} \left| \frac{1}{M-1} \sum_{m=0}^{M-2} (H_m + \xi_m)(H_{m+1} + \xi_{m+1})^* \right| \\ &= \lim_{M\longrightarrow\infty} \left| \frac{1}{M-1} \sum_{m=0}^{M-2} \left(H_m H_{m+1}^* + H_m \xi_{m+1}^* + \xi_m H_{m+1}^* + \xi_m \xi_{m+1}^* \right) \right|. \end{aligned} \tag{4}$$

Since the noise η_m is independent and identically distributed in the classical OFDM systems, the following equations hold:

$$\lim_{M\longrightarrow\infty} \frac{1}{M-1} \sum_{m=0}^{M-2} H_m \xi_{m+1}^* = 0, \tag{5}$$

$$\lim_{M\longrightarrow\infty} \frac{1}{M-1} \sum_{m=0}^{M-2} \xi_m H_{m+1}^* = 0, \tag{6}$$

$$\lim_{M\longrightarrow\infty} \frac{1}{M-1} \sum_{m=0}^{M-2} \xi_m \xi_{m+1}^* = 0. \tag{7}$$

Then, Equation (4) can be rewritten as

$$\begin{aligned} \mathrm{RSRP}_o &= \lim_{M\longrightarrow\infty} \left| \frac{1}{M-1} \sum_{m=0}^{M-2} (H_m + \xi_m)(H_{m+1} + \xi_{m+1})^* \right| \\ &= \lim_{M\longrightarrow\infty} \frac{1}{M-1} \sum_{m=0}^{M-2} H_m H_m^*, \end{aligned} \tag{8}$$

which represents the quality of the channel.

Subsequently, the estimation of noise variance in OFDM systems can be obtained by

$$\widehat{\varepsilon}_o = \frac{1}{M}\sum_{m=0}^{M-1} y_m y_m^* - \mathrm{RSRP}_o. \tag{9}$$

Note that $\widehat{\varepsilon}_o$ will be equal to σ^2 when M goes to infinity in the classical OFDM systems.

2.2. Channel Measurement in FBMC/OQAM. Different from the classical OFDM system, the transmitted signal of each FBMC/OQAM subcarrier gets through a pulse-shaping filter with low spectrum sidelobe, leading to the better spectral utilization and ability of asynchronous transmission. Figure 1 shows the system model of FBMC/OQAM with M subcarriers, in which $g[k]$ is the pulse-shaping filter, and $a_{m,n}$ is the transmitted real-valued symbol at the time-frequency position (m, n). Firstly, the real-valued symbols $a_{m,n}$ are multiplied by $e^{j\pi(m+n)/2}$, and get through the $M/2$ -point upsampling (corresponding to half symbol duration). Then, the obtained signal is filtered by $g[k]$, and the equivalent baseband transmitting signal of FBMC/OQAM can be written as [13, 14]

$$s[k] = \sum_{m=0}^{M-1}\sum_{n\in\mathbb{Z}} a_{m,n} g\left[k - n\frac{M}{2}\right] e^{j2\pi mk/M} e^{j\pi(m+n)/2}. \tag{10}$$

After passing a multipath channel, $r[k]$ is the received signal at the receiver. Then, the demodulation of the received signal can be approximately written as

$$\widehat{a}_{m,n} = \sum_{m=0}^{M-1}\sum_{n\in\mathbb{Z}} r[k] g\left[k - n\frac{M}{2}\right] e^{-j2\pi mk/M} e^{-j\pi(m+n)/2}. \tag{11}$$

Note that the maximum channel delay is usually much smaller than the symbol duration. And it is demonstrated in [13] that Equation (11) can be approximately written as [13]

$$\widehat{a}_{m,n} \approx H_{m,n}\left(a_{m,n} + ja_{m,n}^o\right) + \eta_{m,n}, \tag{12}$$

where $H_{m,n}$ is the channel frequency response at position (m, n). For the slowly time-varying channels, $H_{m,n} \approx H_{m,n+1}$. The term $ja_{m,n}^o$ is the imaginary interference to the symbol $a_{m,n}$, i.e.,

$$ja_{m,n}^o = \sum_{m_0\in\mathbb{Z}}\sum_{n_0\in\mathbb{Z}} a_{m_0,n_0}\zeta_{m,n}^{m_0,n_0}, (m_0, n_0) \neq (m, n), \tag{13}$$

where $\zeta_{m,n}^{m_0,n_0}$ is the imaginary interference factor in FBMC/OQAM and is defined as [13]

$$\zeta_{m,n}^{m_0,n_0} = \sum_{k=-\infty}^{\infty} g\left[k - n\frac{M}{2}\right] g\left[k - n_0\frac{M}{2}\right] e^{j2\pi k(m-m_0)/M} \times e^{j\pi(m+n-m_0-n_0)/2}. \tag{14}$$

Specifically when $(m, n) = (m_0, n_0)$, $\zeta_{m,n}^{m_0,n_0} = 1$. When $(m, n) \neq (m_0, n_0)$, $\zeta_{m,n}^{m_0,n_0}$ is imaginary-valued. It has been revealed that the value of $\zeta_{m,n}^{m_0,n_0}$ is close to zero when $|m - m_0| \geq 2$ or $|n - n_0| \geq 2$ [13], indicating that the imaginary interference to one symbol is mainly determined by its adjacent symbols. Besides, $\eta_{m,n}$ is the noise term in FBMC/OQAM systems and is written as

$$\eta_{m,n} = \sum_{k=-\infty}^{+\infty} \eta[k] g\left[k - n\frac{M}{2}\right] e^{-j2\pi mk/M} e^{-j\pi(m+n)/2}, \tag{15}$$

where $\eta[k]$ is the additional white Gaussian noise (AWGN) in wireless channels and is assumed to have a variance σ^2. It has been proven in [14] that, due to the real orthogonality condition of FBMC/OQAM, $\eta_{m,n}$ is of zero mean and has variance σ^2, while is related to η_{m_0,n_0} for $|m - m_0| \leq 1$ and $|n - n_0| \leq 1$, which is different from classical OFDM systems.

Motivated by the noise correlation mentioned above, pilots as shown in Figure 2 are required if the classical frequency correlation method in (3) is directly used for channel measurement and noise estimation in FBMC/OQAM systems. For the block of Figure 2, the symbols in the first column are set to zero to avoid interference from the previous block, and the symbols in the third column are set to zero to avoid interference from data. The nonzero pilots are in the second column, and it is noted that the pilots are designed to be scattered to ensure that the noise components for nonzero pilots are independent of each other. Assume that $a_{2p,1} = \pm 1$, and $a_{2p+1,1} = 0$ with $p = 0, 1, \cdots, M/2 - 1$. Then, Equation (12) can be rewritten as

$$\widehat{a}_{2p,1} = H_{2p,1} a_{2p,1} + \eta_{2p,1}. \tag{16}$$

By the frequency correlation method, the channel measurement of FBMC/OQAM can be obtained as

$$\mathrm{RSRP}_f = \left|\frac{1}{M/2 - 1}\sum_{m=0}^{M/2-2} \frac{\widehat{a}_{2p,1}}{a_{2p,1}}\left(\frac{a\wedge_{2(p+1),1}}{a_{2(p+1),1}}\right)^*\right|. \tag{17}$$

Subsequently, the estimation of noise variance in FBMC/OQAM systems can be obtained by

$$\widehat{\varepsilon}_f = \frac{2}{M}\sum_{m=0}^{M/2-1} \widehat{a}_{2p,1}\widehat{a}_{2p,1}^* - RSRP_f. \tag{18}$$

As presented above, the classical frequency correlation method can be directly used for channel measurement and noise estimation in FBMC/OQAM systems. However, it is noted that the accuracy of channel measurement depends on the number of samples. For a fixed frequency band, all subcarriers in OFDM can be used for the channel measurement, while only half of the subcarriers in FBMC/OQAM can be used for the channel measurement. Therefore, there exists a loss of channel measurement accuracy when the classical frequency correlation method is directly used in FBMC/OQAM systems.

