Wireless Communications and Mobile Computing

Data Collection in Resource-Limited Networks (WSNs, IoT, Sensor Cloud)

Lead Guest Editor: Ihsan Ali
Guest Editors: Mohammad Hossein Anisi, Suleman Khan, and Rahim Khan

Data Collection in Resource-Limited Networks (WSNs, IoT, Sensor Cloud)

Wireless Communications and Mobile Computing

Data Collection in Resource-Limited Networks (WSNs, IoT, Sensor Cloud)

Lead Guest Editor: Ihsan Ali
Guest Editors: Mohammad Hossein Anisi, Suleman Khan, and Rahim Khan

Chief Editor

Associate Editors

Academic Editors

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9635958, 14 pages
https://doi.org/10.1155/2021/9635958

Research Article

Network Performance Metrics for Energy Efficient Scheduling in Wireless Sensor Networks (WSNs)

Felicia Engmann,[1,2] **Kofi Sarpong Adu-Manu**,[2] **Jamal-Deen Abdulai**,[2] **and Ferdinand Apietu Katsriku**[2]

[1]*School of Technology, Ghana Institute of Management and Public Administration, Accra, Ghana*
[2]*Department of Computer Science, University of Ghana, Legon, Accra, Ghana*

Correspondence should be addressed to Kofi Sarpong Adu-Manu; ksadu-manu@ug.edu.gh

Received 21 May 2021; Revised 26 July 2021; Accepted 9 November 2021; Published 27 November 2021

Academic Editor: Ihsan Ali

In Wireless Sensor Networks, sensor nodes are deployed to ensure continuous monitoring of the environment which requires high energy utilization during the data transmission. To address the challenge of high energy consumption through frequent independent data transmission, the IEEE 802.11b provides a backoff window that reduces collisions and energy losses. In the case of Internet of Things (IoTs), billions of devices communicate with each other simultaneously. Therefore, adapting the contention/backoff window size to data traffic to reduce congestion has been one such approach in WSN. In recent years, the IEEE 802.11b MAC protocol is used in most ubiquitous technology adopted for devices communicating in the IoT environment. In this paper, we perform a thorough evaluation of the IEEE 802.11b standard taking into consideration the channel characteristics for IoT devices. Our evaluation is aimed at determining the optimum parameters suitable for network optimization in IoT systems utilizing the IEEE 802.11b protocol. Performance analysis is made on the sensitivity of the IEEE 802.11b protocol with respect to the packet size, packet delivery ratio (PDR), end-to-end delay, and energy consumption. Our studies have shown that for optimal performance, IoT devices using IEEE 802.11b channel require data packet of size 64 bytes, a data rate of 11Mbps, and an interpacket generation interval of 4 seconds. The sensitivity analysis of the optimal parameters was simulated using NS3. We observed PDR values ranging between 27% and 31%, an average end-to-end delay ranging within 10-15 ms while the energy remaining was between 5.59 and 5.63Joules. The results clearly indicate that scheduling the rate of packet generation and transmission will improve the network performance for IoT devices while maintaining data reliability.

1. Introduction

The proliferation of the Internet has seen an increase in the growth of the Internet of Things (IoT). There is a continuous flow of data seamlessly from trillions of IoT devices that usually go unnoticed and unused. It is expected that by 2021, 28 billion devices will be connected through IoT. IoT devices autonomously form a network that communicates with other devices continuously, hence generating massive amount of data from their deployable environment. IoT devices are adopted for earthquake, healthcare, vehicular tracking, and agricultural monitoring applications. IoTs are similar to Wireless Sensor Networks (WSNs) in their operations. They are made up of a large number of small battery-operated devices that can sense, process, and communicate data wirelessly [1]. Their key benefits are their ability to operate in harsh environments unattended, where human control schemes are difficult or infeasible to implement [2]. The use of IoTs and WSNs has been proven superior to the traditional methods of collecting data from the environment. With IoTs and WSNs, sensor nodes autonomously form interconnected networks that collaborate in data collection [3]. However,

sensor devices can replace traditional devices by sending sensed data to sinks through single-hop or multihop communication. IoT devices mostly communicate via IEEE 802.11 standard due to its flexibility of implementation and scalability.

IoT devices transmit data from different sources such as wearable devices, smart vehicular systems, smartphones, and several laptops. All these devices generate data transmitted through cellular base stations, Wi-Fi access points, and other Road Side Units (RSUs) to a service provider. There are a huge number of transmissions received at each access point creating packet losses due to the limited channel access. Hence, the need to design communication protocols resolves the channel access problems.

The communication protocols designed for IoT and WSN applications sought to address challenges related to collision, overhearing, protocol overheads, and idle listening. The collision occurs when multiple nodes attempt transmitting packets at the same time. The simultaneous multiple transmission from IoTs and WSNs increases the reception cost at the destination node while increasing the cost of retransmission in the source node. Nodes usually listen for transmission on the channel to avoid collisions. Nevertheless, when nodes stay awake listening multiple times for packets destined for another node, overhearing occurs. In overhearing, the broadcast of packets by the wireless medium causes all one-hop neighbors (within its range) of the source node to hear the transmitted packet, aggravating overhearing [4].

Nodes in an idle mode listen to the channel awaiting possible incoming packets. However, a sensor node in an idle state can consume a similar amount of energy as a node that is receiving a packet [5–7]. As a result, turning off the transceiver in idle mode is critical to reducing excessive energy consumption in IoTs and WSNs [2, 8]. However, frequently turning off the transceiver to sleep introduces longer latencies and increase in the exchange of overhead packets. A number of authors have proposed Time Division Multiple Access (TDMA) approaches to mitigate the challenges of overhearing, collision and idle listening [2]. However, an essential factor in TDMA-based protocol is when to update the transmission schedule to prevent conflicts in transmissions from neighboring nodes. Since enforcing fixed transmission schedules might not have proven energy-efficient for networks requiring continuous monitoring, scheduling based on the spatiotemporal differences of data from deployed sensor nodes is proposed and the simulation results are presented in [9].

However, medium access control (MAC) protocols implement some scheduling that allows nodes to communicate in a way that prevents multiple collisions and energy wastage. The IEEE 802.11b MAC protocol defines a standard physical and MAC layer that specifies the access protocol for all nodes in the network. The Distributed Coordination Function (DCF), one of two coordination functions in the MAC layer, supports asynchronous transmissions which as supported in IoT environments [10].

This paper, therefore, evaluates the traditional MAC protocol implemented in the IEEE 802.11b that schedules nodes to sleep and wake up within some optimum periods. Moreover, the MAC protocol of the IEEE 802.11b implements the DCF which mandates a gap of specified minimum delay. The DCF implements a binary exponential back-off the carrier sense multiple access with collision avoidance (CSMA/CA) for effective implementation in both ad hoc and infrastructure systems. Due to the number of simultaneous transmissions, scheduling transmissions of individual nodes is a challenge in IoTs. In this paper, we examine the effect of network parameters such as the payload size, the packet interarrival interval, and the data rate of the IEEE802.11b.

The remaining sections of the paper are discussed as follows. Section 2 presents energy management issues in WSNs. In Section 3, we present related works. In Section 4, simulation results and discussions on the sensitivity analysis of IEEE802.11b are provided. Finally, Section 5 concludes the paper.

2. Energy Management in WSN

In WSNs, sensor nodes sample data from the environment and wirelessly transmit to a base station for onward processing and the outcome communicated via the Internet to remote users. To ensure the continual availability of communication between the sensor nodes and the base station, energy management of the various components of the sensor node, illustrated in Figure 1, is paramount as presented in [11]. However, despite the implementation of these energy management principles, the energy is consumed by the communication subsystem of the sensor node, mainly controlled by the radio, the primary source of energy consumption during transmission as shown in Figure 2.

Energy management includes energy harvesting, energy transfer, and algorithmic schemes (protocols). These energy management approaches restore depleted energy in a sensor node. The rate of discharge of energy through the operations of the radio is mainly faster than the rate of energy replacement by the energy management approaches mentioned in the text. The rate of energy depletion in sensor node makes continuous communication a significant threat to the continuous operation in WSN applications [11]. There is, therefore, the need to control the frequency of communication to reduce the amount of energy consumed by the sensor node, especially, during data transmission.

Duty cycling, an energy conservation method (as represented in Figure 3), schedules the sensor nodes to be turned on or off at regular intervals to reduce the frequency of the operations of the sensor nodes. However, the operations of the sensor node require a trade-off between energy efficiency on one side and latency and throughput gains on the other. To capitalize on the benefits of duty cycles, the authors in [5, 12] proposed power management, topology controls and MAC layer algorithms, and dynamic sleep wake up cycles [13, 14]. However, challenges of regular sleep/wake-ups include the exchange of extra overheads, data losses, and increased latencies and associated energy losses. The possibility of mitigating these challenges includes reducing the frequency of switching to sleep if the node's sleep and wake-up schedule are not mainly based on the network activity [5].

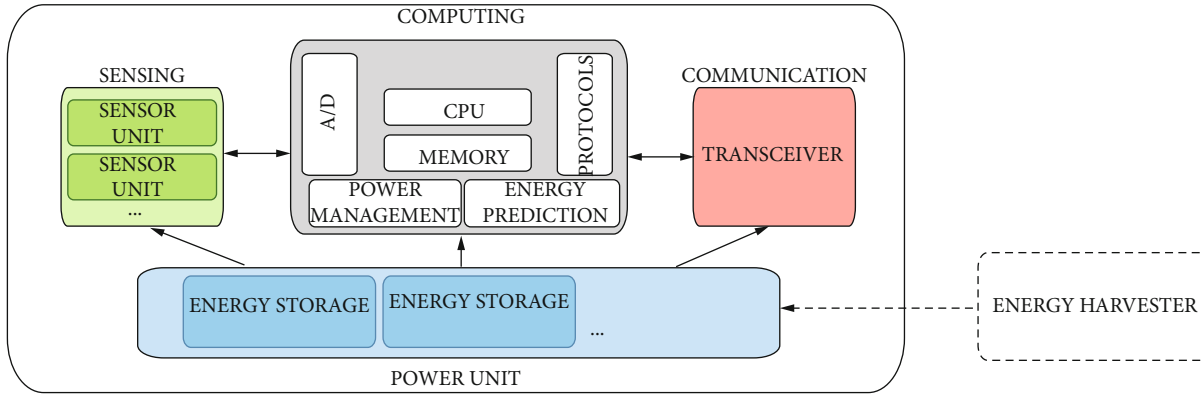


Figure 1: Wireless sensor node.

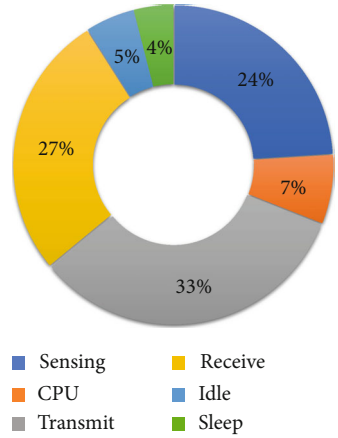


Figure 2: Energy consumption in WSNs.

IoTs introduce a unique challenge of several packets being transmitted simultaneously from randomly placed nodes which may introduce multiple interferences if care is not taken. However, these IoTs have the opportunity to learn and adapt their radios to the activities of its neighboring nodes. An interference and spectrum aware channel access mechanism was proposed by authors in [10]. The channel access approach was proposed because DCF, Point Coordination Function (PCF), and Time Division Multiple Access (TDMA) could not prove efficient in managing interference.

Meanwhile, the DCF implementation with CSMA/CA used in the IEEE 802.11b has been evaluated to decide on optimum parameters for network implementation in the face of these known challenges [15]. In their work, simulations of the DCF with basic mechanisms assumed that nodes arrive in a Poisson distribution interval with fixed payload sizes. However, due to the parameters of the IEEE 802.11b channel, these channel conditions require some standard network parameters to obtain optimum results. These optimum network output include a decrease in the channel delay, increased throughput, and reduction in the energy consumed.

MAC protocols with low duty cycle implementation in WSNs may be classified as either contention-based or TDMA. The energy saving occurs when deployed nodes largely remain in inactive modes until the detection of events. However, when several nodes in proximity detect events and wake-up to sense and transmit, without proper scheduling, extensive collision may occur. Strategies implemented in literature to overcome the challenges of scheduling in WSNs have been implemented in the literature. One such approach is the Sensor MAC (S-MAC), a scheduling method that allows border nodes to adopt a diversified transmission schedule in WSNs [16]. The approach used in SMAC is to mitigate the high energy lost due to the switching of border nodes in virtual clusters. The rational of the one-time scheduling implementation is to reduce the energy the sensor node spends in its frequent switch between listening and sleep mode. TDMA approaches have also been implemented. One of such implementations proposes a tight-time scheduling and increased throughput scheduling with TDMA [17]. The scheduling method reduced the energy and overhead costs incurred by the network during the network setup phase. However, much time and overhead costs were spent in the initial network setup phase. Another TDMA technique was implemented for fast data aggregation at fixed intervals for scheduling when frequent transmissions are required [2]. Since most TDMA implementations require a fixed interval of transmission, the data itself has less effect on the schedule of its packet transmissions. Most TDMA systems are also not effective for larger network sizes since scalability of the scheduling is a challenge. The regular intervals of TDMA, also, introduce several unused slots that increase latency. In our approach, we do not employ the use of tight time synchronization (TTS) for node transmission. Nodes have equal chance to compete for channel access, but due to the limited number of competing nodes accessing the channel, our approach overcame the challenges of collision that TTS is designed to solve.

Since the CSMA/CA protocol implementation of DCF in IEEE802.11b also introduces some basic scheduling, the paper is aimed at obtaining network parameters that could optimize transmissions of multiple packets when no further scheduling is used up.

3. Related Works

In this section, we provide related works that focus on an access method in IEEE802.11b. Many wireless devices in

Figure 3: Energy conservation schemes.

the case of IoTs create a major bottleneck in the wireless channel during the data communication process. Early studies on heavily utilized IEEE802.11b channel were performed to optimize the network performance using link layer information [18]. Congestion in the channel directly impacts on its use; therefore, the authors observed that RTS/CTS which prevents congestions may conflict with node's fair access to the shared channel. Also, even though IEEE802.11b provide users with 4 data rates, which include 1 Mbps, 2 Mbps, 5.5 Mbps, and 11 Mbps, users frequently use the 1 Mbps and 11 Mbps channels. In their work, the authors highlighted the detrimental effect of rate adaptation of channels to network performance if implemented as a response to congestion.

Some researchers have proposed the optimization of the IEEE802.11b channel through adjusting the contention window size [16, 17] or the frame sizes [19–21] in relation to the traffic in the channel. The DCF implementation in 802.11b also suggests the implementation of multirate adaptation channels to mitigate the poor throughput performance of the low data rate channels.

The basic contention-based protocol of MAC is the DCF which is implemented in most WSNs due to its simplicity of implementation, ability to counter the hidden terminal problem, and scalability as seen in IoTs.

In the operation of the DCF Protocol in CSMA/CA, nodes willing to transmit wait a predefined DCF Interframe Space (DIFS) of 50 μs. If the transmitting node's physical layer does not sense any signal within the DIFS, a clear channel assessment (CCA) is sent to the MAC layer. If the recipient node does not sense any packets in transition, it waits a random period known as the contention window. If no signal is sensed, the transmitting node sends a Request-to-Send (RTS) to the receiving node. The RTS contains the MAC Address of the transmitter and receiver. It also has a field that contains the duration of the first MAC fragment. The RTS information allows other transmitters to determine their Network Allocation Vector (NAV) to prevent collisions. Hence, in our work, we consider an optimum interpacket interval that might influence the number of packets that will contend for the channel at every point in time. The receiver node sends a Clear-to-Send (CTS) signal after a short inter-frame space (SIFS) of 10 μs, which contains information that adjusts the neighbor NAV. When a packet is too large, it is divided into several MAC fragments which will be successively sent to the recipient. A diagram of the transmissions of two nodes depicting the explained process above is presented in Figure 4. The work done in this paper, therefore, seeks to also find an optimum payload size combined with its interpacket interval and prevents congestion of the channel if several packets need to be sent in an IoT environment.

A flowchart of the CSMA/CA is presented in Figure 5. The basic CSMA/CA with DCF that coordinates node operation in ad hoc mode permits all nodes to communicate within their transmission range without enforcing association and beacon generation. Because no association rules

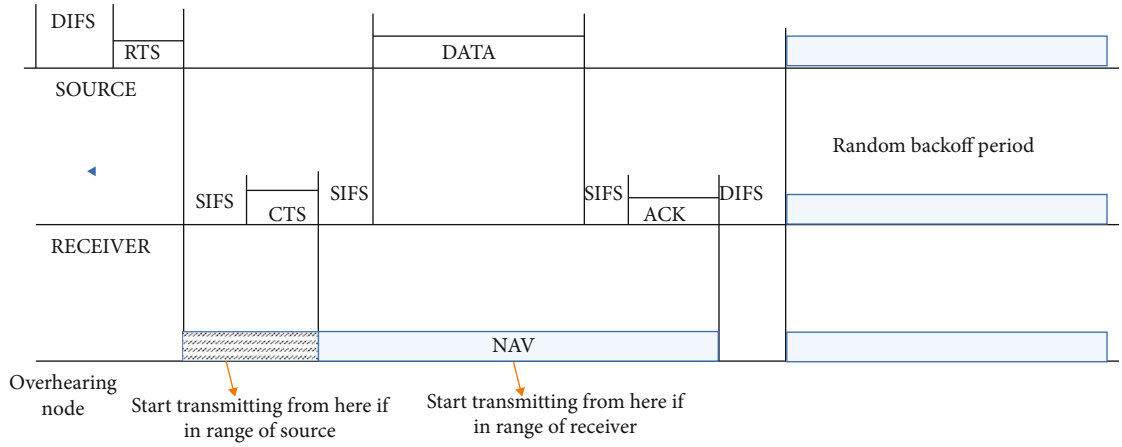


Figure 4: Two devices communicating on an IEEE 802.11b channel with CSMA/CA.

are enforced to coordinate communication, nodes always remain ready to receive messages from neighbors. Hence, nodes in this infrastructure do not sleep but are in a constant idle power consumption state.

4. Simulations and Discussions

An IoT environment typically may have several devices transmitting data to a single receiving device. The energy parameters that relate to the different states of a typical Micaz node used in this paper's simulations are presented in Table 1.

An illustration of a typical IoT system where a number of nodes communicate via a common channel to a sink node is also presented in Figure 6. We observe that the unregulated communication emanating from the sensor nodes causes collision in the channel. Hence, the data arriving at the sink node may be greatly distorted.

From Figure 6, only one node [i.e., node 1] will have access to the channel at a time. The remaining generated packets arriving from nodes 2 to 5 may collide and drop since the channel is busy and not available. The channel is saturated when the queue has more than some maximum value or maxSize of packets. Therefore, any additional packets will be dropped. Packets may also be dropped when they stay longer than Dmax in the channel when competing for channel access, which in IEEE 802.11b implements as FIFO queue.

For our simulations, we assume a first-in-first-out (FIFO) queue implementing a maxSize of 500 packets and a Dmax of 500 seconds (maxSize is the maximum number of packets that can be the queue while Dmax is the maximum time a packet may remain in the queue). The queuing process of the channel as shown in Figure 7 depicts the interval between successive packets en-queuing. The number of packets in the queue may be reduced if the interval between successive en-queuing increases, such that the rate of de-queuing is faster than the interval for en-queuing. Hence, with appropriate scheduling, the optimum number of nodes will occupy the channel within a Dmax while preventing excess collisions.

To obtain the optimum combination of the payload/packet size and interval, the simulations also considered the energy consumption of the nodes. Network performance metrics used included the average end-to-end delay, packet delivery ratio, and energy consumption. The simulations performed on the IEEE 802.11b channel are to generate these optimum parameters to enable a scheduling implementation that reduces the number of packets in the channel per unit time while avoiding collisions and reducing latencies.

4.1. Simulation Parameters. Table 2 shows the network parameters used in the simulation. The NS3 simulator, a discrete event simulator with MAC layer and extensive energy support, implements the scheduling algorithm proposed. Each sensor node communicates with each other directly within its communication range.

General assumptions made on the network include a fixed number of sensor nodes randomly deployed in a fixed-sized network. All nodes are homogenous in size, capabilities, and initial energy. For a single-hop network, all nodes are within a 100 m distance away from the sink. Nodes are not GPRS enabled; hence, they assume the distances from each other using received signal strength indication- (RSSI-) based methods. The simulations adopt the IEEE802.11b channel, corresponding to the likely channel for IoTs. We assume an omnidirectional antenna and suppose there will be no fast or slow fading antenna signal. The network has randomly deployed nodes transmitting directly to a single sink, as shown in Figure 8. Nodes are not required to synchronize with each other or the sink before initiation of communication. Parameters for simulations are as presented in Table 2.

4.2. Simulation Results of the Network Performance of the IEEE802.11b. The primary network parameters used were 32-bytes of payload size transmitted at a data rate of 11 Mbps; simulation time is 1000 s at a packet generation rate of 1 s. The network assumes implementation of the IPV4 base MAC of 255.255.255.0, and hence a maximum of 254 nodes are enough to perform investigations for the simulations. For multiple nodes connected to a single

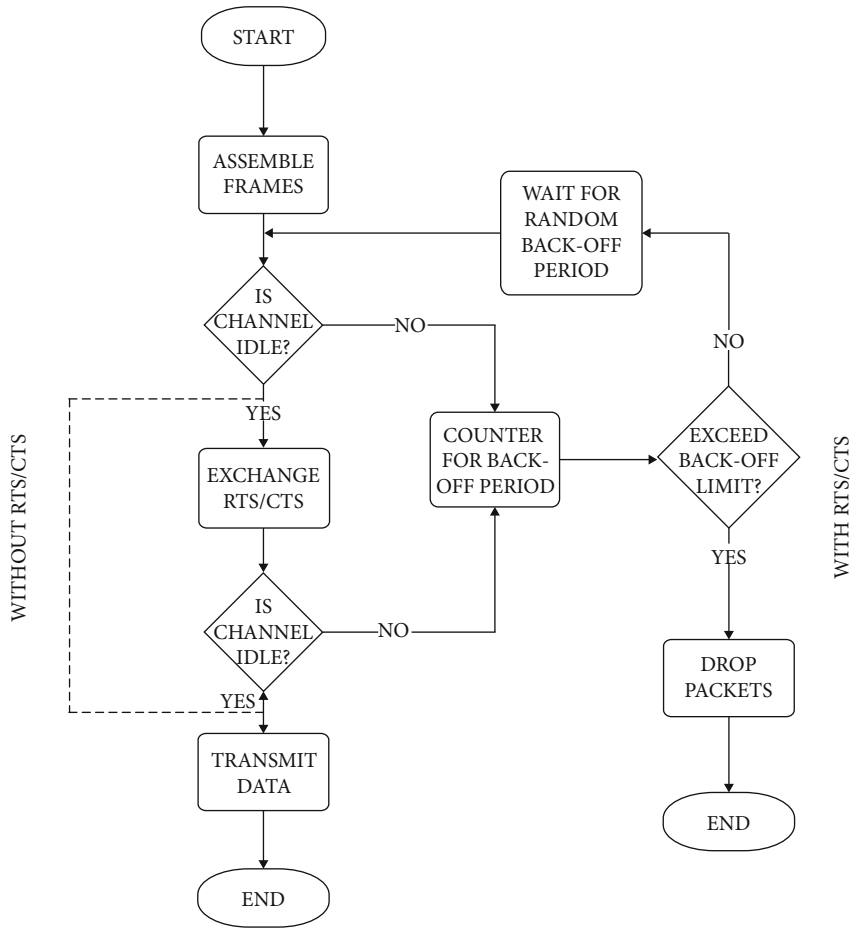


Figure 5: CSMA/CA in IEEE802.11b.

Table 1: Energy measurement of IEEE802.11 used in the simulations [3, 22].

State	Current draw (A)	Power consumption in Watts
TX	0.380	1.14
RX	0.313	0.94
IDLE	0.273	0.82
SLEEP	0.033	0.10

destination, queuing on the channel may be one of the principal causes of collisions that affect network performance.

4.3. Effect of Data Rate on the Node Density. The nodes are deployed randomly in a river sensor network (RSN) of area 100 by 100 square meters. We assume that the deployed nodes are on the surface of the river and make little or no movement, therefore modelled as static. Nodes deployed are assumed to be in the one-hop communication range of the sink. If the communication range is assumed to be 100 m, any node placed within this area will be successfully transmitted without using intermediary nodes for multi hopping. Sensed data is transmitted immediately without an intentional delay; hence, the MAC is assumed to be in an ad hoc mode. Comparing the different data rates of IEEE802.11b, simulations run for a period of 1000seconds with a seed run of 1000. The initial energy on the nodes is 20 J, and the interpacket generation interval is 1 second.

Observations of the graph in Figure 9 indicate simulations for data rates 11 Mbps, 2 Mbps, 5.5 Mbps, and 1 Mbps, recording an increase in the average end-to-end delay of packets generated over the increasing node densities. 1 Mbps channel records the highest delay for all node densities. The delay ranges between 14 ms at the highest and about 10 ms between 140 and 160 nodes. The 11 Mbps channel generally has the best delay. Its average end-to-end delay values range between 11 ms at the highest and 9.8 ms at its lowest. The recorded delay indicates more collisions in the narrow channels of 1 Mbps as opposed to the 11 Mbps channel. The time it takes a packet to access the channel and the retransmission delay are the main factors causing delays in WSN. All data rates have their average delays recording closer values when the network has between 140 and 160 nodes.

The increase in packets causes a corresponding increase in the traffic per unit time. Since the slot time for 802.11b is 20 μs, an increase in the number of nodes increases the competition for the channel. The resulting increase causes collisions and retransmissions, increasing the possible number of dropped packets. The average PDR for 1, 2, 5.5, and

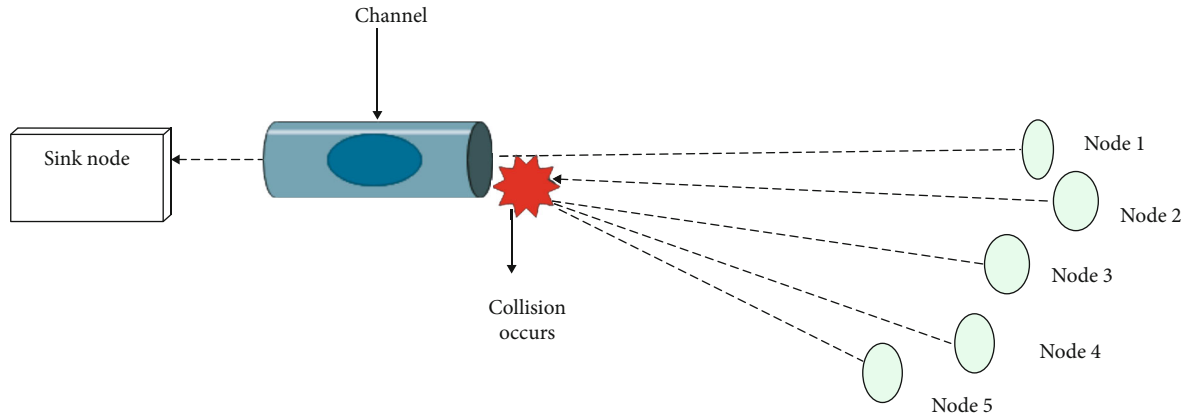


Figure 6: Link level collision for multiple nodes directly communicating to a single source.

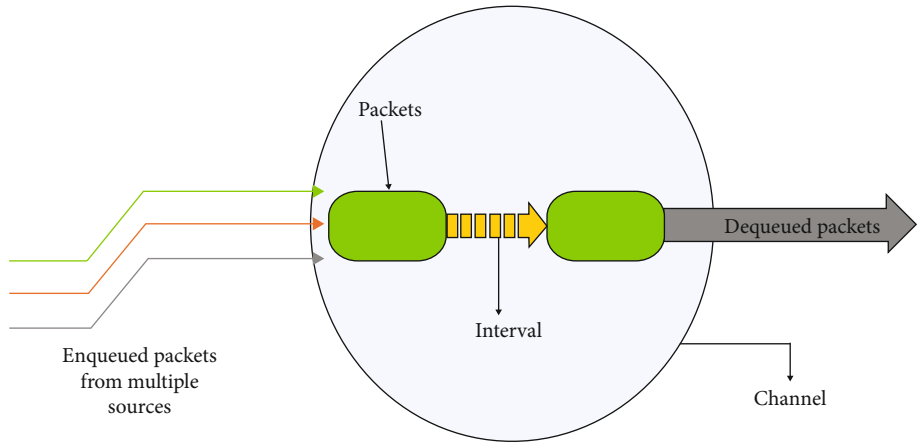


Figure 7: Link level queuing for multiple sources communicating directly with sink.

Table 2: Network simulation parameters.

	Network parameters	Value
Network parameters	Number of nodes	Between 2 and 250
	Number of sink	1
	Initial energy on node	20 J
Packet parameters	Number of packets	Unlimited
	Number of retransmissions	Max 7
	Packet size	32, 64, 128, 512, 1024 bytes
	Packet interval	Between 0.5 and 5 seconds
	Traffic type of packets	CBR
Communication parameters	Sensor communication range	100 m
	Data rate	1 Mbps, 2 Mbps, 5.5 Mbps and 11 Mbps
Transmission slot parameters	Slot time	20 μs
	Queue length	50packets
	Slot length	
	Slot duration	20 microseconds (for 802.11b)
Number of seed runs		1000
Simulation time		Max 1000secs

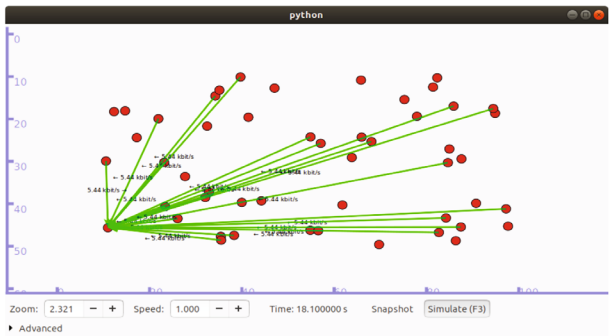


Figure 8: Randomly deployed stationary nodes communicating to a stationary sink.

11 Mbps channels are 8.667, 13, 15.2, and 17.8, as illustrated in Figure 10.

The causes of energy waste for ad hoc MAC include collisions that result in retransmissions of the collided packets and idle listening, which occurs when nodes listen in vain to receive packets. WSNs generally operate in the idle modes for more extended periods and transmit during its active state of the duty cycle. Nevertheless, the energy consumption of the idle state is almost the same as energy for transmission/reception of data. At the same time, it is much higher than the energy consumed during the sleep mode. When the power of the battery source reduces below the energy threshold, nodes remain in an idle state. In the idle state, the node switches to a low-power state and turn-off their transceiver. Therefore, in an idle state, no transmission or reception of data is possible, but the node remains alive to listen to the channel.

When the network density is 140 nodes, the remaining energy on the nodes does not permit the further transmission of packets. The energy remaining permits the sink node to receive; hence generated packets may not be transmitted, reducing PDR and causing energy waste in the network. Observations from Figure 11 show that the 1 Mbps channel has the least remaining energy, while the 11 Mbps channel has the best energy consumption. This minimum energy results from collisions that generate retransmissions when the data rate of the channel is higher. In the next section, the interpacket generation time is varied.

4.4. Effect of Interpacket Generation Interval. Inter packet generation interval (IPI) is the difference between successive packets generated from the same source node in a network. For 2, 10, 50, 100, and 150 nodes, we compare some random intervals from 1 second and below. The PDR obtained is for intervals which are multiples of the delaythreshold of IEEE802.11b represented as Inter-Packet generation Intervals (IPI). For the same simulation period and node density, the ratio of packets transmitted to packets received when the IPI increases do not change significantly. The PDR ranges from 100% to 8% for 2 to 150 nodes, respectively. The effect of IPI on the ratio of packets received and transmitted was not significant in these minor interval differences. However, observations show that an IPI of 1 second gives the best PDR.

The PDR generally decreases with increasing network density, as nodes increase from 100 to 200 nodes. The highest PDR is recorded at 100 nodes when the IPI is 3 seconds. Unlike the PDR at 1 second, which is about 25%, IPI of 3 seconds records a better PDR of 31%. However, observation shows that the PDR for particular network density does not follow a regular pattern. However, the PDR is best for IPI between 3 and 4.5 for all network densities observed, as shown in Figure 12.

Comparing the average end-to-end delay, intervals of 0.5 and 1 performed better with most minor delays, as illustrated in Figure 13. The delay was 0.3 ms for 2 nodes and increases 10 ms at 100 nodes when the IPI is either 0.5 seconds or 1 second. On the other hand, 0.9 seconds recorded the highest delay at 100 nodes. The graph suggests that the network reaches saturation after 100 seconds. Further observations of node densities concentrate on nodes from 100 nodes and beyond.

0.5 seconds is the maximum threshold for average E2E delay in the channel; after which packets remaining in the channel are dropped.

The average remaining energy, presented in Figure 14, is dependent on the total number of nodes available in the network. Therefore, when the network is saturated and the node battery reaches its low energy threshold, the per-node remaining energy remains relatively stable for an increasing number of nodes. This regular graph represents the fairness of energy and load distribution during the network lifetime. However, energy consumption depends on the number of transmissions/reception of packets and other channel access conditions. The network saturates after 140 nodes; hence, the total remaining energy remains almost the same for all network densities. IPI of 1.5 and 3.5 seconds recorded the highest remaining energy for 100 nodes, but the remaining energy levels were highest for IPI between 3.0 and 5.0 seconds for all network densities. This buttressed the trends observed in the PDR and Average E2E delay graphs that the network performance improves with higher IPI. However, the particular IPI chosen must be according to the network density.

4.5. Effect of Packet Size (Payload) on Network Performance. Generally, the energy consumption of a wireless network depends mainly on the transmission of packets and the network's reliability. Smaller packets may increase the chances of reliable channel transmissions since fewer errors may occur in the channel. However, smaller packet sizes increase the overhead of network protocols and error-correcting codes, leading to less energy-efficient transmissions. In WSN, the collision rates experienced during transmission are closely related to the size of the packet. Analysis in [19] suggests an increase in the payload size decreases the MAC layer failure rate.

The effect of the payload size on the network performance investigated using network densities of 100 by 20 nodes to 200 is shown in Figure 15. The best PDR is recorded in all network densities for payload size 64 bytes, while 32 bytes had the worst PDR. The observed trend confirms a decrease in MAC failure as packet size increases.

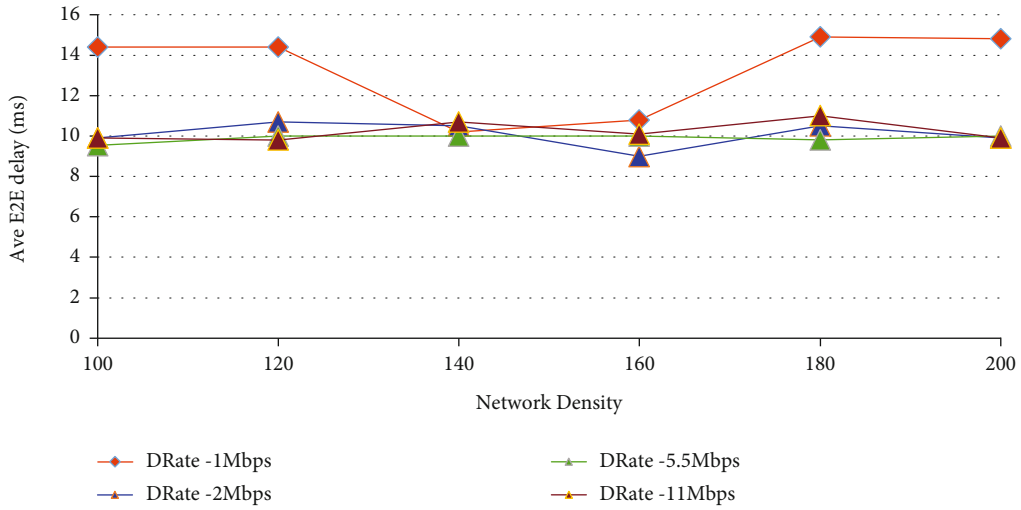


Figure 9: Average end-to-end delay of 1, 2, 5.5, and 11 Mbps channel in IEEE 802.11b.

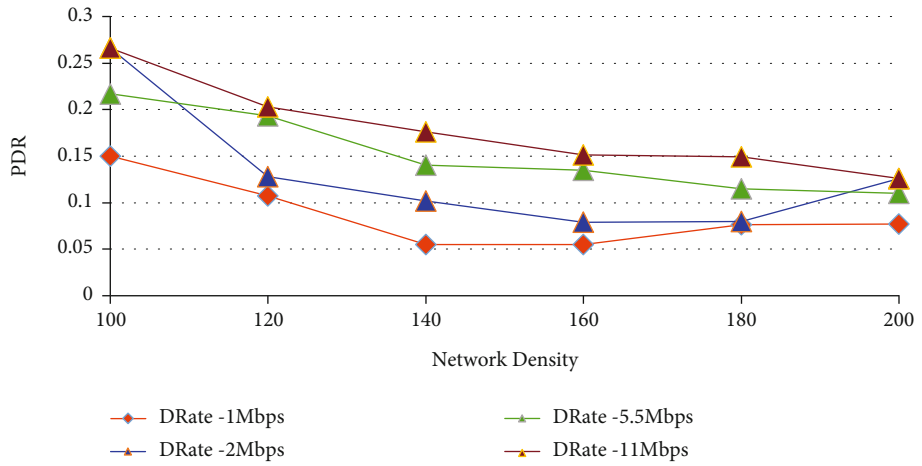


Figure 10: Packet Delivery Ratio of 1, 2, 5.5, and 11 Mbps channel in IEEE 802.11b.