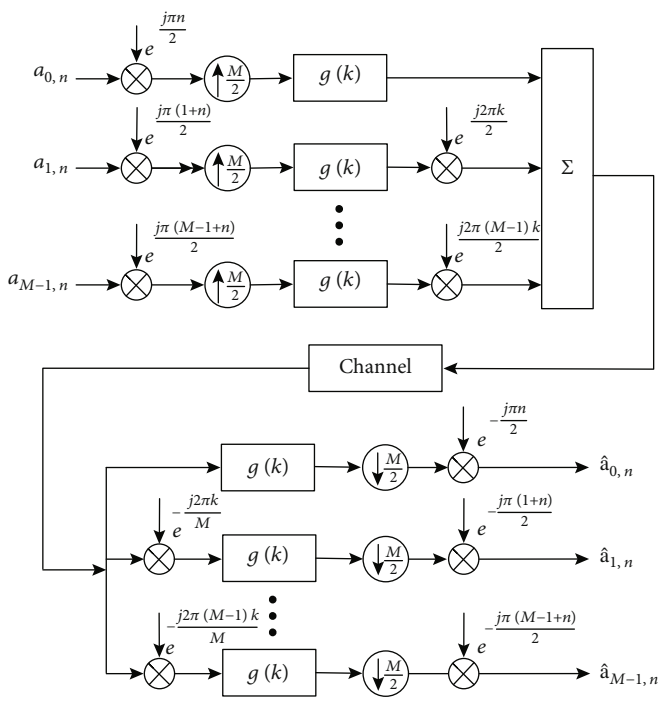


Figure 1: System model of FBMC/OQAM.

3. Proposed Scheme for Channel Measurement in FBMC/OQAM

In this section, we propose a novel scheme for channel measurement in FBMC/OQAM systems. Firstly, the imaginary interference factor is analyzed and it is noted that its amplitude is symmetric on the time-frequency position. Then, on the basis, a reversed-order repeated block is designed to remove the correlation among symbols so that all subcarriers can be used for the channel measurement in FBMC/OQAM systems, by the frequency correlation method, improving the RSRP accuracy.

3.1. Analysis of Imaginary Interference Factor. According to (14), the imaginary interference factor is defined as

$$\zeta_{m,n}^{m_0,n_0} = \sum_{k=-\infty}^{\infty} g\left[k - n\frac{M}{2}\right] g\left[k - n_0\frac{M}{2}\right] e^{j2\pi k(m-m_0)/M} \times e^{j\pi(m+n-m_0-n_0)/2}, \tag{19}$$

which is the interference coefficient to the symbol a_{m_0,n_0} from the symbol $a_{m,n}$. It is easily seen that the value of $\zeta_{m,n}^{m_0,n_0}$ only depends on the filter $g[k]$ and the time-frequency position (m, n). Note that the pulse-shaping filter $g[k]$ in FBMC/OQAM is real-valued and symmetric [15, 16]. Due to the fact that $g[k]$ is real-valued and symmetric, it can be easily demonstrated that the following equations hold

$$\zeta_{m_0,n_0}^{m_0,n_0-1} = -\zeta_{m_0,n_0}^{m_0,n_0+1}, \tag{20}$$

$$\zeta_{m_0,n_0}^{m_0-1,n_0} = -\zeta_{m_0,n_0}^{m_0+1,n_0}, \tag{21}$$

Pilot			Data						
0	1	0	$a_{0,3}$	$a_{0,4}$	$a_{0,5}$	$a_{0,6}$	$a_{0,7}$	$a_{0,8}$	$a_{0,9}$
0	0	0	$a_{1,3}$	$a_{1,4}$	$a_{1,5}$	$a_{1,6}$	$a_{1,7}$	$a_{1,8}$	$a_{0,9}$
0	−1	0	$a_{2,3}$	$a_{2,4}$	$a_{2,5}$	$a_{2,6}$	$a_{2,7}$	$a_{2,8}$	$a_{0,9}$
0	0	0	$a_{3,3}$	$a_{3,4}$	$a_{3,5}$	$a_{3,6}$	$a_{3,7}$	$a_{3,8}$	$a_{0,9}$
0	1	0	$a_{4,3}$	$a_{4,4}$	$a_{4,5}$	$a_{4,6}$	$a_{4,7}$	$a_{4,8}$	$a_{0,9}$
0	0	0	$a_{5,3}$	$a_{5,4}$	$a_{5,5}$	$a_{5,6}$	$a_{5,7}$	$a_{5,8}$	$a_{0,9}$

Figure 2: Scattered pilot for channel measurement and noise estimation in FBMC/OQAM.

$$\zeta_{m_0,n_0}^{m_0-1,n_0-1} = \zeta_{m_0,n_0}^{m_0-1,n_0+1} = \zeta_{m_0,n_0}^{m_0+1,n_0-1} = \zeta_{m_0,n_0}^{m_0+1,n_0+1}. \tag{22}$$

Furthermore, let $\bar{k} = k - n(M/2)$; then, Equation (19) can be rewritten as

$$\begin{aligned}\zeta_{m,n}^{m_0,n_0} &= \sum_{\bar{k}=-\infty}^{\infty} g[\bar{k}]\, g\left[\bar{k} + n\frac{M}{2} - n_0\frac{M}{2}\right] \times e^{j2\pi\left(\bar{k}+n(M/2)\right)(m-m_0)/M} e^{j\pi(m+n-m_0-n_0)/2} \\ &= \sum_{\bar{k}=-\infty}^{\infty} g[\bar{k}]\, g\left[\bar{k} + n\frac{M}{2} - n_0\frac{M}{2}\right] \times e^{j2\pi\bar{k}(m-m_0)/M} e^{j\pi(m-m_0)n} e^{j\pi(m+n-m_0-n_0)/2}.\end{aligned} \tag{23}$$

According to Equation (23), for adjacent subcarriers, i.e., $m - m_0 = \pm 1$, sign of $\zeta_{m,n}^{m_0,n_0}$ is opposite for the odd index n and the even index n. When the isotropic orthogonal transform algorithm (IOTA) is used for pulse shaping in FBMC/OQAM systems [14], the values of $\zeta_{m,n}^{m_0,n_0}$ are depicted in Tables 1 and 2 for the odd index and the even index,

respectively. Note that, as mentioned above, the value of $\zeta_{m,n}^{m_0,n_0}$ is close to zero when $|m-m_0| \geq 2$ and $|n-n_0| \geq 2$ [13]; therefore, Tables 1 and 2 only show the values of $\zeta_{m,n}^{m_0,n_0}$ for $|m-m_0| \leq 1$ and $|n-n_0| \leq 1$.