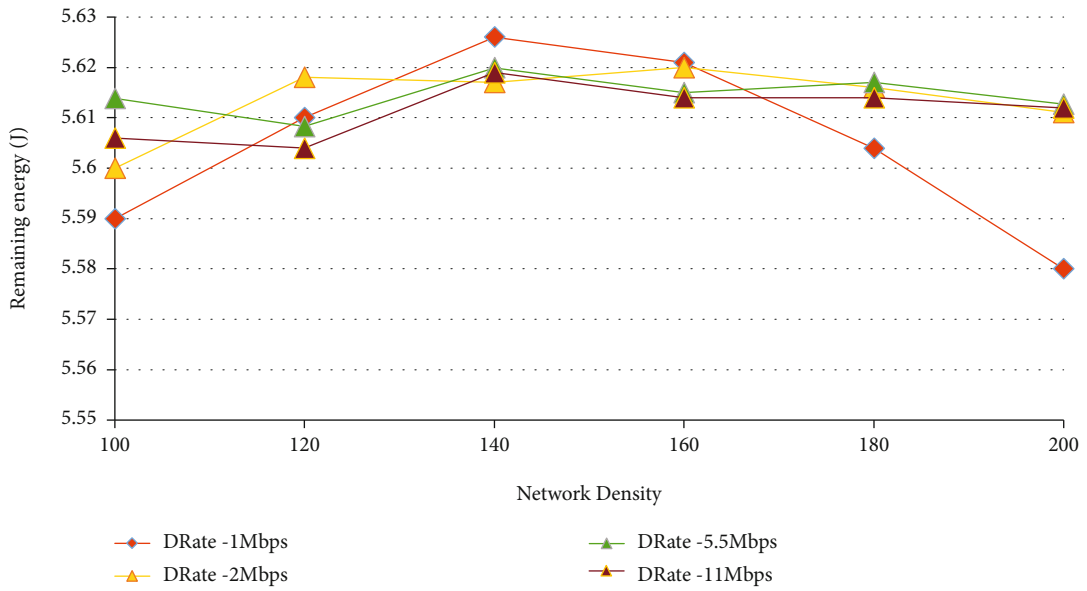


Figure 11: Remaining Energy of the 1, 2, 5.5, and 11 Mbps channel in IEEE 802.11b.

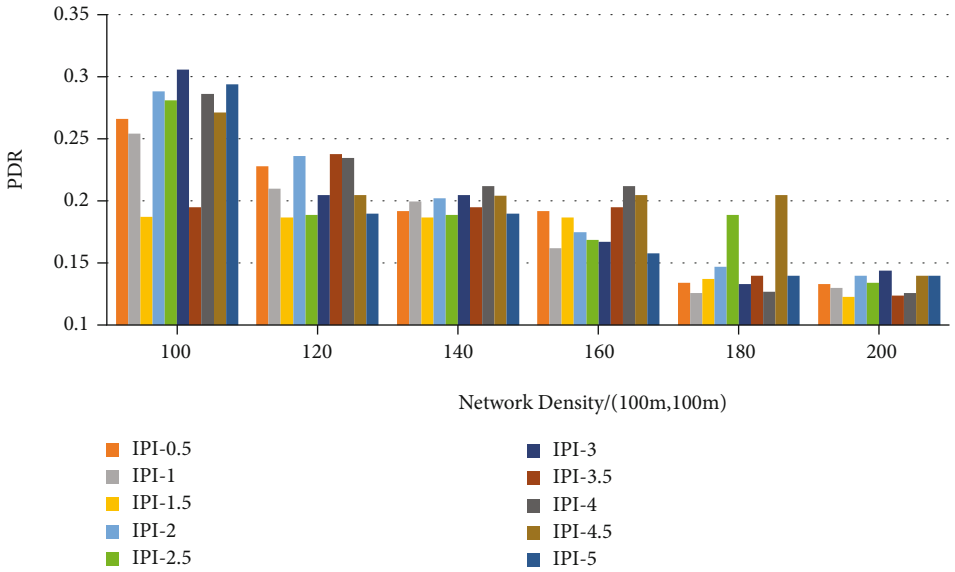


Figure 12: Packet delivery ratio of interpacket generation interval over 100 to 200 nodes.

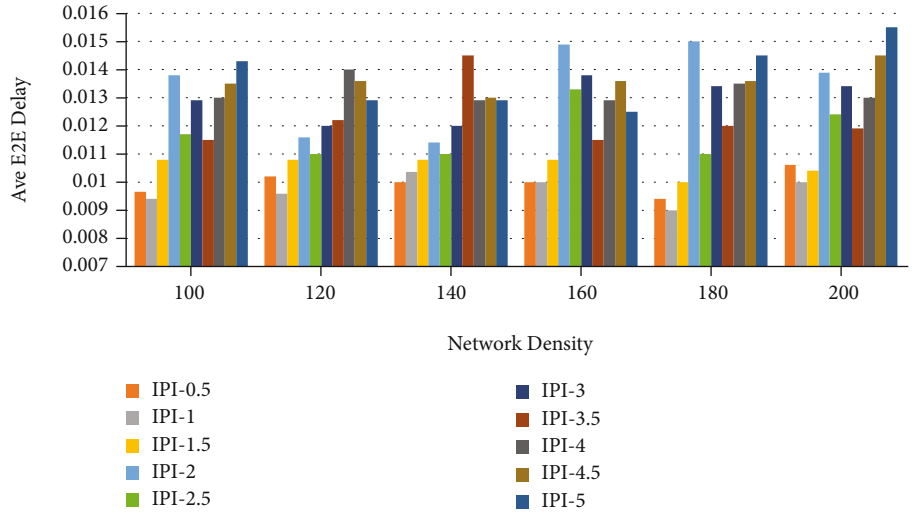


Figure 13: Average end-to-end delay over 100 to 200 nodes.

However, a slight decrease in PDR when payload increase from 64 bytes to 512 bytes suggests that other network factors other than payload negatively affect the network's performance. The graph of PDR versus node densities presented in Figure 15 shows the effect of the payload sizes on the network.

The average end-to-end delay on payload sizes seen in Figure 16 shows that larger payload sizes recorded higher delays than smaller payload sizes. 1024 bytes had the highest average delay of 19.405, while the payload size of 32 bytes had an average delay of 13.25 ms for packet sizes between 100 and 200 nodes. Interestingly, 64 and 128 bytes recorded the slightest delay for all network densities, with average delays of 10.93 ms and 10.88 ms. The increase in delay results from the rise in overall packets in the network per second. An increase in packets increases collisions, causing a doubling of the contention window after retransmissions. Larger contention window sizes increase the overall network end-to-end delay.

The graph trends recorded in Figure 17 confirm the energy consumption values recorded in [22, 23]. The authors asserted that the increase in payload size increases the energy consumption of the sensor network. Payload size of 1024 bytes experiences the highest energy consumption. On the other hand, the payload size of 32-byte recorded the highest remaining energy and the least remaining energy on the individual nodes. The highest remaining energy is recorded at 140 nodes and sees a slight decline when it reaches the lowest at the same 5.58 J at the payload of 1024-byte.

When the number of packets generated per second increases in the network, the number of collisions increases. After a saturation point of about 140 nodes transmitting per second, the contention window of the CSMA/CA reaches its

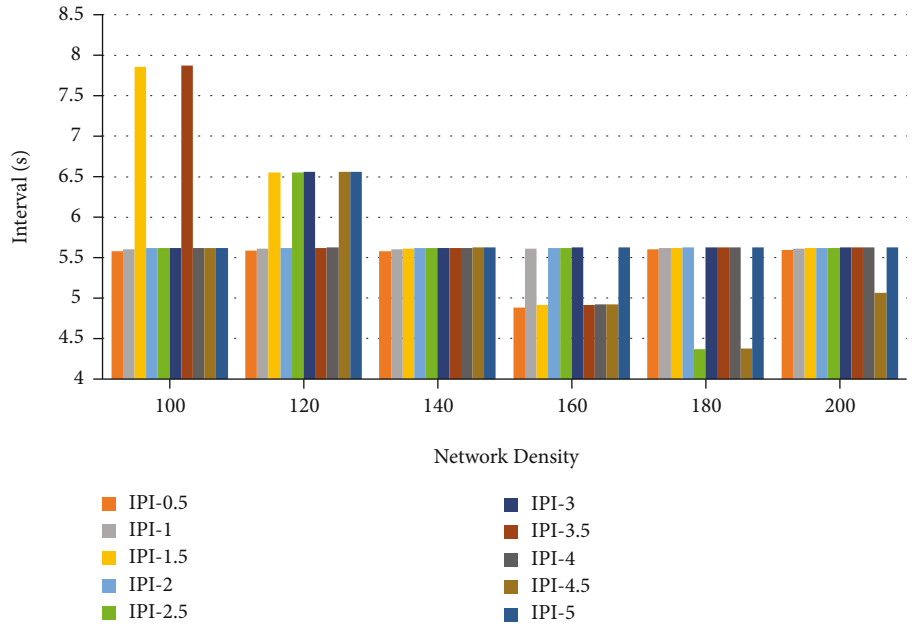


Figure 14: Average remaining energy on individual nodes in the network for different interpacket generation intervals.

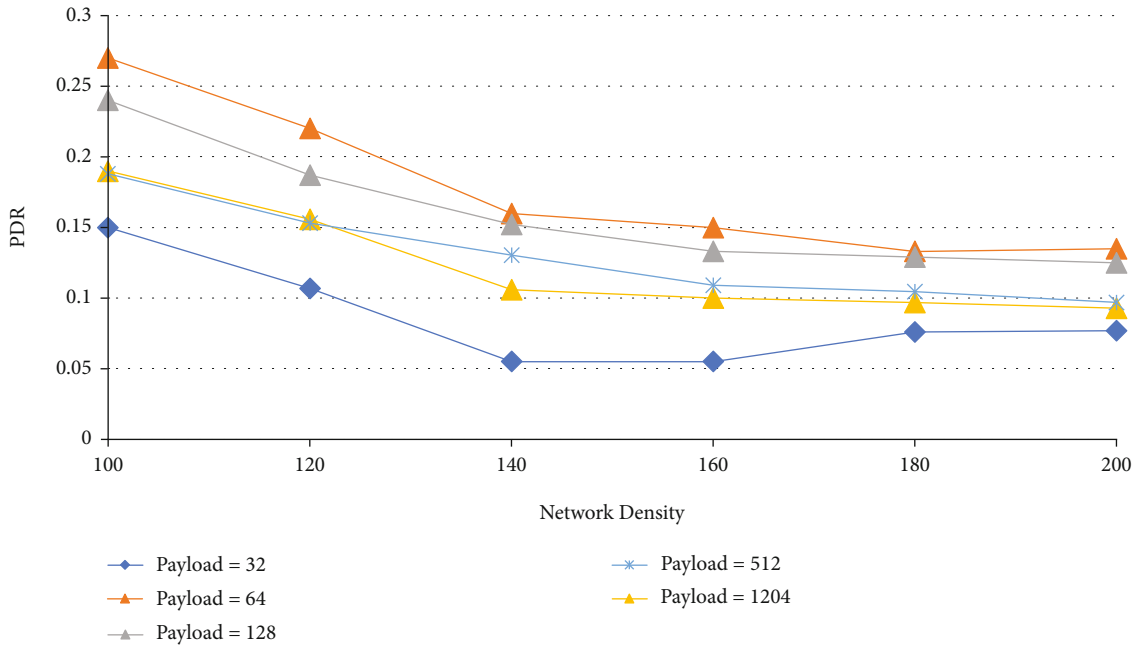


Figure 15: The effect of packet sizes on the packet delivery ratio.

maximum and, consequently, the maximum random backoff. After several retransmissions, the energy on a node in the network has less than their minimum threshold of energy. Further increase in packets per unit time does not increase the contention window; hence, delay remains relatively constant. Other factors like packet transmission mainly contribute to the further decrease in remaining energy. Observation of the graph for average end-to-end delay shows that after 180 nodes, the remaining energy increases slightly for all payloads and remains relatively constant for increasing network density.

As presented in Table 3, IoT and WSN applications that require continuous monitoring of the environment but are energy constrained may use these sensor node and channel parameters for packet scheduling. Applications such as animal tracking, water quality monitoring, and vehicular monitoring among others [24], precision agriculture and underground applications [25] for continuous monitoring

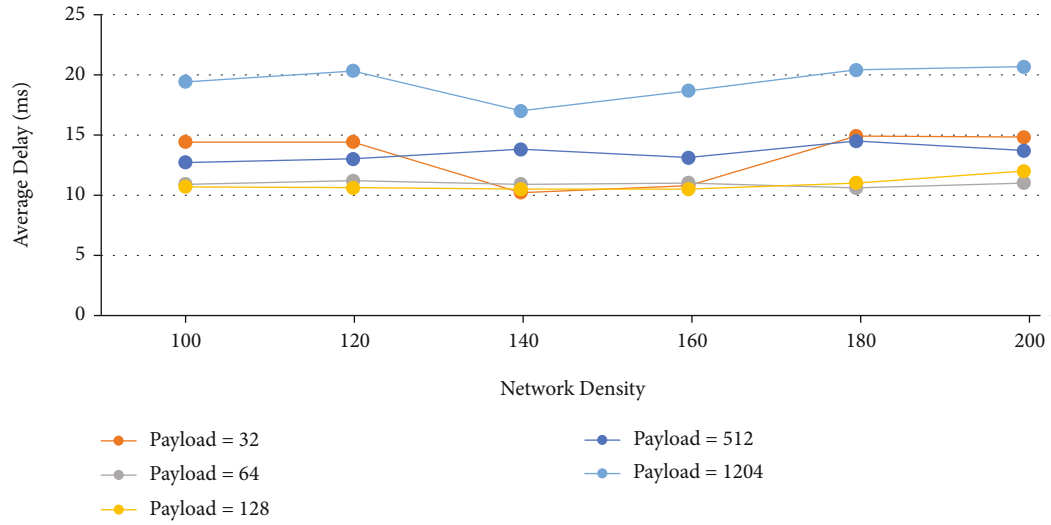


Figure 16: Average end-to-end delay of different payload sizes in IEEE802.11b.

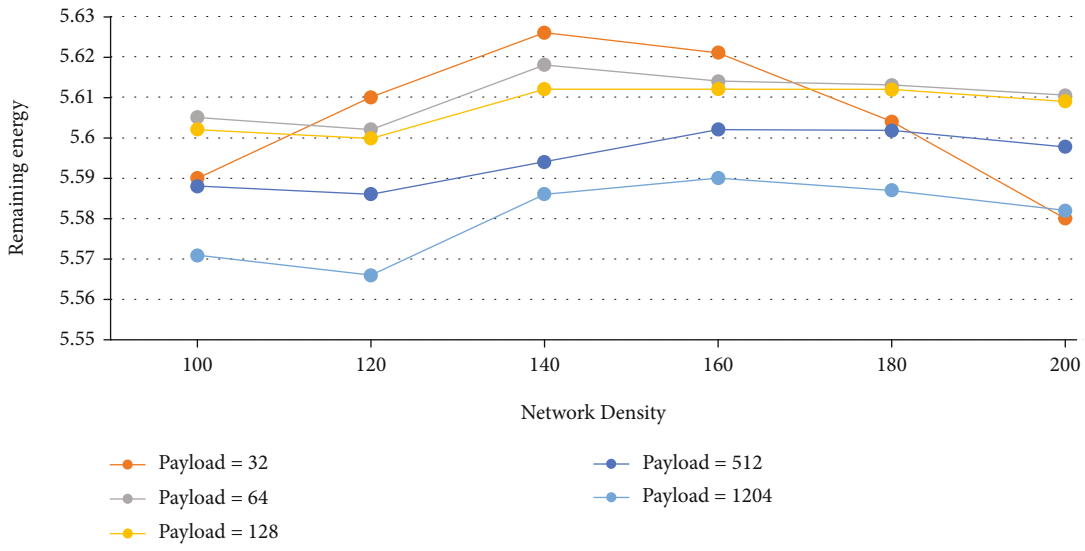


Figure 17: Average remaining energy for IEEE802.11b for different payload sizes.

Table 3: Optimal values for packet scheduling in IEEE802.11b.

Parameter	Optimal value
Packet size of data	64 bytes
Data rate	11 Mbps
Interpacket generation interval	4 seconds

have data streams with spatiotemporal properties that suggest that time series models for predicting future values are beneficial for energy savings and optimum network conditions.

The importance of transmission scheduling is seen in real-time applications such as water quality, structural, air quality, animal habitat monitoring, and tracking of endangered species. These applications, for continuous monitoring, have implementation challenges such as cost, energy efficiency, and communication issues. Water quality monitoring, with a real-time deployment discussed in the paper by [4], will receive continuous supply of the freshwater data without continuous transmission from the nodes. This work is similar to the implementation in Fiji [26], where nodes transmit continuously for an hour and sleep for 15 minutes. However, using the spatiotemporal difference between consecutive data streams, nodes here will transmit within a lesser period and, without turning nodes to sleep, achieve higher energy and network performance gains.

5. Conclusion

Scheduling of packets generated from sensor nodes influences transmission. In this paper, we performed sensitivity analysis of the IEEE 802.11b to determine the suitable parameters to adopt when used for IoT systems. From our simulation, we observed that nodes transmit packet immediately after packet generation without significant delay. The

effect of the scheduled packet transmission on the network was tested to determine the optimal threshold values in relation to (1) varying the inter-packet transmission interval, (2) varying the packet size, (3) and the data rate. Our findings revealed that for IoT systems implemented on IEEE802.11b, attention should be paid to the size of the payload and the data rate. These parameters have detrimental effect on the energy consumed by IoT devices and the node density. We can conveniently conclude that the optimal values obtained in Table 3 are suitable for IEEE802.11b channels where continuous monitoring with reduced data streams is essential for energy optimization.

Data Availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Conflicts of Interest

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

References

[1] W. Li, F. C. Delicato, and A. Y. Zomaya, "Adaptive energy-efficient scheduling for hierarchical wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 9, no. 3, pp. 1–34, 2013.

[2] S. Kumar and H. Kim, "Energy efficient scheduling in wireless sensor networks for periodic data gathering," *IEEE Access*, vol. 7, pp. 11410–11426, 2019.

[3] F. Engmann, K. S. Adu-Manu, J.-D. Abdulai, and F. A. Katsriku, "Applications of prediction approaches in wireless sensor networks," in *Wireless Sensor Networks-Design, Deployment and Applications*, IntechOpen, 2021.

[4] K. S. Adu-Manu, F. A. Katsriku, J.-D. Abdulai, and F. Engmann, "Smart river monitoring using wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8897126, 19 pages, 2020.

[5] F. Alfayez, M. Hammoudeh, and A. Abuarqoub, "A Survey on MAC Protocols for Duty-cycled Wireless Sensor Networks," *Procedia Computer Science*, vol. 73, pp. 482–489, 2015.

[6] S. Kumar and S. Chauhan, "A survey on scheduling algorithms for wireless sensor networks," *International Journal of Computers and Applications*, vol. 20, no. 5, pp. 7–13, 2011.

[7] R. C. Carrano, D. Passos, L. C. S. Magalhaes, and C. V. N. Albuquerque, "Survey and taxonomy of duty cycling mechanisms in wireless sensor networks," *IEEE Communication Surveys and Tutorials*, vol. 16, no. 1, pp. 181–194, 2014.

[8] F. Engmann, F. A. Katsriku, J.-D. Abdulai, K. S. Adu-Manu, and F. K. Banaseka, "Prolonging the lifetime of wireless sensor networks: a review of current techniques," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 8035065, 23 pages, 2018.

[9] C. Liu, K. Wu, and J. Pei, "An energy-efficient data collection framework for wireless sensor networks by exploiting spatio-temporal correlation," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 7, pp. 1010–1023, 2007.

[10] A. R. Kumar, "Smart network access for 802.11 based internet of things," in *2015 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, pp. 315–320, Kumaracoil, India, 2015.

[11] R. Gomathi and N. Mahendran, "An efficient data packet scheduling schemes in wireless sensor networks," in *2nd International Conference on Electronics and Communication Systems, ICECS 2015*, pp. 542–547, Coimbatore, India, 2015.

[12] J. A. Khan, H. K. Qureshi, and A. Iqbal, "Energy management in wireless sensor networks: a survey," *Computers and Electrical Engineering*, vol. 41, pp. 159–176, 2015.

[13] O. Khader, A. Willig, and A. Wolisz, "Distributed wakeup scheduling scheme for supporting periodic traffic in WSNs," in *2009 European Wireless Conference*, Aalborg, Denmark, 2009.

[14] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 3, pp. 493–506, 2004.

[15] B. N. Bhandari, R. V. R. Kumar, R. Banjari, and S. L. Maskara, "Sensitivity of the IEEE 802.11 b MAC protocol performance to the various protocol parameters," in *2004 International Conference on Communications, Circuits and Systems (IEEE Cat. No. 04EX914)*, Chengdu, China, 2004.

[16] D. Saha, M. R. Yousuf, and M. A. Matin, "Energy efficient scheduling algorithm for S-MAC protocol in wireless sensor network," *International Journal of Wireless & Mobile Networks*, vol. 3, no. 6, pp. 129–140, 2011.

[17] F. Royo, M. Lopez-Guerrero, T. Olivares, and L. Orozco-Barbosa, "A fast network configuration algorithm for TDMA wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, Article ID 513527, 10 pages, 2010.

[18] A. P. Jardosh, K. N. Ramachandran, K. C. Almeroth, and E. M. Belding-Royer, "Understanding congestion in IEEE 802.11 b wireless networks," in *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, Berkeley, CA, 2005.

[19] S. Kurt, H. U. Yildiz, M. Yigit, B. Tavli, and V. C. Gungor, "Packet size optimization in wireless sensor networks for smart grid applications," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2392–2401, 2016.

[20] M. Yigit, H. U. Yildiz, S. Kurt, B. Tavli, and V. C. Gungor, "A survey on packet size optimization for terrestrial, underwater, underground, and body area sensor networks," *International Journal of Communication Systems*, vol. 31, no. 11, article e3572, 2018.

[21] M. C. Vuran and I. F. Akyildiz, "Cross-layer packet size optimization for wireless terrestrial, underwater, and underground sensor networks," in *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, Phoenix, AZ, USA, 2008.

[22] H.-Y. Zhou, D.-Y. Luo, Y. Gao, and D.-C. Zuo, "Modeling of node energy consumption for wireless sensor networks," *Wireless Sensor Network*, vol. 3, no. 1, pp. 18–23, 2011.

[23] I. Koutsopoulos and M. Halkidi, "Distributed energy-efficient estimation in spatially correlated wireless sensor networks," *Computer Communications*, vol. 45, pp. 47–58, 2014.

[24] K. S. Adu-Manu, C. Tapparello, W. Heinzelman, F. A. Katsriku, and J. D. Abdulai, "Water quality monitoring using wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 13, no. 1, pp. 1–41, 2017.

[25] K. S. Adu-Manu, N. Adam, C. Tapparello, H. Ayatollahi, and W. Heinzelman, "Energy-harvesting wireless sensor networks (EH-WSNs)," *ACM Transactions on Sensor Networks*, vol. 14, no. 2, pp. 1–50, 2018.

[26] A. N. Prasad, K. A. Mamun, F. R. Islam, and H. Haqva, "Smart water quality monitoring system," in *2015 2nd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, Nadi, Fiji, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5572084, 20 pages
https://doi.org/10.1155/2021/5572084

Research Article

Presenting a Reliable Routing Approach in IoT Healthcare Using the Multiobjective-Based Multiagent Approach

Saeed Javid and A. Mirzaei

Department of Computer Engineering, Ardabil Branch, Islamic Azad University, Ardabil, Iran

Correspondence should be addressed to A. Mirzaei; a.mirzaei@iauardabil.ac.ir

Received 28 January 2021; Revised 28 July 2021; Accepted 31 August 2021; Published 29 September 2021

Academic Editor: Ihsan Ali

Developments in information and related technologies have led to a wider use of the Internet of things (IoT). By integrating both virtual and physical worlds, IoT creates an integrated communication framework of interrelated things and operating systems. With the advent of IoT systems based on digital remote care, transferring medical data is becoming a daily routine. Healthcare is one of the most popular IoT applications and tries to monitor patients' vital signs during the day for weeks and to eliminate the need for hospitalization. In a healthcare system, many sensors are installed to collect the patient's information, including environmental monitoring sensors and vital and unstructured message sensors in order to reduce the patients' expenses. The IoT network contains flexible sensors in dynamically changing environments where sensors collect environmental information and send it to nursing stations for healthcare applications. Due to the wireless nature of IoT networks, secure data transmission in the healthcare context is very important. Data collected from sensors embedded in healthcare devices may be lost for various reasons along the transmission path. Therefore, establishing a secure communication path in IoT networks in the context of healthcare is of great importance. In this paper, in order to provide a reliable data transfer protocol in the context of healthcare, a reliable routing using multiobjective genetic algorithm (RRMOGA) method is presented. The contribution of this paper can be summarized in two steps: (i) using a multiobjective optimization approach to find near-optimal paths and (ii) using reliable agents in the network to find backup paths. The simulation outcomes reveal that the proposed approach, based on the use of the multiobjective optimization approach, tries to find optimal paths for information transfer that improve the main parameters of the network. Also, the use of secure agents leads to a secure information transfer in the network in the context of healthcare. Experimental results show that the proposed method has achieved reliability and data delivery rates, 99% and 99.9%, respectively. The proposed method has improved network lifetime, delivery rate, and delay by 14%, 2%, and 5.6%, respectively.

1. Introduction

Recently, the advent of the Internet of things (IoT) has led to a paradigm effect on all regions of human-machine relations [1]. This novel technology has gained popularity in industries, healthcare systems, user interface development, and other areas [2]. IoT is the interconnection of devices, applications, sensors, and network connections that enhances these entities for data collection and data exchange. By integrating the physical and virtual worlds, IoT provides an integrated communication framework for interrelated devices and operating systems [3, 4].

Since IoT networks provide a lot of benefits for monitoring of patients in hospitals and healthcare centers, IoT healthcare applications have received more attention [5]. In the context of healthcare, IoT is one of the most fundamental aspects of the IoT application because it helps healthcare applications make full use of the IoT and cloud resources. The cloud environment provides standard protocols to support the connection between medical equipment and computational resource and transmission of medical data from embedded sensors on smart devices to a network of fog computing [6]. With the advent of IoT systems based on digital healthcare, medical data transmission is becoming

a common task. Therefore, developing an efficient method to support a reliable and secure routing and aggregation of patient biomedical data from the Internet of things is an essential challenge [7]. In healthcare programs, objects collect information about patients and send it to remote nursing stations using communication networks, especially the Internet. Analysis of information in nursing stations can lead to timely treatment for patients and can also prevent potential risks for patients [8]. Given that some patients may be in critical condition, the rapid and reliable transfer of data to the nursing station can be one of the main challenges in IoT-based healthcare systems [9, 10].

On the other hand, system failure can be another reason for the lack of transfer of data to nursing stations in IoT healthcare programs [11]. System failures may be due to hardware crashes, defective software performance, power leakage, or environmental hazards. In addition, as the number of nodes in an IoT system in healthcare applications increases, the likelihood of errors can also increase and disrupt system performance. Data loss due to system failure in care settings can have a negative effect on the use of healthcare programs and eventually can cause irreparable injuries in patients' bodies [12–14].

In this paper, in order to provide a reliable data transfer protocol in the context of healthcare, a reliable routing using multiobjective genetic algorithm (RRMOGA) method is presented. RRMOGA uses a combination of the multiobjective genetic algorithm to find the optimal paths between nodes in the IoT network and nursing stations and select reliable factors based on the weight of each path in the network to find support paths. Backup routes are introduced to transfer data in case of failure and failure to receive the original data from the optimal route. In the multiobjective genetic algorithm, in the proposed method, a multicriteria objective function is used to improve the main objectives in the network. The criteria related to the evaluation function in the proposed method include the distance between objects, the distance from the node to the destination, the quality of the link, and the degree of reliability. In fact, the node with the lowest hop count to the source node and the shortest straight distance to the access point, with the highest link quality and highest degree of reliability is selected as the optimal and most reliable node in each hop in the network. Also, the backup node is selected as the second optimal node in neighboring nodes with the values of evaluation criteria in the second category, as a reliable factor in the backup path. The selection of optimal nodes at each step leads to global optimization of reliable routing in healthcare systems on the IoT platform. The application of multiobjective genetic optimization algorithms has been proven in many fields of research, including the energy-efficient routing in wireless sensor networks [15], resource allocation in the cloud-fog-IoT infrastructure [16], service placement, and load distribution in edge computing [17]. In general, the contribution of this paper can be summarized in four steps:

(i) Simulating IoT networks according to standard parameters to extract used data in the proposed method

(ii) Using a multiobjective genetic algorithm optimization approach to find optimal paths

(iii) Using reliable agents in the network to find support paths

(iv) Evaluating the proposed method through known criteria and comparing it with the-of-are methods

In the proposed method, each device as a factor receives a certain amount of reliability based on the number of packets received and the number of packets sent. Getting dynamic network reliability is considered as one of the innovations of this paper. Also, at each step of transferring information from each node to find the next hop, a multiobjective genetic algorithm evaluates the distance between nodes relative to each other, the next hop distance to the destination, the delay between nodes, the quality of links between nodes, and the degree of the reliability value of the nodes in the next hop through fitness function. The use of the multiobjective evaluation function in a genetic algorithm is another innovation in this paper that combines the parameters of quality of service and reliability value. Finally, in order to send packets from the source node, based on the values of the evaluation function, the main path and an alternative path between the source node and the next hop are created. Then, the routing process in the main and alternative route is done using the multiobjective evaluation function. In the end, the two routes are returned as the main and the alternative route. The use of two paths that is the most optimal path in terms of evaluation function values as the main path and the other one as an alternative path has been proposed as another innovation in the proposed method in this paper. The rest of the paper is organized as follows:

In Section 2, the related work will be reviewed. Section 3 will describe the proposed method. Section 4 will provide simulations and evaluation of results. Section 5 concludes and discusses the article results.

2. Background

Patient-centered care services (PCCs) are an emerging healthcare model focused on patients' individual medical needs, originally developed by the Picker/Commonwealth program created by the Picker Institute in 1988. In fact, the node with the lowest hop count to the source node and the shortest straight distance to the access point, with the highest link quality and highest degree of reliability is selected as the optimal and most reliable node in each hop in the network.. They have the necessary training and support to make decisions and participate in self-care. While many initiatives have provided evidence of PCC success on a smaller scale, due to its conflict with the existing hospital-based model, the potential of PCC on a larger scale has not yet been realized in order to eliminate the need for hospitals or clinics. Conversely, the PCC uses these institutions in a common model of patient care using the Internet of things. The Internet of things is the convergence of telecommunications, sensors, actuators, cloud computing, and big data over the Internet to provide specific services [18].

The Internet of things can be customized to meet the challenges of modern healthcare. The healthcare system can be classified into three main areas: (a) large healthcare institutions (e.g., hospitals), (b) small clinics and pharmacies, and (c) nonclinical environments (e.g., homes care, communities, and rural areas without healthcare). To understand the role of the Internet of things in healthcare, firstly, the function of each region must be understood. Because the Internet of things has the ability to perform specific operations at a lower cost and more reliably, higher or more timely, it ultimately affects operations in other areas, resulting in a more stable and self-sustaining healthcare ecosystem [19].

Prior studies in the field of healthcare in the IoT indicate the importance of using applications and the use of the IoT in the field of health since it has an important role in promoting and developing health services and benefits [20–25]. The structure of the Internet of things and the healthcare monitoring systems are described in this section. An IoT-based healthcare monitoring system includes the four main elements of IoT medical equipment, information, and communication technology in the healthcare monitoring system, Internet services, and medical data management and processing [26, 27].

IoT applications in healthcare control systems refer to the use of information and digital communication technology, such as computers and mobile devices, for health management. An Internet application of objects in a healthcare system is also called an electronic health system or mobile healthcare monitoring system (MHCMS) and includes various health services [28].

IoT applications in medical fields such as intelligent sensor control, patient data transfer from a remote point to a clinic or hospital, integration of medical devices, and the possibility of data exchange between them, improves medical experiments in providing care. It also promote interaction between physicians about the effect of the drug, management and controlling various connecting devices, the possibility of transmitting IoT information medically by physicians, accurate diagnosis of other health problems and control patterns (heart rate, temperature, blood pressure, blood sugar levels in the body and gastrointestinal tract), the possibility of transmission and the information used by the physician to process and perform the appropriate medical activity. Devices connected to you are vital throughout the day, transmitting wirelessly to medical devices such as computers and smartphones. Overall, there is a tremendous potential for Internet-connected healthcare control systems and smart medical equipment for the well-being of the people. IoT healthcare systems have been used in various cases such as reducing preparation time for work in emergency rooms [29], personal health records [30], remote monitoring of patient health [31], and communicate specialists [32].

2.1. Related Works. Due to the widespread use of healthcare and monitoring applications in the context of IoT networks, the importance of secure data transmission in the network has been increasingly considered. Therefore, many researchers have tried to provide a reliable routing protocol in IoT networks for healthcare applications, which we will mention in the article.

Woo et al. have developed a secure oneM2M-based IoT network for individual hygiene equipment. In order to have a dedicated application sensor, in this paper, the protocol between ISO/IEEE 11073 protocol messages has converted to oneM2M messages at the access point embedded in the devices and the management server. The oneM2M-based IoT system is built for personal hygiene devices and has been evaluated in various experiments [11]. The result of experiments has shown that the oneM2M-based IoT system transmits secure messages but the disadvantage of this method is the lack of attention to other factors of service quality. Critical messages may be transmitted over the network that needs to be sent immediately, but such situations have not been considered in the oneM2M-based IoT system.

Soufiene et al. have categorized healthcare information in the IoT into two categories: priority emergency and crucial health signals. Emergency signals have the highest priority among signals and should be sent securely as soon as possible. Crucial health signals are data that physicians request to continue monitoring patients. This paper uses direct transmission to forward important data using multistep communication to deliver crucial health information [33]. The main idea of this method is the rapid transmission of critical data in the IoT network, while the loss of these critical messages can cause great damage. The main disadvantage of this method is that it does not consider the amount of reliability to the devices in the multistep path specified in the network to send critical data.

The Sarwesh et al. node location placement method and reliable routing mechanism in the environment of the IoT network have been extended. In this method, sensor nodes and relay nodes have been utilized. Sensor nodes have been defined as responsible for gathering signals, and relay nodes have been defined as responsible for data combination and path reliability control. In the node placement method, the density of the intermediate nodes varies due to the traffic region, to avoid the problem of energy holes. In the routing method, to decrease the number of retransmissions, the route is calculated efficiently and reliably [34]. The main disadvantage of this method is that it does not use a backup path to send packets. If something happens on the main route and the data is not sent, the delay in sending the message to resend the data will be very high.

Gochhayat et al. have proposed a new distributed key administration structure for the IoT platform. The method presented in this paper resourcefully provides IoT systems by assigning more cryptographic processing resources to local agents. These agents are coordinated with other peers to create a distributed key as well as an authentication mechanism for network things. Specifically, the presented method utilizes the benefits of mobile agents and deploys them in diverse subnets if necessary for the authentication process [35]. The disadvantage of this method is the lack of reliability evaluation for the distributed key controller agents.

Conti et al. have developed the REMI method, a trustworthy multihop routing method for IoT networks. The core

purpose of this article is to facilitate effective transmission in power efficiency networks like the Internet of things, ensuring that messages are received from arbitrary source nodes, regardless of network dimensions and the existence of defective nodes. REMI uses a multistep clustering method that speeds up the transmission of messages across the network [36]. The disadvantage of this method is energy-based routing in the IoT network, while other parameters of service quality and, most importantly, reliability do not play a role in the next hop selection.

Lyu and Yi conducted an in-depth study on IoT communication and network trustworthiness in a complex framework. In this article, an artificial bee colony optimization algorithm has been used to obtain the near optimal path from each node to the nearest head cluster. The implementation of this method in the IoT framework shows that the presented method can efficiently decrease the amount of data forwarded to the base station by sensor nodes through cluster head union. Data collection efficiency, energy balance, and reliability as well as network lifetime are also improved. [37]. The disadvantage of this method is the use of a single objective fitness function based on reliability in the bee colony optimization algorithm, and other service quality parameters are not considered as the goals of the fitness function.

In 2019, Asghari et al. presented a medical monitoring plan for the cloud-based IoT platform in which patients' medical conditions are predicted by extracting their physiological data collected from IoT devices, and then, other medical results are obtained. The disease diagnosis model is used to analyze patients' medical data to provide a hygienic/medical prescription. After the results are confirmed by the medical team, it is sent to the patient. The patient then identifies their nonsurgical needs such as location, cost, and time to find the most appropriate combination of health/medical services based on their preferences. Experimental results show that the proposed scheme for achieving effective disease diagnosis for combined health/medical prescriptions is successful [38]. This method greedily tries to send data to patients in need of surgery, while greedy selection may lead to local optimization and routing may not be as efficient.

In 2020, Asghari et al. proposed a privacy-aware cloud service combination approach to optimize service quality in the IoT environment by providing a hybrid IoT-based cloud service concept model on the privacy level calculation model and an algorithm. They have proposed a new hybrid evolution using the shuffled frog leaping algorithm (SFLA) genetic algorithm (GA), known as SFLA-GA. The proposed algorithm is used to optimize the composition of the proposed services in terms of the accumulation of different quality factors as the value of appropriateness. Also, to help users choose the right combination service, they are categorized in terms of the level of privacy that is achieved through a computational model. The simulation results showed that the proposed method improves the fitness compared to other contemporary algorithms [39]. Using an optimal path and not considering the backup path for emergencies can be considered a disadvantage of this method.

In 2020, Asghari et al. critical data is collected through IoT monitoring objects, and then, data analysis has been done through various machine learning methods such as the decision tree (J48), sequential minimum optimization (SMO), multilayer perceptron (MLP), and Naïve Bayesian (NB). Classification is required to identify the level of potential risks of physiological and behavioral changes in the elderly. Experimental results confirm that the SMO, MLP, and NB classifiers perform almost closely in terms of accuracy, precision, sensitivity, and F-measure [40]. This method tries to prioritize the data sensed in the IoT network and does not pay attention to the quality of service parameters.

3. The Proposed Method

In this paper, in order to provide a reliable data transfer protocol in the context of healthcare, a multiobjective method is presented. Given that the selection of reliable agents among the devices in the IoT network in order to meet the quality of service criteria is a complex issue, in this paper, the multiobjective genetic optimization algorithm is used. The multiobjective genetic optimization algorithm has tried to find the optimal path by creating a tradeoff between the quality of service criteria. The witness of the optimality of the selected algorithm in the proposed method is acquired results and comparison with other works in terms of quality of service criteria and improvements.. The proposed method uses a multistep information transfer approach to send information to access points. Access points in the proposed network are connected to nursing stations to send information in order to be analyzed by physicians and specialists. The proposed method uses a multiobjective genetic algorithm to find the optimal nodes in each step of information transfer. The proposed method tries to determine the optimal path for packet transmission based on the selection of reliable agents using the multiobjective genetic optimization algorithm. The proposed method, also, uses the backup path to deal with possible failures in the optimal path. The backup path is in the second optimal solution in the repository of the multiobjective genetic optimization algorithm which can be considered as innovation and significance of the proposed method. Since the fitness function used in the proposed method uses the criteria of the distance between nodes, distance to access point, link quality, and reliability degree, the optimal nodes are selected as reliable factors in the proposed method. In order to ensure the reliability of the proposed method, in the first step, two reliable nodes are selected and the path from the optimal node is selected as the main path and the second optimal node is selected as the backup path. The prerequisites of the proposed method will be described below. The overall architecture of the proposed method is shown in Figure 1.

3.1. Proposed IoT Network Model. As mentioned in the proposed method, to find a secure path in the context of IoT networks, four parameters are discussed: distance between nodes, distance between intermediate nodes to the base station node, link quality, and degree of reliability. These parameters play a decisive role in the process of selecting the next hop nodes, and improving these parameters will improve other goals in the network. The proposed method

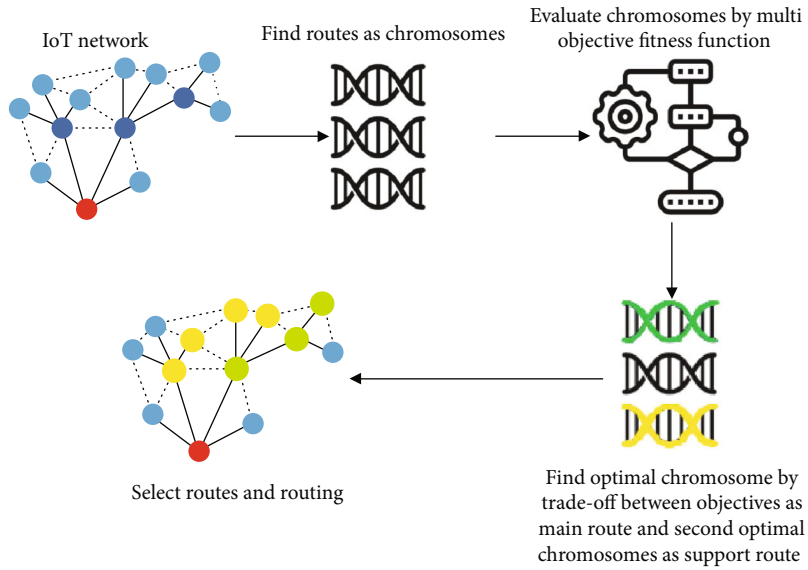


Figure 1: Overall architecture of the proposed method.

tries to find a reliable path in the IoT network by optimizing these parameters in the form of a multiobjective fitness function, and then, the parameters used in the proposed method are modeled. A flowchart of the proposed method is presented in Figure 2.

3.1.1. Distance between Nodes $D(K, K+1)$. In the IoT network, intermediate nodes must send sensed data from the environment to base stations. When the source node, the node that currently holds the packets, sends the data, it tries to send the packets to the nearest node in its neighborhood. Sending packets in the shortest distance between nodes not only reduces the delay of data transfer between the source and the base station but also increases the reliability of the route. The distance model between nodes is expressed in equation (1) as follows:

$$D(K, K+1) = \min\left(\sum_{K=1}^{N} \left(\sqrt{x_K - x_{K+1}}\right)^2\right), \quad (1)$$

where $D(K, K+1)$ is the distance between the current nodes and the nearest node in the next hop, K is the number of neighboring nodes of the current node, N is the total number of nodes in the IoT network, x_K is the current node position, and x_{K+1} is the position of the next node.

3.1.2. Distance of the Next Node to the Base Station $D(K, BS)$. In IoT networks, the equipment can act as a relay between source and destination nodes. In order to select the optimal route between the source node and the base station, the nearest route can be less general and more reliable. The distance from the next node to the base station can determine the convergence of the path to the base station. In fact, if the distance from the next node to the base station is less than the distance from the current node to the base station, then, the selected path converges in the direction of the base station, and then, the current node from its neighbors is the node with the shortest distance to the base station. Now, we can select a safe path. The model of the distance between the next node and the base station is expressed as equation (2) as follows:

$$D(K+1, BS) = \min\left(\sum_{K=1}^{N} \left(\sqrt{x_{K+1} - x_{BS}}\right)^2\right), \quad (2)$$

where BS represents the base station and x_{BS}represents the base station position.

3.1.3. Link Quality. In the proposed method, the link quality is considered as a criterion for estimating the number of steps required to transfer data to the base station. The link quality is used with a metric called ETX (expected transfer) to indicate the estimated number of steps between a node and the base station at the time of data aggregation. In other words, for a clustered node, ETX is the estimated total cost of collecting data from the cluster member nodes that belong to it and transferring of the aggregated data to the hole in several steps. Obviously, the lower the ETX for a node, the fewer the steps required to send data and the better the link quality. Link quality modeling is shown in equation (3) as follows:

$$ETX(K+1, BS) = \min \sum_{i=1}^{K} \sum_{j=1}^{BS} H(S, K) + \left(\frac{D(K+1, BS)}{R} + 1\right), \quad (3)$$

where $H(S, K)$ is the number of goats passed from the source node to the current node and R represents the range of nodes in the IoT network.

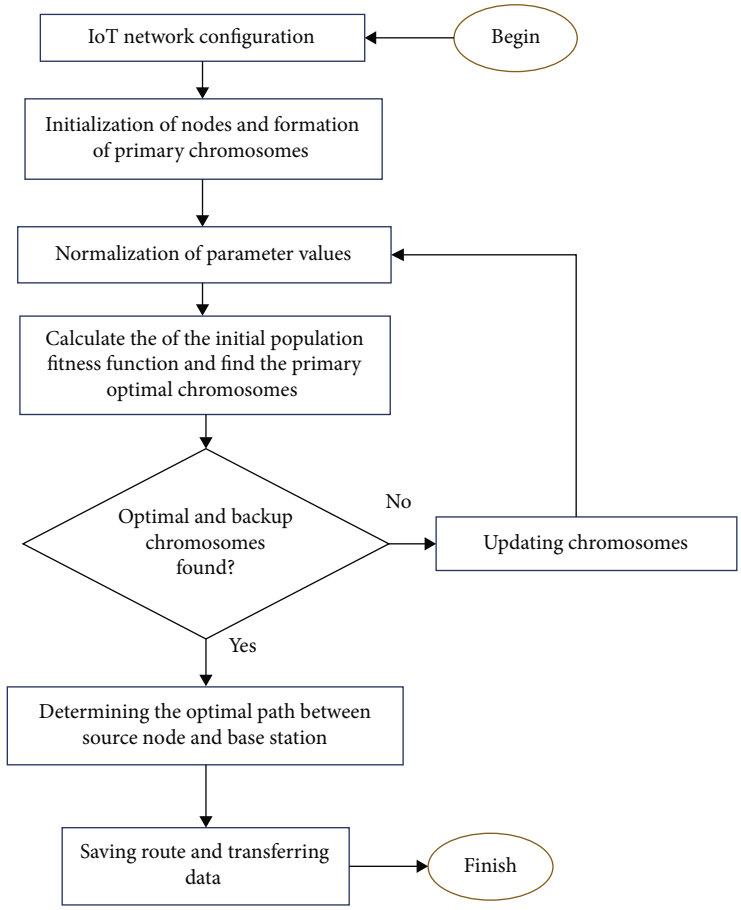


Figure 2: Flowchart of the proposed method.

3.1.4. The Degree of Reliability of the Route. In the proposed method, in order to find a safe path for transferring information between the destination node and the base station, the degree of reliability of the nodes in each step is checked. In order to transmit packets to neighboring nodes, the current node must select the node with the highest degree of security in the neighborhood to create a secure path between the destination node and the base station. The degree of reliability in the proposed method is quantified by equation (4) as follows:

$$P(K, K+1) = \min\left(\sum_{i=1}^{M}\left(1 - \frac{\mathrm{PRR}_{(K,K+1)}}{\mathrm{PDP}_{(K,K+1)}}\right)\right), \quad (4)$$

where $\mathrm{PRR}_{(K,K+1)}$ is the rate of receiving packets between the current node and the neighboring node and $\mathrm{PDP}_{(K,K+1)}$ is the rate of sending packets between the current node and the neighboring node.

The proposed method, based on the mentioned parameters, selects the optimal node in the neighborhood and the IoT network.

3.2. Multiobjective Genetic Algorithm. The genetic algorithm begins with the random production of the initial population as some random solutions to the problem, and in an iterative process, these solutions evolve and approach the optimal solutions. Thus, genetic algorithms are greatly influenced by the operators that create the newly evolved population. A genetic algorithm is a random search technique that increases the likelihood of selecting optimal solutions. This algorithm includes a population of solutions which are considered as chromosomes. Each solution, to be presented as a chromosome, requires a coding method in which genes are formed based on the application of the problem and then together form chromosomes [41].

Each iteration in a genetic algorithm involves many operators that begin by assessing the fitness of chromosomes. Subsequent operators that are important include crossover, mutation, and selection. The crossover operator enables the genetic algorithm to generate a new generation with solutions resulting from the interconnection of chromosomes from two selected parents, resulting in the production of one or more offspring through the integration of specific gene houses from the parents. The crossover operator selects the cut and merges the location of the two chromosomes based on the probability of P_cross. Equation (1) shows the selection of the crossover point on the chromosomes [41].

$$\mathrm{Crossover} = (P\_\mathrm{cross} * (G_{\max} - G_{\min})), \quad \forall i = 1, 2, \cdots, M-1, \quad (5)$$

where crossover is the chromosome configuration, the $G_{\max}$ parameter is the maximum number of genes per chromosome, the $G_{\min}$ parameter is the minimum number of genes per chromosome, and $P_$cross is a random number in the range (0,1). The fitting operator is used to diversify the initial population. The mutation operator sporadically updates some of the genes on the chromosome to increase chromosome fitness.

The mutation operator selects the location of the gene on the chromosome based on the probability of $P_$mutate. Equation (2) shows the choice of gene location on the chromosome [41].

$$\text{Mutation} = (P\_mutate * (G_{\max} - G_{\min})), \quad R \text{ is a rand in } [0, 1], \tag{6}$$

where the mutation parameter indicates the gene to be mutated and $P_$mutate is a random value in the range [0,1].

The selection operator also transmits the best chromosomes from the current population to the next generation. In genetic algorithms, a number of possible solutions are first randomly generated as the initial population. In chromosome coding and gene representation, selection plays an important role in the production of the initial population. In the next step, the degree of competency of each initial population is measured based on the evaluation function. The evaluation function is a measure of the evolution of existing chromosomes. Then, a number of solutions are selected as the parent based on the degree of competency and create new children as new solutions. New children are transferred to a new population and create a new generation, and the same process is repeated for the new generation. The iteration of the genetic algorithm continues until the termination condition is met.

3.3. Initial Population Coding. The first step in using a genetic algorithm is to encode and generate a random primary population. Coding refers to how genes are quantified on chromosomes, where the arrangement of genes together will lead to a solution for the routing problem. In genetic algorithm-based routing methods, chromosomes include nodes that exist between the source and destination nodes. These nodes are selected as relays in the path between the source and destination nodes, which is considered as a solution. Table 1 shows an overview of coding based on intermediate relay nodes.

As shown in Figure 1, each of the genes in the chromosomes in the proposed method represents an intermediate node between the source node and the destination node. The number of genes on all chromosomes must be the same. Therefore, the number of genes must be part of the number of nodes in the network. In the proposed method, the number of available genes is obtained from equation (7) as follows:

$$N_M = \frac{x_S - x_{BS}}{\sum_{i=1}^{N} (R_i/N)} + 1. \tag{7}$$

Table 1: How to code primary chromosomes based on intermediate nodes.

N_1	N_2	N_3	N_4	N_5	N_6	…	N_M
10	14	21	11	8	7	…	4

According to equation (7), it can be said that the number of genes in the proposed method is determined by dividing the direct distance between the source sphere and the base station by the average range of nodes in the network. Since the existing IoT network nodes are heterogeneous, the communication range of the nodes may be varied. Therefore, in the proposed method, the average communication range of all nodes in the network is used.

3.4. Proposed Fitness Function. In order to formally define the evaluation function in the proposed method, we assume that the current node is represented by index i and the neighboring nodes are represented by index j. For this purpose, considering the modeling of optimization methods for multiobjective genetic algorithms, we consider the following limitations in the proposed method.

(i) The sum of the distances between nodes in the optimal path should not be more than a fixed value. This limit is set so that the number of steps in the path is not large. If the threshold of distances between nodes is too large, all the nodes in the network may be in the same path, which greatly affects the performance of the proposed method

(ii) The number of steps of sending data between nodes in the optimal path should not exceed the quality of the estimated link. This constraint prevents looping and adding additional nodes in the path and ensures that the proposed method finds the shortest path in the network to transfer data from the destination node to the base station

(iii) The initial degree of reliability of the nodes is equal to a fixed value and not more than that. Reliability is updated as data transfer between nodes increases. Given that in the proposed method, the main focus is on increasing the degree of reliability between nodes in the IoT, nodes must have a constant and uniform initial degree of reliability in order to ensure fairness in data transmission in the proposed methods

(iv) The total delay of packet transfer on the route should not be more than a fixed value. Critical messages on the network must reach the base station within the timeframe; otherwise, they may disrupt the intended applications in the IoT

Since the goals in the network may be conflicting and the improvement of one goal reduces the optimization of the other one, the usual criteria for IoT are insufficient. Thus, multiobjective criteria are considered to find the right path

from the source node to the base station. The multiobjective genetic algorithm optimization method tries to create a balance between the goals in the network that may be contradictory or compatible, and more importantly, the optimality of all goals is considered. Therefore, the mentioned parameters are considered for the multiobjective function in order to find the most optimal path between the source node and the base station. Finally, the summative evaluation function is shown in equation (8) as follows:

$$\begin{aligned} F = \text{Min} \sum_{i}^{n}\sum_{j}^{k} & D(K+1, BS) - D(K+1, BS) \\ & - ETX(K+1, BS) + P(K, K+1) \end{aligned}$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{j=1}^{k} D(K+1, BS) \geq \alpha \\ & \sum_{i=1}^{n} D(K+1, BS) \leq \beta \\ & \sum_{i=1}^{k} ETX(K+1, BS) \leq \gamma \\ & \sum_{i=1}^{k} ETX_i \leq \delta \\ & \sum_{i=1}^{k} \text{delay}(S, BS) \leq \omega. \end{aligned} \tag{8}$$

In equation (8), i is the number of nodes, j is the number of neighboring nodes, α is the upper limit of the distance between nodes, β is the upper limit of the distances between nodes and the base station, γ is the total initial confidence of the nodes, δ is the maximum number of steps in the optimal path, and ω is the maximum delay in the IoT network. According to equation (8), in each round of the proposed multiobjective genetic algorithm in the path selection step, the node that minimizes the value of the F function is selected as the first optimal neighbor node. Also, the node that fits in the second place is selected as the optimal node in the backup path. The data transfer process is then sent from the current node to the base station. If an error occurs in the optimal path and a node leaves the transfer process, the data transfer will start from the backup path. The selected path based on balancing the goals of the IoT network is the shortest path with the shortest distance between nodes which will cause the least amount of delay.

On the other hand, due to the fact that the values of the parameters used in the fitness function do not match and the scale of these parameters is different from each other, combining these parameters in the form of a function can result in incorrect results. In other words, since the values of the parameters are in different ranges with different scales, the parameter that has a higher value can have more parameters than the values with lower values. Therefore, it is necessary to convert the parameters and map the values related to the parameters in a certain range, which is called normalization. Normalization maps values related to different parameters in the range of [0,1] so that the parameters do not have a negative effect on the results. The problems related to the scale of these parameters are eliminated. In fact, after normalization, the values of all parameters are relatively between zero and one and various units for different parameters will not affect the results of weight determination for chromosomes. Various normalization methods have been introduced in journals—the most famous of which is the Z -score normalization [33]. In the proposed method, Gaussian normalization has been used in order to eliminate the negative effects of parametric values on the scale. The Gaussian normalization is introduced in equation (9) as follows:

$$\text{Normalized} - \text{data} = \frac{\text{data} - \text{mean}(\text{data})}{\text{std}(\text{data})}. \tag{9}$$

In equation (9), normalized data of parameter values after normalization process, data of parameter values before normalization, mean (data) mean of values of each parameter, and std (data) of the standard deviation of values of each parameter. The output from the normalization phase will be used as the input of a multiobjective genetic algorithm.

4. Implementing the Proposed Method

The data used in this paper are obtained from IoT network simulation in MATLAB software. This article does not use standard datasets stored in data warehouses. Instead, a standard network with 100 nodes is implemented using standard network parameters. The implemented nodes are randomly distributed in the network space. These nodes have the ability to receive and send data. Each node has a defined communication range in which it can communicate with other nodes. By transferring information in the network, the values of the variables used in the proposed method are extracted from the network and used to apply the proposed model.

We start the implementation of the proposed method with the initial configuration of nodes in the IoT network and the distribution of nodes in the monitored environment in MATLAB software version 2019. We consider the monitored environment to be 100×100 space in which 100 sensor nodes are randomly scattered. The number of sensors can be adjusted according to different scenarios, and this number can be changed in order to compare the proposed method with other methods available in publications. Other parameters related to network configuration are considered according to the standards mentioned in the publications. The proposed network is implemented in an environment of 100∗100. The proposed IoT network parameters are shown in Table 2. Figure 3 also shows the initial configuration of the wireless sensor network based on the proposed scenario.

As shown in Table 1, the IoT network is formed according to the random distribution of nodes and based on the values of the parameters in Table 2. A base station has been installed in the monitored area. In the first step of the hole

Table 2: Initial parameters of the proposed sensor network.

Parameter	Value
Network dimensions	100∗100
Number of nodes	100
Base station coordinates	(100,100)
The initial degree of reliability of the nodes	0
The initial probability of selecting a neighbor node	0.01
Maximum number of rounds	100
Data package length	40
Number of packages sent per hop	10
Length of the routing package	100
Radiation range	5

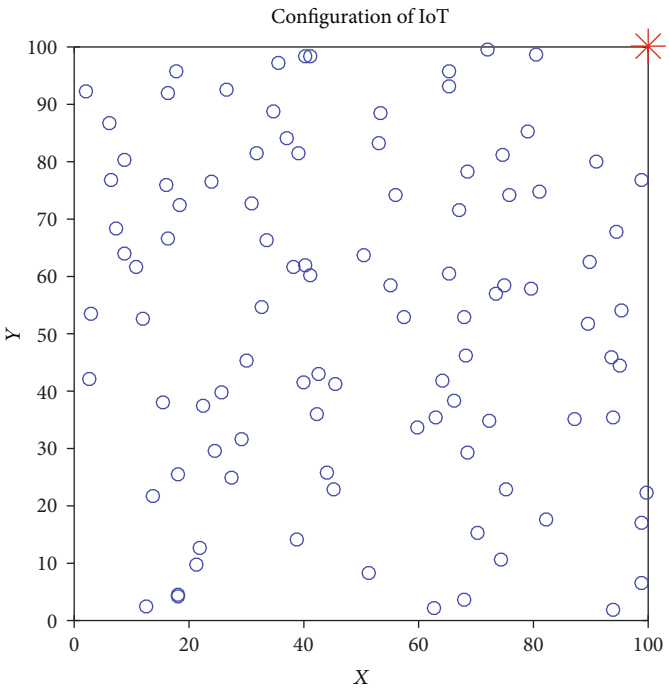


Figure 3: Initial configuration of the proposed IoT network.

node simulation, by sending a routing package in the form of a "Hello" message, it tries to obtain information about the Hasim node in the network. Each node that receives this message immediately sends RREP routing response packets to begin the process of route selection and network security.

4.1. Primary Chromosome Quantification. As mentioned in previous chapters, the proposed method uses a multiobjective genetic algorithm to route in the IoT network. One of the differences between the proposed method and the simple genetic algorithm is the use of a multiobjective fitness function. In order to implement the proposed method in the proposed scenario, the initial solutions must first be quantified. Since the solutions of the proposed method are applied as chromosomes to the proposed genetic algorithm, they must be encoded in the standard chromosome defined for the proposed method. The chromosomes of the proposed method, as discussed in Section 3, are arranged based on the number of nodes in the path. Thus, in Table 3, a part of the initial population is shown in the proposed method.

As shown in Table 3, the initial population is adjusted according to the number of optimal path steps based on equation (7). Therefore, in the next step of implementing the proposed method, we will implement the proposed genetic algorithm on the initial population.

4.2. Implementation of the Proposed Genetic Algorithm. To implement the proposed method, in this section, we first evaluate the initial population. The initial population, as shown in Table 3, consists of a number of chromosomes whose size is predicted to be equal to the number of steps. Each of these genes takes on a value that represents the node in the optimal path. According to the objectives of this paper, the proposed multiobjective genetic algorithm uses a multiobjective fitness function to evaluate the initial population. In addition to the distance between nodes and the distance between nodes with the base station, the proposed proportionality function also examines the link quality and the degree of path reliability. Table 4 shows the values of general fitness and partial factors for the initial population.

As shown in Table 4, for each chromosome, the values of the overall fitness function and the fitness function are calculated based on multiple evaluation factors. Table 4 shows only 5 chromosomes from the original population, and to show the values of the other chromosomes in the original population, we use the distribution graph of the values of the evaluation function in the search space. Figure 4 shows the distribution of the values of the overall proportionality function of the initial population, which is arranged in descending order.

As shown in Figure 3, the values of the overall fitness function are calculated for the initial population. The overall fitness function represents the overall quality of a chromosome and examines the proposed routing from all aspects of the network. According to Table 1, it can be seen that some of the randomly selected primary solutions have a good proportion value but some other chromosomes in the initial population do not have a good proportion value and should be removed from the initial population. The figure shows that the most optimal chromosome has a proportionality value of about 0.0308. In the next step, the selected chromosomes are transferred to the fitting operator. In the fitting operator, a shear point is used and the fitness values are checked for each of the fitted populations and the population with the best fitness value is transferred to the jump operator as the output population. In the proposed method, we use this operator for creating an expert population among the new population. In fact, in the proposed method, the new population obtained from the fitting operator will have near-optimal proportion values and chromosomes that have nonoptimal proportion function values will not enter this population. In addition to generating new populations and diversifying the generation of previous chromosomes, this operator accelerates the convergence of new populations toward optimal points in terms of research objectives.

Table 3: Sample of the initial population.

	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9	G_{10}	G_{11}	G_{12}	G_{13}	G_{14}	G_{15}
CH1	4	3	17	15	17	8	20	7	14	10	9	1	17	22	24
CH2	1	4	6	3	2	8	13	15	9	23	11	14	12	23	18
CH3	3	1	2	10	11	3	15	18	7	6	13	22	2	10	11
CH4	14	1	3	11	7	2	5	21	6	8	14	20	9	11	10
CH5	18	3	14	10	25	4	9	2	20	12	18	5	11	10	22

Table 4: The values of the initial population fitness function.

Chromosome number	Value of overall fitness function	Value of fitness function for node intradistance	Value of fitness function for distance between nodes and base station	Value of fitness function for link quality	Value of fitness function for reliability
1	0.0316	0.0164	0.0111	0.0076	0.4430
2	0.1189	0.0659	0.0652	0.0659	0.4431
3	0.0551	0.0275	0.0185	0.0183	0.4430
4	0.0217	0.0330	0.0222	0.0282	0.4431
5	0.1770	0.0275	0.0185	0.0127	0.4431

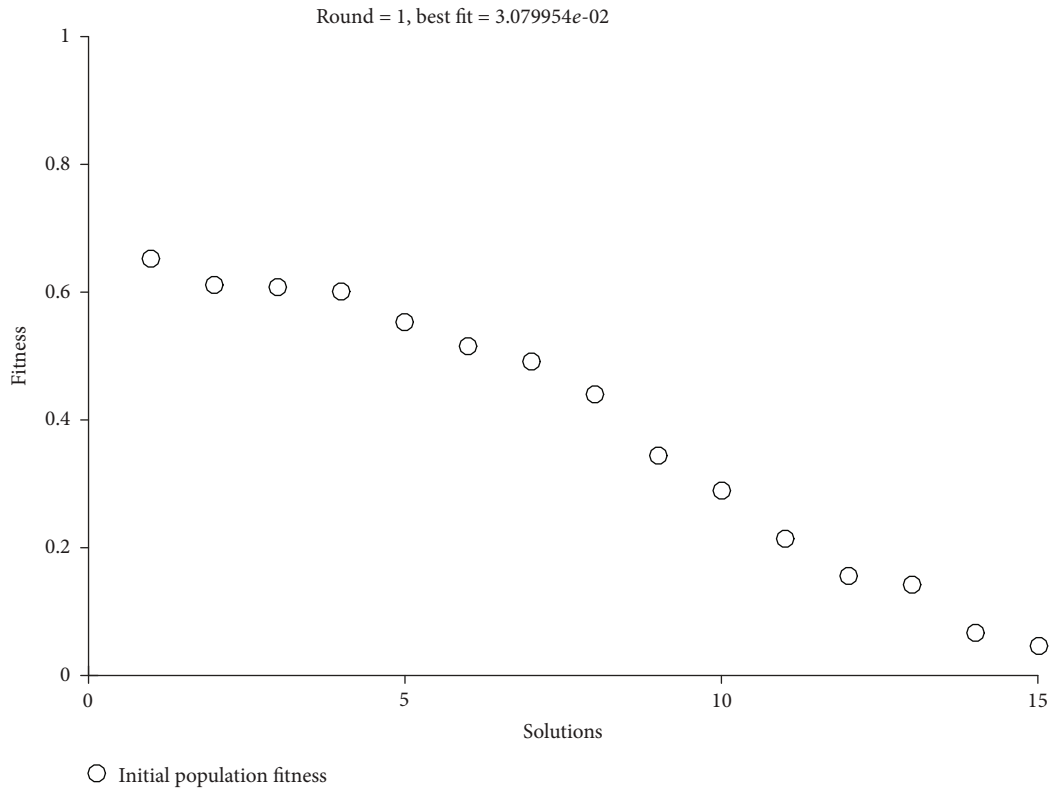


Figure 4: Values of the overall fitness function of the initial population.

After fitting the initial population and producing a new population close to optimal, in the next step, the jump operator is applied to the new population and slightly increases the diversity of the new population compared to the previous population. Due to the fact that the proposed method uses the adaptive fitness operator, variations in the fitted population may lead to inverse results in the values of the fitness function. Figure 5 shows a graph of the dispersion of the fitness function in the initial population and the new population resulting from crossover and mutation.

As shown in Figure 5, the values of the overall fitness function of the solutions are optimized for fitted chromosomes from the original population. Given that the proposed proportionality function is introduced as a minimization

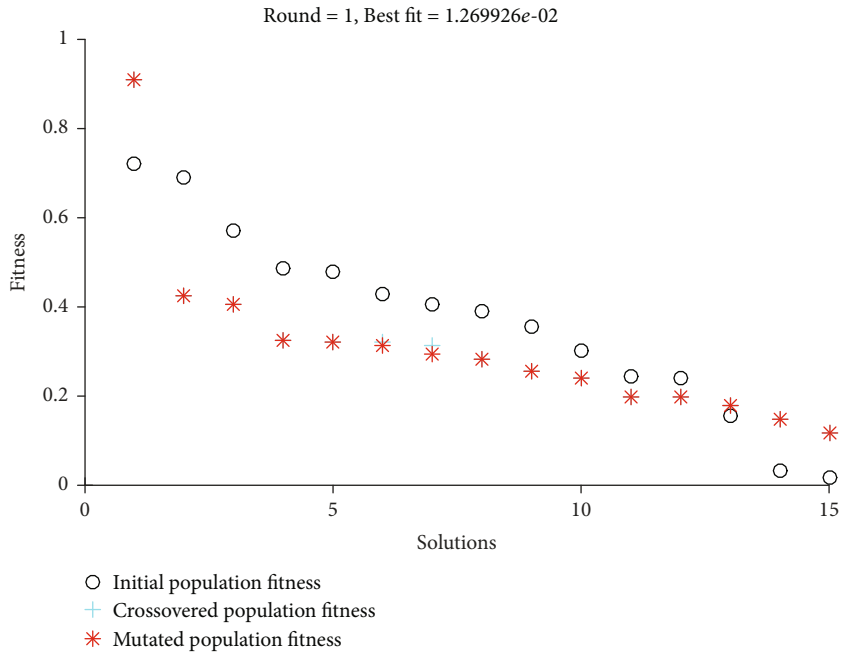


FIGURE 5: Values of the fitness function of the initial population after the application of genetic operators.

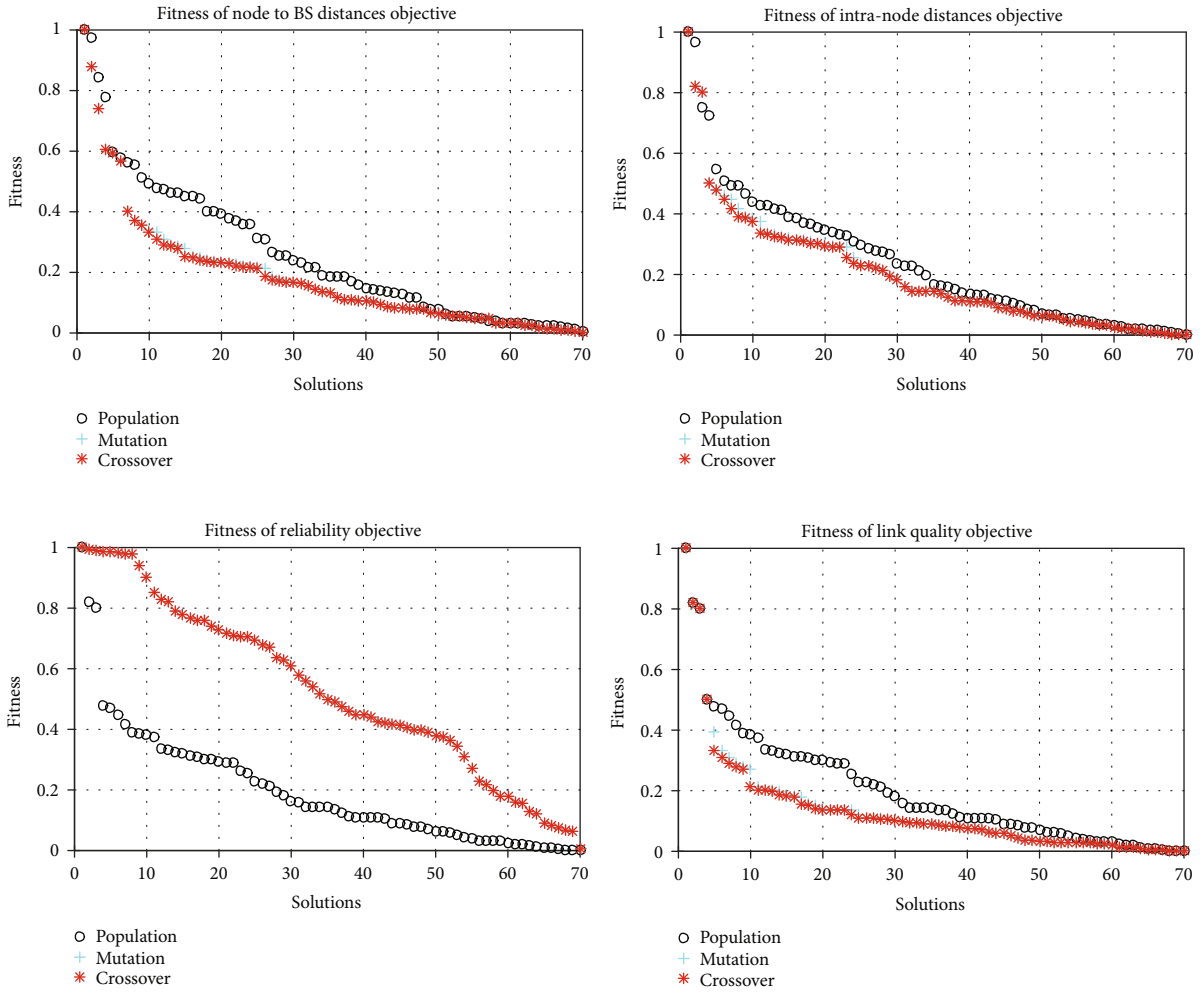


FIGURE 6: Fitness values on multiple criteria after applying the crossover and mutation operator.

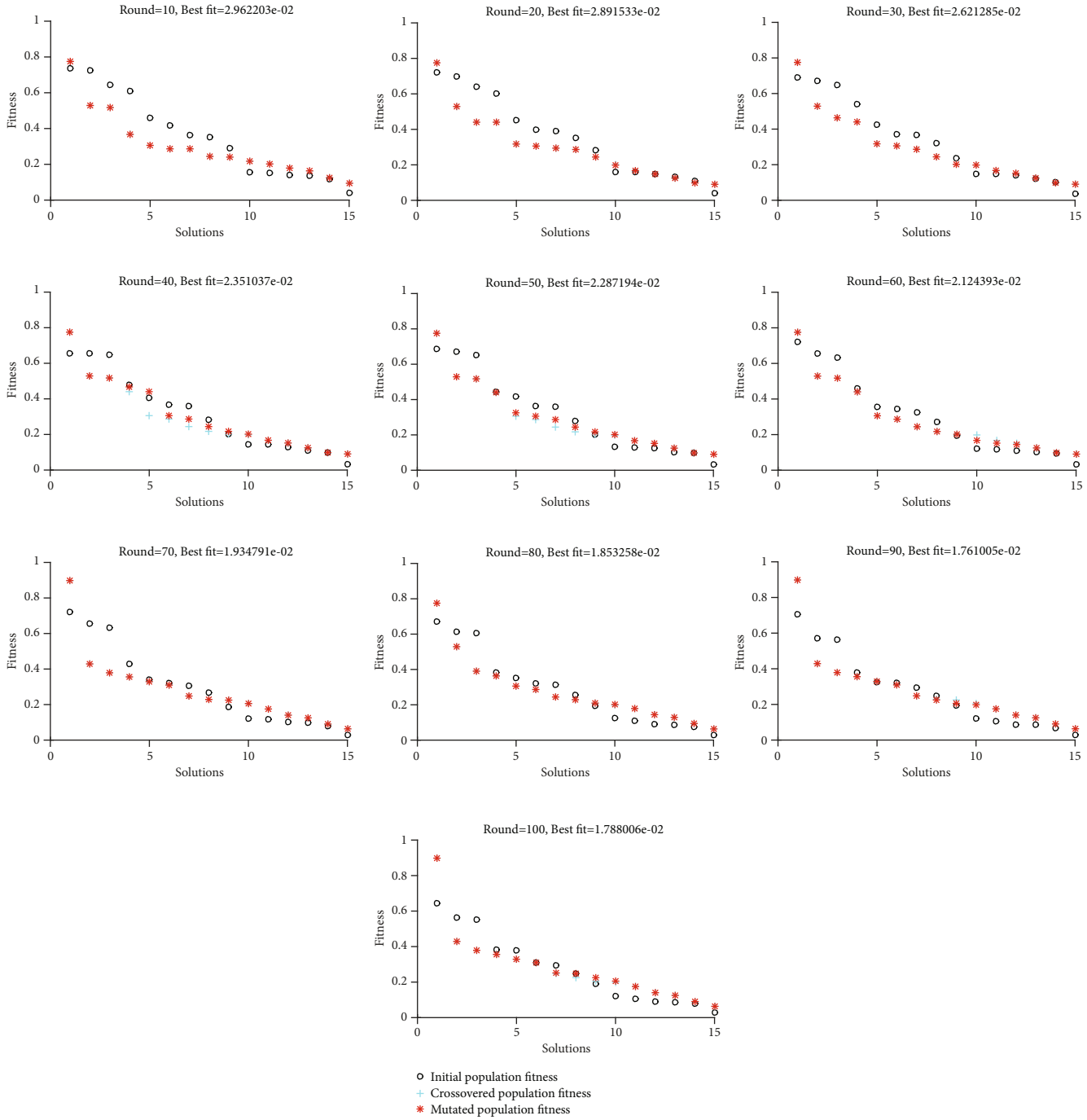


Figure 7: Repetition process in the proposed genetic algorithm.

problem consisting of three minimization components, the lower the value of the overall proportionality function, the higher the quality for the obtained chromosomes. Hence, the value of the overall fitness function in the proposed method approaches the optimum value by decreasing the values for each chromosome. Figure 4 also shows that the value of the fitness function for chromosomes is reduced and improved by applying adaptive fitting operators. Due to the fact that the proposed method uses a multiobjective fitness function, the fitting operator affects each of the criteria used in the fitness function. Therefore, Figure 6 shows the changes of adaptive crossover and mutation operators on the criteria used in routing in the proposed method.

As shown in Figure 6, the values of the fitness function for multiple criteria in the network are improved by the fitting operator. Since this criterion is the distance between nodes, the distance of nodes to the base station, and the quality of the link concerning the proposed proportionality function as a minimization criterion, according to Figure 5, the diagrams related to these criteria can be seen as it is less

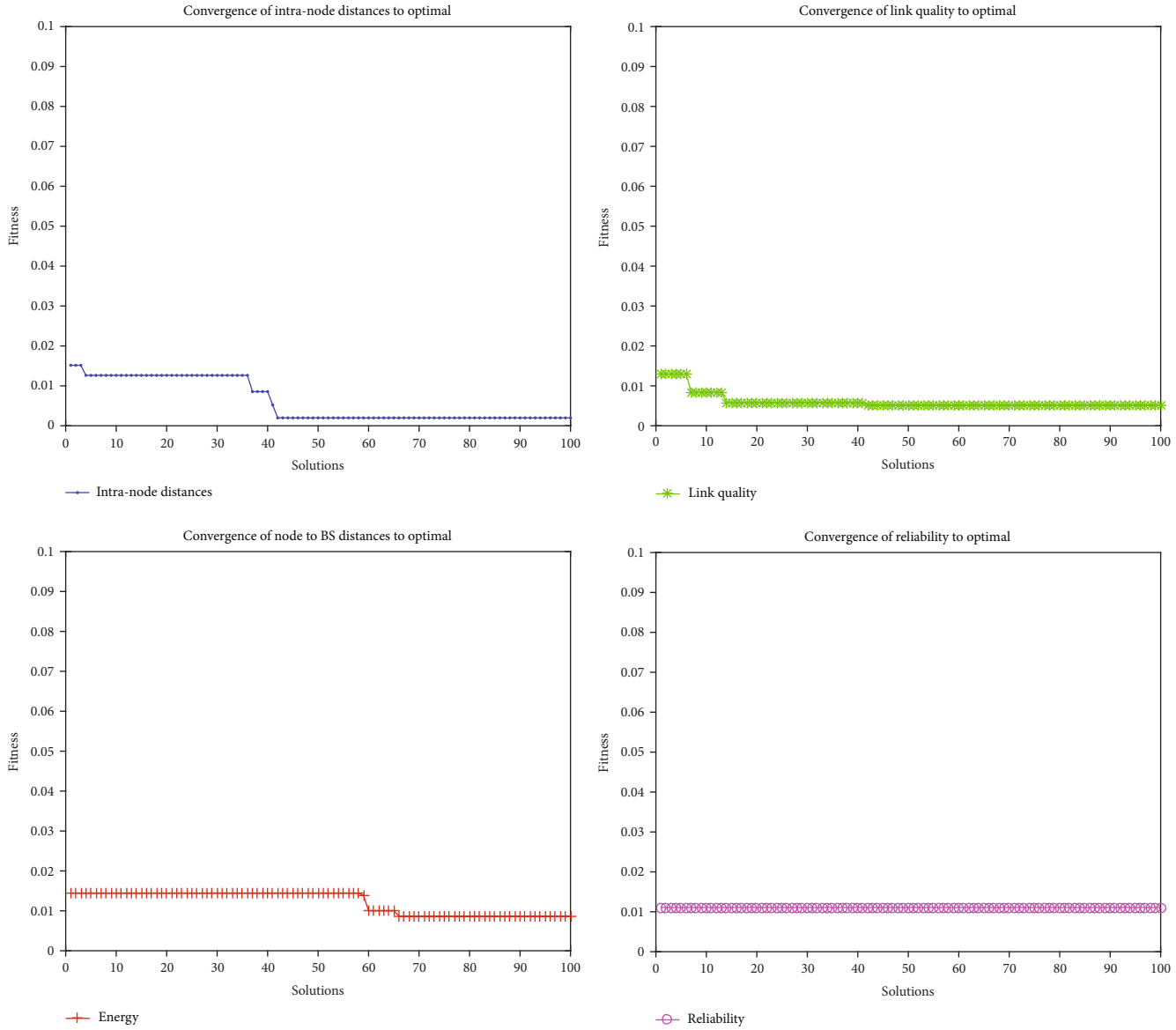


Figure 8: Convergence of the proposed genetic algorithm towards optimal points for the purposes of the fitness function.