3.2. Reversed-Order Repeated Block. In this subsection, a reversed-order repeated block is presented to eliminate the correlation among adjacent subcarrier, utilizing the properties in (20)–(22), Tables 1 and 3. Assume that $a_{m,n}$ with $m = 0, 1, \cdots, M-1, n = 0, 1, \cdots, N-1$ are the transmitted symbols in FBMC/OQAM. Then, these symbols are transmitted repeatedly in the next block, satisfying $a_{m,N+n} = a_{m,N-1-n}$ with $n = 0, 1, \cdots, N-1$, as shown in Figure 3 in which $M = 6$ and $N = 5$ are taken as an example. Note that the symbol interval between $a_{m,n}$ and $a_{m,n+1}$ is still half of a quadrature amplitude modulation (QAM) symbol, exactly as the symbol interval of normal FBMC/OQAM systems.

After the modulation at the FBMC/OQAM transmitter and demodulation at the FBMC/OQAM receiver, demodulations of original symbols can be obtained as

$$\hat{a}_{m,n} = H_{m,n}\left(a_{m,n} + ja_{m,n}^{o}\right) + \eta_{m,n}, n = 0, 1, \cdots, N-1, \quad (24)$$

and demodulations of repeated symbols can be obtained as

$$\begin{aligned}\hat{a}_{m,2N-1-n} &= H_{m,2N-1-n}\left(a_{m,2N-1-n} + ja_{m,2N-1-n}^{o}\right)\\ &= H_{m,n}\left(a_{m,n} + ja_{m,2N-1-n}^{o}\right) + \eta_{m,2N-1-n}, n = 0, 1, \cdots, N-1.\end{aligned} \quad (25)$$

According to (13), the terms $ja_{m,n}^{o}$ and $ja_{m,2N-1-n}^{o}$ can be, respectively, written as

$$ja_{m,n}^{o} = \sum_{n_0=n-1}^{n+1}\sum_{m_0=m-1}^{m+1} a_{m_0,n_0}\zeta_{m,n}^{m_0,n_0}, (m_0, n_0) \neq (m, n), \quad (26)$$

$$ja_{m,2N-1-n}^{o} = \sum_{n_0=2N-2-n}^{2N-n}\sum_{m_0=m-1}^{m+1} a_{m_0,n_0}\zeta_{m,2N-1-n}^{m_0,n_0}, (m_0, n_0) \neq (m, 2N-1-n), \quad (27)$$

where $n = 0, 1, \cdots, N-1$.

Then, utilizing the properties of $\zeta_{m,n}^{m_0,n_0}$ in (20)–(22), Tables 1 and 3, and considering $a_{m,N+n} = a_{m,N-1-n}$ with $n = 0, 1, \cdots, N-1$ as shown in Figure 3, it can be easily demonstrated that the following equation holds:

$$ja_{m,n}^{o} + ja_{m,2N-1-n}^{o} = 0. \quad (28)$$

On this basis, it can be obtained as

$$\frac{\hat{a}_{m,n} + \hat{a}_{m,2N-1-n}}{2} = H_{m,n}a_{m,n} + \frac{\eta_{m,n} + \eta_{m,2N-1-n}}{2}. \quad (29)$$

From Equation (29), it is easily realized that the imaginary interference in FBMC/OQAM is removed by the linear combination. Figure 4 depicts the diagram of the imaginary interference cancelation in FBMC/OQAM, and compared with the classical scheme, the key differences are the employment of repeated symbols at the transmitter and the operation of the linear combination at the receiver. At the transmitter, the employment of repeated symbols does not increase additional add operations and multiplication operations. At the receiver, the linear combination only requires add operations, instead of multiplication operations. As a result, the complexity of the proposed scheme is almost the same as the classical scheme.

Table 1: $\zeta_{m,n}^{m_0,n_0}$ with odd time index n.

	n_0-1	n_0	n_0+1
m_0-1	$0.2280j$	$0.4411j$	$0.2280j$
m_0	$0.4411j$	1	$-0.4411j$
m_0+1	$0.2280j$	$-0.4411j$	$0.2280j$

Table 2: Simulation parameters.

Modulation	4QAM
Sampling frequency (MHz)	9.14
Number of paths	6
Power profile (in dB)	-6, 0, -7, -22, -16, -20
Delay profile (μs)	-3, 0, 2, 4, 7, 11

3.3. Analysis of Noise Distribution. It is noteworthy that, although the intersymbol interference is removed in Equation (29), the classical frequency correlation method is still not applicable for the channel measurement in FBMC/OQAM if the noise in (29) is not independent. Note that the noises at the receive subcarriers are correlated in the classical FBMC/OQAM structure without symbol repetition [14], due to its only real-filed orthogonality condition.

In this subsection, the distribution of noise term $\eta_{m,n} + \eta_{m,2N-1-n}/2$ in (29) is analyzed, and it will be demonstrated theoretically that the noises $\eta_{m,n} + \eta_{m,2N-1-n}/2$ are independent for different m, which is a key difference from the existing concept of noise distribution in FBMC/OQAM systems.

According to Equation (15), $\eta_{m,n}$ and $\eta_{m,2N-1-n}$ can be, respectively, written as

$$\eta_{m,n} = \sum_{k=-\infty}^{+\infty}\eta[k]g\left[k - n\frac{M}{2}\right]e^{-j2\pi mk/M}e^{-j\pi(m+n)/2}, \quad (30)$$

$$\eta_{m,2N-1-n} = \sum_{k=-\infty}^{+\infty}\eta[k]g\left[k - (2N-1-n)\frac{M}{2}\right]e^{-j2\pi mk/M} \times e^{-j\pi(m+2N-1-n)/2}, \quad (31)$$

where $n = 0, 1, \cdots, N-1$. Note that $\eta[k]$ is the AWGN noise in wireless channels and satisfies

$$\mathbb{E}\{\eta[k]\} = 0, \quad (32)$$

$$\mathrm{Var}\{\eta[k]\} = \sigma^2, \quad (33)$$

$$\mathrm{Cov}\{\eta[k], \eta[k_0]\} = 0, k \neq k_0, \quad (34)$$

Table 3: $\zeta_{m,n}^{m_0,n_0}$ with even time index n.

	n_0-1	n_0	n_0+1
m_0-1	-0.2280j	-0.4411j	-0.2280j
m_0	0.4411j	1	-0.4411j
m_0+1	-0.2280j	0.4411j	-0.2280j

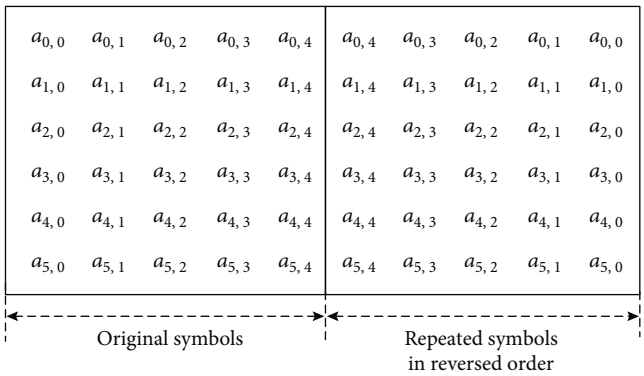


Figure 3: The reversed-order block for FBMC/OQAM.

where $\mathbb{E}\{\cdot\}$, $\mathrm{Var}\{\cdot\}$, and $\mathrm{Cov}\{\cdot\}$ are the operators of the expectation, variance, and covariance, respectively. As a result, the mean of $\eta_{m,n}$ is obtained as

$$\begin{aligned}\mathbb{E}\left[\eta_{m,n}\right] &= \mathbb{E}\left\{\sum_{k=-\infty}^{+\infty}\eta[k]g\left[k-n\frac{M}{2}\right]e^{-j2\pi mk/M}e^{-j\pi(m+n)/2}\right\}\\ &= \sum_{k=-\infty}^{+\infty}\mathbb{E}\{\eta[k]\}g\left[k-n\frac{M}{2}\right]e^{-j2\pi mk/M}e^{-j\pi(m+n)/2}=0.\end{aligned} \tag{35}$$

Note that Equation (35) holds for any m and n; hence, the mean of $\eta_{m,2N-1-n}$ is also zero.