Table 5: Final optimal chromosome.

	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9	G_{10}	G_{11}	G_{12}	G_{13}	G_{14}	G_{15}
CH1	1	3	2	6	7	8	9	10	5	19	23	17	19	22	24
CH2	1	5	85	77	99	41	92	37	97	89	96	74	25	14	32

active than the main chromosomes. Also, since this degree of confidence in the proposed proportionality function is expressed as a maximization criterion, according to Figure 5, it can be seen that the graph related to this criterion is higher than the main chromosomes after the operators are applied. The proposed genetic algorithm is based on the new population as well as the fitting and mutation operators until the stop condition is met. In this research, the condition for stopping is 100 repetition generations, which is shown in Figure 7, the results of the general fitness function in every ten repetition steps.

As shown in Figure 7, the proposed genetic algorithm continues to operate with respect to the genetic operators of fitting and mutation, increasing the values of the overall fitness function from 0.0289 to about 0.0178. These values indicate the evolution of the population in the process of replicating the proposed genetic algorithm and determining the best population based on the proposed method. The

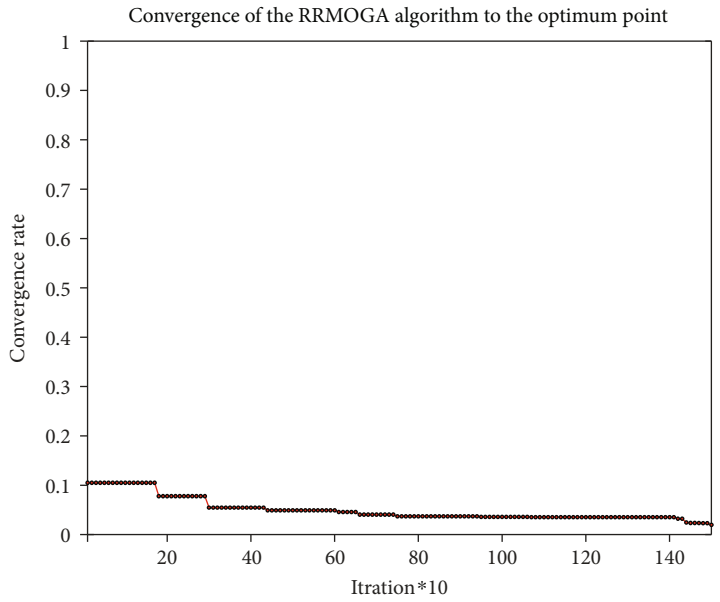


Figure 9: Convergence of the multiobjective genetic algorithm towards the optimal point.

Table 6: Average values of evaluation criteria.

Criteria	Lifetime	Delay	Data delivery rate	Reliability
Average	1400	0.0523	99.9	99.44

proposed genetic algorithm approaches the optimal points step by step and, in Figure 8, converges the four criteria of the distance between nodes, node distance to the base station, link quality, and degree of reliability used in the fitness function in the multiobjective genetic algorithm in the last generation of iteration.

As shown in Figure 8, the proposed method converges to the optimal point based on the objectives introduced during the algorithm. Thus, the final population found by the proposed multiobjective genetic algorithm is the near-optimal solution. Table 5 shows the final optimal chromosome as the primary pathway and the second optimal chromosome as the backbone for routing.

As shown in Table 5, the final solution of the problem is determined based on the proposed genetic algorithm and the criteria specified in the proposed method. Finally, the general tendency of the multiobjective genetic algorithm method towards the optimal point during the IoT network application time is shown in Figure 9.

4.3. Evaluating the Proposed Method. The evaluation of the proposed method is done in order to evaluate the quality of the proposed method and to provide the improvement created by the proposed method based on the initial problem. Given the popularity of IoT networks, there are a variety of criteria to evaluate. Evaluation criteria in IoT networks vary according to different applications. Therefore, in this study, given that we seek to increase the degree of reliability, throughput, and data delivery rate and to reduce network latency, we will be satisfied with these four criteria. Thus, we first show the main values of the evaluation criteria in Table 6.

As shown in Table 6, the proposed method has averaged the appropriate evaluation criteria. Figure 10 shows the average energy consumption of nodes in the IoT network.

As shown in Figure 10, the average energy consumption of nodes in the proposed IoT network increases in a balanced way according to the selection of optimal nodes in each step. Thus, it can be concluded that the energy of the sensor nodes embedded in the equipment is exhausted at almost the same time and the lifetime that the network can have its maximum value. Network lifetime refers to the period of time that the network is available and does not interfere with the aggregation of network data. Therefore, the time of energy depletion in a number of nodes in the network in a way that we are not able to continue the operation of the network with the remaining nodes can be considered as the lifetime of the network. Figure 11 shows the network lifetime in the proposed method.

Figure 11 shows that in the proposed network, the network lifetime reaches 1400 cycles and the death of the first node occurs after 1023 repetitions. In wireless sensor networks, with the death of a node, all the paths leading to this node are blocked and they have to use backup paths. Thus, by using a backup path in the proposed method, data transfer in the IoT network continues its normal process until no optimal path can be found for data transfer, and consequently, the degree of reliability of the proposed routing protocol will be high. Figure 12 shows the reliability diagram of the proposed method during data transfer in the proposed IoT network.

As shown in Figure 12, the reliability of the proposed method is high owing to the use of the optimal path and the choice of backup path throughout the simulation time. Therefore, it can be said that the proposed method has provided a reliable approach for the transfer of medical data in the healthcare system.

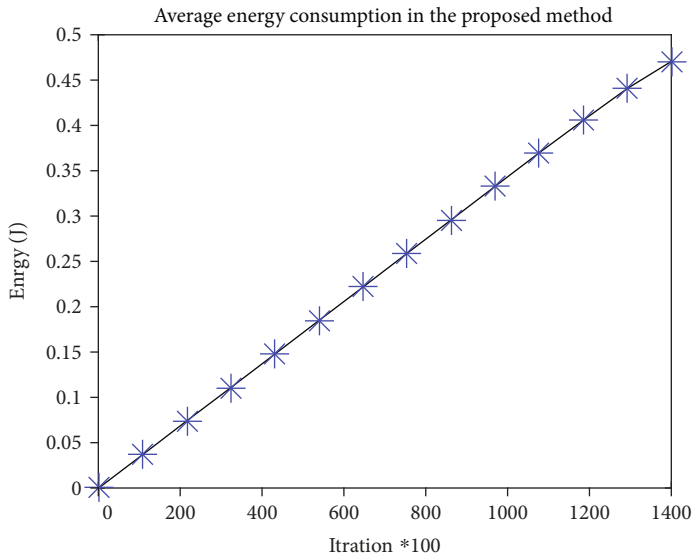


Figure 10: Average of energy consumption in the network.

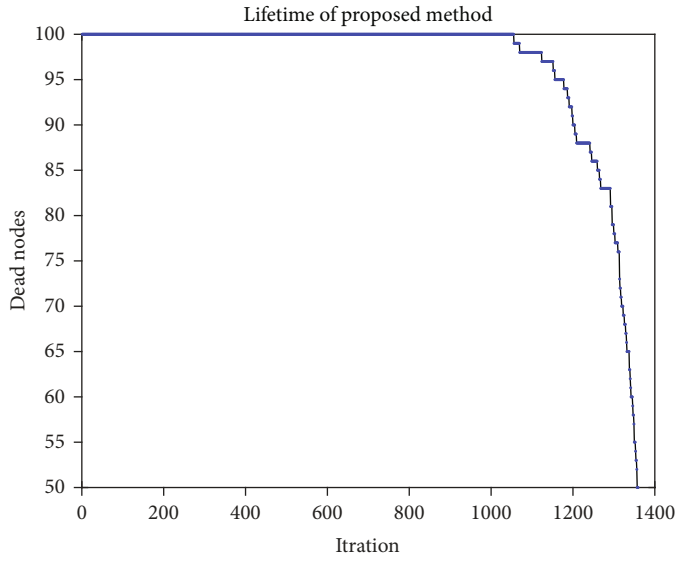


Figure 11: Node lifetime in the proposed IoT network.

Another criterion used to evaluate the proposed method is the data delivery rate criterion in the proposed IoT network. The data delivery rate criterion is the rate of packets sent per unit of time and received at the destination. In other words, the rate of data collected that has safely reached the nursing station per unit of time in the IoT network is based on the healthcare system and is called the network data delivery rate. Figure 13 shows the passing criteria of the proposed method.

As shown in Figure 13, the proposed method tries to escape bottlenecks and create a secure path to send data to the hole nodes with respect to optimal routing. Therefore, the data delivery rate in the proposed network is high and its average is 99.9%. The last criterion utilized for evaluation in the proposed method is the end-to-end delay of nodes in the network. Given that the transfer time is for fixed nodes, the main reason for the delay in the end-to-end transfer is the distance between nodes. Since in the proposed method, according to the multiobjective fitness function, data transfer between nodes occurs in the shortest distance, the end-to-end delay will be the least. The total network latency, which consists of the sum of the end-to-end latency between nodes, is shown in Figure 14.

As shown in Figure 14, in the proposed method, the end-to-end latency between nodes increases in a balanced way indicating the minimum distance and minimum delay for sending packets between nodes.

4.4. Comparing the Proposed Method with Previous Works. The development of the use of the IoT infrastructure in

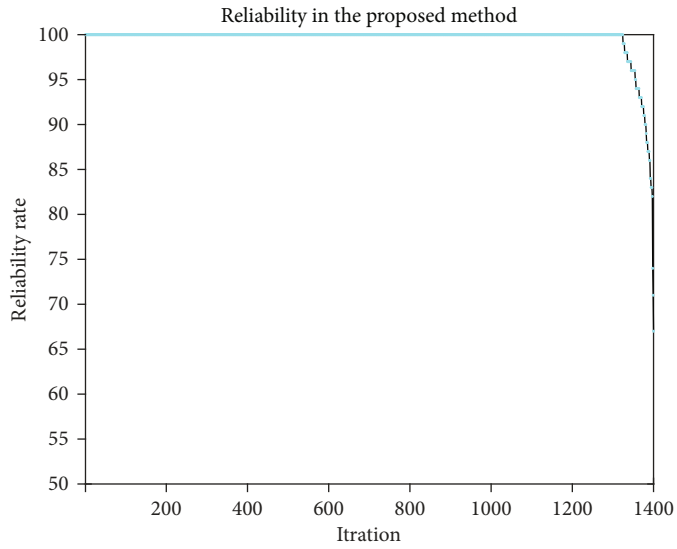


Figure 12: Degree of confidence in the proposed method.

Figure 13: Data delivery rate of the proposed method.

healthcare systems has led to extensive research into routing in IoT networks. In the continuation of this part of the article, in order to validate the proposed method, we will compare it with the previous methods in terms of reliability, lifetime, latency, and data delivery rate in the network. The most important criterion that the proposed method focuses on is the degree of reliability in the information transmission path. So, we compare RRMOGA with RRDLA [42], DETR [43], and REER [44] in terms of reliability. Figure 15 shows a comparison of the degree of reliability in the proposed method with other previous methods.

As shown in Figure 15, the proposed method has a higher degree of reliability than the previous methods due to the selection of reliable agents and the use of a backup path to deal with possible failures in the optimal path.

We compare the proposed method with the basic methods of AODV [45], LABILE, and REL [46] algorithms in terms of the mentioned criteria. Figure 16 represents a comparison between the proposed method and previous methods considering the length of live nodes in the network.

As shown in Figure 16, the proposed method has more live nodes than the previous methods. Therefore, it can be concluded that due to the use of the multiobjective fitness function, the network life in the proposed method is longer than in other previous methods. Furthermore, in Figure 17, the proposed method is compared with previous methods in terms of the data delivery rate.

As shown in Figure 17, the proposed method has a higher delivery rate than other related methods. The data delivery rate in the proposed method is higher due to the

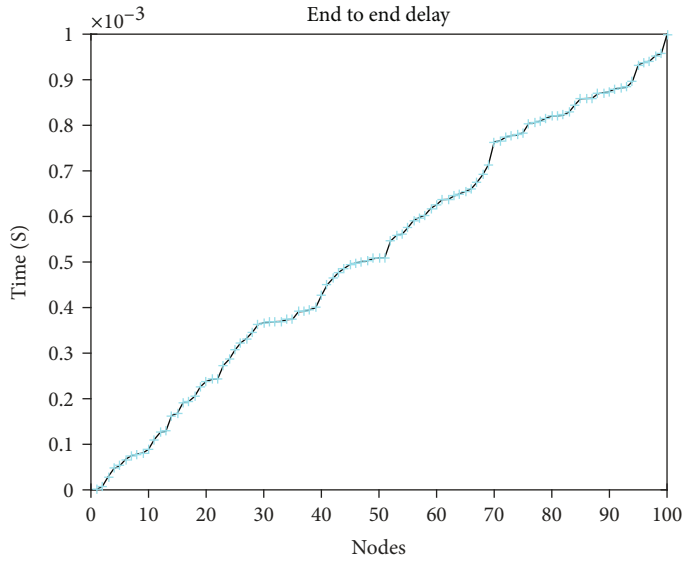


Figure 14: End-to-end delay of nodes in the proposed IoT network.

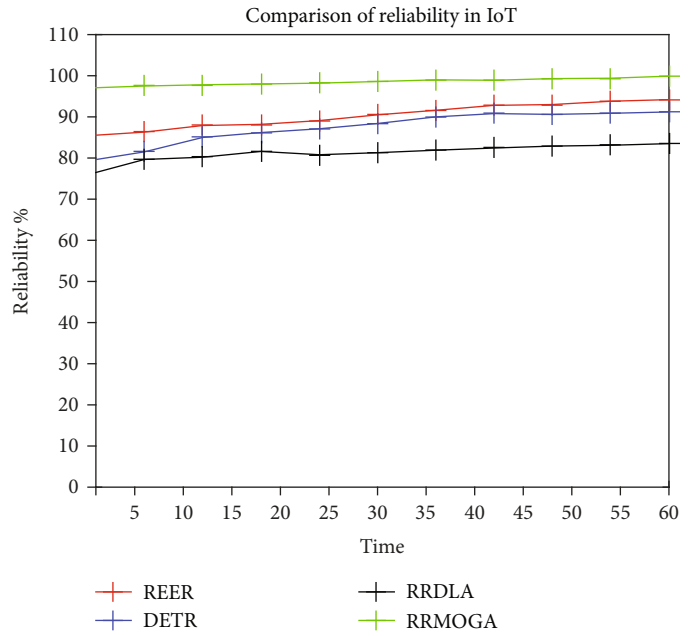


Figure 15: Comparison between the proposed method and previous methods in the aspect of reliability.

balance between the basic factors in the network based on the multiobjective proportionality function. Another criterion that shows the improvement of the proposed method compared to other existing methods is the overall delay of the method in data transfer. Figure 18 compares the proposed method with previous methods in terms of packet transfer delay.

As shown in Figure 18, the proposed method has less overall latency than other related methods. The proposed method has lower latency than the previous methods due to the selection of optimal nodes based on multiple objective functions in each step of packet transmission.

5. Discussion and Conclusion

As mentioned, this paper presents a reliable routing approach for IoT healthcare applications. Healthcare systems are trying to send data safely through highly reliable routes on the network. Therefore, this paper tries to achieve this goal by presenting an approach based on a multiobjective genetic algorithm. The main innovation of this paper is the integration of important goals of the network in the fitness function, which tries to find the optimal path by creating a tradeoff between these goals. According to the findings in experiments and by the implementation of the

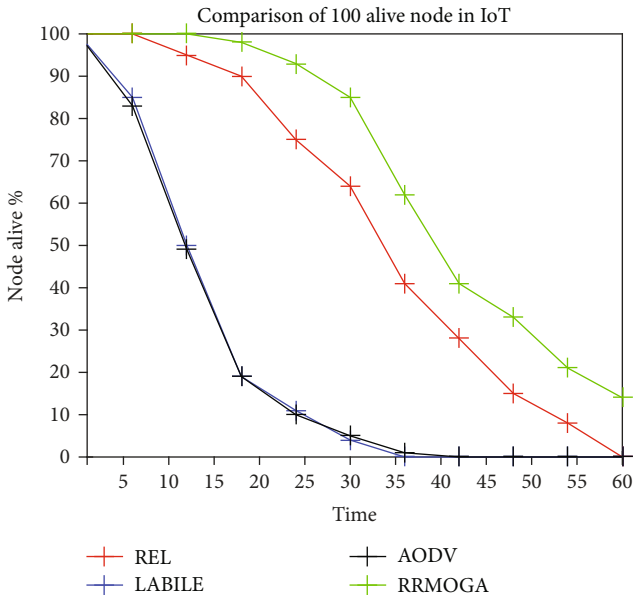


Figure 16: Comparison between the proposed method and previous methods in the aspect of live nodes.

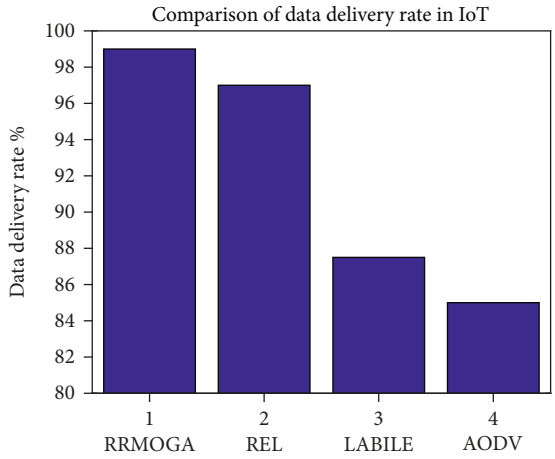


Figure 17: Comparison between the proposed method and previous methods in terms of the data delivery rate.

Internet of things, it can be seen that the method proposed in this article has not sacrificed other goals for reliability. The present method tries to improve reliability along with other objectives such as throughput, data delivery rate, network lifetime, and delay. The multiobjective fitness function used in the proposed method has created a tradeoff in the mentioned parameters. Previous methods have focused on one of the goals in the network, and therefore, optimal local routes have been found in each method and other basic parameters in the network have been ignored. For this reason, the proposed method not only has increased reliability but also has improved other basic network parameters. Findings show that the proposed method has been able to achieve better results compared with previous methods by relying on the multiobjective fitness function.

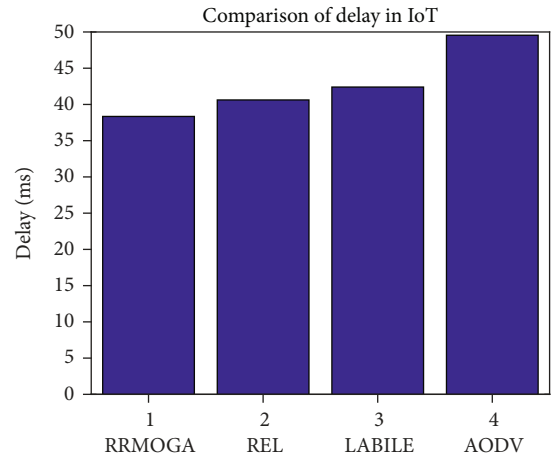


Figure 18: Comparison between the proposed method and previous methods in terms of data transfer delay.

In this paper, we have used reliable agents to send packets. Therefore, in each step, the proposed method selects an optimal reliable route and an alternative reliable route for sending packets. Both routes are composed of optimal agents. If data transmission is difficult in the near-optimal initial reliable agents, the proposed method uses an alternative route for sending data packets. Therefore, the data delivery rate in the proposed method is high, which the results of experiments and simulations prove the claim.

In this research, the secure routing protocol in IoT networks in the healthcare context is presented based on a multiobjective method as well as a multiobjective genetic algorithm. In the proposed method, chromosomes are selected as paths between the source node and the base station. Their proportionality function is considered based on network factors including distance between nodes, the base station, link quality, and degree of reliability. The proposed method, in each step, considers the optimal path and optimal nodes in order to transfer data in the network. Each node is considered as a factor of IoT-connected equipment in the IoT network in the healthcare context. The path between the source node and the base station is identified as an interconnected set of agents. Support factors are also selected as backup pathways, including the second optimal chromosome in the multiobjective genetic algorithm, to increase reliability. The simulation results of the proposed method show that the proposed method, in addition to having a longer network life and data delivery rate compared with other previous methods, has also reduced the data transmission delay. The proposed method, via using the balance between objectives in the proportionality function, is able to provide a secure approach to information transfer with minimum latency. Therefore, by balancing several objectives in network routing, the proposed method has been able to achieve good results in comparison to previous methods. The simulation results show that the proposed method, due to the use of reliable agents, has achieved reliability and data delivery rates, 99% and 99.9%, respectively, which have significantly improved compared to previous methods.

Data Availability

The data used to support the findings of this study have been adjusted according to standard setting and the results have been extracted from simulation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Balaji, K. Nathani, and R. Santhakumar, "IoT technology, applications and challenges: a contemporary survey," *Wireless Personal Communications*, vol. 108, no. 1, pp. 363–388, 2019.

[2] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare Internet of things: a survey of emerging technologies," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1121–1167, 2020.

[3] S. Banka, I. Madan, and S. Saranya, "Smart healthcare monitoring using IoT," *International Journal of Applied Engineering Research*, vol. 13, no. 15, pp. 11984–11989, 2018.

[4] A. Darwish, A. E. Hassanien, M. Elhoseny, A. K. Sangaiah, and K. Muhammad, "The impact of the hybrid platform of Internet of things and cloud computing on healthcare systems: opportunities, challenges, and open problems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 4151–4166, 2019.

[5] A.-M. Rahmani, N. K. Thanigaivelan, T. N. Gia et al., "Smart e-health gateway: bringing intelligence to Internet-of-things based ubiquitous healthcare systems," in *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 826–834, Las Vegas, NV, USA, 2015.

[6] L. M. Dang, M. J. Piran, D. Han, K. Min, and H. Moon, "A survey on Internet of things and cloud computing for healthcare," *Electronics*, vol. 8, no. 7, p. 768, 2019.

[7] A. Shehab, M. Elhoseny, K. Muhammad et al., "Secure and robust fragile watermarking scheme for medical images," *IEEE Access*, vol. 6, pp. 10269–10278, 2018.

[8] A. Sharma and R. Kumar, "An optimal routing scheme for critical healthcare HTH services—an IOT perspective," in *2017 Fourth International Conference on Image Information Processing (ICIIP)*, pp. 1–5, Shimla, India, 2017.

[9] P. Chanak and I. Banerjee, "Congestion free routing mechanism for IoT-enabled wireless sensor networks for smart healthcare applications," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 3, pp. 223–232, 2020.

[10] M. Lazarevska, R. Farahbakhsh, N. M. Shakya, and N. Crespi, "Mobility supported energy efficient routing protocol for IoT based healthcare applications," in *2018 IEEE Conference on Standards for Communications and Networking (CSCN)*, pp. 1–5, Paris, France, 2018.

[11] M. W. Woo, J. Lee, and K. Park, "A reliable IoT system for personal healthcare devices," *Future Generation Computer Systems*, vol. 78, pp. 626–640, 2018.

[12] A. Sharma and R. Kumar, "Computation of the reliable and quickest data path for healthcare services by using service-level agreements and energy constraints," *Arabian Journal for Science and Engineering*, vol. 44, no. 11, pp. 9087–9104, 2019.

[13] K. Wang, Y. Shao, L. Xie, J. Wu, and S. Guo, "Adaptive and fault-tolerant data processing in healthcare IoT based on fog computing," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 263–273, 2020.

[14] K. Machado, D. Rosário, E. Cerqueira, A. Loureiro, A. Neto, and J. de Souza, "A routing protocol based on energy and link quality for Internet of things applications," *Sensors*, vol. 13, no. 2, pp. 1942–1964, 2013.

[15] J. Jia et al., "Energy efficient coverage control in wireless sensor networks based on multi- objective genetic algorithm," *Computers & Mathematics with Applications*, vol. 57, no. 11-12, pp. 1756–1766, 2009.

[16] M. Abbasi, E. Mohammadi Pasand, and M. R. Khosravi, "Workload allocation in iot-fog-cloud architecture using a multi-objective genetic algorithm," *Journal of Grid Computing*, vol. 18, no. 1, pp. 43–56, 2020.

[17] A. M. Maia, Y. Ghamri-Doudane, D. Vieira, and M. Franklin de Castro, "An improved multi-objective genetic algorithm with heuristic initialization for service placement and load distribution in edge computing," *Computer Networks*, vol. 194, p. 108146, 2021.

[18] J. M. Corrigan, E. K. Swift, and M. P. Hurtado, *Envisioning the National Health Care Quality Report*, 2001.

[19] B. Farahani, F. Firouzi, and K. Chakrabarty, "Healthcare iot," in *Intelligent Internet of Things: From Device to Fog and Cloud*, pp. 515–545, Springer, 2020.

[20] M. Haghi Kashani, M. Madanipour, M. Nikravan, P. Asghari, and E. Mahdipour, "A systematic review of IoT in healthcare: applications, techniques, and trends," *Journal of Network and Computer Applications*, vol. 192, p. 103164, 2021.

[21] Q. Ai, W. Meng, F. Bensaali, X. Zhai, L. Liu, and N. Alaraje, "Editorial for FGCS special issue: intelligent IoT systems for healthcare and rehabilitation," *Future Generation Computer Systems*, vol. 125, pp. 770–773, 2021.

[22] J. Khan, J. P. Li, A. U. Haq et al., "Efficient secure surveillance on smart healthcare IoT system through cosine-transform encryption," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 1, pp. 1417–1442, 2021.

[23] A. Ullah, M. Azeem, H. Ashraf, A. A. Alaboudi, M. Humayun, and N. Z. Jhanjhi, "Secure healthcare data aggregation and transmission in IoT—A survey," *IEEE Access*, vol. 9, pp. 16849–16865, 2021.

[24] S. Goyal, N. Sharma, B. Bhushan, A. Shankar, and M. Sagayam, "Iot enabled technology in secured healthcare: applications, challenges and future directions," in *Cognitive Internet of Medical Things for Smart Healthcare: Services and Applications*, pp. 25–48, Springer, 2021.

[25] B. Pradhan, S. Bhattacharyya, and K. Pal, "IoT-based applications in healthcare devices," *Journal of Healthcare Engineering*, vol. 2021, Article ID 6632599, 18 pages, 2021.

[26] K. T. Kadhim, A. M. Alsahlany, S. M. Wadi, and H. T. Kadhum, "An overview of patient's health status monitoring system based on Internet of things (IoT)," *Wireless Personal Communications*, vol. 114, no. 3, pp. 2235–2262, 2020.

[27] H. K. Bharadwaj, A. Agarwal, V. Chamola et al., "A review on the role of machine learning in enabling IoT based healthcare applications," *IEEE Access*, vol. 9, pp. 38859–38890, 2021.

[28] T. J. Oh, J. E. Lee, S. Kim, S. Yoo, and H. C. Jang, "Mobile healthcare system provided by primary care physicians improves quality of diabetes care," *CardioMetabolic Syndrome Journal*, vol. 1, no. 1, pp. 88–97, 2021.

[29] S. Lavanya, G. Lavanya, and J. Divyabharathi, "Remote prescription and I-home healthcare based on IoT," in *2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT)*, pp. 1–3, Coimbatore, India, 2017.

[30] J. Hong, P. Morris, and J. Seo, "Interconnected personal health record ecosystem using IoT cloud platform and HL7 FHIR," in *2017 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 362–367, Park City, UT, USA, 2017.

[31] V. Pardeshi, S. Sagar, S. Murmurwar, and P. Hage, "Health monitoring systems using IoT and raspberry pi—a review," in *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 134–137, Bengaluru, India, 2017.

[32] K. T. Kadhim, A. M. Alsahlany, S. M. Wadi, and H. T. Kadhum, "Monitoring vital signs of human hear based on IOT," *Al-Furat Journal of Innovations in Electronics and Computer Engineering*, vol. 1, no. 2, pp. 9–13, 2020.

[33] B. . soufiene, A. A. Bahattab, A. Trad, and H. Youssef, "PEERP: an priority-based energy-efficient routing protocol for reliable data transmission in healthcare using the IoT," *Procedia Computer Science*, vol. 175, pp. 373–378, 2020.

[34] P. Sarwesh, N. S. V. Shet, and K. Chandrasekaran, "Effective integration of reliable routing mechanism and energy efficient node placement technique for low power IoT networks," *International Journal of Grid and High Performance Computing*, vol. 9, no. 4, pp. 16–35, 2017.

[35] S. P. Gochhayat, C. Lal, L. Sharma et al., "Reliable and secure data transfer in IoT networks," *Wireless Networks*, vol. 26, 2020.

[36] M. Conti, P. Kaliyar, and C. Lal, "REMI: a reliable and secure multicast routing protocol for IoT networks," in *Proceedings of the 12th International Conference on Availability, Reliability and Security*, pp. 1–8, New York, NY, USA, 2017.

[37] Y. Lyu and P. Yin, "Internet of things transmission and network reliability in complex environment," *Computer Communications*, vol. 150, pp. 757–763, 2020.

[38] P. Asghari, A. M. Rahmani, and H. Haj Seyyed Javadi, "A medical monitoring scheme and health-medical service composition model in cloud-based IoT platform," *Transactions on Emerging Telecommunications Technologies*, vol. 30, no. 6, 2019.

[39] P. Asghari, A. M. Rahmani, and H. H. S. Javadi, "Privacy-aware cloud service composition based on QoS optimization in Internet of things," *Journal of Ambient Intelligence and Humanized Computing*, vol. 2020, pp. 1–26, 2020.

[40] M. Hosseinzadeh, J. Koohpayehzadeh, M. Y. Ghafour et al., "An elderly health monitoring system based on biological and behavioral indicators in Internet of things," *Journal of Ambient Intelligence and Humanized Computing*, vol. 2020, pp. 1–11, 2020.

[41] Y. Wang, X. Geng, F. Zhang, and J. Ruan, "An immune genetic algorithm for multi-echelon inventory cost control of IOT based supply chains," *IEEE Access*, vol. 6, pp. 8547–8555, 2018.

[42] H. Mostafaei, "Energy-efficient algorithm for reliable routing of wireless sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 7, pp. 5567–5575, 2019.

[43] Z. Liu, L. Dai, L. Xue, X. Guan, and C. Hua, "Reliability considered routing protocol in wireless sensor networks," in *Proceedings of the 30th Chinese Control Conference*, pp. 5011–5016, Yantai, China, 2011.

[44] Xiang-Yang Li, Yu Wang, Haiming Chen, Xiaowen Chu, Yanwei Wu, and Yong Qi, "Reliable and energy-efficient routing for static wireless ad hoc networks with unreliable links," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 10, pp. 1408–1421, 2009.

[45] T. Abbas et al., "Performance analysis of ad hoc on-demand distance vector routing protocol for MANET," in *2020 IEEE Student Conference on Research and Development (SCOReD)*, pp. 194–199, Batu Pahat, Malaysia, 2020.

[46] P. V. Kallapur and T. Nargis, "Performance analysis of LABILE and REL protocol using Castalia simulator," in *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, pp. 339–343, Kerala, India, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 1542657, 9 pages
https://doi.org/10.1155/2021/1542657

Research Article

Intelligent Detection System Enabled Attack Probability Using Markov Chain in Aerial Networks

Inam Ullah Khan,[1] **Asrin Abdollahi**,[2] **Ryan Alturki**,[3] **Mohammad Dahman Alshehri**,[4] **Mohammed Abdulaziz Ikram**,[5] **Hasan J. Alyamani**,[6] **and Shahzad Khan**[7]

[1]*Department of Electronic Engineering, School of Engineering and Applied Sciences, Isra University, Islamabad, Pakistan*
[2]*Department of Electrical Engineering, University of Kurdistan, Sanandaj, Iran*
[3]*Department of Information Science, College of Computer and Information Systems, Umm AlQura University, Makkah, Saudi Arabia*
[4]*Department of Computer Science, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia*
[5]*Computer Science Department, University College in Al-Jamoum, Umm Al-Qura University, Saudi Arabia*
[6]*Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia*
[7]*Department of Computer Science, Abdul Wali Khan University Mardan, 23200 Mardan, KPK, Pakistan*

Correspondence should be addressed to Asrin Abdollahi; a.abdollahi@eng.uok.ac.ir

Received 13 April 2021; Accepted 18 August 2021; Published 9 September 2021

Academic Editor: Ihsan Ali

The Internet of Things (IoT) plays an important role to connect people, data, processes, and things. From linked supply chains to big data produced by a large number of IoT devices to industrial control systems where cybersecurity has become a critical problem in IoT-powered systems. Denial of Service (DoS), distributed denial of service (DDoS), and ping of death attacks are significant threats to flying networks. This paper presents an intrusion detection system (IDS) based on attack probability using the Markov chain to detect flooding attacks. While the paper includes buffer queue length by using queuing theory concept to evaluate the network safety. Also, the network scenario will change due to the dynamic nature of flying vehicles. Simulation describes the queue length when the ground station is under attack. The proposed IDS utilizes the optimal threshold to make a tradeoff between false positive and false negative states with Markov binomial and Markov chain distribution stochastic models. However, at each time slot, the results demonstrate maintaining queue length in normal mode with less packet loss and high attack detection.

1. Introduction

Flying ad hoc networks have changed human life where wireless communication is utilized as a backbone technology. Flying networks have remote nodes that can switch along with all three directions [1]. The term "flying ad hoc network" represents a complex pattern of mobility with constantly changing physical structures [2].

Secure communication channels must be designed to improve connectivity within the network. Dealing with false data injection attacks, an intruder can take data during remote surgery, which can lead to the death of a patient. Also, in defense operations, aerial vehicles are used to trigger false data attacks in the surrounding environment, which causes very serious destruction [3]. However, the probability of Poisson distribution is used for the detection of ping of

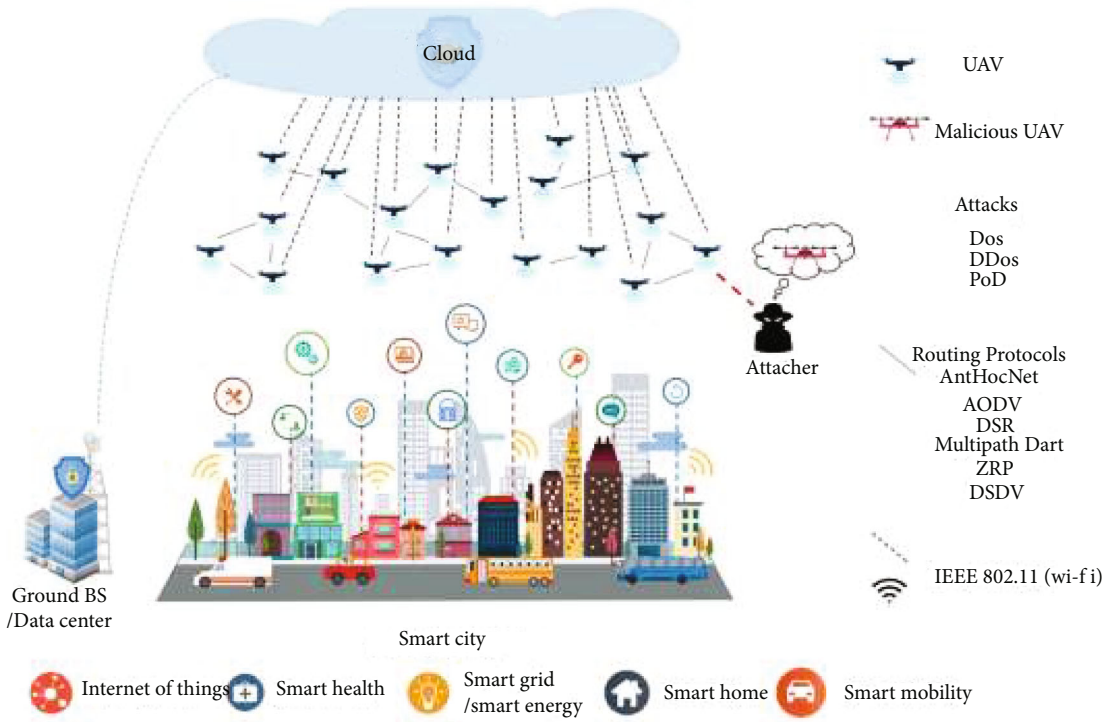


Figure 1: Intruder/attacker within IoT network.

death attacks that secure data packets [4]. Several security attacks are recorded from 1982 till now in different industries. In 2014, a special type of attack occurred, which was later on called Trojan, where the main target was petroleum pipeline networks [5].

Detecting cybercrimes over the internet can be identified using an intrusion detection system by using different techniques and tools [6]. However, a swarm of drones can protect an entire IoT network [7]. Detection of security attacks is a major problem; this research study formulates the scenario on quadcopter using open-source software [8]. Therefore, a tree-based strategy can easily portray the moves of intruders/attackers; also, for risk evaluation, a game-theory scheme is used [9]. Every technology is just made to facilitate mankind; for this purpose, aerial vehicles can be used to safeguard women [10].

The proposed scheme is focused on reducing queue data packets in flying networks which is a tough task to tackle. The aim of this paper is to explore the IDS model and how it can be improved using a Markov chain approach. Using flying networks, the IDS architecture has been developed to reduce data packets in the queue at different stages. Figure 1 explains the concept of an intruder within an IoT network to demonstrate a practical scenario. The main contributions of this study include some important points, which are given below.

(1) For the identification of security threats, an intrusion detection system is modeled

(2) Denial-off-service, distributed denial of service, and ping of death attacks are simulated in flying vehicles

(3) Markov chain distribution is used to enhance security countermeasures

The rest of the article is organized as follows. In Section 1, brief literature relevant to the problem is studied. The proposed scheme is elaborated in Section 3, followed by simulation results and theoretical analysis in Section 4. The future research directions and paper's conclusion are given in Section 5.

2. Literature Survey

Every new technology is first used by military, later on, it becomes commercialized. However, in US, due to flying vehicles, some accidents take place, like if a drone comes in the way of airplane while landing. Apart from that, many other issues occur due to the technical fault in quadcopters or small UAV's. As communication plays an important role but major issue in day-to-day life is to secure transmission links [11]. The demand of aerial vehicles is increasing on daily basis. In normal flying systems, there is the concept of pilot but drone is basically unmanned which makes them unsafe or unprotected [12]. The two popular areas like machine learning and software-define-networks can provide a pathway to address the challenges related to security in terms of internet of everything [13, 14].

Wireless vision (Wi-Vi) sensors are put in service for self-controlled flying vehicles [15]. The indoor scenario is very much mature using the wireless network; therefore, channel state information can give accurate data about location coordinates [16]. A novel framework is introduced,

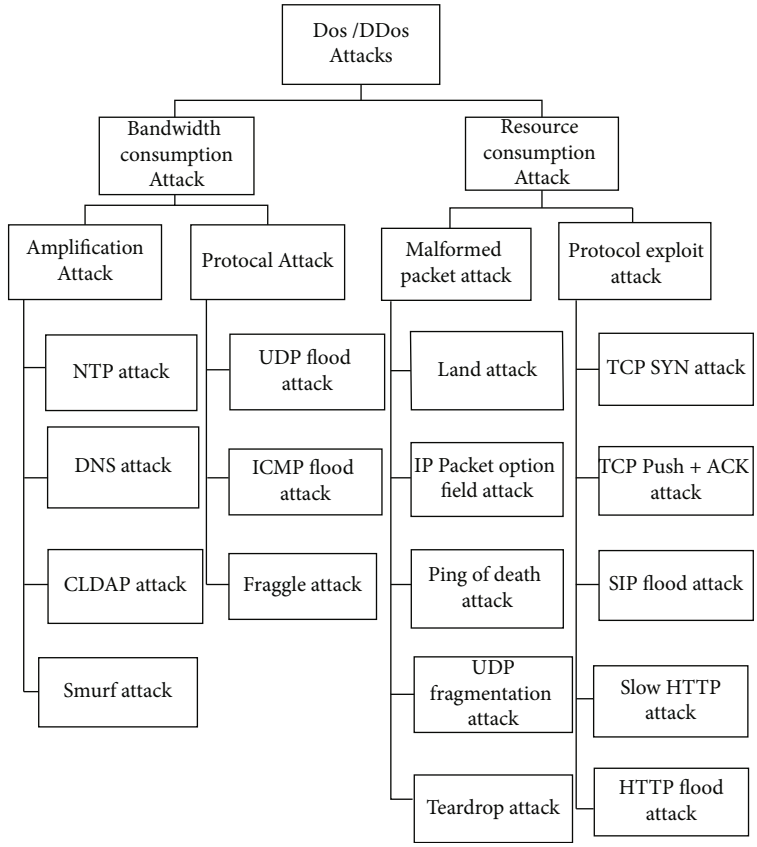


Figure 2: Classification of DoS/D-DoS security threats.

which has flying vehicle-enabled IoT using a 5G communication network. Human safety is the prime focus of every technology. In this context, if the flying drone having sensitive information is hijacked or attacked, it may result a big threat to the environment. Flying thing-based architecture is initiated which gives a solution mechanism for security and privacy to secure U2U communication [17]. Heuristic computational drone-based projects must be having pragmatic results in civil and military fields [18]. While working on false alarm threat, intrusion detection system can be utilized [4]. Furthermore, the classifications of DoS/D-DoS security threats are shown in Figure 2.

3. Intelligent Detection System (Proposed Scheme)

The proposed study is having physical topology with thirty drones ($N = 30$) and one ground station. Two major scenarios are mentioned either "no attack" or "with attack." Assuming that our internetwork is secure and there is no intruder inside the system. For this purpose, aerial vehicles send data packets having an average length which is cited as λ_l. Apart from that, aerial network modeling can be concluded for generating information of arrival data net which lined up in the entry to pinpoint land station. Figure 3 shows the physical structure of IDS in land station where malicious data packets can be removed easily.

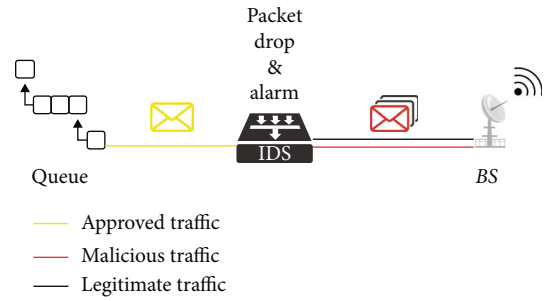


Figure 3: Physical structure of IDS in land station.

The evolution of queue length is calculated using the following equation:

$$Q(t+1) = Q(t) + \lambda(t) - \mu(t), \tag{1}$$

where $Q(t) > 0$, $Q(t)$ is queue length, $\lambda(t)$ is arrival rate, $\mu(t)$ is out rate or departed data rate.

The above metric values can be either constant or random. Furthermore, the randomness can be generated using Poisson distribution. The four probabilistic options are practically demonstrated in Figure 4 as mentioned.

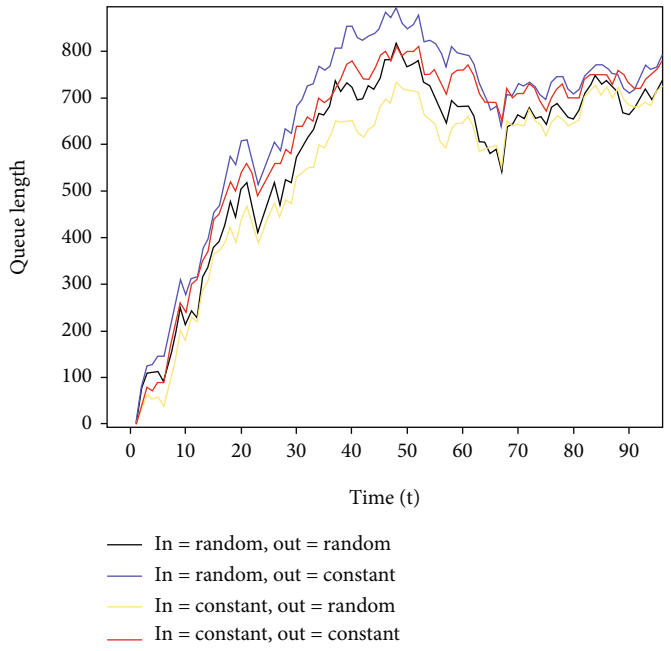


Figure 4: Probabilistic experimentation for generating queue.

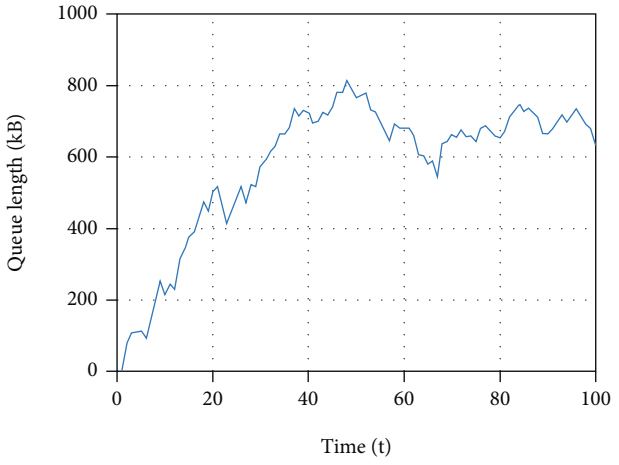


Figure 5: Queue length with Poisson variable.

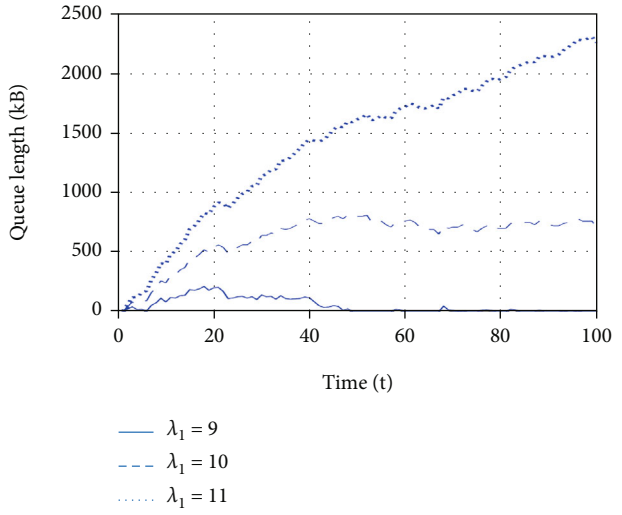


Figure 6: Impact λ_l on the queue length without μ_{avg}.

For $t = 100$ sec, the Poisson random variable with queued length is followed in Figure 5.

Inside the flying networks, once in a while, there might be no unwanted nodes to attack on the dynamic networks. But in the proposed network simulation, the input rate and flying nodes (N) are shown in Figure 6, which shows a high rise while flipping λ_l. By achieving the optimal results in between input and output queue rates which is presented in Figure 7. In the simulated work, by utilizing throughput, metric value can be effective in terms of outcome.

$$\mu_{\text{avg}} = \frac{N\lambda_l}{2}. \tag{2}$$

3.1. Markov Chain Distribution. Markov chain is a fundamental part of stochastic processes that use memory distribution in discrete-time steps that recall discrete-time Markov chain (DTMC). Suppose $X = \{X_t : t = 0, 1, 2, \cdots, T\}$ be the state of Markov chain stochastic process at time t with finite state spaces $S = \{1, 2\}$, where "1" represents "no attack level" which means normal, and "2" stands for the attack level.

$$P = (X_t = S_t \mid X_{t-1} = S_{t-1}, \cdots, X_0 = S_0) = P(X_t = S_t \mid X_{t-1} = S_{t-1}). \tag{3}$$

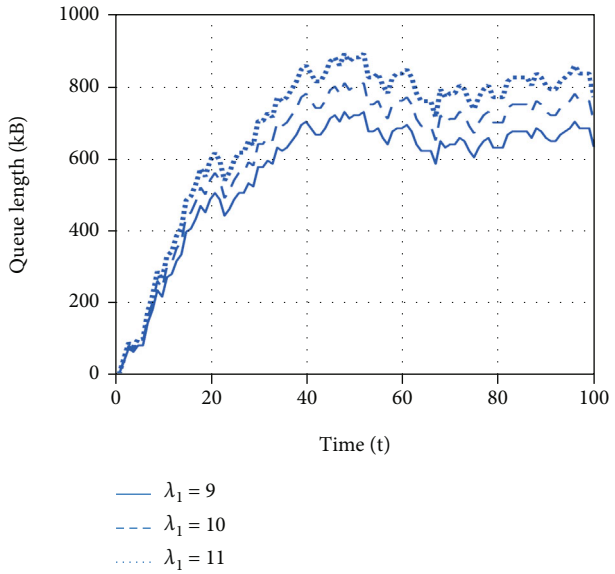


FIGURE 7: Impact λ_l on the queue length with μ_{avg}.

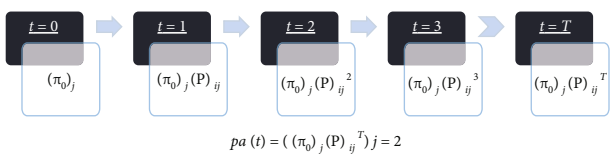


FIGURE 8: Attack probability at each time slot for desired Markov chain.

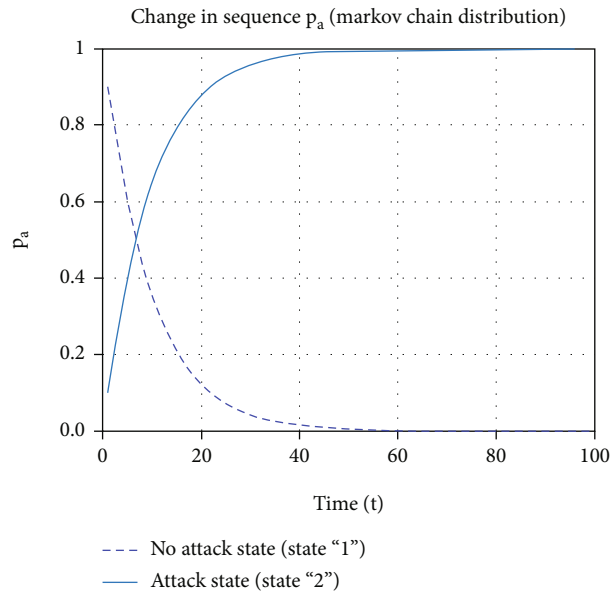


FIGURE 9: Attack probability versus time in Markov chain distribution.

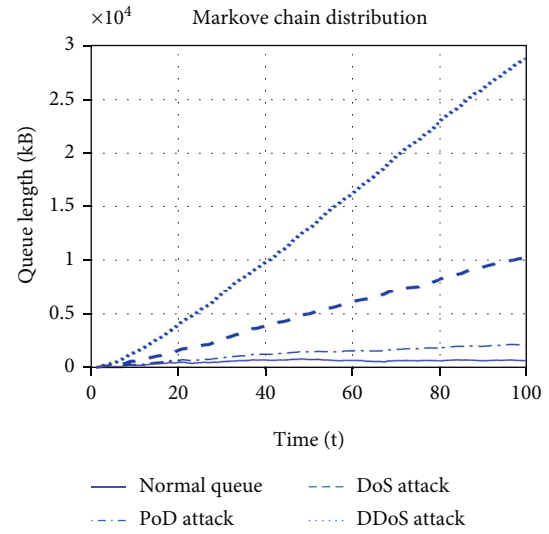


FIGURE 10: Queue length generation using cyberattacks for Markov chain.

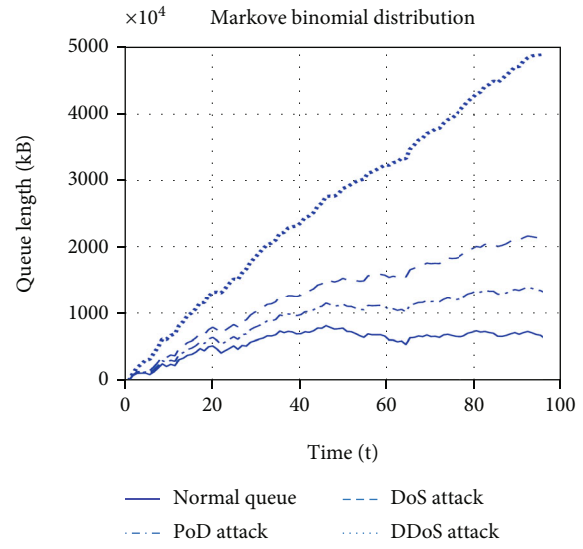


FIGURE 11: Queue length generation using cyberattacks for Markov binomial chain.

Equation (3) shows the formulation of Markov chain where for distribution X_t just having dependency on X_{t-1}. Finding the probability of being in state "1" or "2" at time t. In DoS, the attacker injects illegal packets to the network security systems by spoofing one node and attempts to increase the numbers of packets by utilizing the ratio $1 + \gamma$ p_a (γ is a positive constant). Apart from that modeling probability is p_a being changed in the first scenario where Markov chain with following transition matrix where α and β, respectively, is p_{a_0}, and 1 is proposed in the matrix.

$$(P)_{ij} = \left(p_{ij}\right) = \begin{bmatrix} 1-\alpha & \alpha \\ 1-\beta & \beta \end{bmatrix} = \begin{bmatrix} p_{a_0} & 1-p_{a_0} \\ 0 & 1 \end{bmatrix}. \quad (4)$$

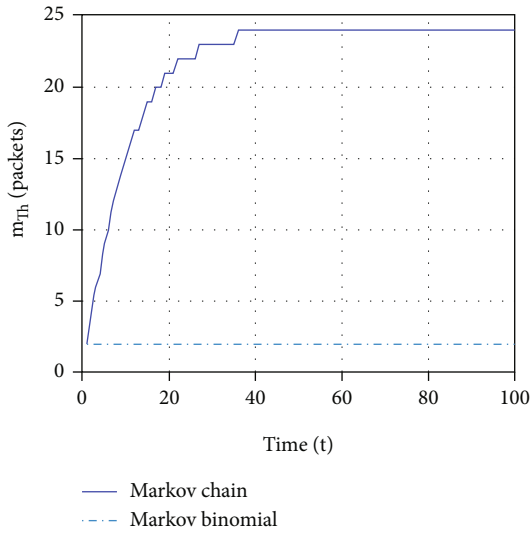


Figure 12: Certain level versus time in Markov chain distribution.

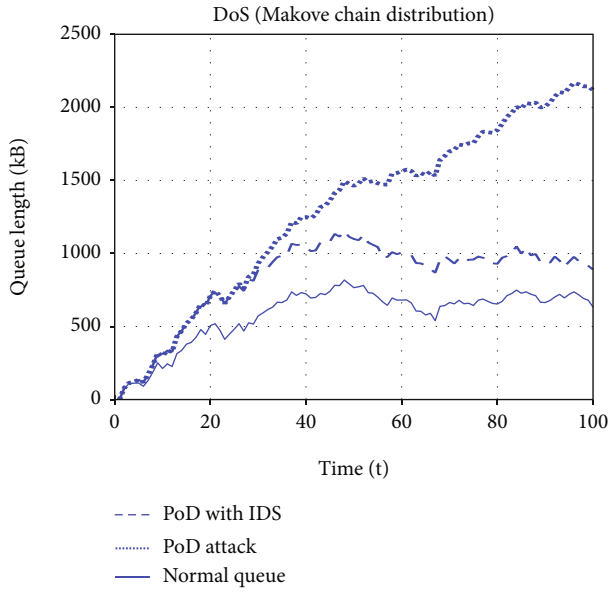


Figure 13: Queue length using PoD (Markov chain) with IDS.

Attack probability of being in state "2" at time "t" is proofed mathematically as

$$\begin{aligned} P(X_t = m) &= P(X_t = m, \cdots, X_0 = n) \\ &= \sum_{m,k,n \in S} P(X_t = m \mid X_{t-1} = k, \cdots, X_0 = m) \times P(X_{t-1} = k) \\ &= \sum_{m,k,h,n \in S} P(X_t = m \mid X_{t-1} = k) \times P(X_{t-1} = k \mid X_{t-2} \\ &= h, \cdots, X_0 = n) \times P(X_{t-2} = h) = \sum_{m,k,h,g,n \in S} P(X_0 = n) \\ &\times p_{ng} \times \cdots \times p_{hk} \times p_{km} = \left((\pi_0)_j (P)_{ij}^t\right)_{j=2}. \end{aligned} \tag{5}$$

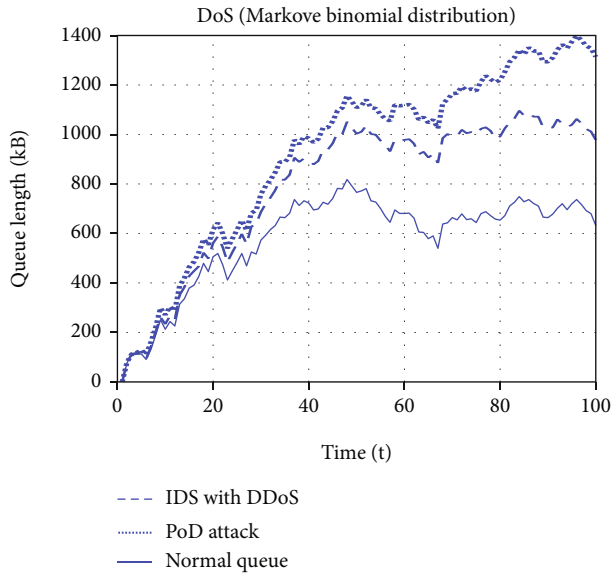


Figure 14: Queue length using PoD (Markov binomial) with IDS.

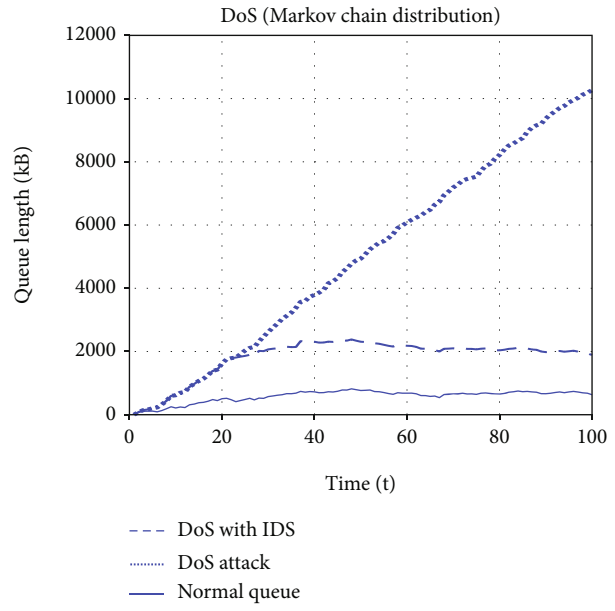


Figure 15: Queue length using DoS (Markov chain) with IDS.

Whereas,

$$\begin{aligned} (\pi_0)_i = (P(X_0 = n))_i &= P\left(\text{Attacker choose state } n = '1' \text{ to start}\right) \\ &= \begin{bmatrix} \pi_{0_1} \\ \pi_{0_2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \end{aligned} \tag{6}$$

However, the attack probability p_a at each time slot will change in sequence using random variables according to

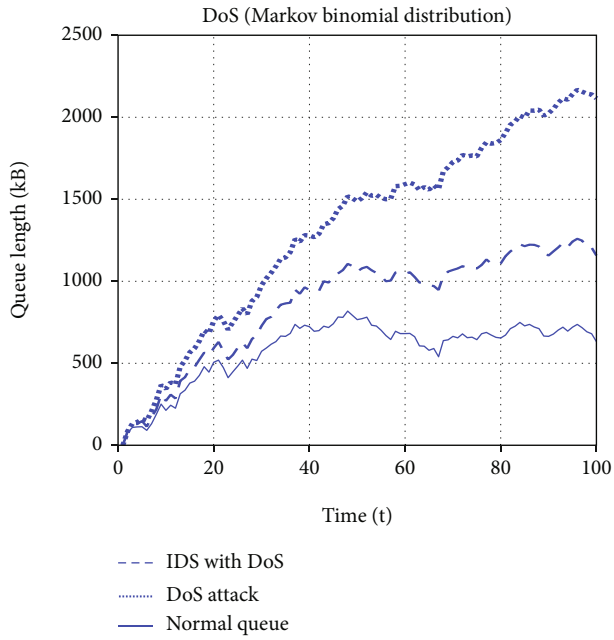


Figure 16: Queue length using DoS (Markov binomial) with IDS.

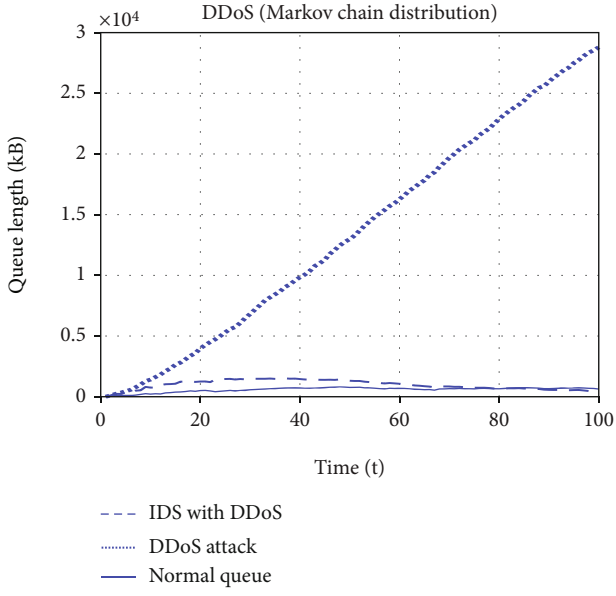


Figure 17: Queue length using DDoS (Markov) with IDS.

DTMC in blocks, and attack probability is shown in below Figure 8.

3.2. Markov Binomial Distribution. Binomial distribution is memory less scheme with having probability p_a, where attack at each time slot is stationary which can be symbolized as p_{a_0}. By simulating Markov binomial distribution, $\alpha = \beta = p_{a_0}$ are shown in below matrix from

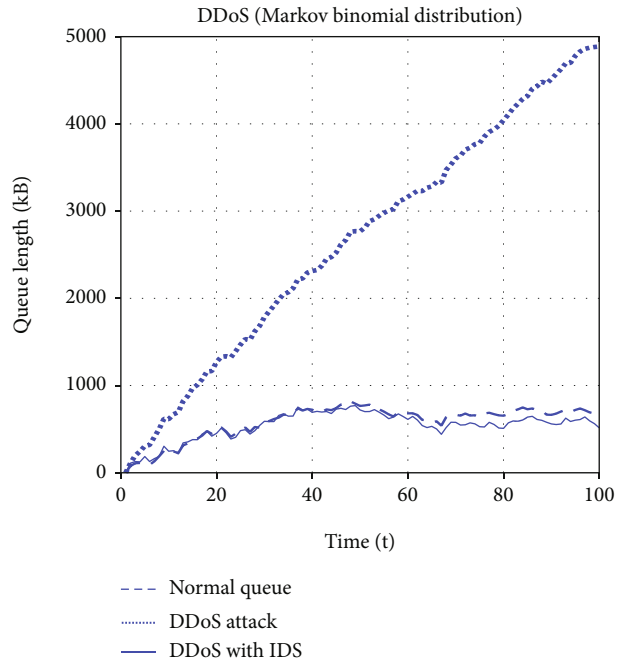


Figure 18: Queue length using DDoS (Markov binomial) with IDS.

$$(P)_{ij} = \left(p_{ij}\right) = \begin{bmatrix} 1 - p_{a_0} & p_{a_0} \\ 1 - p_{a_0} & p_{a_0} \end{bmatrix}. \quad (7)$$

Figure 8 elaborates attack probability changes with the passage of time, and the results are discussed using Figure 9 in which two states are discussed. Where state "1" is used for no attack, and state "2" represents attack.

4. Simulation Results

4.1. Without IDS. In simulation results, the attacker is attempting to use various flooding attacks such as DoS, DDoS, and PoD. Due to the aforementioned attacks, the ground BS is heavily buffered in a queue. The length of the queue for various attacks in order to impact the effect of attack probability on queue length for Markov chain and Markov binomial distribution, respectively, is shown in Figures 10 and 11. Markov chain distribution p_a attack is changing with the stream of time; due to that, queue length will escalate. Where using binomial distribution p_a attack probability must be constant; because of this reason, queue length will become very less in comparison with Markov chain distribution.

4.2. With IDS. Optimization of connection links will reshape the entire planet; therefore, safety of this society needs countermeasures to make the information-age secure. For the security of modeled smart IoT network having drones to stabilize path-flying things, detection system is launched to detect some cyber threats. Due to high network performance, the detection system attempts to trade-off between false positive and false negative probability. This concept

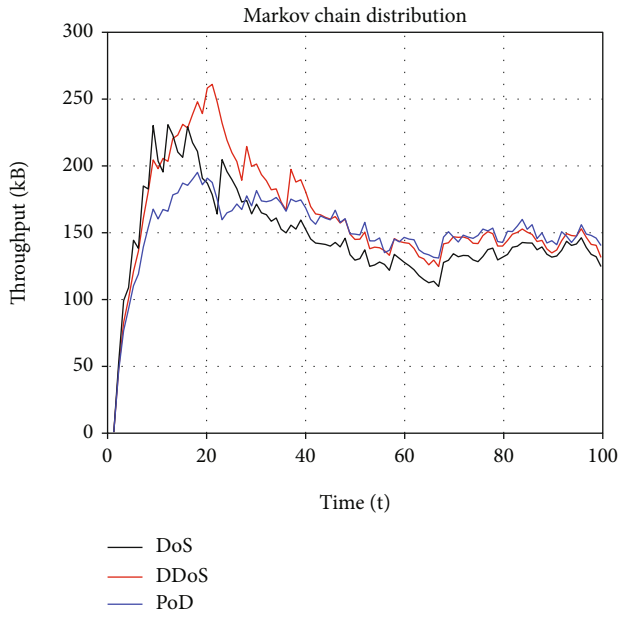


Figure 19: Average throughput using Markov chain.

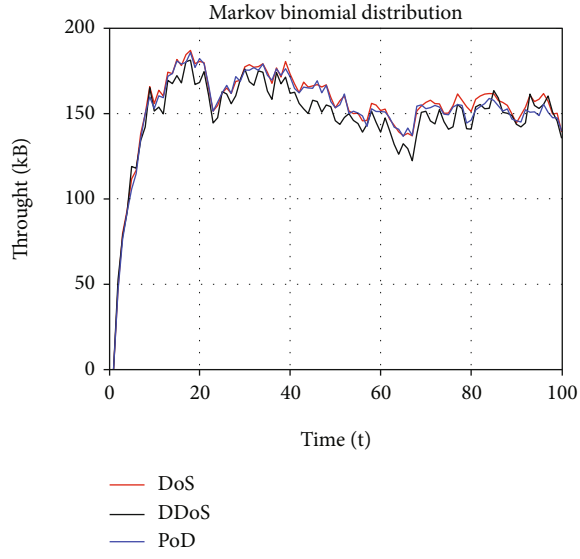


Figure 20: Average throughput using Markov binomial distribution.

assists researchers to have interconnectivity having maximum missed detection probability along with minimum false alarm prospects. The proposed IDS based on certain level m_{Th} try to prevent gateway queue lengths from rapidly increasing and maintaining them at the predictable level. Figure 12 shows the optimized certain level value per time for Markov chain and Markov binomial distributions.

The PoD with Markov chain using queue length is shown in Figure 13; while for Markov binomial distribution using security, attacks are discussed using Figure 14. Respectively, in Figures 15 and 16, the same techniques are utilized for DoS attack. However, similar schemes are incorporated for D-DoS threat in Figures 17 and 18. Throughput study of security attacks using Markov distribution and Markov binomial are having great impact on the data analysis which is shown in Figures 19 and 20.

5. Conclusion

Aerial ad hoc networks use to perform variety of tasks which include monitoring and collection of data from IoT networks. In flying networks, our main focus is to protect ground station from security attacks. While communication comprises drone-2-drone and land-station-2-aerial-vehicles which use IEEE 802.11 wireless technology to improve transmission routes. Intrusion detection system is the optimal way to deal with cyber threats. The proposed intrusion detection monitors incoming packets and filters them using Markov distribution. Markov chain stochastic process assists to find the gateway approach for flying vehicles. Intelligent intrusion detection controls flying networks to filter queue length data packets. The possibility of missed detection and false alarm is easily minimized. While buffer queue length will be maintained to normal level as demonstrated in the simulations. However, in future, machine learning techniques can be used to improve the aerial network security.

Data Availability

All the data is available in the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by Taif University Researchers Supporting Project number (TURSP-2020/126), Taif University, Taif, Saudi Arabia.

References

[1] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint relaying for flooding broadcast messages in mobile wireless networks," in *Proceedings of the 35th annual Hawaii international conference on system sciences*, pp. 3866–3875, Big Island, HI, USA, 2002.

[2] I. U. Khan, I. M. Qureshi, M. A. Aziz, T. A. Cheema, and S. B. H. Shah, "Smart IoT control-based nature inspired energy efficient routing protocol for flying ad hoc network (FANET)," *IEEE Access*, vol. 8, pp. 56371–56378, 2020.

[3] M. Ahmed and A. K. Pathan, "False data injection attack (FDIA): an overview and new metrics for fair evaluation of its countermeasure," *Complex Adaptive Systems Modeling*, vol. 8, no. 1, p. 4, 2020.

[4] A. Abdollahi and M. Fathi, "An intrusion detection system on ping of death attacks in IoT networks," *Wireless Personal Communications*, vol. 112, no. 4, pp. 2057–2070, 2020.

[5] D. Pliatsios, P. Sarigiannidis, T. Lagkas, and A. G. Sarigiannidis, "A survey on SCADA systems: secure protocols, incidents, threats and tactics," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1942–1976, 2020.

[6] W. A. Al-Khater, S. Al-Maadeed, A. A. Ahmed, A. S. Sadiq, and M. K. Khan, "Comprehensive review of cybercrime detection techniques," *IEEE Access*, vol. 8, pp. 137293–137311, 2020.

[7] M. Albalawi and H. Song, "Data security and privacy issues in swarms of drones," in *2019 Integrated communications, navigation and surveillance conference (ICNS)*, pp. 1–11, Herndon, VA, USA, 2019.

[8] J. Chen, Z. Feng, J. Wen, B. Liu, and L. Sha, "A container-based DoS attack-resilient control framework for real-time UAV systems," in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1222–1227, Florence, Italy, 2019.

[9] S. Garg, G. S. Aujla, N. Kumar, and S. Batra, "Tree-based attack–defense model for risk assessment in multi-UAV networks," *IEEE Consumer Electronics Magazine*, vol. 8, no. 6, pp. 35–41, 2019.

[10] R. Mohan, C. V. Raj, P. Aswathi, and R. R. Bhavani, "UAV based security system for prevention of harassment against woman," in *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, pp. 874–879, Kannur, 2017.

[11] M. Podhradsky, C. Coopmans, and N. Hoffer, "Improving communication security of open source UAVs: encrypting radio control link," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1153–1159, Miami, FL, USA, 2017.

[12] C. Rani, H. Modares, R. Sriram, D. Mikulski, and F. L. Lewis, "Security of unmanned aerial vehicle systems against cyber-physical attacks," *Journal of Defense Modeling and Simulation*, vol. 13, no. 3, pp. 331–342, 2016.

[13] F. Restuccia, S. D'Oro, and T. Melodia, "Securing the internet of things in the age of machine learning and software-defined networking," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4829–4842, 2018.

[14] F. E. Salamh, U. Karabiyik, M. Rogers, and F. Al-Hazemi, "Drone disrupted denial of service attack (3DOS): towards an incident response and forensic analysis of remotely piloted aerial systems (RPASs)," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 704–710, Tangier, Morocco, 2019.

[15] A. Sehrawat, T. A. Choudhury, and G. Raj, "Surveillance drone for disaster management and military security," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 470–475, Greater Noida, 2017.

[16] C. Wang, L. Zhu, L. Gong et al., "Accurate sybil attack detection based on fine-grained physical channel information," *Sensors*, vol. 18, no. 3, p. 878, 2018.

[17] T. Lagkas, V. Argyriou, S. Bibi, and P. Sarigiannidis, "UAV IoT framework views and challenges: towards protecting drones as "things"," *Sensors*, vol. 18, p. 4015, 2018.

[18] Z. Zaheer, A. Usmani, E. Khan, and M. A. Qadeer, "Aerial surveillance system using UAV," in *2016 Thirteenth International Conference on Wireless and Optical Communications Networks (WOCN)*, pp. 1–7, Hyderabad, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7154705, 13 pages
https://doi.org/10.1155/2021/7154705

Research Article

Quasi-Identifier Recognition Algorithm for Privacy Preservation of Cloud Data Based on Risk Reidentification

Huda O. Mansour[1,2], **Maheyzah M. Siraj**[2], **Fuad A. Ghaleb**[1], **Faisal Saeed**[3], **Eman H. Alkhammash**[4] and **Mohd A. Maarof**[1]

[1]*Faculty of Engineering, School of Computing, Universiti Teknologi Malaysia (UTM), Johor 81310, Malaysia*
[2]*Department of Computer Science, Faculty of Computer Science and Information Technology, University of Kassala, Kassala 31111, Sudan*
[3]*College of Computer Science and Engineering, Taibah University, Medina, Saudi Arabia*
[4]*Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Fuad A. Ghaleb; abdulgaleel@utm.my

Received 30 April 2021; Revised 26 June 2021; Accepted 9 August 2021; Published 26 August 2021

Academic Editor: Ihsan Ali

Cloud computing plays an essential role as a source for outsourcing data to perform mining operations or other data processing, especially for data owners who do not have sufficient resources or experience to execute data mining techniques. However, the privacy of outsourced data is a serious concern. Most data owners are using anonymization-based techniques to prevent identity and attribute disclosures to avoid privacy leakage before outsourced data for mining over the cloud. In addition, data collection and dissemination in a resource-limited network such as sensor cloud require efficient methods to reduce privacy leakage. The main issue that caused identity disclosure is quasi-identifier (QID) linking. But most researchers of anonymization methods ignore the identification of proper QIDs. This reduces the validity of the used anonymization methods and may thus lead to a failure of the anonymity process. This paper introduces a new quasi-identifier recognition algorithm that reduces identity disclosure which resulted from QID linking. The proposed algorithm is comprised of two main stages: (1) attribute classification (or QID recognition) and (2) QID dimension identification. The algorithm works based on the reidentification of risk rate for all attributes and the dimension of QIDs where it determines the proper QIDs and their suitable dimensions. The proposed algorithm was tested on a real dataset. The results demonstrated that the proposed algorithm significantly reduces privacy leakage and maintains the data utility compared to recent related algorithms.

1. Introduction

In the modern information age, many companies are using external sources of data for processing, storing, or obtaining some services such as data mining. Unlimited computational resources, reduced costs, nonburden of maintenance, and nondiligence to learn the skills of proficiency in certain services, all of these were temptations to advance to the modern change. However, there are still security and privacy concerns that hinder the use of the features offered by the cloud [1]. Numerous studies clarified that attackers often reveal the information from third-party services or third-party clouds [2]. For example, one of the security breaches in October 2014 was a breakthrough for Dropbox. The attackers stole 700 user passwords to obtain cash values of its Bitcoins (BTC). In 2015, a lot of users' information, which exceeds 4 million, such as the user's name, date of birth, address, e-mail, phone number, and other sensitive data, were leaked through the TalkTalk service provider in the UK. In 2016, Time Warner, one of the largest cable television companies in the United States, has announced that about 32 million passwords and e-mail of the users have been stolen via an attacker. In 2017, more than 200 million data of the users containing users' names, phone numbers,

e-mail addresses, home addresses, and other data have been disclosed through the API of McDelivery Company in India [2, 3]. A fresh security violation in Google displayed that any administrator of the server who has access to the secret information can misuse it easily. The worst problem is that administrator of the honest-but-curious server can violate privacy without being discovered [4].