Then, for the sake of simplicity, define

$$\chi_{m,n}=\frac{\eta_{m,n}+\eta_{m,2N-1-n}}{2}. \tag{36}$$

As mentioned above, $\eta_{m,n}$ and $\eta_{m,2N-1-n}$ satisfy the Gaussian distribution. Since $\chi_{m,n}$ is the linear combination of $\eta_{m,n}$ and $\eta_{m,2N-1-n}$, it can be realized that $\chi_{m,n}$ also satisfies the Gaussian distribution. The mean of $\chi_{m,n}$ can be obtained as

$$\mathbb{E}\left\{\chi_{m,n}\right\}=\mathbb{E}\left\{\frac{\eta_{m,n}}{2}\right\}+\mathbb{E}\left\{\frac{\eta_{m,2N-1-n}}{2}\right\}=0. \tag{37}$$

The variance of $\chi_{m,n}$ can be obtained as

$$\begin{aligned}\mathrm{Var}\left\{\chi_{m,n}\right\} &= \mathbb{E}\left\{\chi_{m,n}\chi_{m,n}^*\right\}-\mathbb{E}\left\{\chi_{m,n}\right\}\mathbb{E}\left\{\chi_{m,n}^*\right\}\\ &= \mathbb{E}\left\{\chi_{m,n}\chi_{m,n}^*\right\}=\frac{1}{4}\left[\mathbb{E}\left\{\eta_{m,n}\eta_{m,n}^*\right\}+\mathbb{E}\left\{\eta_{m,n}\eta_{m,2N-1-n}^*\right\}\right.\\ &\quad\left.+\mathbb{E}\left\{\eta_{m,2N-1-n}\eta_{m,n}^*\right\}+\mathbb{E}\left\{\eta_{m,2N-1-n}\eta_{m,2N-1-n}^*\right\}\right].\end{aligned} \tag{38}$$

Note that $\mathbb{E}\{\eta_{m,n}\eta_{m_0,n_0}^*\}$ can be written as

$$\begin{aligned}\mathbb{E}\left\{\eta_{m,n}\eta_{m_0,n_0}^*\right\} &= \mathbb{E}\left\{\left(\sum_{k=-\infty}^{+\infty}\eta[k]g\left[k-n\frac{M}{2}\right]e^{-j2\pi mk/M}e^{-j\pi(m+n)/2}\right)\right.\\ &\quad\left.\cdot\left(\sum_{k=-\infty}^{+\infty}\eta^*[k]g\left[k-n_0\frac{M}{2}\right]e^{j2\pi m_0k/M}e^{j\pi(m_0+n_0)/2}\right)\right\}\\ &= \mathbb{E}\left\{\sum_{k=-\infty}^{+\infty}\eta[k]\eta^*[k]g\left[k-n\frac{M}{2}\right]g\left[k-n_0\frac{M}{2}\right]\right.\\ &\quad\left.\times e^{-j2\pi(m-m_0)k/M}e^{-j\pi(m+n-m_0-n_0)/2}\right\}.\end{aligned} \tag{39}$$

According to the definition of the imaginary interference factor in Equation (14), we can rewrite (39) as

$$\mathbb{E}\left\{\eta_{m,n}\eta_{m_0,n_0}^*\right\}=\sigma^2\zeta_{m_0,n_0}^{m,n}. \tag{40}$$

Note that it has been realized in [13] that $\zeta_{m_0,n_0}^{m,n}$ is a pure imaginary value when $(m,n)\neq(m_0,n_0)$. And when $(m,n)=(m_0,n_0)$, the value of $\zeta_{m_0,n_0}^{m,n}$ is equal to 1, i.e.,

$$\zeta_{m,n}^{m,n}=1. \tag{41}$$

Furthermore, it has been revealed that the value of $\zeta_{m_0,n_0}^{m,n}$ is close to zero when $|m-m_0|\geq 2$ and $|n-n_0|\geq 2$ [13]. In addition, the term $\eta_{m,n}\eta_{m,2N-1-n}^*$ is the conjugate of $\eta_{m,2N-1-n}\eta_{m,n}^*$; hence, it can be obtained as $\mathbb{E}\{\eta_{m,n}\eta_{m,2N-1-n}^*\}+\mathbb{E}\{\eta_{m,2N-1-n}\eta_{m,n}^*\}=0$. Then, $\mathrm{Var}\{\chi_{m,n}\}$ can be obtained as

$$\begin{aligned}\mathrm{Var}\left\{\chi_{m,n}\right\} &= \frac{\mathbb{E}\left\{\eta_{m,n}\eta_{m,n}^*\right\}+\mathbb{E}\left\{\eta_{m,2N-1-n}\eta_{m,2N-1-n}^*\right\}}{4}\\ &= \frac{\sigma^2+\sigma^2}{4}=\frac{\sigma^2}{2}.\end{aligned} \tag{42}$$

Then, the covariance of $\chi_{m,n}$ and χ_{m_0,n_0} with $(m,n)\neq(m_0,n_0)$ is obtained.

$$\begin{aligned}\mathrm{Cov}\left\{\chi_{m,n},\chi_{m_0,n_0}\right\} &= \mathbb{E}\left\{\chi_{m,n}\chi_{m_0,n_0}^*\right\}-\mathbb{E}\left\{\chi_{m,n}\right\}\mathbb{E}\left\{\chi_{m_0,n_0}^*\right\}=\mathbb{E}\left\{\chi_{m,n}\chi_{m_0,n_0}^*\right\}\\ &= \frac{1}{4}\left[\mathbb{E}\left\{\eta_{m,n}\eta_{m_0,n_0}^*\right\}+\mathbb{E}\left\{\eta_{m,n}\eta_{m_0,2N-1-n_0}^*\right\}\right.\\ &\quad\left.+\mathbb{E}\left\{\eta_{m,2N-1-n}\eta_{m_0,n_0}^*\right\}+\mathbb{E}\left\{\eta_{m,2N-1-n}\eta_{m_0,2N-1-n_0}^*\right\}\right].\end{aligned} \tag{43}$$

Substituting Equations (39) and (40) into (43), it can be obtained as

$$\mathrm{Cov}\left\{\chi_{m,n},\chi_{m_0,n_0}\right\}=\frac{1}{4}\left[\sigma^2\zeta_{m_0,n_0}^{m,n}+\sigma^2\zeta_{m_0,2N-1-n_0}^{m,n}+\sigma^2\zeta_{m_0,n_0}^{m,2N-1-n}+\sigma^2\zeta_{m_0,2N-1-n_0}^{m,2N-1-n}\right]. \tag{44}$$

Note that, according to the definition of $\zeta_{m_0,n_0}^{m,n}$ in (14), the value of $\zeta_{m_0,n_0}^{m,n}$ only depends on the values of $(m-m_0)$ and $(n-n_0)$, instead of individual m, m_0, n, or n_0. Based

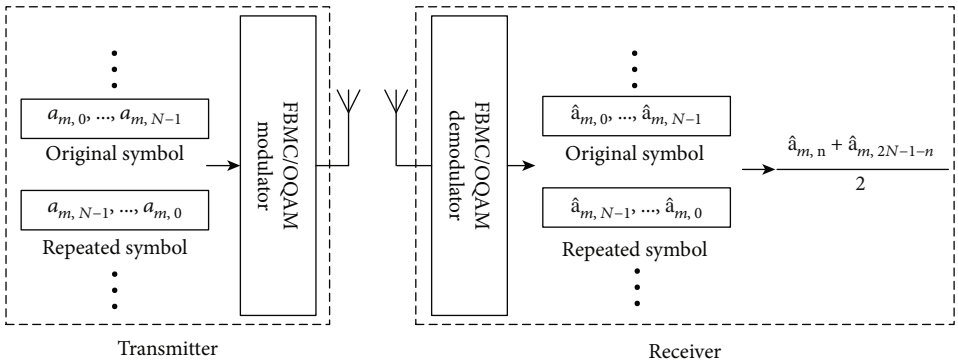


Figure 4: The scheme of the imaginary interference cancelation in FBMC/OQAM.

on the properties in (20)–(22), Tables 1 and 3, it can be obtained as

$$\mathrm{Cov}\left\{\chi_{m,n}, \chi_{m_0,n_0}\right\} = 0. \tag{45}$$

Therefore, it can be concluded that by the linear combination on the reversed-order block in Figure 3, the noise for the demodulated symbol is independent to each other. Note that it has been pointed out in [14] that the noise is correlated to that of its adjacent symbols in normal FBMC/OQAM without the reversed-order block.

Above all, Equation (29) in FBMC/OQAM will be equivalent to that of OFDM systems as Equation (2). Then, the classical frequency correlation method can be directly employed for channel measurement and noise estimation in FBMC/OQAM systems.

4. Proposed Channel Measurement in FBMC/OQAM

In this section, the channel measurement in FBMC/OQAM is presented based on the signal model as Equation (29), by using the classical frequency correlation method. Then, an improved frequency correlation method is proposed to improve the channel measurement accuracy.

4.1. Channel Measurement Using Classical Frequency Correlation. Figure 5 shows the proposed pilot design for channel measurement in FBMC/OQAM, in which $a_{m,0}$ with $m = 0, 1, \cdots, M-1$ are the pilots and $a_{m,n}$ with $n = 1, 2, \cdots, N-1$ are the data symbols. Note that all symbols in Figure 5 are complex-valued because the imaginary interferences among them can be eliminated completely by the linear combination in (29). As a consequence, compared with Figure 2, a lower pilot overhead is achieved in Figure 5.

According to (29), it can be obtained as

$$\frac{\hat{a}_{m,0} + \hat{a}_{m,2N-1}}{2a_{m,0}} = H_{m,0} + \frac{\eta_{m,0} + \eta_{m,2N-1}}{2a_{m,0}}. \tag{46}$$

For simplicity, denote $\hat{a}_{m,0} + \hat{a}_{m,2N-1}/2a_{m,0}$ as $\phi_{m,0}$, and denote $\eta_{m,0} + \eta_{m,2N-1}/2a_{m,0}$ as $\xi_{m,0}$. Then, (46) can be rewritten as

$$\phi_{m,0} = H_{m,0} + \xi_{m,0}, m = 0, 1, \cdots, M-1. \tag{47}$$

Note that, as analyzed in Section 3.2, the noise term $\xi_{m,0}$ satisfies independent identical distribution and its variance is $\sigma^2/2$ when pilots $a_{m,0}$ have the unit transmitting power.

Afterwards, the classical frequency correlation method can be directly used to obtain RSRP in FBMC/OQAM systems,

$$\mathrm{RSRP}_p = \left|\frac{1}{M-1}\sum_{m=0}^{M-2}\phi_{m,0}\phi_{m+1,0}^*\right|. \tag{48}$$

Subsequently, the estimation of noise variance can be obtained by

$$\hat{\varepsilon}_p = 2\left(\frac{1}{M}\sum_{m=0}^{M-1}\phi_{m,0}\phi_{m,0}^* - \mathrm{RSRP}_p\right). \tag{49}$$

Different from Figure 2 in which only half of subcarriers are valid, all subcarriers are used for the channel measurement in Figure 5, indicating the improvement of accuracy in theory.

4.2. Improved Frequency Correlation Method. In this subsection, we propose an improved frequency correlation method to further improve the accuracy of channel measurement in FBMC/OQAM systems, by assuming the channel frequency responses quasi-invariant for several successive subcarriers. Subsequently, we give the theoretical analysis to show its superiority for the accuracy of channel measurement.

The frequency correlation method utilizes the properties of quasi-invariant channels for adjacent subcarrier and independent distribution of noises. It is realized that, by the linear combination on the reversed-order block, the noises in (47) are independent for different subcarrier in FBMC/OQAM systems. By assuming the channels to be quasi-invariant for several successive subcarriers, more samples can be used to achieve the channel measurement by the statistical methods. Different from the conventional frequency correlation method, the correlation operations are performed on the several successive subcarriers in the improved frequency correlation method, instead of only two subcarriers, as shown in Figure 6. Hence, the improved frequency

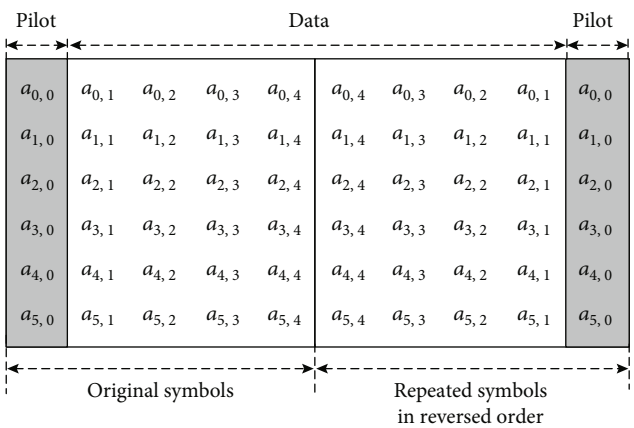


Figure 5: Proposed pilot design for channel measurement in FBMC/OQAM.

correlation method can be expressed as

$$\mathrm{RSRP}_i = \left| \frac{1}{N_c} \sum_{l=1}^{N_r-1} \sum_{m=0}^{M-l-1} \phi_{m,0} \phi_{m+l,0}^* \right|, \tag{50}$$

where N_r is the number of subcarriers whose channel frequency responses are quasi-invariant. N_c is the number of sum terms in (50), i.e.,

$$N_c = (M-1) + (M-2) + \cdots + (M-N_r+1) = \frac{(2M-N_r)(N_r-1)}{2}. \tag{51}$$