Three kinds of the disclosure can cause privacy leakage, identity disclosure, attribute disclosure, and membership disclosure [5]. In attribute disclosure and identity disclosure, the intruder identifies that the tuple of the target individual is found in the released dataset and he aims to acquire some private/sensitive data about that individual from the released dataset [6]. Serious issues that lead to identity disclosure are quasi-identifier (QID) value linking and the attacker's knowledge background. The QIDs are the dataset attributes that if each of them is considered separately does not distinguish the individual, but when several attributes are combined they can give a distinctive identification of individuals [7]. For example, when looking at the attributes of date of birth, gender, and ZIP code together, one can reidentify the individuals as stated in [8]. Reidentification of the individuals through linking their QIDs leads to what are called linking attacks. Therefore, the careless publication of QIDs will lead to leakage of privacy [9].

One of the popular practices to avoid privacy leakage is anonymization. The anonymization can be performed via several types of transformations, by removing the values, changing the structure, replacing the values by taxonomy, and combining the values. The anonymization-based methods use one or a combination of operations to accomplish an optimum level of concealment [10]. A commonly utilized privacy criterion of anonymization is k-anonymity introduced by Sweeney [8]. The k-anonymization model is aimed at making any record in the released dataset that cannot be distinguished from at least $(k-1)$ other records [1, 11]. To avoid the linking attacks, k-anonymization can be used. The effective method to determine the real QIDs is the primary issue for privacy-preserving methods based on k-anonymity or other anonymization models seek to prevent QID linking. While most of the current methods neglected this issue or just determine QIDs manually, this reduces the validity of the anonymization method as well as negatively affects the usefulness of anonymous data [9]. This study is aimed at overcoming the identity disclosure resulting from QID linking and reducing the leakage of privacy by proposing a QID recognition (QIR) algorithm based on risk rate reidentification. The proposed algorithm comprises two main stages: (1) attribute classification (or QIDs Recognition) and (2) QID dimension identification. The algorithm works based on the reidentification of risk rate for all attributes and the dimension of QIDs where it determines the proper QIDs and their suitable dimensions. Figure 1 shows the cause-effect diagram of privacy leakage. The dark boxes in Figure 1 explain the privacy leakage causes addressed by the proposed QID recognition (QIR) algorithm in this study. As shown in Figure 1, it is essential to properly identify the QID attributes to overcome the identity disclosure to reduce the leakage of privacy resulting from QID linking. This paper is made up of 5 sections. Section 2 describes the state of the art of privacy-preserving data mining (PPDM) over the cloud, whereby some of the current methods and algorithms that address the issue of identification QIDs accurately to avoid identity disclosure are presented. A detailed description of the proposed algorithm has been provided in Section 3. Section 4 demonstrates the experimental evaluation, discussion, and comparison with related work. Section 5 concludes this work.

2. Related Work

The research of privacy-preserving outsourced data focuses on anonymization-based methods [12–18], cryptographic-based methods [19–24], hybrid methods [2, 25–27], and methods that seek to improve the data utility [26, 28, 29]. Some recent studies have demonstrated the privacy requirements of incremental datasets [30–32] and multiple sensitive attributes [33–35]. However, most of these studies neglected the issue of identification of the right QIDs, despite its importance in the success of the anonymity process. Few of these studies have attempted to introduce methods so that identification of the QIDs is required in the anonymization process, as presented in the next section.

Huang and others [36] introduce a new method that depends on the hypergraph to find a group of related views and QID set. This method maps the group of related views into a hypergraph and includes all paths available between every two nodes instead of finding the group of related views. The weakness of this method is that the QID group produced may include so many attributes. Further, it has high computational complexity resulting from the process of degeneration of the common graph from the hypergraph.

Omer and Mohamad [37] introduce a new method to select a quasi-identifier (QID) to achieve k-anonymity. Selective and decompose algorithms depend on nominating multiple attributes as a set and then generating power set $P(S)$ for them. Following that, the distinct values of the power set $P(S)$ elements were computed and listed in a table. Finally, the candidate element from the power set is the element with the maximum distinct value. The main problem in this method is selecting the primary nominate set of attributes, where the accuracy of the selection depends on the user experience [9]. Furthermore, it is impractical to generate $P(S)$ if the number of attributes is big (e.g., more than 8).

Y. J. Lee and K. H. Lee [38] examine the factors and the likelihood of an individual reidentified for medical information through inferable QIDs. The QIDs were considered as database variables that enable the reidentification of individuals by linking their QIDs with available external information or a specific individual. They selected five factors to form QID attributes to prevent patient privacy violations. The factors were selected based on their influence on the likelihood of reidentification and the possibility of inferring it from background knowledge. One of the disadvantages of this study is that the QIDs that can be extracted to reidentify patients' records may exceed 5. Besides, the paper focused only on the problem of reidentification of patients' records and avoiding leakage of privacy in the medical

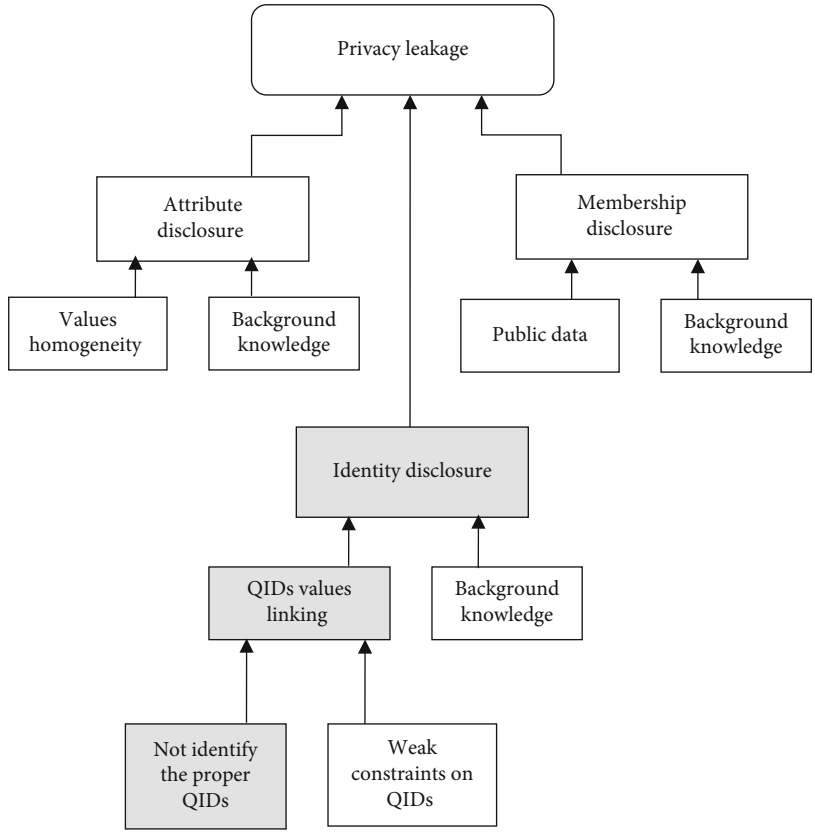


Figure 1: Privacy leakage causes addressed by the QIR algorithm.

records, lacking a public method that could be used for general data publishing. Bampoulidis and others [7] assume that some QIDs are more important than others (i.e., in data mining/analysis) and, therefore, should be distorted as little as possible in the anonymization process. They present a tool to address the issue of QIDs by utilizing a local recoding algorithm for k-anonymity. The tool outperforms the ARX (data anonymization tool) in terms of dataset quality. The major problem with this method is that it depends on the user in defining the QID attributes, giving priority to each attribute, as the user relies on his personal experience in determining the QID attributes, which are usually not accurate.

Kaur and Agrawal [10] study the impact of QIDs on the anonymization process. They gave new ways to consider before choosing the quasi-identifiers. The reidentification risks have been examined using different QIDs, diverse parameters, and different sizes of a data sample. The results of their work showed that when making the variance in selecting the QIDs for anonymization operation, note that the risk of reidentification increases when the number of QIDs increases, and it decreases when using QIDs that contain fewer categories. Although it is good to take into account these observations before starting the anonymity process, it should be noted that these observations extracted by the study are not fixed and may change from one dataset to another.

Wong and others [39] do not reveal the complete set of quasi-identifiers (QID) to the data collector before and after the data anonymization process. They believed that the QIDs can be both sensitive values and identifying values; they allow the respondents/data owners to hide sensitive QID attributes from other parties. The first issue with this method is that the QID attributes that respondents consider them are sensitive which may contain data that are very useful in mining or may adversely affect mining outcomes. The second issue is if respondents submit inaccurate data, there is no guarantee of the usefulness of the results obtained from data analysis.

Sei and others [40] consider that some QIDs are regarded as sensitive QIDs and they propose novel privacy models, namely, $(l1, \cdots, lq)$ – diversity and $(t1, \cdots, tq)$ – closeness, and a method that can treat sensitive QIDs. Their proposed method comprises two algorithms: anonymization and reconstruction algorithms that can treat sensitive QIDs. Although this method can perform anonymity while preserving the quality of the data, it suffers from the problem of the Wong [39] method; this is because there is no effective method to accurately determine which of the QID attributes is considered sensitive QIDs.

Victor and Lopez [41] offer a (k, n, m) anonymity method for sensitive/private data based on the k-anonymity. The graph algorithms were used to perform QIDs and are moreover been improved by selecting similar QIDs based on the composite and derived attributes. The set of QIDs obtained from the methods in [36, 41] may include too many attributes, which increases the information loss in models based on generalizations like the k-anonymity [9].

3. The Proposed QID Recognition Algorithm

There are two main stages involved in the QID recognition algorithm (QIR) to prevent privacy leakage of outsourced data. *First*, classify the dataset attributes into quasi-identifiers (QIDs), sensitive attributes (SAs), and nonsensitive attributes (NSs). That is, each attribute in the dataset is classified into one of the aforementioned groups (QIDs, SAs, or NSs). In the attributes' classification (QID recognition) stage, the IDs (identifier attributes) are usually removed from the dataset by the data owner. The quasi-identifiers (QIDs) are the attributes that, when linked together, define the individual, for example, age, gender, and ZIP. The sensitive attributes (SAs) are the attributes that explain sensitive/private information about an individual such as medical information, financial records, and location. Meanwhile, the nonsensitive attributes (NSs) are the other attributes in the dataset that do not fall under the previously mentioned categories, as they do not help reidentify the identity of the individual, for example, state and religious attributes. In the basic privacy models (such as *k*-anonymity [7–9, 11–13, 18, 28], l-diversity [40, 42], and t-closeness [34, 43]), the attributes of a dataset were categorized into two groups: sensitive and nonsensitive. Meanwhile, most of the recent researchers such as in [9, 44–47] divide the dataset attributes into three types: QID, SA, and NS (not including identifiers) directly. Accordingly, the classification of dataset attributes in this study is divided into three types of QID, SA, and NS (not including identifiers) utilizing the same definitional meaning of each category as in the previous work in [9, 44–47].

Second, determine the actual dimension of QIDs that should be used in an anonymization operation that will achieve optimum case. If the set of QIDs contains too many attributes, the loss of information caused by generalization will be exacerbated. Nonetheless, sometimes the minimal set of QID does not imply the most appropriate privacy protection setting because the method does not consider what attributes the adversary could potentially have [37]. Therefore, we need a mechanism that determines the appropriate dimension of the QIDs to avoid these problems. In the QID dimension determining stage, the proposed algorithm performs this task. Figure 2 illustrates the general procedure of the two main phases of the QIR algorithm. The following subsections explain these two stages in more detail.

3.1. QID Recognition Stage. In this stage, the algorithm classifies the attributes depending on the reidentification risk rate for each attribute in the dataset, and then, the risk rate of the attribute is compared to the threshold values of the classification. As shown in Figure 2, the attribute classification stage comprises four main activities. These activities include (1) dataset preprocessing, (2) computing risk rate for all attributes, (3) selecting the classification thresholds, and (4) classifying the attributes according to the selected thresholds.

In the first activity, the dataset is preprocessed which includes filling the missing values, fixing the inconsistencies in the dataset, and data normalization. Then, in the second activity, the risk rate is computed according to the g-distinct which is adopted in computing the reidentification risk rate [48]. A detailed description of the g-distinct method is presented in the next section. In the third activity, the classification thresholds were selected based on the maximum and minimum risk of reidentification as follows. These thresholds are denoted by β and α in this study; α threshold represents the maximum risk of reidentification of the individual while β represents the minimum risk of reidentification. The threshold values can be determined by the user or the data owner after calculating the reidentification risk for all attributes. Based on percentages of the highest and lowest attribute risk, one can choose the α value to be less than the highest risk value and choose the β value to be less than the lowest risk value. The nature of the data and the degree of importance of each attribute affect the selection of the threshold values. So, these thresholds are adjustable and differ from one dataset to another. For instance, let the dataset (D) contain attributes $(A_1, A_2, \cdots, A_n)$, i.e., $D = A_1, A_2, \cdots, A_n$; let $\beta = 0.05\%$ and $\alpha = 30\%$. Let Rrisk_{A_i} be the reidentification risk of attribute A_i and $\text{Rrisk}_{A_i} = 35\%$. As $\text{Rrisk}_{A_i} > \alpha$, then A_i is classified as SA. Suppose Rrisk_{A_3} and Rrisk_{A_5} are 23 and 0.01, respectively, then A_3 is classified as QID while A_5 will be classified as NS, respectively. Reidentification risk rate of attribute A_i computes the degree that makes the records distinguished based on this attribute. Finally, the fourth activity includes classifying the attributes according to the selected thresholds using rules represented by *if-else* testaments (see Algorithm 1, lines 27–39). In the following subsection, a detailed description of computing the reidentification risk rate (g-distinct) is presented. More explanation of the QID recognition stage is also presented.

3.2. g-Distinct. The g-distinct is adopted in computing the reidentification risk rate [48]. A person or record in any dataset is said to be unique if he/she or it has a combination of attributes that is not for someone/record else. The person/record is g-distinct if their combination of attributes is matching to g-1 or less than other people/records in the dataset [48]. Thus, uniqueness is the base situation of 1-distinct. In general, g-distinct is the total of the number of subgroups with i individuals, which is computed as

$$h_n(g) = \sum_{i=1}^{g} i.f_n(i), \tag{1}$$

where $f_n(i)$ refers to the expected number of subgroups with i individuals that can be derived from a given aggregated group and g represents the whole number of individuals in a subgroup. That is, g is associated with the g-distinct to represent the number of distinguished individuals in the subgroup. For example, when we say 3-distinct, it means that three individuals have common QID characteristics out of the total number of people g in the subgroup. The sum of all g-distinct of individuals in a specific attribute represents the reidentification risk rate that the attribute potential to cause it. We can compute the general risk of the whole

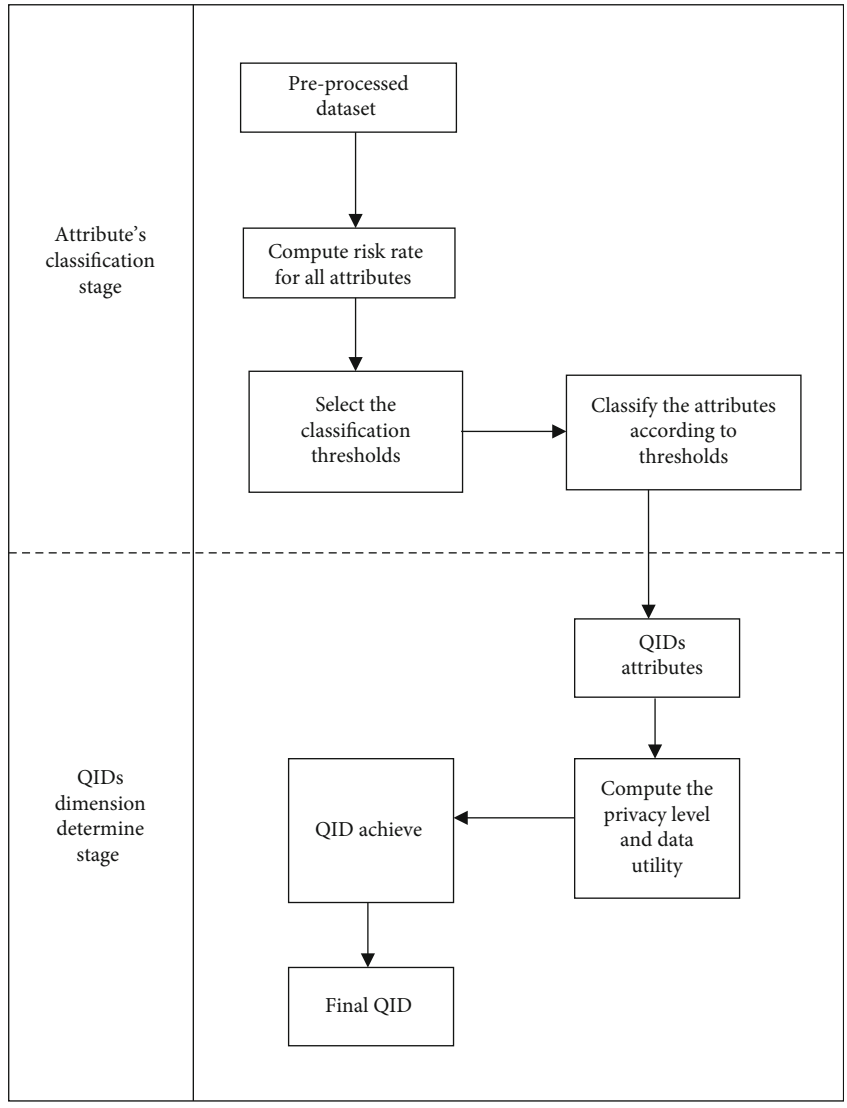


Figure 2: The general procedure of the proposed QIR algorithm.

dataset through equation (2) where b is the number of possible subgroups.

$$R_n^j(g) = \left(\frac{j}{n}\right) b^{1-n} \left(b^n - (b-1)^n\right). \quad (2)$$

Finally, the attribute classification stage returns the reidentification risk rate for each attribute in the dataset. Based on the resulting reidentification risk rates, the dataset attributes are classified to sensitive and nonsensitive according to the rate of the reidentification risk for each attribute in addition to threshold values β, α. The outcomes of this stage will be input into the QID dimension identification stage to determine the dimension of QIDs that is suitable to achieve optimal privacy requirements. The practical steps of the classification stage are explained by Algorithm 1. Lines 2–16 in Algorithm 1 are to compute the g-distinct for all dataset attributes while lines 18–26 are to calculate the reidentification risk rate based on the attributes' g-distinct. Finally, lines 28–40 addressed the process of attribute classification using the reidentification risk rate of each attribute to produce three categories of attributes: QIDs, SAs, and NSs.

The importance of this stage of the proposed algorithm represented by Algorithm 1 is that it contributes to reducing the attribute disclosure resulting from linking the QID values due to a weakness/failure in defining the QID characteristics correctly. This contribution helps in minimizing the leakage of information and avoiding privacy violations.

3.3. QID Dimension Identification Stage. This stage of the algorithm is aimed at determining the best dimension of QIDs that will achieve optimum cases. The optimum case gives high privacy with a high/reasonable percentage of preserving data quality. In other words, it has high privacy gain (PG) with high/reasonable nonuniform entropy (NUE). Algorithm 2 describes the implementation steps for this stage. The algorithm takes a sample of data with the QID that has the highest reidentification risk rate. Following that, the QIR calculates the PG and NUE base on k-anonymity

Input: dataset D, β, α.
Output: classified dataset.
1: //*Compute g -distinct for all dataset tuples for each attribute.*
2: $Dg_{\text{Attr}} \longleftarrow g$-distinct of the attribute (Attr)
3: $n \longleftarrow$ attribute domain
4: $m \longleftarrow$ tuple domain
5: Attr $\epsilon\, n$
6: $g \epsilon\, m$
7: $tv \longleftarrow$ attribute value of a specific tuple
8: $\text{Attr}_{Dg}[i][j] = 0$
9: **For** $i \coloneqq 1$ to n.length **do**
10: **For** $j \coloneqq 1$ to m.length **do**
11: $Dg_{\text{Attr}[i]} = 1/\int (tv)_j$
12: Attr_$Dg\,[i][j]$ = Attr_$Dg\,[i][j] + Dg_{\text{Attr}(i)}$;
13: $j = j + 1$;
14: **End**
15: $i = i + 1$;
16: **end**
17: //*Compute reidentification risk rate for all dataset attributes.*
18: Rrisk_Attr$[i] = 0$
19: $\text{Rrisk}_{\text{Attr}} \longleftarrow$ reidentification risk rate of Attr
20: **For** $i \coloneqq 1$ to Attr_$Dg\,[i]$.length **do**
21: **For** $j \coloneqq 1$ to m.length **do**
22: $\text{Rrisk}_{\text{Attr}}[i] = \text{Rrisk}_{\text{Attr}}[i] + Dg_{\text{Attr}[i][j]}$
23: $j = j + 1$;
24: **End**
25: $i = i + 1$;
26: **End**
27: //*Classify the attributes based on risk rate and threshold values.*
28: QIDs $[\,] = 0$
29: SAs$[\,] = 0$
30: NSs $[\,] = 0$
31: **For** $i \coloneqq 1$ to $\text{Rrisk}_{\text{Attr}}\,[i]$.length **do**
32: **If**($\text{Rrisk}_{\text{Attr}}\,[i]$in range($\beta$))
33: QIDs$[i]$ = QIDs $[\,] + \text{Rrisk}_{\text{Attr}}[i]$;
34: **Else If**($\text{Rrisk}_{\text{Attr}}\,[i]$in range($\alpha$))
35: SAs$[i]$ = SAs$[\,] + \text{Rrisk}_{\text{Attr}}\,[i]$;
36: **Else**
37: NSs $[i]$ = SAs $[\,] + \text{Rrisk}_{\text{Attr}}[i]$;
38: $i = i + 1$;
39: **end**
40: **Return**(QIDs$[\,]$, SAs$[\,]$, NSs$[\,]$)

Algorithm 1: Attribute classification.

through equations (3) and (4). In the next step, the QID number is increased, and PG and NUE are calculated again and so on until all QIDs are finished.

Finally, the algorithm determines the optimum case that gives high privacy with a high/reasonable percentage of preserving data quality. The best QID dimension is the QIDs with the optimum case. Algorithm 2 provides the executive steps of this stage; lines 5–12 implement the anonymization by k-anonymity on a sample of the dataset. It begins with QID that has the highest reidentification risk rate. After that, the algorithm calculates the privacy gain (PG) and nonuniform entropy (NUE) through equations (3) and (4). Then, the QID number is increased; PG and NUE have been calculated repeatedly until all the QIDs are finished. Lastly, in lines 13–15, the algorithm determines the best QID dimension (QidD) that achieves the optimum case to be involved in the anonymization process.

It was observed in study [9] that in most cases, when the QID dimension is large, the data loss increases. However, when the QID dimension is small, the privacy protection is not applied optimally because one cannot know what the actual QIDs an attacker possesses [37]. Therefore, determining an appropriate QID dimension is important to reduce data loss.

3.4. Performance Measures. Two performance evaluation measures were used in this study: the privacy gain (PG) and the nonuniform entropy (NUE). More explanation and the derivation of these measures are presented in the following subsections.

Input: dataset sample d, QIDs [ ], privacy parameter k.
Output: optimal dimension of QIDs.
1: QidD ⟵ dimension of QIDs
2: QidD ϵ QIDs [ ]
3: Optimal_QidD ⟵ Optimal dimension of QIDs
4: QidD[ ] = 0
5: **For** i = 1 to QIDs [ ].length **do**
6: QidD[i] = QidD[ ] + QIDs[i];
7: Anonymized$_{\text{data}}$[i] = k-anonymity(d, QidD[i], k);
8: PG [i] = Privacy_gain(Anonymized_data[i]);
9: NUE [i] = Nonuniform_Entroy (Anonymized_data [i]);
10: Difference[i] = PG [i] − EIL[i];
11: $i = i + 1$;
12: **end**
13: **If**((PG[ ] == **max**)&&(NUE[ ] == **max**))
14: Optimal_QidD[ ] = QidD[i];
15: **Return**(Optimal$_{\text{QidD}}$[ ]).

Algorithm 2: QID dimension identification.

3.4.1. The Privacy Gain. To evaluate the privacy level for the proposed algorithm, equation (3) and Definition 1 are used as follows.

$$\text{PG} = A_{t(\text{gen})-}A_{b(\text{gen})}, \tag{3}$$

where $A_{t(\text{gen})}$ is anonymity after generalization (gen) and $A_{b(\text{gen})}$ is anonymity before generalization [27, 49, 50].

Definition 1. Anonymity quasi-identifier: a quasi-identifier qid is an anonymity quasi-identifier if $|\text{QIG}(\text{qid})| = \min_{\text{qid}' \epsilon \text{QID}}|\text{QIG}(\text{qid}')|$, where $||$ represents the size of a QI group [50].

3.4.2. Nonuniform Entropy. In the context of data deidentification, the nonuniform entropy is to compare the frequencies of attribute values in the transformed dataset according to frequencies in the input dataset; it was originally introduced as a model for measuring the loss of information [51]. When a dataset D is transformed into another dataset D', nonuniform entropy is defined as

$$\Delta\left(D, D'\right) = \sum_{x \in D} -\log\left(\frac{f(D, x)}{f\left(D', x\right)}\right). \tag{4}$$

4. Experimental Evaluation

In this section, the experimental evaluation of our implementation algorithm will be presented in terms of PG and NUE. In Dataset Setup, we describe the datasets we have used for running the experiments and the experimental environment setup. In Experimental Results, we present the first set of experiments and provide the results from our algorithm. In Performance Benchmark and Discussion, we provide benchmark and discussion results of our algorithm against a close recent algorithm introduced by Omer and Mohamad [37].

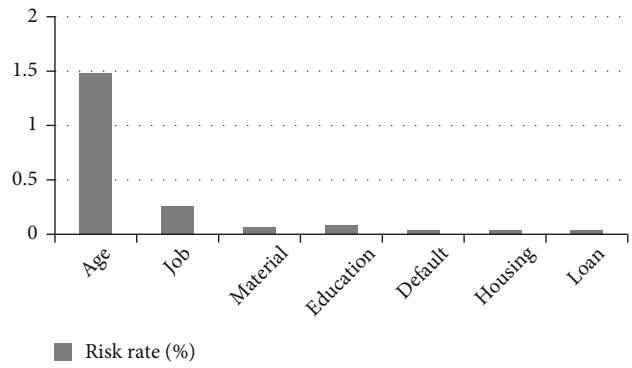


Figure 3: Risk rate of the bank dataset attributes.

4.1. Dataset Setup. Two real-life datasets from the University of California–Irvine were used in this study to demonstrate the performance of the proposed algorithms. The first is the bank direct marketing dataset [52]. The bank dataset consists of 17 attributes and 45,211 tuples and does not include any missing values. The dataset attributes are divided into three divisions which are (1) data of bank clients: age, job, marital status, education, default, balance, housing, and loan; in this paper, we will consider these attributes because these attributes are significant for bank clients and reidentification purposes; (2) data related to the last contact of the current campaign; and (3) other attributes like the campaign and days. The second dataset is the adult dataset [53] used as a standard for anonymization algorithm evaluation [7] consisting of 48,842 census records and 15 attributes.

ARX data anonymization software is open source introduced and developed by Prasser et al. [54] for data anonymization; we used it to implement the algorithms as explained in the following sections. The experiments were executed on

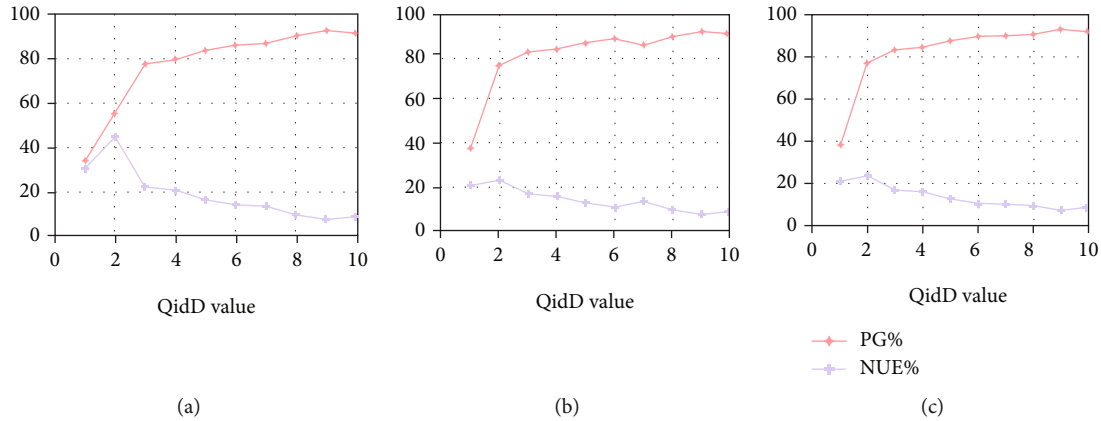


Figure 4: (a–c) The best QidD selection for the bank dataset by QIR on different k values.

Table 1: Classification of the bank dataset.

Classification	Threshold value $\alpha = 30, \beta = 0$	Attributes
SAs	Rrisk > α	Balance
QIDs	$\beta \leq$ Rrisk < α	Age, job, education, and marital status
NSs	Rrisk < β	Default, housing, and loan

Table 2: Classification of the adult dataset.

Classification	Threshold value $\alpha = 0.2, \beta = 0.01$	Attributes
SAs	Rrisk > α	Capital gain, capital loss
QIDs	$\beta \leq$ Rrisk < α	Hours-per-week, work-class, age, native-country, education, education-num, occupation, marital-status, relationship, and race
NSs	Rrisk < β	Sex, income

Table 3: Experimental results for selecting the best QidD in the adult dataset.

QID value	$k = 5$		$k = 15$		$k = 25$	
	PG %	NUE %	PG %	NUE %	PG %	NUE %
1	33.8	30.41	38.35	21.05	38.35	21.05
2	55.34	44.65	76.86	23.13	76.86	23.13
3	77.94	22.05	83.17	16.82	83.17	16.82
4	79.53	20.46	84.39	15.6	84.39	15.6
5	83.62	16.37	87.48	12.51	87.48	12.51
6	85.91	14.08	89.56	10.43	89.56	10.43
7	86.65	13.34	86.65	13.34	89.69	10.3
8	90.51	9.48	90.51	9.48	90.51	9.48
9	92.68	7.31	92.68	7.31	92.68	7.31
10	91.59	8.4	91.59	8.4	91.59	8.4

a machine with an Intel Core i7 2.7 GHz processor with 8 GB RAM, under Windows 10.

4.2. Experimental Results. The first experiment is to classify the dataset attributes according to their risk rate. Figures 3 and 4 illustrate the risk rate for bank attributes and adult attributes, respectively.

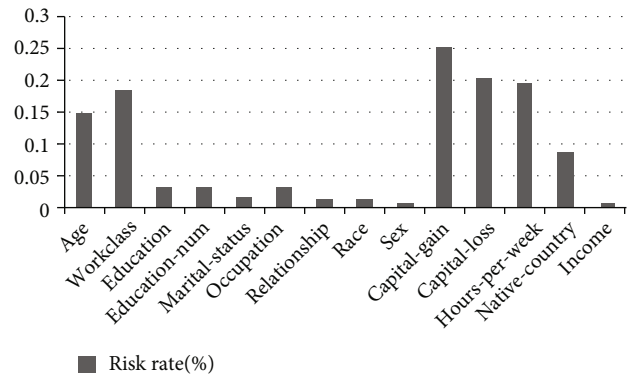


Figure 5: Risk rate of the adult dataset attributes.

For the bank dataset, we identify α and β as $\alpha = 30, \beta = 0$. Table 1 demonstrates bank attribute classification. In the adult dataset, we add $\alpha = 0.2, \beta = 0.01$ to classify the attributes. Table 2 demonstrates the classification of the adult dataset. Because the "balance" attribute has a risk of 52.04 %, which is large compared to other attributes, it is excluded

Table 4: Experimental results for selecting the best QidD in the bank dataset.

QidD	QID	$k = 5$		$k = 15$		$k = 25$	
		PG %	NUE %	PG %	NUE %	PG %	NUE %
1	Age	**23.89**	**45.28**	**36.12**	**17.27**	**36.12**	**17.27**
2	Age, job	21.83	36.65	21.83	36.65	21.83	36.65
3	Age, job, marital status	15.83	40.35	16.67	37.18	17.94	32.37
4	Age, job, marital status, education	14.88	35.93	14.88	35.93	16.43	29.26

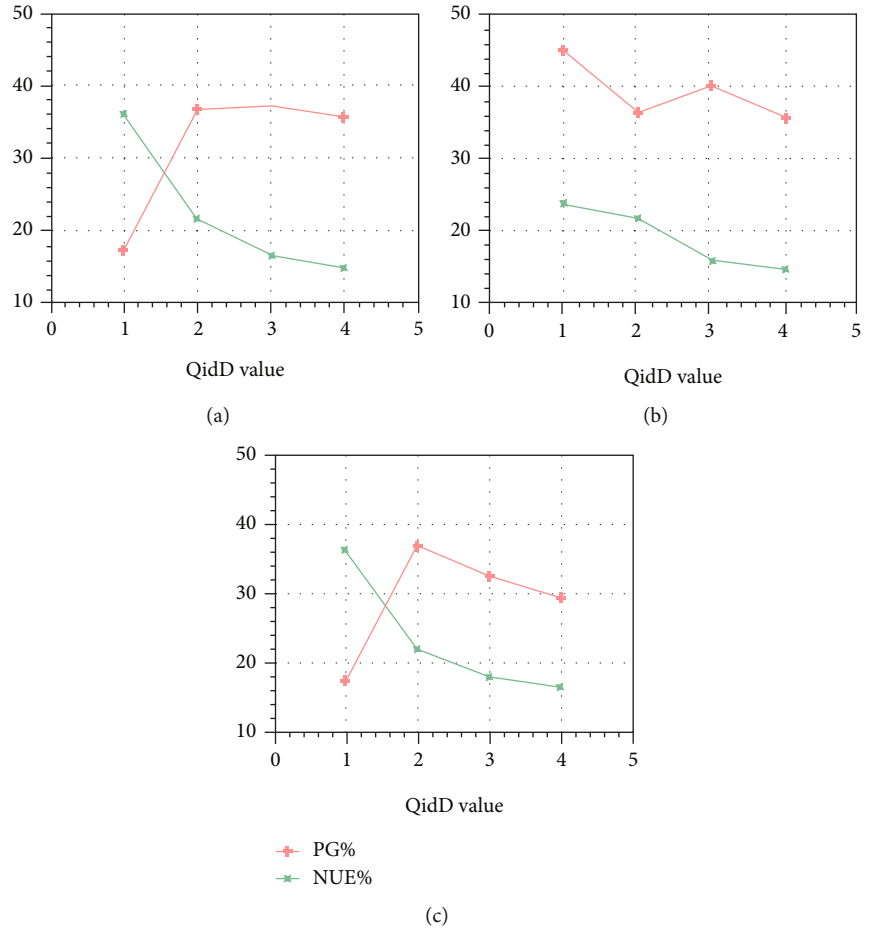


Figure 6: (a–c) The best QidD selection for the bank dataset by QIR on different k values.

from Figure 3 to highlight the difference between the attributes that have relatively small risk values.

After calculating the risk rate of each attribute in the dataset, the attribute is classified according to the selected threshold α and β as was explained in QID Recognition Stage. Tables 1 and 2 show the classification results of the bank dataset and the adult dataset, respectively, according to the selected classification thresholds α and β for each dataset. After the classification stage, the best dimension of QIDs that achieves optimum case should be determined. In the bank dataset, the QID dimension (QidD) is four (QidD = 4) while in the adult dataset QidD is 10 (QidD = 10). For each dataset, the initial value of QID dimension is set to one (QidD = 1) to be used as input into the proposed QID dimension identification algorithm (as explained in Algorithm 2) Identification of QID dimension begins with the initial value of QidD, and it is incremented until the maximum number of QID dimension. Identification of QID dimension begins also with a sample size equal to 10% of the dataset with k-anonymity of 5, and it is incriminated until $k = 25$ for each QidD value (sample size is changeable). Then, the privacy gain (PG) and the nonuniform entropy (NUE) are calculated for each sample and each new QidD until QidD values reach four (QidD = 4) for the bank dataset and QidD = 10 for the adult dataset.

Finally, the proposed algorithm returns the QidD that achieves the optimum case to be as the best dimension will be used in the anonymization process. Table 3 demonstrates the results of finding the best QidD for the adult dataset.

According to Table 3, we observed that QidD = 2 is the optimum case that increases the privacy gain as well as the NUE. Moreover, we can notice that the privacy level also

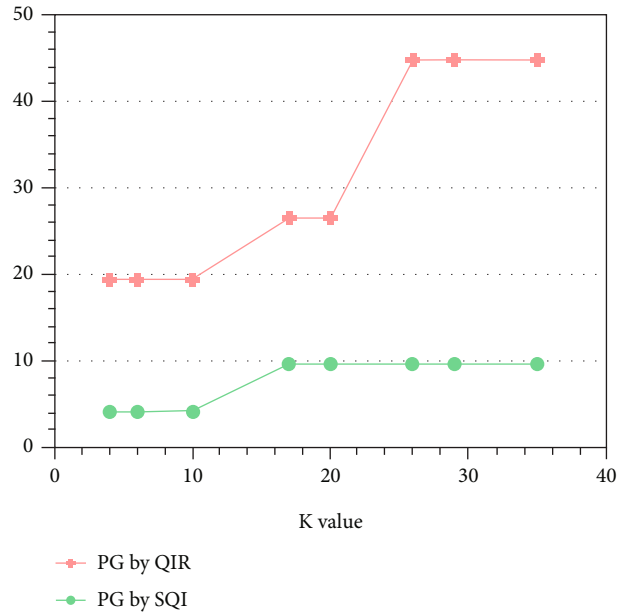


Figure 7: PG at several k values. Dataset: adult dataset 48,842 tuples. QidD of QIR = 2 (work class, HPW). QidD of SQI = 1 (age).

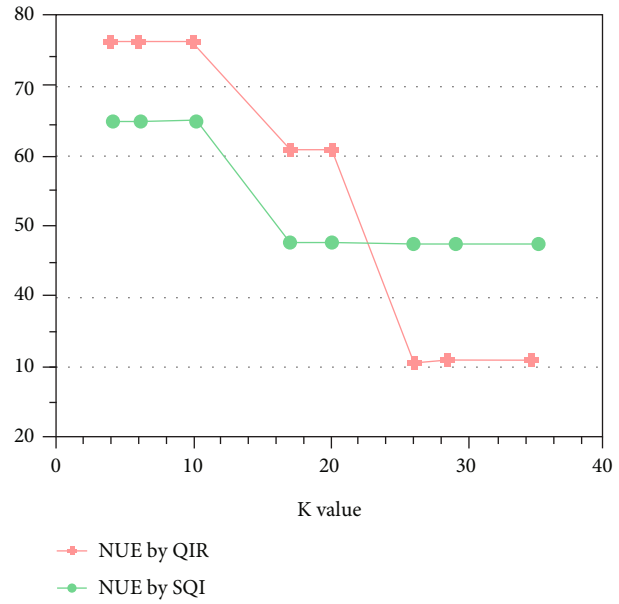


Figure 8: NUE at several k values. Dataset: adult dataset 48,842 tuples. QidD of QIR = 2 (work class, HPW). QidD of SQI = 1 (age).

increases when QidD value increases. The privacy gain reaches 91.59% when QidD is 10. On the other hand, NUE decreases, and accordingly, the data utility decreases when QidD increases. Figures 4(a)–4(c) demonstrate the selection of the best QidD for the bank dataset by the proposed QIR algorithm on different k-anonymity values, 5, 15, and 25, respectively. In the bank dataset, the proposed algorithm's selected QID attributes are work-class and hours-per-week (HPW). These two attributes achieve the highest reidentification risk; thus, they must be involved in the anonymization process (see Figure 5).