Subsequently, the estimation of noise variance can be obtained by

$$\widehat{\varepsilon}_i = 2\left(\frac{1}{M} \sum_{m=0}^{M-1} \phi_{m,0} \phi_{m,0}^* - \mathrm{RSRP}_i \right). \tag{52}$$

In theory, when the subcarrier number M goes to infinity, the channel measurement is accurate by the statistical method. However, M is limited in practice, leading to an interference. To evaluate the accuracy of the improved frequency correlation method, the variance of the residual interference is analyzed, in which the AWGN channel is assumed for simplicity. According to (50), the residual interference term can be easily written as

$$\Theta_i = \frac{1}{N_c} \sum_{l=1}^{N_r-1} \sum_{m=0}^{M-l-1} \left(\xi_{m,0} + \xi_{m+l,0}^* + \xi_{m,0} \xi_{m+l,0}^* \right). \tag{53}$$

Note that, for the classical frequency correlation method in (48), the residual interference term is

$$\Theta_p = \frac{1}{M-1} \sum_{m=0}^{M-2} \left(\xi_{m,0} + \xi_{m+1,0}^* + \xi_{m,0} \xi_{m+1,0}^* \right). \tag{54}$$

Since the noise term $\xi_{m,0}$ satisfies independent identical distribution, it is easily demonstrated that the expectations of Θ_p and Θ_i are equal to zero, i.e., $E[\Theta_i] = E[\Theta_p] = 0$. Then, the variances of Θ_p and Θ_i can be obtained as

$$\mathrm{Var}\left[\Theta_p\right] = \mathrm{E}\left[\Theta_p \Theta_p^*\right] - \mathrm{E}\left[\Theta_p\right] E\left[\Theta_p^*\right] = \frac{2\sigma^2 + \sigma^4}{2(M-1)}, \tag{55}$$

$$\begin{aligned} \mathrm{Var}[\Theta_i] = \mathrm{E}[\Theta_i \Theta_i^*] - \mathrm{E}[\Theta_i]\mathrm{E}[\Theta_i^*] &= \frac{12(M-N_r+1) + 4(N_r-2)(2N_r-3)}{3(2M-N_r)^2(N_r-1)} \sigma^2 \\ &+ \frac{1}{(2M-N_r)(N_r-1)} \sigma^4. \end{aligned} \tag{56}$$

From Equations (55) and (56), when $N_r = 2$, it can be obtained as $\mathrm{Var}[\Theta_i] = \mathrm{Var}[\Theta_p]$ and the proposed frequency correlation method is equivalent to the classical frequency correlation method. When $N_r > 2$, $\mathrm{Var}[\Theta_i]$ will be smaller than $\mathrm{Var}[\Theta_p]$, and the proposed frequency correlation method can achieve better RSRP measurement performance than the conventional frequency correlation method. Especially when $M \gg N_r > 2$, the variance of Θ_i can be rewritten as

$$\begin{aligned} \mathrm{Var}[\Theta_i] &\approx \lim_{M/N_r \longrightarrow \infty} \left(\frac{12(M-N_r+1)}{3(2M-N_r)^2(N_r-1)} \sigma^2 + \frac{1}{(2M-N_r)(N_r-1)} \sigma^4 \right) \\ &= \lim_{M/N_r \longrightarrow \infty} \frac{4(M-N_r+1)\sigma^2 + (2M-N_r)\sigma^4}{(2M-N_r)^2(N_r-1)} \\ &= \lim_{M/N_r \longrightarrow \infty} \frac{4\sigma^2 + 2\sigma^4}{(4M-4N_r)(N_r-1)} = \lim_{M/N_r \longrightarrow \infty} \frac{2\sigma^2 + \sigma^4}{2(M-1)(N_r-1)} \\ &= \frac{1}{N_r-1} \lim_{M/N_r \longrightarrow \infty} \frac{2\sigma^2 + \sigma^4}{2(M-1)}. \end{aligned} \tag{57}$$

According to (55), we have

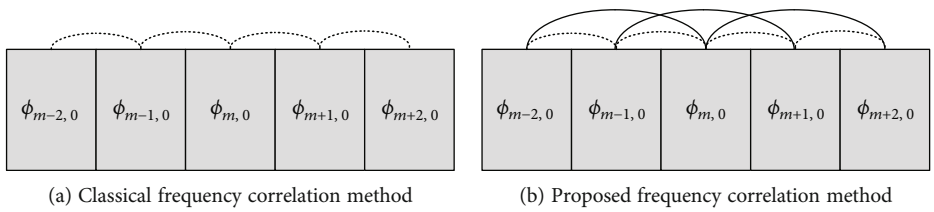
$$\mathrm{Var}[\Theta_i] = \frac{1}{N_r-1} \mathrm{Var}\left[\Theta_p\right], M \gg N_r > 2. \tag{58}$$

From Equation (58), the variance of Θ_i is approximately $1/N_r - 1$ of that of Θ_p when M/N_r goes to infinity, which indicates that the proposed frequency correlation method with M pilots can achieve similar accuracy of the classical frequency correlation method with $(N_r - 1)M$ pilots.

5. Simulation Results

In this section, numerical simulations are given to validate the proposed schemes. In our simulations, we consider an FBMC/OQAM system with 2048 subcarriers, in which the IOTA function is taken as the pulse-shaping filter [15]. Note that only a partial frequency band, i.e., 180 subcarriers is assigned for one user to achieve channel measurement and noise estimation. In addition, the detailed parameters for the multipath channel are listed as Table 3 [13].

Figure 7 shows the BER of FBMC/OQAM with the proposed repeated block under the AWGN channel, and for



(a) Classical frequency correlation method (b) Proposed frequency correlation method

Figure 6: Proposed frequency correlation method.

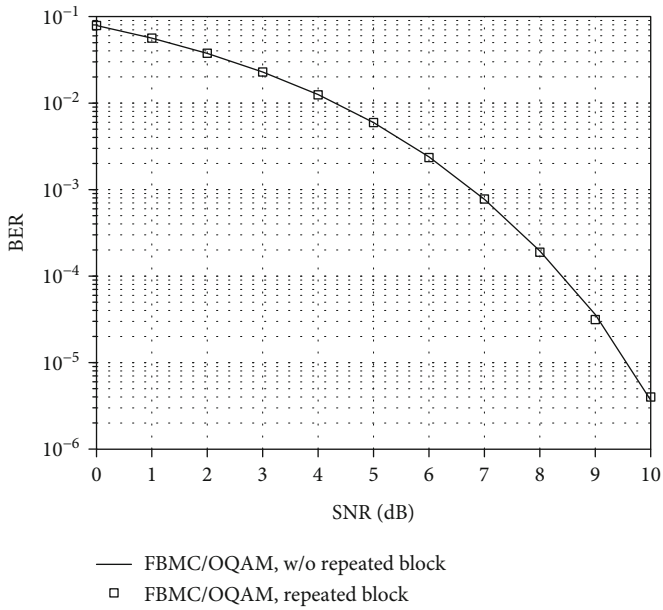


Figure 7: BER comparison in FBMC/OQAM, with or without repeated block.

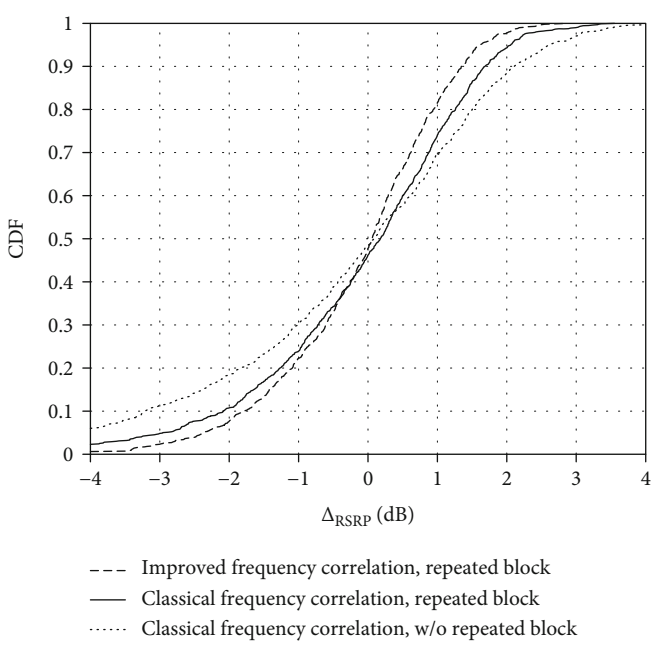


Figure 8: CDF of the proposed schemes under the AWGN channel, SNR = −6 dB.

comparison, the performance of FBMC/OQAM without the repeated block is also given. From the results, FBMC/OQAM with the proposed repeated block achieves the same BER performance as the FBMC/OQAM without repeated block, validating the effectiveness of the proposed scheme. As mentioned above, the imaginary interferences among

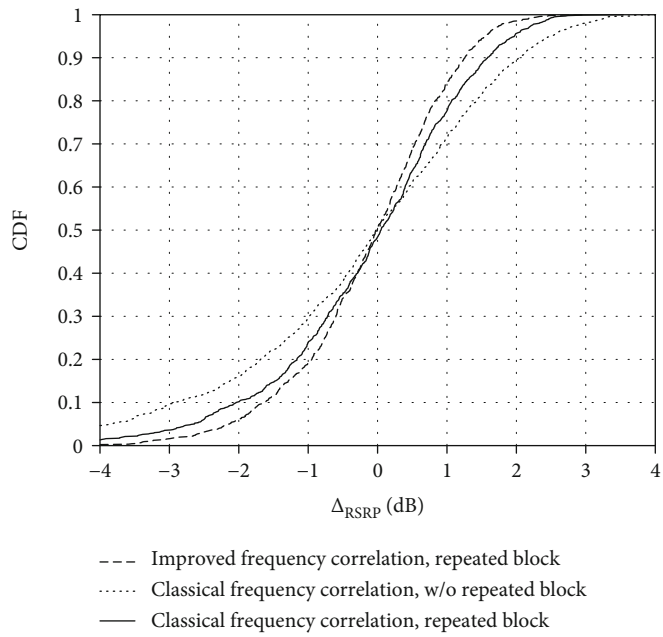


Figure 9: CDF of the proposed schemes under the multipath channel, SNR = −6 dB.

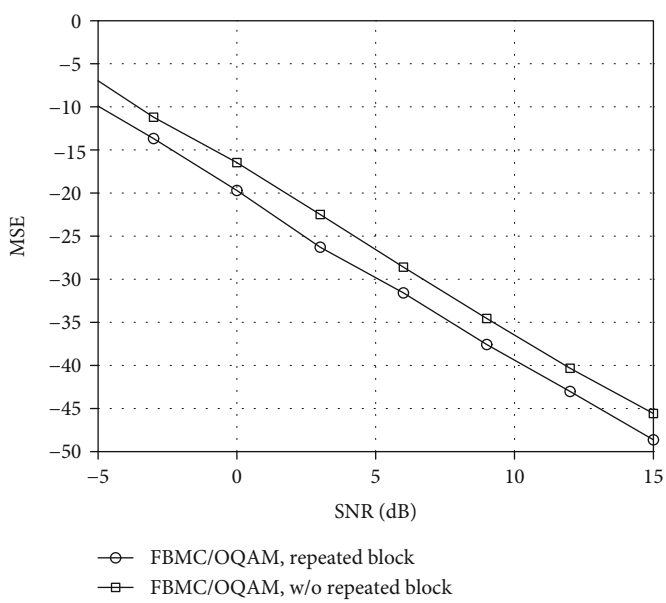


Figure 10: MSE of noise estimation in the proposed schemes under the AWGN channel.

symbols can be eliminated completely by the linear combination in (46); hence, complex-valued symbols can be directly transmitted if the reversed-order repeated block (Figure 5) is adopted. Note that the symbol interval is always $T/2$ in FBMC/OQAM. Therefore, FBMC/OQAM with the proposed repeated block can transmit $M \times N$ complex-valued symbols during the time $2N \times T/2 = NT$, while FBMC/OQAM without the repeated block can transmit $M \times 2N$ real-valued symbols during the time $2N \times T/2 = NT$. Therefore, the data rate is maintained in the proposed scheme. In addition, compared with FBMC/OQAM without repeated block, only the operation of the linear combination is required in FBMC/OQAM with repeated block; hence, the increase of complexity is negligible.

Figures 8 and 9 depict the cumulative distribution function (CDF) to evaluate the accuracy of channel measurement in FBMC/OQAM systems, in which the value Δ_{RSRP} represents the offset between the estimated RSRP and the ideal RSRP [17], i.e.,

$$\Delta_{RSRP} = 10 \log \left(RSRP_p\right) - 10 \log \left(RSRP_{ideal}\right). \quad (59)$$

When Δ_{RSRP} is closer to 0, the RSRP measurement is

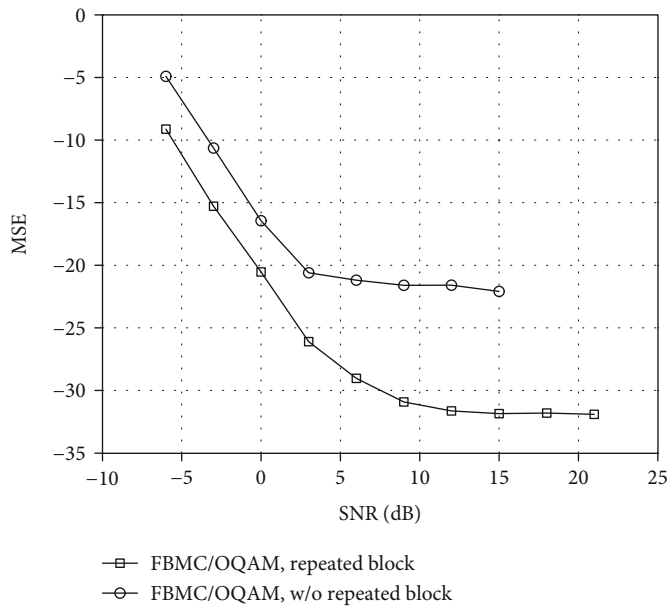


Figure 11: MSE of noise estimation in the proposed schemes under the multipath channel.