To determine the best QidD in the bank dataset, track Table 4 and Figures 6(a)–6(c); it is clear that when QidD = 1 the proposed algorithm achieves the optimum case as it gives high privacy in several cases of k values. It can be also observed in Table 4 that NUE drops from 45.28% when $k = 5$ to 17.27% when k increases above 15. It is also noticeable in the bank database that privacy decreases as the value of QidD increases which is normal with the level of privacy provided.

4.3. Performance Benchmark and Discussion. To evaluate the proposed QIR algorithm, we compare it based on k-anonymity against recent similar work SQI algorithm [37]. The comparison was conducted in terms of their privacy gain (PG) and nonuniform Entropy (NUE). Multiple k values and different dataset sizes of the adult dataset will be used. In Figures 7 and 8, the privacy provided by QIR is more than the privacy achieved by SQI, where the improvement average exceeds 23%. Although SQI outperformed the QIR in data utility represented by NUE at $k = 26, 29, 35$, with a privacy rate of 9.57%, this is considered a deficiency because QIR provided data utility higher than that with much higher privacy at $k = 4, 6, 10, 17$, and 20.

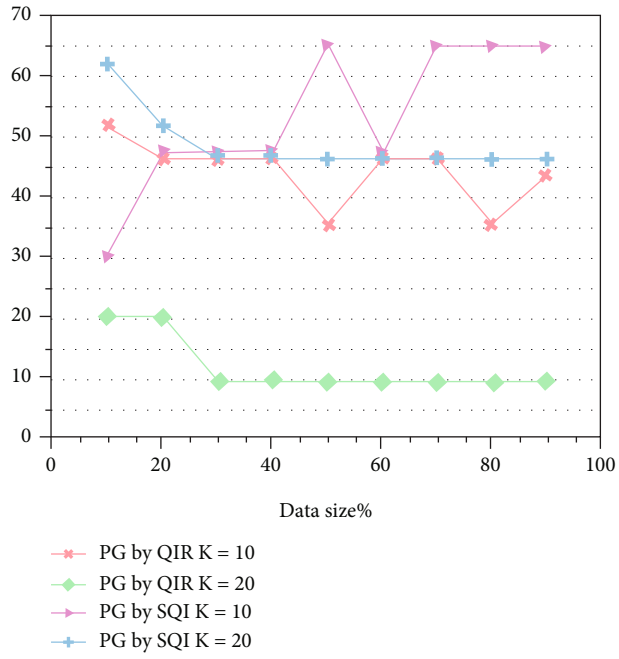


Figure 9: PG at several data sizes. Dataset: adult dataset. QidD of QIR = 2 (work class, HPW). QidD of SQI = 1 (age). $k = 10, 20$.

In Figures 9 and 10, it can be observed that at 10% of the dataset and $k = 10$ the privacy achieved by the proposed QIR algorithm is more than double the privacy achieved by the SQI algorithm with slight increases in data utility, that is, the proposed QIR algorithm outperforms the SQI algorithm in terms of preserving privacy and data utility. With data size 20% and $k = 20$, NUE obtained by SQI and QIR is 30.27 and 31.66%, respectively, while the privacy given by SQI is 20.52% and that by QIR is 51.82 which is twice more than that achieved by SQI. Similar results were obtained at

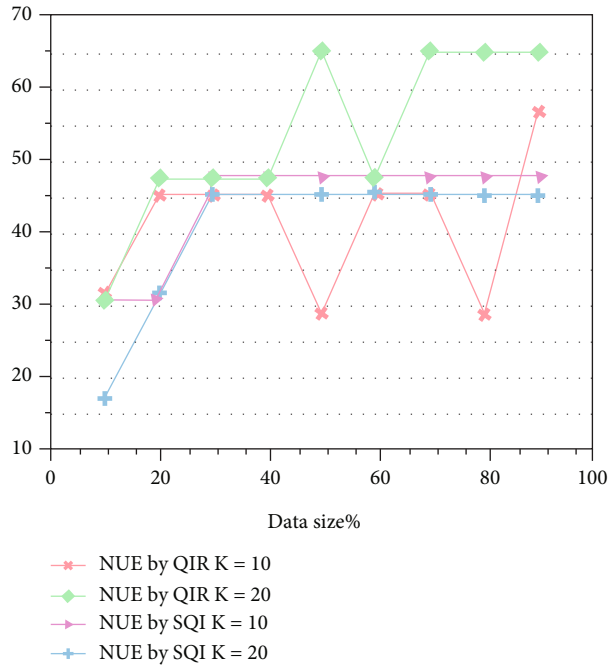


Figure 10: NUE at several data sizes. Dataset: adult dataset. QidD of QIR = 2 (work class, HPW). QidD of SQI = 1 (age). $k = 10, 20$.

$k = 20$ and data size = 30% and 90%, respectively. In most cases, when data size increases the privacy decreases, and therefore, the data utility increases.

Generally, for the whole adult data, results of the experiments at $k = 10$ and $k = 20$ show that the average privacy percentage presented by SQI is 10.17% with 48.62% data utility, while the average privacy percentage offered by the proposed QIR is 46.49% with 41.04% data utility. Also, for the whole adult dataset and all k values experimented, the average privacy provided by SQI is 7.51% against 54.13% data utility, while the average privacy percentage achieved by QIR is 30.67% against 55.46% data utility; hence, using QIR for identification of the real QIDs is considered more ideal.

5. Conclusions

Accurate identification of QIDs is an important issue for the success and validity methods of privacy-preserving outsourced data that seek to avoid privacy leakage caused by QID linking. This paper is aimed at classifying dataset attributes before the anonymization process and determining the proper QIDs that should be involved in the anonymity operation. A new algorithm is proposed based on the calculation of the reidentification risk for dataset attributes to classify attributes to SAs, QIDs, and NSs based on prespecified thresholds. In addition to attribute classification, the algorithm determines the actual dimension of QIDs that is required in the anonymization process depending on the amount of privacy provided versus a loss of the quality of the data. The experiment results indicated that the proposed identification algorithm has better performance and is more perfect in terms of privacy provided against data utility when compared with other works. Although the proposed algorithm is suitable to be used with any method or privacy model concerned with QID attributes, in this paper, we have relied on the k-anonymity model.

Data Availability

All data used in this article are available in the machine learning repository at the University of California, Irvine (UCI): https://archive.ics.uci.edu/ml/datasets/.

Conflicts of Interest

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

Acknowledgments

The authors would like to acknowledge Taif University Researchers Supporting Project (number TURSP-2020/292) Taif University, Taif, Saudi Arabia.

References

[1] J. Domingo-Ferrer, O. Farràs, J. Ribes-González, and D. Sánchez, "Privacy-preserving cloud computing on sensitive data: a survey of methods, products, and challenges," *Computer Communications*, vol. 140–141, pp. 38–60, 2019.

[2] S. Aldeen Yousra and S. Mazleena, "A new heuristic anonymization technique for privacy preserved datasets publication on cloud computing," *Journal of Physics: Conference Series*, vol. 1003, p. 012030, 2018.

[3] C. Bradford, "7 most infamous cloud security breaches - StorageCraft," *storagecraft*, 2020, https://blog.storagecraft.com/7-infamous-cloud-security-breaches/.

[4] B. Chen, P. Cheung, P. Cheung, and Y. Kwok, "Cypherdb: a novel architecture for outsourcing secure database processing," *IEEE Transactions on Cloud Computing*, vol. 6, no. 2, pp. 372–386, 2018.

[5] B. C. M. Fung, K. Wang, R. Chen, and P. S. Yu, "Privacy-preserving data publishing," *ACM Computing Surveys*, vol. 42, no. 4, pp. 1–53, 2010.

[6] S. A. Abdelhameed, S. M. Moussa, and M. E. Khalifa, "Privacy-preserving tabular data publishing: a comprehensive evaluation from web to cloud," *Computers & Security*, vol. 72, pp. 74–95, 2018.

[7] A. Bampoulidis, I. Markopoulos, and M. Lupu, "PrioPrivacy: a local recoding K-anonymity tool for prioritised qasi-identifiers," in *IEEE/WIC/ACM International Conference on Web Intelligence - Companion Volume*, pp. 314–317, ACM: New York, 2019.

[8] L. Sweeney, "Achieving k-anonymity privacy protection using generalization and suppression," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 571–588, 2002.

[9] Y. Yan, W. Wang, X. Hao, and L. Zhang, "Finding quasi-identifiers for k-anonymity model by the set of cut-vertex," *Engineering Letters*, vol. 26, no. 1, 2018.

[10] G. Kaur and S. Agrawal, "Differential privacy framework," in *Impact of Quasi-identifiers on Anonymization, vol. 46*, Springer, Singapore, 2019.

[11] D. Wei, K. Natesan Ramamurthy, and K. R. Varshney, "Distribution-preserving k-anonymity," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 11, no. 6, pp. 253–270, 2018.

[12] P. R. Bhaladhare and D. C. Jinwala, "Novel approaches for privacy preserving data mining in k-anonymity model," *Journal of Information Science and Engineering*, vol. 32, no. 1, pp. 63–78, 2016.

[13] M. S. Simi, K. S. Nayaki, and M. S. Elayidom, "An extensive study on data anonymization algorithms based on K-anonymity," *IOP Conference Series: Materials Science and Engineering*, vol. 225, p. 012279, 2017.

[14] H. Kaur, N. Kumar, and S. Batra, "ClaMPP: a cloud-based multi-party privacy preserving classification scheme for distributed applications," *The Journal of Supercomputing*, vol. 75, no. 6, pp. 3046–3075, 2019.

[15] G. G. Dagher, B. C. M. Fung, N. Mohammed, and J. Clark, "SecDM: privacy-preserving data outsourcing framework with differential privacy," *Knowledge and Information Systems*, vol. 62, no. 5, pp. 1923–1960, 2020.

[16] A. F. Westin, "Privacy and freedom," *American Sociological Review*, vol. 33, no. 1, p. 173, 1968.

[17] M. Templ, *Statistical disclosure control for microdata*, Springer International Publishing, Cham, 2017.

[18] W. Mahanan, W. A. Chaovalitwongse, and J. Natwichai, "Data anonymization: a novel optimal k-anonymity algorithm for identical generalization hierarchy data in IoT," *Service Oriented Computing and Applications*, vol. 14, no. 2, pp. 89–100, 2020.

[19] S. Mayil, M. Vanitha, C. Science, J. J. College, and T. St, "A survey on privacy preserving data mining techniques for clinical decision support system," *International Research Journal of Engineering and Technology*, vol. 5, no. 5, pp. 6054–6056, 2016.

[20] N. Uttarwar and M. A. Pradhan, "K-NN data classification technique using semantic search on encrypted relational data base," in *2016 International Conference on Computing Communication Control and automation (ICCUBEA)*, Pune, India, 2017.

[21] K. El Makkaoui, A. Beni-Hssane, A. Ezzati, and A. El-Ansari, "Fast Cloud-RSA scheme for promoting data confidentiality in the cloud computing," *Procedia Computer Science*, vol. 113, pp. 33–40, 2017.

[22] W. Wang, L. Chen, and Q. Zhang, "Outsourcing high-dimensional healthcare data to cloud with personalized privacy preservation," *Computer Networks*, vol. 88, pp. 136–148, 2015.

[23] K. El Makkaoui, A. Beni-Hssane, and A. Ezzati, "Speedy Cloud-RSA homomorphic scheme for preserving data confidentiality in cloud computing," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 12, pp. 4629–4640, 2019.

[24] D. Chandravathi and P. V. Lakshmi, "Privacy preserving using extended Euclidean algorithm applied to RSA-homomorphic encryption technique," *VOLUME-8 ISSUE-10, AUGUST 2019, REGULAR ISSUE*, vol. 8, no. 10, pp. 3175–3179, 2019.

[25] P. Shyja Rose, J. Visumathi, and H. Haripriya, "Research paper on privacy preservation by data anonymization in public cloud for hospital management on big data," *International Journal of Control Theory and Applications*, 2016.

[26] Y. A. A. S. Aldeen and M. Salleh, "Privacy preserving data utility mining architecture," in *Smart Cities Cybersecurity and Privacy*, pp. 253–268, Elsevier Inc., 2019.

[27] Y. A. A. S. Aldeen and M. Salleh, "Techniques for privacy preserving data publication in the cloud for smart city applications," in *Smart Cities Cybersecurity and Privacy*, pp. 129–145, Elsevier Inc., 2019.

[28] Y. A. A. S. Aldeen and M. Salleh, "A hybrid K-anonymity data relocation technique for privacy preserved data mining in cloud computing," *Journal of Internet Computing and Services*, vol. 17, no. 5, pp. 51–58, 2016.

[29] H. Lee, S. Kim, J. W. Kim, and Y. D. Chung, "Utility-preserving anonymization for health data publishing," *BMC Medical Informatics and Decision Making*, vol. 17, no. 1, p. 104, 2017.

[30] Y. A. A. S. Aldeen, M. Salleh, and Y. Aljeroudi, "An innovative privacy preserving technique for incremental datasets on cloud computing," *Journal of Biomedical Informatics*, vol. 62, pp. 107–116, 2016.

[31] S. R. P. Reddy, K. V. S. V. N. Raju, and V. V. Kumari, "Personalized privacy preserving incremental data dissemination through optimal generalization," *International Journal of Engineering & Technology*, vol. 7, no. 2.20, p. 197, 2018.

[32] R. V. Sudhakar and T. C. M. Rao, "Security aware index based quasi–identifier approach for privacy preservation of data sets for cloud applications," in *Cluster Computing*, pp. 1–11, Springer, 2020.

[33] S. A. Onashoga, B. A. Bamiro, A. T. Akinwale, and J. A. Oguntuase, "KC-Slice: a dynamic privacy-preserving data publishing technique for multisensitive attributes," *Information Security Journal: A Global Perspective*, vol. 26, no. 3, pp. 121–135, 2017.

[34] R. Wang, Y. Zhu, T.-S. Chen, and C.-C. Chang, "Privacy-preserving algorithms for multiple sensitive attributes satisfying t-closeness," *Journal of Computer Science and Technology*, vol. 33, no. 6, pp. 1231–1242, 2018.

[35] S. Srijayanthi, T. Sethukarasi, and A. Thilagavathy, "Efficient anonymization algorithm for multiple sensitive attributes," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 1, pp. 4961–4963, 2019.

[36] L. Huang, J. Song, Q. Lu, X. Liu, and C. Zhang, "Hypergraph-based solution for selecting quasi-identifier," *International Journal of Digital Content Technology and its Applications*, vol. 6, no. 20, pp. 597–606, 2012.

[37] A. M. Omer and M. M. Bin Mohamad, "Simple and effective method for selecting quasi-identifier," *Journal of Theoretical and Applied Information Technology*, vol. 89, no. 2, pp. 512–517, 2016.

[38] Y. J. Lee and K. H. Lee, "Re-identification of medical records by optimum quasi-identifiers," in *2017 19th international conference on advanced communication technology (ICACT)*, pp. 428–435, PyeongChang, Korea, 2017.

[39] K. S. Wong, N. A. Tu, D. M. Bui, S. Y. Ooi, and M. H. Kim, "Privacy-preserving collaborative data anonymization with sensitive quasi-identifiers," in *2019 12th CMI Conference on Cybersecurity and Privacy (CMI)*, Copenhagen, Denmark, 2019.

[40] Y. Sei, H. Okumura, T. Takenouchi, and A. Ohsuga, "Anonymization of sensitive quasi-identifiers for l-diversity and t-closeness," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 4, pp. 580–593, 2019.

[41] N. Victor and D. Lopez, "Privacy preserving sensitive data publishing using (k, n, m) anonymity approach," *Journal of communications software and systems*, vol. 16, no. 1, pp. 46–56, 2020.

[42] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, "L-diversity: privacy beyond *k* -anonymity," in *22nd International Conference on Data Engineering (ICDE'06)*, pp. 24–24, Atlanta, GA, USA, 2006.

[43] N. Li, T. Li, and S. Venkatasubramanian, "t-Closeness: privacy beyond *k*-anonymity and l-diversity," in *2007 IEEE 23rd International Conference on Data Engineering*, pp. 106–115, Turkey, 2007.

[44] H. Y. Tran and J. Hu, "Privacy-preserving big data analytics a comprehensive survey," *Journal of Parallel and Distributed Computing*, vol. 134, pp. 207–218, 2019.

[45] K. Patel and G. B. Jethava, "Privacy preserving techniques for big data: a survey," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 194–199, Coimbatore, India, 2018.

[46] E. E. Brown, "Improving privacy preserving methods to enhance data mining for correlation research," in *Southeast-Con 2017*, pp. 3–6, Concord, NC, USA, 2017.

[47] X. Jiang, A. D. Sarwate, and L. Ohno-Machado, "Privacy technology to support data sharing for comparative effectiveness research: a systematic review," *Medical Care*, vol. 51, 8 Supplement 3, pp. S58–S65, 2013.

[48] K. Benitez and B. Malin, "Evaluating re-identification risks with respect to the HIPAA privacy rule," *Journal of the American Medical Informatics Association*, vol. 17, no. 2, pp. 169–177, 2010.

[49] X. Zhang, C. Liu, S. Nepal, C. Yang, W. Dou, and J. Chen, "Combining top-down and bottom-up: scalable sub-tree anonymization over big data using MapReduce on cloud," in *2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, pp. 501–508, Melbourne, VIC, Australia, 2013.

[50] X. Zhang, C. Liu, S. Nepal, C. Yang, W. Dou, and J. Chen, "A hybrid approach for scalable sub-tree anonymization over big data using MapReduce on cloud," *Journal of Computer and System Sciences*, vol. 80, no. 5, pp. 1008–1020, 2014.

[51] F. Prasser, R. Bild, and K. A. Kuhn, "A generic method for assessing the quality of de-identified health data," *Studies in Health Technology and Informatics*, vol. 228, pp. 312–316, 2016.

[52] S. Moro, P. Cortez, and P. Rita, *UCI Machine Learning Repository: Bank Marketing Data Set*, 2014, https://archive.ics.uci.edu/ml/datasets/Bank+Marketing.

[53] R. Kohavi and B. Becker, *Adult Census Income | Kaggle*, 2016, https://www.kaggle.com/uciml/adult-census-income.

[54] F. Prasser, K. A. Kuhn, and J. Eicher, "Flexible data anonymization using ARX—current status and challenges ahead," *Software: Practice and Experience*, vol. 50, no. 7, pp. 1277–1304, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9958003, 10 pages
https://doi.org/10.1155/2021/9958003

Research Article

Communication Delay Modeling for Wide Area Measurement System in Smart Grid Internet of Things Networks

Mohammad Kamrul Hasan,[1] **Shayla Islam**,[2] **Muhammad Shafiq**,[3] **Fatima Rayan Awad Ahmed**,[4] **Somya Khidir Mohmmed Ataelmanan**,[4] **Nissrein Babiker Mohammed Babiker**,[5] **and Khairul Azmi Abu Bakar**[1]

[1]*Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia*
[2]*Institute of Computer Science and Digital Innovation, UCSI University Malaysia, 56000 Kuala Lumpur, Malaysia*
[3]*Cyberspace Institute of Advanced Technology, Guangzhou University, China*
[4]*Computer Science Department, College of Computer Engineering & Science, 1, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia*
[5]*Information System Department, College of Science and Arts, Bisha University, P.O Box 551, Bisha 61922, Saudi Arabia*

Correspondence should be addressed to Muhammad Shafiq; srsshafiq@gmail.com

Received 15 March 2021; Revised 21 May 2021; Accepted 27 June 2021; Published 14 July 2021

Academic Editor: Rahim Khan

We present communication frameworks, models, and protocols of smart grid Internet of Things (IoT) networks based on the IEEE and IEC standards. The measurement, control, and monitoring of grid being achieved through phasor measurement unit (PMU) based wide area measurement (WAM) framework. The WAM framework applied the IEEE standard C37.118 phasor exchange protocol to collect grid data from various substation devices. The existing frameworks include the IEC 61850 protocol and programmable logic controllers (PLCs) based supervisory control and data acquisition (SCADA) system. These protocols have been selected as per the smart grid configuration and communication design. However, the existing frameworks have severe synchronization errors due to the communication delays of IoT networks in the smart grid. Therefore, this article designs the timing mechanism and a delay model to reduce the timing delay and boost real-time measurement, monitoring, and control performance of the smart grid WAM applications. The result shows that the proposed model outperformed the existing WAM system.

1. Introduction

The electrical substation is a standout among the most critical parts of the electrical transmission and distribution frameworks. As indicated by [1], the substations are classified into switchyard substation, client substation, framework substation, and conveyance substation. The switchyard substations are situated at generations, which interface the generators to the utility matrices that give the off-site capacity to the plants. Generator switchyards tend to have substantial establishments that are commonly built and developed by the power plant operators and are liable to arranging, financing, and development accomplishments not quite the same as those of predictable substation projects—the substation functions as the essential source of electric power supply for heavy consumers such as factories. The specific necessities and business case, for this kind of usages, rely more on the customer's essentials than utility needs. In the meantime, the intelligent grid substations are network connected from the generation to the distribution through a communication framework to maintain the power grid intelligent devices and sensors deployed at the various components. The main purpose of the network-connected framework is to collect the grid data from various substation components, so that

the power grid becomes smart with the facility of ICT infrastructure. To accomplish the smart grid operation, the communication infrastructure plays a massive role in connecting, monitoring, and controlling. The communication infrastructure plays a massive role in connecting, monitoring, and controlling the smart grid system. Communication infrastructure is the key factor to have real-time accessibility on the grid. This is because proper communication architecture can deliver a large amount of operation data, communications, processing, and control for the remote monitoring for grids and the consumptions through the communication infrastructure. Figure 1 presents a hierarchical distributed communication infrastructure for the smart grid that includes smart homes and the communication architecture from the perspective of generation, transmission, distribution, and consumption [2]. The architecture considers SCADA, PMUs, massive-scale SMs, and SMDs to obtain and measure the voltage amplitude, current phasor, and power quality of generation and distribution information at a high sampling level of effectiveness in controlling, estimating, and ongoing checking of the keen matrix. SCADA is one of the mainstream frameworks that use its uncommon convention and equipment configuration to control and estimate the substation framework. The PLC-based SCADA system has been implemented in the existing power grid automation.

For the automation purpose the various components of substations have been connected to the wide area measurement system. Choosing the right convention and correspondence system is essential to accomplish the best level of effectiveness in controlling, estimating, and ongoing checking of the keen matrix. SCADA is one of the mainstream frameworks that use its uncommon convention and equipment configuration to control and estimate the substation framework.

The existing automation architecture is based on PLCs with the SCADA system. This SCADA framework incorporates the IEDs, the correspondence protocols like IEC 61850 that have detailed data demonstrating the information out of devices that use this protocol. The IEC 61850's strategy transport embraced connecting estimations devices, similar to instrument transformers, actuators, protection systems, and circuit breakers, to the IEDs of the sound amid an essential station. Indeed, these devices are associated with the controller of the inlet by cable connections. The IEC 61850 randomly proposes a captivated technique bus to curtail the cabling and upkeep cost. In any case, the quality does not maintain with the devoted topology of the bus bay: one technique transport will interconnect all the circle devices of the station, or a process bus per bay will connect the devices specifically.

The main contribution of this paper is as below:

(1) The communication framework and the delay modeling with the timing mechanism

(2) The key determinant performance factors are total measured delay, and real-time delays are modeled and analyzed the performance focusing propagation delays, receive delays, initial delays, skews, and phasor offsets parameters

(3) The comparison evaluation with the C37.118 standard

2. Smart Grid Communication Systems and Protocols

The IEC 61850 communication protocol has expressively improved the substation automation system (SAS), by introducing the control structure in a communication framework to measure, control, and monitoring [3]. The IEC 61850 protocol suits two arranged communication phases: the substation bay and the system bay that reliably communicates the SCADA framework to the strait embellishment of the basic substation. The instrumentations are transformers; the protection unit and MU to IEDs are connected to the system bay. The two foundations have specific necessities, and consequently, the advances and plans need to send the two bays are routinely intriguing. Conventionally, the substation bay is executed utilizing a performing fiber optical gigabit Ethernet, with ring topology to redesign the receptiveness of the correspondence. The strategic bay needs to deal with a more significant proportion of information starting from transformers or merging unit (MU) that sent on the substation. Generally, a high-speed gigabit Ethernet (1/10) network system is incorporated to accomplish the communication infrastructure. The IEC 61850-9-2 LE represents a course of action of guidelines to empower the interfacing of instruments and MUs to the bay system [4]. The IEC 61850-9-2LE discusses the rate of SV that is focused on protection applications: the estimation contraptions reliable with this profile need to test the current or the voltage multiple times for each system cycle (50 Hz in EU, 60 Hz in the USA). The instruments made for also asking for applications, like quality control monitoring, test the characteristics multiple times for the lattice cycle. The MUs should be synchronized with one another to achieve precision underneath four μs. The synchronization of the equipment should be set up using a 1-PPS committed banner, generally using a GPS recipient and enclosed to each contraption [5]. A progressive package number is utilized to recover the testing cycle of each instrument. In any case, the progressing standard for the propelled substations for this standard perseveres through the synchronization challenges.

The IEEE C37.118 standard uses synchrophasors in estimating smart grid devices and components. This is to manage the synchrophasor estimation. At the required significances, the configuration parameters analyses in different conditions and displays immaterial precision [5]. The standard depicts by class P suggests security applications require helpful reaction time. Anyway, class M gives the base element of accuracy for the estimation applications, which do not require an intelligent reaction time yet perfect precision over class P. The contrasts between the two classes are recognizable in the execution of work in progress changes of amplitudes or stages. For this kind of test, the implementation required for the P class, concerning reaction time, is on an elementary level higher than M class and is accessible on the detailing pace of synchrophasors in any case, and it relies just upon the straightforward rehash of the framework [3, 6]. The circled current and voltage transformers should completely synchronize. The IEEE C37.118 synchronization calculation involves a synchronization exactness of end

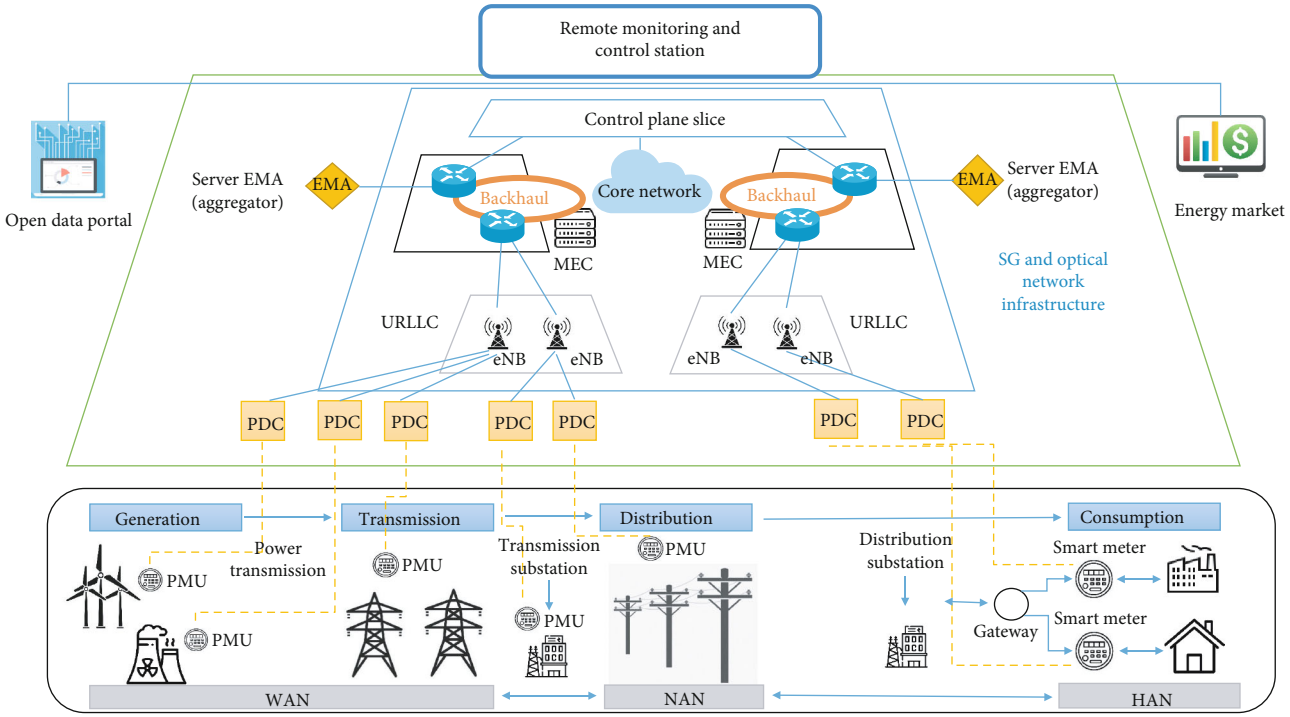


Figure 1: A hierarchical distributed communication infrastructure: where energy management agent (EMA), eNB is the macrocell base station, SM denoted massive scale, SMD states the supervisory control and data acquisition measurement devices, N3 is the communication interface between the router and macrocell base station, URLLC is the ultrareliable, MU is the merging unit, and low-latency communications, MEC used for mobile edge computing [2].

gadgets underneath the microsecond, enough to estimate the synchrophasors. Moreover, the voltage and current on the framework ought to be assessed on different occasions for each cycle (50 Hz), to give an adequate number of tests. Right now, the estimation focus point can make up to two megabits for every second of development, which the system structure ought to precisely sensible with no adversity came in liberal stacked conditions.

IEEE C37.118 is messages that are structured with the data of information, configuration, header, and control message. Information messages are utilized to send supportable estimations. Data from different PMUs are transmitted in a special message related to a particular timestamp [7, 8]. Configuration messages are machine-arranged and contain information about the game-plan parts and data sorts. Configuration messages are of three explicit categories: CFG-1, CFG-2, and CFG-3. CFG-1 finds out for kind of data and point of confinement of PMU/PDU. CFG-2 displays the synchrophasor estimations, which are starting at now being transmitted/declared. In any case, CFG-3 resembles CFG-2, yet contains some extra flexibility and information about PMU qualities and estimations. Header messages contain expressive information sent by the PMU by techniques for PDC. Request messages are used to control the endeavor of contraption sending synchrophasor estimations. Like this, data, configuration, and request messages are passed on in machine-clear methodology while the header explains information in the fathomable connection. Moreover, data, configuration, and header are sent by the data sources while control messages are received by the data sources. The IEEE C37.118 correspondence design is introduced in Figure 2 [8]. IEEE C37.118 encounters the nonattendance of standard data names checks autodisclosure and self-depiction without learning of configuration message, customization weakens interoperability and joining support, and no apparent security framework or even more all it depends on GPS for the outside timestamps [9–16].

3. Recent Issues and Challenges of Communication Systems in Smart Grid

The significant advancement has been accomplished for the communication infrastructure with the special technologies that include artificial intelligence, Internet of Things (IoT), 4G, and 5G data access [15–29]. This advancement enriched the communication infrastructure for smart grid protection, measurement, and control system in a sophisticated manner. However, the smart grid maintains high-level communication with specific devices with the applications to support substation automation. The above discussion can be discussed with the specific illustration of WAM for SCADA and PMU system issues. The issues and the challenges of SCADA and PMU based WAM system for substation automation are discussed below.

3.1. Issues and Challenges of PLC and SCADA System for the Digital Substation. The network infrastructure of the WAM system is designed with the consideration of multiple device

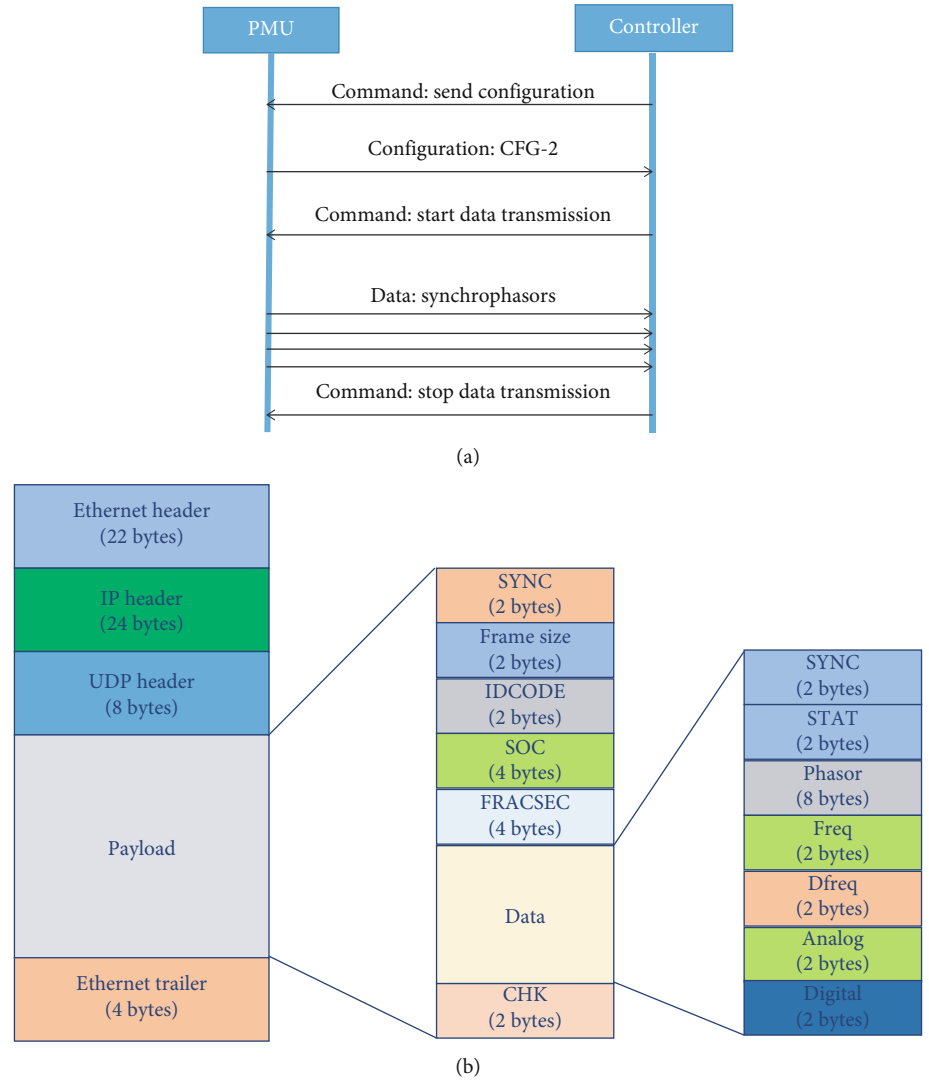


Figure 2: Functionalities of IEEE C37.118 protocol (a) and data structure (b) [8].

associations to access substation devices, reduce maintenance cost, and support the execution of additional substation mechanization applications [25–29]. The coordination of sequential protocol substation devices onto an Ethernet structure is a significant test. There are different devices called Terminal Servers, Serial Device Servers, or Console Servers. These devices symbolize short sequential data messages into TCP/IP bundles to send over Ethernet, and each sequential stream is related to a reliable TCP/IP session. Terminal Servers may be connected to be worked at the client or server zone to change over Serial-IP streams back to the sequential plan. Then again, some expert servers or remote PCs may interface clearly to Serial-IP streams over IP/Ethernet affiliations. The execution over wide area networks (WANs) can be an issue with Serial-IP organize fuse in perspective on limited information exchange limit, e.g., 56 kbps or fragmentary T1 (under 1.5 Mbps) layout hand-off or gave electronic organizations. Various SCADA has that usage serial remote devices have short reviewing breaks; except for on the off chance that they get a response from a remote IED in less than 100 milliseconds, they may expect a framework issue. Some Serial-IP frameworks cannot dependably achieve this low dormancy. One segment that impacts orchestrate torpidity can be the traditional overhead of TCP/IP epitome. As showed up in Figure 3, sequential SCADA messages may be only two or three bytes in length; anyway, TCP/IP tradition headers increase the length of the Serial-IP packages by demand of degree. This can be relieved with a SCADA frame forwarding technique that uses layout hand-off-based embodiment with only two or three bytes of the header to multiplex sequential SCADA development on a WAN framework. SCADA frames implemented together using IP-based WAN infrastructure. The SCADA frame forwarding packet formats have been designed with the link header, IP header, TCP header, message and data information field, and FCS. This frame forwarding is used to reduce the overhead by saving 90% of bandwidth and also reduce the latency on the wide area network.

Generically, cybersecurity is another essential requirement from the extruders hacking for the entire internal

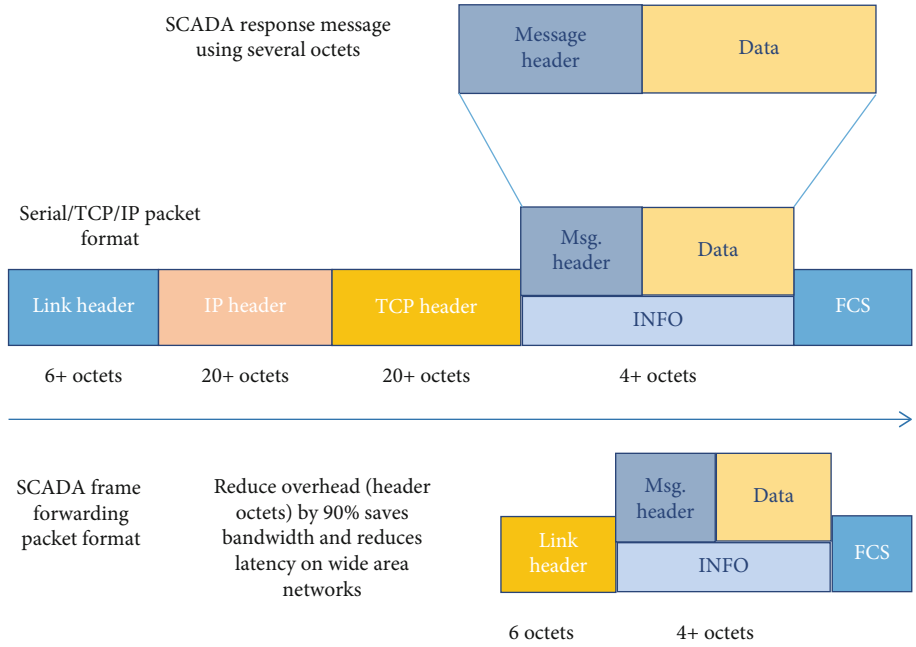


Figure 3: Serial IP protocol overhead vs. SCADA frame forwarding [29–31] (reproduced under the Creative Commons Attribution License/public domain).

systems and smart grid processes. Researchers have been emphasizing the effort on cybersecurity-related projects to comply with current industry standards. Another critical issue is network reliability. For the high network, reliability is required in SCADA and other operational systems in grid operations. In addition, substation networking becomes advanced and integrated. Therefore, the network outages can affect many systems and control elements that need to be minimized.