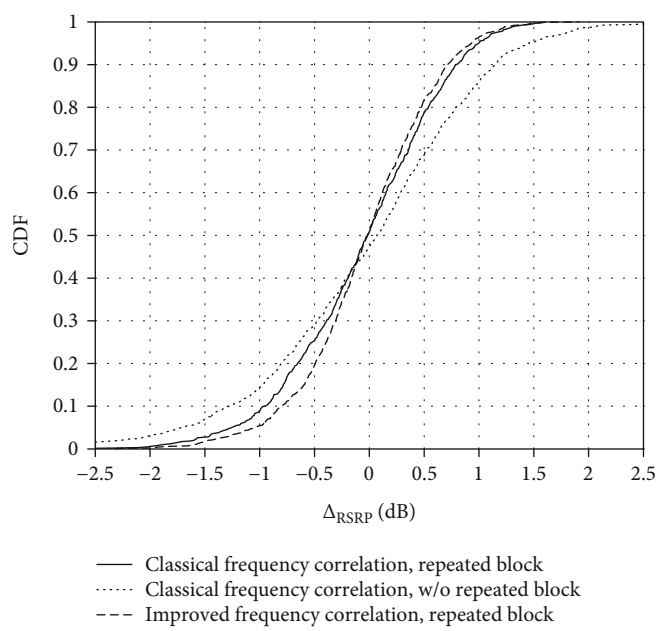


Figure 12: CDF of the proposed schemes under the LTE EPA channel, SNR = −6 dB.

more accurate. Otherwise, the accuracy of channel measurement is poor. In simulations, the AWGN channel is adopted in Figure 8 and the multipath channel is adopted in Figure 9, where SNR is set to be -6 dB [18]. From the simulation results, by the classical frequency correlation method, FBMC/OQAM with the proposed repeated block can achieve better RSRP accuracy than FBMC/OQAM without the repeated block. The reason is that, for the repeated block in Figure 5, all subcarriers can be utilized for the channel measurement since the imaginary interference can be completely removed by the linear combination in (46), while only half of subcarriers in Figure 2 can be utilized for the channel measurement. In addition, by the improved frequency correlation method, the RSRP accuracy can be further improved due to the increase of samples as mentioned in Section 4.2.

To evaluate the estimation accuracy of noise variance, Figures 10 and 11 show the mean square error (MSE) under the AWGN and multipath channels, respectively. From the simulation results, FBMC/OQAM with the proposed repeated block can achieve better estimation accuracy than FBMC/OQAM without the repeated block, due to the fact that all subcarriers are utilized for estimation in the proposed scheme. Especially, compared with Figure 10, there

exists the performance floor under the multipath channel in Figure 11. The reason is that the essence of the frequency correlation method is the correlation of adjacent subchannels that are assumed identical. However, multipath channels are frequency-selective. Therefore, there exists a bit of interference in the frequency correlation method. Note that the performance floor is above 23 dB for the conventional scheme while it is below 30 dB for the proposed scheme, indicating the superiority of the proposed scheme.

For more adequate evaluation, extended pedestrian A (EPA) model is also used in simulations, and Figure 12 depicts the CDF of the proposed scheme under the EPA channel, where SNR is set to be -6 dB. From the simulation results, by the classical frequency correlation method, FBMC/OQAM with the proposed repeated block can achieve better RSRP accuracy than FBMC/OQAM without the repeated block. In addition, by the improved frequency correlation method, the RSRP accuracy can be further improved. The conclusion is consistent to Figures 8 and 9.

6. Conclusions

In this paper, we addressed the channel measurement for link reliability in the FBMC/OQAM-based IoT networks, in which the block repetition was adopted. Firstly, The noise distribution was derived mathematically, which indicated that the noises at the receiver subcarriers are independent. Then, it was shown that the classical frequency correlation method could be applied to achieve the channel measurement, based on the noncorrelation of noises at the receiver subcarriers. Subsequently, an advanced frequency correlation method was proposed to improve the accuracy of the channel measurement in FBMC/OQAM-based wireless networks. Numerical simulations showed that the accuracy of channel measurement could be significantly improved by the proposed schemes.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was financially supported in part by the National Science Foundation of China with Grant number 62001333.

References

[1] S. Sesia, M. Baker, and I. Toufik, *LTE: The UMTS Long Term Evolution: From Theory to Practice*, Wiley Publishing, 2009.

[2] T. Zhou, C. Tao, S. Salous, L. Liu, and Z. Tan, "Implementation of an LTE-based channel measurement method for high-speed railway scenarios," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 1, pp. 25–36, 2016.

[3] D. Kong, J. Li, K. Luo, and T. Jiang, "Reducing pilot overhead: channel estimation with symbol repetition in MIMO-FBMC systems," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7634–7646, 2020.

[4] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-Complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2021.

[5] R. Nissel, S. Schwarz, and M. Rupp, "Filter bank multicarrier modulation schemes for future mobile communications," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1768–1782, 2017.

[6] D. Kong, X. Zheng, Y. Zhang, and T. Jiang, "Frame repetition: a solution to imaginary interference cancellation in FBMC/OQAM systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1259–1273, 2020.

[7] D. Kong, X. Zheng, Y. Tian, T. Jiang, and Y. Zhang, *Alamouti Code Based on Block Repetition in FBMC/OQAM Systems*, Appear to Digital Communications and Networks, 2021.

[8] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; Measurements*, 2011, http://www.3gpp.org/ftp/Specs/archive/36_series/36.214/36214-a30.zip.

[9] G. Huang, A. Nix, and S. Armour, "DFT-based channel estimation and noise variance estimation techniques for single-carrier FDMA," in *IEEE Vehicular Technology Conference*, pp. 1–5, Ottawa, Canada, 2010.

[10] 3GPP TSG RAN WG4 63, TDocs R4-63AH-0145, *Further considerations for the feasibility research of CSI-RS based RSRP measurement*, Samsung, 2012.

[11] 3GPP TSG RAN WG4 63AH, TDocs R4-63AH-0084, *Further link level performance analysis of CSI-RS based RSRP measurements*, Intel Corporation, 2012.

[12] D. Kong, X.-G. Xia, P. Liu, and Q. Zhu, "MMSE channel estimation for two-port demodulation reference signals in new radio," *Science China Information Sciences*, vol. 64, pp. 169303:1–169303: 2, 2021.

[13] C. Lélé, J. P. Javaudin, R. Legouable, A. Skrzypczak, and P. Siohan, "Channel estimation methods for preamble-based OFDM/OQAM modulations," in *European Wireless Conference*, pp. 59–64, Paris, France, 2007.

[14] D. Kong, D. Qu, and T. Jiang, "Time domain channel estimation for OQAM-OFDM systems: algorithms and performance bounds," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 322–330, 2014.

[15] A. Zafar, L. Zhang, P. Xiao, and M. A. Imran, "Spectrum efficient MIMO-FBMC system using filter output truncation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2367–2381, 2018.

[16] J. Du and S. Signell, "Time frequency localization of pulse shaping filter in OFDM/OQAM systems," in *International Conference on Information, Communication and Signal Processing*, pp. 1–5, Singapore, 2007.

[17] 3GPP TSG RAN WG1 88, TDocs R1-1703097, *DL signals for mobility measurements in NR and mobility schemes*, Nokia, 2017.

[18] 3GPP TSG RAN WG1 87, TDocs R1-1612123, *DL RRM measurement over narrowband reference signals*, MediaTek, 2016.