3.2. Issues and Challenges of IEC 61850-90-5 Communication System. The current communication convention IEC 61850 for substation computerization inside the grid framework today is being utilized with the help of up to 12 RS232/485 and 6 Ethernet ports, SCADA server (see Figure 4). Although there are timing and control issues inside the present framework applications, the framework has been confronting difficulties utilizing IEC 61850. One point to be communicated about is that the usage of IEC 61850 inside a repetition setting, at both the gear and correspondences levels, and its impact on the general inertness of a framework, including the potential effect and issues concerning control and assurance in such an excess domain. One must ensure that the methodology will give appropriate reaction times and adaptability for future expansion (s), without precluding the adaptability to deal with considerably further developed conceivable control outcomes that may be required during a future keen network.

The IEC 61850-90-5 is inherited from the IEC 61850 main version that had been implemented for substation digitalization. The IEC 61850 is a network correspondence convention that will, in general, show power framework parts, thought of interconnected administrations, and correspondence conventions just as frameworks. It was sketched out because of the interoperability and combination between power control structure, device configurations, self-association, and item autorevelation, the persistent quality framework through retransmission, reduced substation cost through multicast, and multicast, and multicast, and multicast remote network communication system, and reinforce for the machine-to-machine interchanges. Even though IEC 61850 furthermore has weaknesses, including the nonappearance of security constituent and confined correspondence, IEC 61850-90-5 gets all of the features of IEC 61850 while in like manner vanquishing its constraints. The critical differences between IEC 61850 and IEC 61850-90-5 showed up in Figure 5. IEC 61850-90–5 joins a security segment considering Group Domain of Interpretation (GDOI) and besides allows the transmission of time-essential shows over wide-area organizes by relying upon transport and framework layer conventions.

3.3. Issues and Challenges of IEEE C37.118 Protocol with SCADA and WAM System. Using high-level data-efficient communication system, many smart grids established the communication framework and IEEE C37.118 and SCADA. It is identified that the existing grid system uses fiber optic-WiMAX technologies as the backbone of AMI traffic is to control and access between the control center and the grid [16]. Using WiMAX provides a high communication data rate from the control center to the smart grid substations. On the other hand, researchers also investigated LTE-based fiber-optic technology to provide high communication data rates from the control center to the smart grid substations [17, 29, 30, 32–37]. Depending on the GPS timestamps, the nonappearance of GPS lopsidedness influences the accuracy

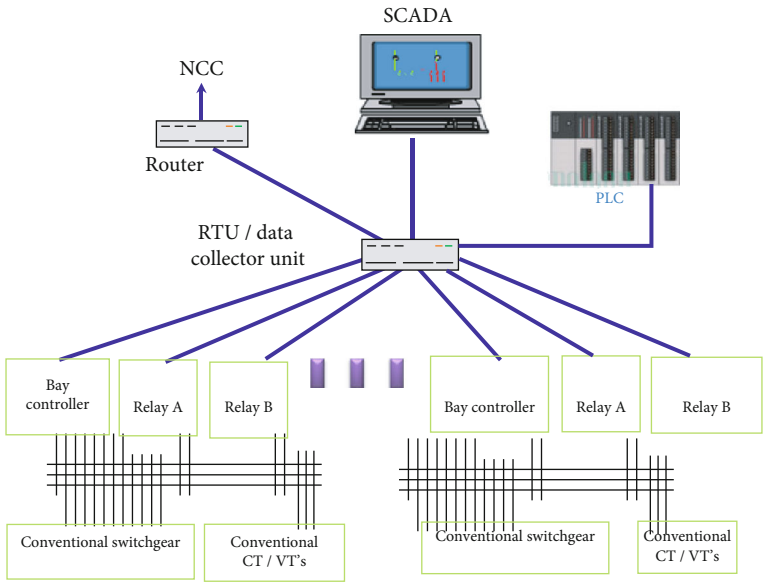


Figure 4: Diagram of IEC 61850 in SCADA system.

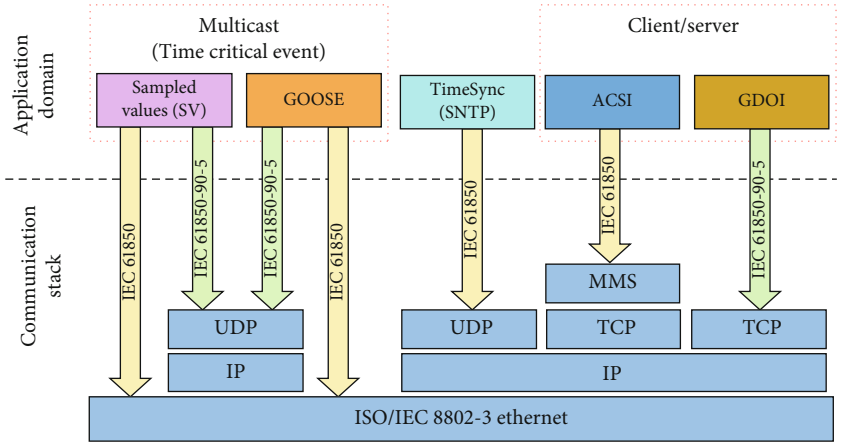


Figure 5: Difference between IEC 61850 and IEC 61850-90-5 protocol stack [14].

of the lattice-observing framework because of offbeat and mistake events in estimating the ongoing information [17, 18, 38, 39]. The nonattendance of standard data names stays away from autorevelation and self-association without the learning of configuration messages: vender specific features and customization cripple interoperability and blend support. There is no inherent security component to ensure digital assault [4, 12, 16–18, 27, 30, 33]. The other significant issue is the implementation of the PMUs using GPS in the smart grid. The study suggests that if the GPS connectivity is missing a little longer, the measurement and control will be complicated; therefore, the effect will be on the efficiency.

However, it is observed that latency, phasor delays, and internal processing delays are still key concerns in smart grid communications. Therefore, delay modeling is critical to calibrate the existing WAM systems for smart grid applications [30, 34–37, 39]. The next section will discuss and propose the communication delay modeling by focusing and consider the few delay parameters that can be calibrated to the existing system to enhance the grid measurement performance.

4. Proposed Delay Modeling for the Smart Grid Applications

A crucial prerequisite for every data transmission of the smart grid communication system's proper operation is timing and phase synchronization. The method by which a receiver finds the correct pulse width of time at which to sample the incoming signal is known as timing synchronization. Phase synchronization is the process by which a receiver adjusts the frequency and phase of its local carrier oscillator to match that of the transmitter. However, due to the synchronization issue of the existing IEEE C37.118 in the WAM system timing synchronization requires delay

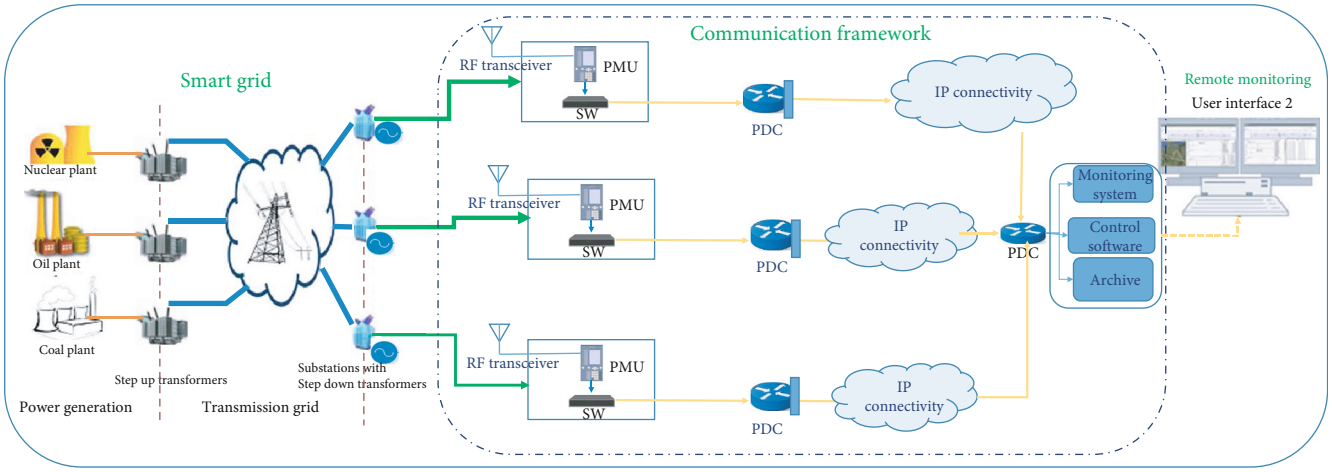


FIGURE 6: Communication diagram for wide area measurement (WAM) system, where an external timing source is used using GPS system for each of the PMUs in the set of substations, and PDCs are connected to monitor the smart grid remotely.

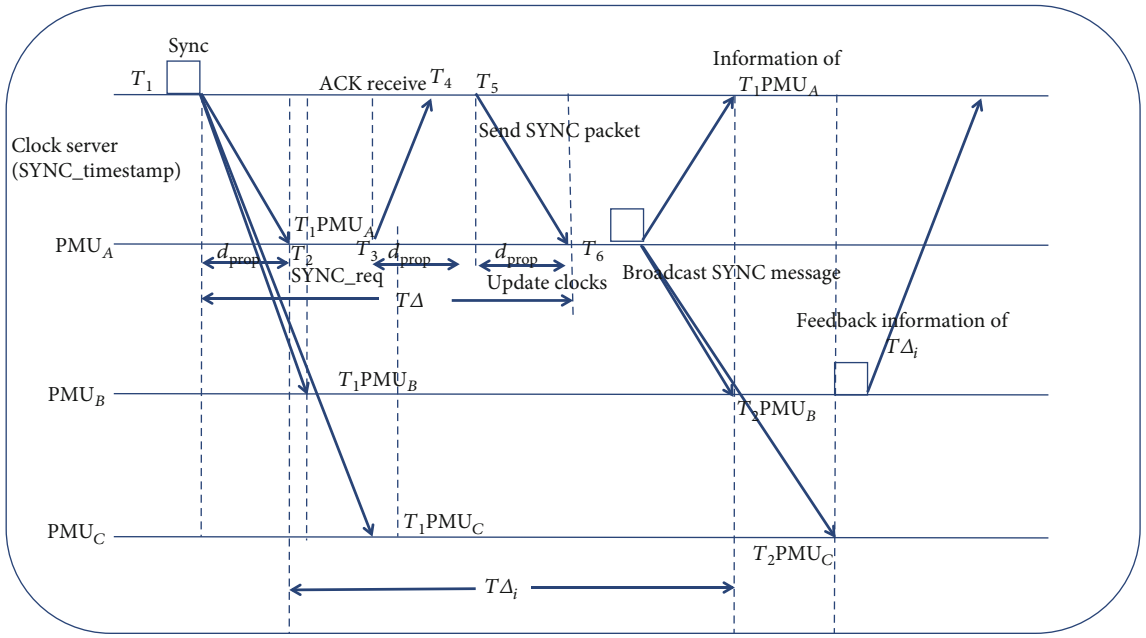


FIGURE 7: The timing diagram for the delay measurement and modeling [4, 16, 29, 33].

modeling. In order to delay modeling, a timing synchronization diagram is designed along with the network architecture. The communication delay is modeled using the following proposed timing diagram of existing, where the number of PMU, PDC, and the external timestamp source is considered (see Figure 6). Figure 7 demonstrates the timing diagram [4]. The timing diagram is used for the proposed delay modeling.

The delay is estimated from the above timing diagram, the time differences between the external timing source to all PMUs, PDCs, and the entire WAM communication system. To measure the total delay of the system, the first need is to consider small partial delays, propagation delays, receive delays and the phase errors, and the skew errors. The following equation is for the entire timing measurement where all sorts of delays are considered. The list of symbols that are used in Figure 7, and Eqn. (1) to Eqn. (6) are defined in Table 1.

$$T^{\text{timing\_signal}}_{\text{PMU}A\longrightarrow i} = \Phi^{\text{skew\_error}}_{\text{PMU}A\longrightarrow i}\left\{T^{\text{PMU}_{A\longrightarrow i}}_{1} + \frac{\alpha}{2} + \frac{\rho^{\text{PMU}_{A\longrightarrow i}}_{\text{prop\_uplink}} - \rho^{\text{PMU}_{A\longrightarrow i}}_{\text{prop\_downlink}}}{2}\right\} + \widehat{T}^{\text{phase\_error}}_{\text{PMU}_{A\longrightarrow i}} + \frac{\bar{\mathscr{D}}^{\text{delay}}_{\text{PMU}_A} - \bar{\mathscr{D}}^{\text{delay}}_{\text{PMU}_i}}{2}, \tag{1}$$

$$\bar{\mathscr{D}}^{\text{total\_delay}}_{A\longrightarrow i} = \left[\left(\bar{\rho}^{\text{PMU}_{A\longrightarrow i}}_{\text{prop}} - \mathscr{D}^{\text{PMU}_{A\longrightarrow B}}_{\text{rec},i}\right) + \left(\mathscr{D}^{\text{PMU}_{A\longrightarrow i}}_{\text{rec},i} - \mathscr{D}^{\text{PMU}_{B\longrightarrow C}}_{\text{rec},i}\right) + \left(\mathscr{D}^{\text{PMU}_{A\longrightarrow i}}_{\text{rec},i} - \mathscr{D}^{\text{PMU}_{i\longrightarrow A}}_{\text{rec},i}\right)\right]. \tag{2}$$

Table 1: Definition of the symbols that are used in Eqn. (1) to Eqn. (6).

Symbol	Definition
Sync_start	Synchronization start message frame
Ack	Acknowledgement
$T^{\text{timing\_signal}}_{\text{PMU}A\longrightarrow i}$	Timing signal of PMU A to the i^{th} number
$\Phi^{\text{skew\_error}}_{\text{PMU}A\longrightarrow i}$	Phase skew error of PMU A to the i^{th} number
α	Initial receive delay
$\rho^{\text{PMU}_{A\longrightarrow i}}_{\text{prop\_uplink}}$	Propagation delay for uplink
$\rho^{\text{PMU}_{A\longrightarrow i}}_{\text{prop\_downlink}}$	Propagation delay for downlink
$\bar{\mathcal{D}}^{\text{delay}}_{\text{PMU}_A}$	Receive delay at PMU_A
$\bar{\mathcal{D}}^{\text{delay}}_{\text{PMU}_i}$	Receive delay at PMU_i
$\Phi^{\text{skew\_error}}_{\text{PMU}A\longrightarrow i}$	Total skew error
$\widehat{T}^{\text{phase\_error}}_{PMU_{A\longrightarrow i}}$	Total phase offset
$\gamma^{\text{desired\_delay}}_{\text{PMU}_{A\longrightarrow i}}$	Desired signal of PMU_A to PMU_i

The total skew or the related frequency error can be expressed in Eqn. (3).

$$\Phi^{\text{skew\_error}}_{\text{PMU}A\longrightarrow i} = T^{\text{timing\_signal}}_{\text{PMU}_A} - T^{\text{timing\_signal}}_{\text{PMU}_i}. \tag{3}$$

To measure the absolute delay of the entire communication system, the phase error needs to be estimated. Consequently, the total estimated phase errors can be expressed in Eqn. (4) considering Eqns. (1), (2), and (3). Therefore, the real-time delay can be modeled in Eqn. (5) using Eqn. (4),

$$\bar{\rho}^{\text{PMU}_{A\longrightarrow i}}_{\text{prop}} = \frac{\rho^{\text{PMU}_{A\longrightarrow i}}_{\text{prop\_uplink}} - \rho^{\text{PMU}_{A\longrightarrow i}}_{\text{prop\_downlink}}}{2}, \tag{4}$$

$$\widehat{T}^{\text{Measured\_delay}}_{\text{PMU}_{A\longrightarrow i}} \cong \frac{1}{M}\sum_{i=1}^{M} \widehat{T}^{phase_{error}}_{\text{PMU}_{A\longrightarrow i}} + \bar{\mathcal{D}}^{\text{total}_{\text{delay}}}_{A\longrightarrow i} + \Phi^{\text{skew}_{\text{error}}}_{\text{PMU}A\longrightarrow i}, \tag{5}$$

$$D^{\text{PMU}_{A\longrightarrow i}}_{\text{real-time\_delay}} = \gamma^{\text{desired\_delay}}_{\text{PMU}_{A\longrightarrow i}} - \widehat{T}^{\text{Measured}_{\text{delay}}}_{\text{PMU}_{A\longrightarrow i}}. \tag{6}$$

5. Result and Discussion

The performance analysis was accomplished by applying the Monte Carlo simulation approach in MATLAB [16, 38]. The performance of the proposed delay model is simulated and then compared with the existing IEEE C 37.118-based WAM system. The performance was measured with the integration of the proposed delay model into the existing IEEE C37.118 and then compared with IEEE C37.118 assessed without the proposed delay model. The total measured delay and the real-time delay are considered as the performance metric of delay measurement. The simulation scenario considered 50 PMUs, and the 50 PMUs, 20 PDCs, 20 trials, and the designed communication topology. The communication network is shown in Figure 8. The initial random receive delay reference was set to 15 μs, 20 μs, and 30 μs. The number of samples was 10,000. Setting the initial delay at 10 micro-

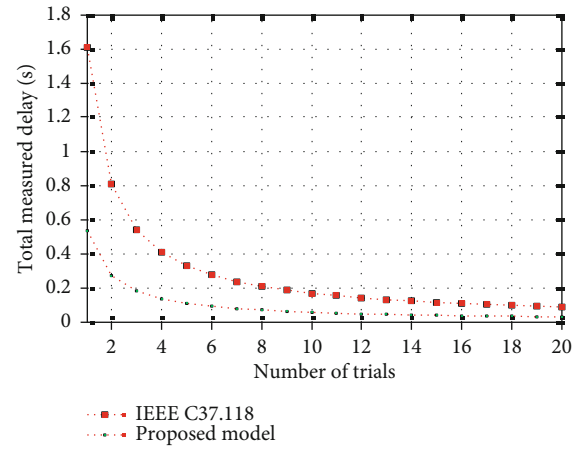


Figure 8: Total delay measurement in terms of number of trials at the initial 10 uc delay consideration.

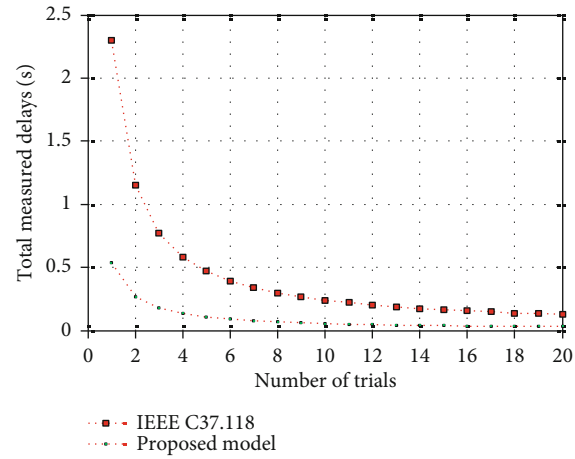


Figure 9: Total delay measurement in terms of number of trials at the initial 30 uc delay consideration.

seconds (μs) and the network bandwidth with an external timing server set to10 Mbps, the total delay is represented in Figure 8. In addition, with the setting of 30 μs initial delays, the total delay measurement is shown in Figure 9. It is visible that the total measured delay is marginally acceptable using the proposed model. However, the existing IEEE C37.118 shows the extra delays.

Figure 10 shows the delay performance of the proposed model function of the number of occurrences and the exchanges of phasor signals among different PMUs and PDCs in the communication framework. This figure shows the actual delay in terms of asymmetric bandwidth rates; if the bandwidth increases from 1 Mbps to 10 Mbps, the communication delays are lower. For the convenience of the analysis, the receive delay processing time was varied from 10 μs to 30 μs, and the simulation processing time was assumed 1 hour. Therefore, the study suggests that the delays can be lower if the bandwidth reaches higher to higher. The observations of the proposed model result show that the communication delays are less than 0.5 seconds while bandwidth

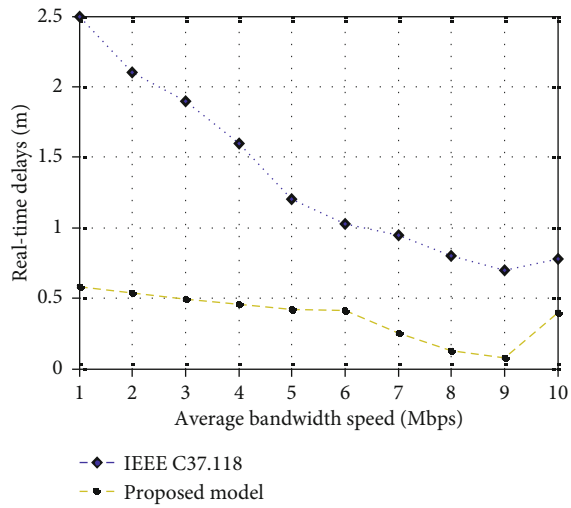


Figure 10: Real-time delay measurement in terms of average bandwidth speed.

reaches 5 Mbps, whereas the C37.118 has a communication delay of more than 1.7 seconds.

The proposed model performs better than the existing C37.118 IEEE synchrophasors standard because of the design consideration of the timing controlling and parameters of delay modeling of the communication framework. The main parameter selection considers the propagation delays, receive delays, frequency errors, and phase errors so that it can finally estimate the real-time precisely communication delays. Therefore, in terms of delay modeling, the proposed model performs better in comparison with the C37.118 IEEE standard.

6. Conclusions

This article has investigated the smart grid automation systems and their protocols. This is to trace out the traditional communication system's limitations and advantages in smart grids. The algorithms discussed the functionalities and characteristics. The main contributions of this paper are the proposed real-time delay measurement model and then the performance analyzing over the existing IEEE C37.118. The main achievement of the proposed model is the mitigation of the communication delays in the wide area measurement system. The result of the proposed model suggests that the unexpected delays can affect the efficiency of the smart grid application. The performance of the proposed model was found better in enhancing the precise real-time measurement and monitoring the prefault situation and the exigence occurrence of the smart grid applications. The future work of this study is to calibrate the delay parameter into the real-time smart grid WAM system.

Data Availability

The data used to support the findings of this study are available in the manuscript.

Conflicts of Interest

The authors declare no conflict of interest regarding this paper.

Authors' Contributions

Data, methodology, simulation, result analysis, software, validation, and manuscript writing are done by Mohammad Kamrul Hasan. Review and editing are done by Shayla Islam, Muhammad Shafiq, Fatima Rayan Awad Ahmed, Somya Khidir Mohmmed Ataelmanan, Nissrein Babiker Mohammed Babiker, and Khairul Azmi Abu Bakar.

Acknowledgments

This research is supported by the research grant of the Universiti Kebangsaan Malaysia (UKM) under the grant GGPM 2020-028 and FRGS/1/2020/ICT03/UKM/02/6.

References

[1] J. D. McDonald, *Electric Power Substations Engineering*, CRC Press, 2017.

[2] J. S. Choi, "A hierarchical distributed energy management agent framework for smart homes, grids, and cities," *IEEE Communications Magazine*, vol. 57, no. 7, pp. 113–119, 2019.

[3] J. Claveria and A. Kalam, "The influence of IEC 61850 standard: implementation and development of a functional substation automation simulator," *Australian Journal of Electrical and Electronics Engineering*, vol. 17, no. 1, pp. 28–35, 2020.

[4] M. K. Hasan, S. H. Yousoff, M. M. Ahmed, A. H. A. Hashim, A. F. Ismail, and S. Islam, "Phase offset analysis of asymmetric communications infrastructure in smart grid," *Elektronika ir Elektrotechnika*, vol. 25, no. 2, pp. 67–71, 2019.

[5] A. Bhargav, S. Ahmad, S. Kumari, A. Sahu, S. Luthra, and A. Gupta, "Technical evaluation and optimization of phasor measurement unit using CSIR-NPL PMU calibrator system to ensure reliability," *MAPAN*, vol. 35, no. 1, pp. 117–124, 2020.

[6] D. Fan and V. Centeno, "Phasor-based synchronized frequency measurement in power systems," *IEEE Transactions on Power Delivery*, vol. 22, no. 4, pp. 2010–2016, 2007.

[7] K. E. Martin, "Synchro-phasor standards development-IEEE C37. 118 & IEC 61850," in *2011 44th Hawaii International Conference on System Sciences (HICSS)*, pp. 1–8, Kauai, HI, USA, January 2011.

[8] K. Narendra, *Role of phasor measurement unit (PMU) in wide area monitoring and control*, Erlphase Power Technology Ltd, Winnipeg, Canada, Tutorial-Cbip, 2007.

[9] L. Montini, T. Frost, G. Dowd, and V. Shankarkumar, "Precision time protocol version 2 (PTPv2) management information base," *Information Base (No. RFC 8173)*, IETF, 2017.

[10] S. Lee, S. Lee, and C. Hong, "An accuracy enhanced IEEE 1588 synchronization protocol for dynamically changing and asymmetric wireless links," *IEEE Communications Letters*, vol. 16, no. 2, pp. 190–192, 2012.

[11] M. K. Hasan, R. A. Saeed, R. A. Alsaqour, A. F. Ismail, H. A. Aisha, and S. Islam, "Cluster-based time synchronisation scheme for femtocell network," *International Journal of Mobile Communications*, vol. 13, no. 6, pp. 567–598, 2015.

[12] Y. Mostofi and D. C. Cox, "Mathematical analysis of the impact of timing synchronization errors on the performance of an OFDM system," *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 226–230, 2006.

[13] R. Khan, K. McLaughlin, D. Laverty, and S. Sezer, "Analysis of IEEE C37.118 and IEC 61850-90-5 synchrophasor communication frameworks," *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5, Boston, MA, USA, July 2016.

[14] A. K. Das and S. Zeadally, "Data security in the smart grid environment," in *Pathways to a Smarter Power System*, pp. 371–395, Academic Press, 2019.

[15] D. Bian, M. Kuzlu, M. Pipattanasomporn, S. Rahman, and D. Shi, "Data security in the smart grid environment," *IET Communications*, vol. 13, no. 8, pp. 1025–1033, 2019.

[16] M. K. Hasan, M. M. Ahmed, A. H. A. Hashim, A. Razzaque, S. Islam, and B. Pandey, "A novel artificial intelligence based timing synchronization scheme for smart grid applications," *Wireless Personal Communications*, vol. 114, no. 2, pp. 1067–1084, 2020.

[17] J. Zhang, X. Luo, X. Fu, X. Wang, C. Guo, and Y. Bai, "Experimental study on the influence of satellite spoofing on power timing synchronization," *International Journal of Network Security*, vol. 22, no. 6, pp. 954–960, 2020.

[18] M. Z. Gunduz and R. Das, "Cyber-security on smart grid: threats and potential solutions," *Computer Networks*, vol. 169, article 107094, 2020.

[19] M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: a survey," *Sustainable Cities and Society*, vol. 60, article 102177, 2020.

[20] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "IoT malicious traffic identification using wrapper-based feature selection mechanisms," *Computers & Security*, vol. 94, article 101863, 2020.

[21] R. Ahmad, E. A. Sundararajan, N. E. Othman, and M. Ismail, "Handover in LTE-advanced wireless networks: state of art and survey of decision algorithm," *Telecommunication Systems*, vol. 66, no. 3, pp. 533–558, 2017.

[22] M. Shafiq, Z. Tian, Y. Sun, X. Du, and M. Guizani, "Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city," *Future Generation Computer Systems*, vol. 107, pp. 433–442, 2020.

[23] I. Shgluof, M. Ismail, and R. Nordin, "An enhanced system information acquisition scheme for CSG femtocells in 3GPP LTE/LTE a systems," *Wireless Personal Communications*, vol. 96, no. 3, pp. 3995–4011, 2017.

[24] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "Corrauc: a malicious bot-iot traffic detection method in iot network using machine learning techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242–3254, 2021.

[25] F. Yang and X. Zhang, "Performance analysis for OFDM smart-grid networks under AWSαSN with imperfect synchronization," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, Dublin, Ireland, June 2020.

[26] M. Akhtaruzzaman, M. K. Hasan, S. R. Kabir, S. N. H. S. Abdullah, M. J. Sadeq, and E. Hossain, "HSIC bottleneck based distributed deep learning model for load forecasting in smart grid with a comprehensive survey," *IEEE Access*, vol. 8, pp. 222977–223008, 2020.

[27] M. K. Hasan, M. M. Ahmed, S. S. Musa et al., "An improved dynamic thermal current rating model for PMU-based wide area measurement framework for reliability analysis utilizing sensor cloud system," *IEEE Access*, vol. 9, pp. 14446–14458, 2021.

[28] A. H. Aman, N. Shaari, and R. Ibrahim, "Internet of things energy system: smart applications, technology advancement, and open issues," *International Journal of Energy Research*, vol. 45, no. 6, pp. 8389–8419, 2021.

[29] M. K. Hasan, M. M. Ahmed, A. F. Ismail, S. Islam, A. H. Hashim, and R. Hassan, "Timing synchronization framework for wide area measurement system in smart grid computing," in *2020 Global Conference on Wireless and Optical Technologies (GCWOT)*, pp. 1–5, Malaga, Spain, October 2020.

[30] M. K. Hasan, M. M. Ahmed, Z. Janin, S. Khan, A. H. Abdalla, and S. Islam, "Delay analysis of two-way synchronization scheme for phasor measurement unit based digital smart grid applications," in *2018 IEEE 5th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, pp. 1–6, Songkla, Thailand, November 2018.

[31] "Evolving to a strategic substation network architecture," June 2021, http://www.bomara.com/Garrett/wp_substation_architecture.htm.

[32] N. Nurelmadina, M. K. Hasan, I. Memon et al., "A systematic review on cognitive radio in low power wide area network for industrial IoT applications," *Sustainability*, vol. 13, no. 1, p. 338, 2021.

[33] K. Yu, M. Arifuzzaman, Z. Wen, D. Zhang, and T. Sato, "A key management scheme for secure communications of information centric advanced metering infrastructure in smart grid," *IEEE transactions on instrumentation and measurement*, vol. 64, no. 8, pp. 2072–2085, 2015.

[34] A. K. Ahasan Habib, S. M. Motakabber, M. I. Ibrahimy, and M. K. Hasan, "Active voltage balancing circuit using single switched-capacitor and series LC resonant energy carrier," *Electronics Letters*, vol. 56, no. 20, pp. 1036–1039, 2020.

[35] N. Fatima, T. A. Qasuria, and M. A. Ibrahim, "A brief review on smart grid residential network schemes," *Sains Malaysiana*, vol. 49, no. 12, pp. 2989–2996, 2020.

[36] A. Albayati, N. F. Abdullah, A. Abu-Samah, A. H. Mutlag, and R. Nordin, "A serverless advanced metering infrastructure based on fog-edge computing for a smart grid: a comparison study for energy sector in Iraq," *Energies*, vol. 13, no. 20, article 5460, 2020.

[37] N. Safitri and A. S. Yunus, "Integrated arrangement of advanced power electronics through hybrid smart grid system," *TELKOMNIKA*, vol. 18, no. 6, pp. 3202–3209, 2020.

[38] D. D. Chowdhury, "Synchronization for smart grid infrastructure," in *NextGen Network Synchronization*, pp. 181–207, Springer, Cham, 2021.

[39] M. Liu, I. Dassios, G. Tzounas, and F. Milano, "Model-independent derivative control delay compensation methods for power systems," *Energies*, vol. 13, no. 2, p. 342, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9953416, 14 pages
https://doi.org/10.1155/2021/9953416

Research Article

DMTC: Optimize Energy Consumption in Dynamic Wireless Sensor Network Based on Fog Computing and Fuzzy Multiple Attribute Decision-Making

Abbas Varmaghani,[1] Ali Matin Nazar,[2] Mohsen Ahmadi,[3] Abbas Sharifi,[4] Saeid Jafarzadeh Ghoushchi,[3] and Yaghoub Pourasad[5]

[1]*Department of Computer Engineering, Islamic Azad University of Hamadan, P.O. Box: 8415683111, Hamedan, Iran*
[2]*Institute of Port, Coastal and Offshore Engineering, Ocean College, Zhejiang University, Zhoushan 316021, Zhejiang, China*
[3]*Department of Industrial Engineering, Urmia University of Technology (UUT), P.O. Box: 57166-419, Urmia, Iran*
[4]*Department of Mechanical Engineering, Urmia University of Technology (UUT), P.O. Box: 57166-419, Urmia, Iran*
[5]*Department of Electrical Engineering, Urmia University of Technology (UUT), P.O. Box: 57166-419, Urmia, Iran*

Correspondence should be addressed to Abbas Sharifi; abbas.sharifi@mee.uut.ac.ir

Received 11 March 2021; Revised 31 May 2021; Accepted 18 June 2021; Published 1 July 2021

Academic Editor: Ihsan Ali

Advances in wireless technologies and small computing devices, wireless sensor networks can be superior technology in many applications. Energy supply constraints are one of the most critical measures because they limit the operation of the sensor network; therefore, the optimal use of node energy has always been one of the biggest challenges in wireless sensor networks. Moreover, due to the limited lifespan of nodes in WSN and energy management, increasing network life is one of the most critical challenges in WSN. In this investigation, two computational distributions are presented for a dynamic wireless sensor network; in this fog-based system, computing load was distributed using the optimistic and blind method between fog networks. The presented method with the main four steps is called Distribution-Map-Transfer-Combination (DMTC) method. Also, Fuzzy Multiple Attribute Decision-Making (Fuzzy MADM) is used for clustering and routing network based on the presented distribution methods. Results show that the optimistic method outperformed the blind one and reduced energy consumption, especially in extensive networks; however, in small WSNs, the blind scheme resulted in an energy efficiency network. Furthermore, network growth leads optimistic WSN to save higher energy in comparison with blinded ones. Based on the results of complexity analysis, the presented optimal and blind methods are improved by 28% and 48%, respectively.

1. Introduction

A sensor network consists of many sensor nodes interacting strongly with the physical environment, which receives and responds to environmental information through the sensor. The connection between these nodes is wireless. Each node works independently and has specific capabilities and a certain energy level. To perform the placement operation in some methods, several nodes are equipped with higher capabilities such as higher radio range, more energy, auxiliary equipment for movement, and a GPS receiver [1]. According to the data collection methods, the wireless sensor network can be divided into two categories: homogeneous sensor networks, including base stations and sensor nodes equipped with the same capabilities (e.g., computing power and capacity). They have the same memory. Data collection in these types of networks is based on the data structure. Heterogeneous sensor networks have a base station (complex sensor nodes equipped with advanced processing and communication capabilities) compared to conventional sensor nodes [2].

Sensor distribution (i.e., the location of sensors in the target area) is one of the leading design issues in wireless sensor networks. The location of a sensor may affect the implementation of system requirements and network performance

metrics [3]. Careful placement of the sensors can be an effective optimization tool to achieve the desired design goals. For example, the total coverage is directly related to how the sensors are adequately positioned to cover the desired area on the wireless sensor network. The sensors should not be too close to each other, not overlap, and not be overused. They also should not be too far apart to prevent the formation of coverage gaps in the network. A good distribution makes it possible to perform better in gathering information and communication [3, 4]. Some distribution methods also use stationary sensors to support the sensor location's dynamic adaptation, making it possible to reconfigure a dynamic distribution and improve network performance to minimize energy consumption [5]. During the design process of the network infrastructure, the creation of routes is affected by the sensors' energy limit because wireless transmission is directly related to the second (and higher) power of distance [6]. Using multistep delivery methods will result in less power loss, but using this method will cause problems in topology management and access control to the transmission environment [7]. Therefore, because in most networks, the sensors are randomly located in the network, it is not possible to use multipath methods [8].

Clustering in network routing can significantly affect the overall scalability of the system, lifespan, and energy efficiency [9]. Hierarchical routing is one of the most efficient ways to reduce energy consumption within a cluster and reduce the number of responses sent to the base station [10, 11]. In contrast, a single-level network may overload the gates as traffic congestion increases. In addition, a single-level architecture is not scalable for a large set of nodes because sensors are usually not able to communicate over long distances. In addition, clustering can stabilize network topology along routes and reduce overhead and overall topology maintenance costs. It means that the nodes are protected only when connected to CHs. Furthermore, they are not affected by changes in levels between CHs [12]. CHs can also implement an optimized management strategy that will drive network performance and battery and network life. A CH can schedule intracluster activities so that the nodes switch to sleep mode (low power consumption) and reduce the energy dissipation rate. Nodes can also be used in a rotating order to specify a time for sending and receiving information. As a result, data retrieval is prevented [13].

One of the main goals of wireless sensor network designs is to make data transmission work to extend the network's lifespan and prevent connections from failing through power management methods. Routing protocols in such networks are affected by some challenging factors. Tolerance against error and the ability to organize and expand have been the reasons for the success of wireless sensor networks in applications [14]. Creating an efficient architecture in distributing information between nodes can meet the time lost in the abnormal data filter. These wireless sensor networks can be created within fog computing, distributing the computational load between several nodes [15] effectively. The sensor node is first used to identify the data, and CHs should evaluate it. The tendency toward the adoption of wireless sensor networks has intensified in recent years due to its extensive applications in a variety of industries. A wireless sensor network [16] is formed by linking a large number of sensor nodes. Prior to its actual application, the designed methodology must be tested. Having a live sensor network environment, on the other hand, is not always feasible. In that case, simulation is the only way to test the study before moving on with real-world implementation. To date, a wide range of modeling tools for WSN networks are accessible, some of which are dedicated to wireless sensor networks and others which are applicable to both wireless and wired networks [16]. The distance between the data center and the data source is the fundamental downside of cloud computing. Fog computing is a cloud computing technology that addresses these issues. It is one of the paradigms for distributed service computing. It makes full use of terminal devices' diverse computational features. It has paravirtualized architecture as well [17]. With strict energy and processing resource constraints, distributed detection is a critical challenge for WSNs. The appropriate threshold in most detection cases is determined by the noise power, which is subject to considerable variability in practice [18]. Fog computing adds to the power and benefits of cloud computing and services by extending data generation and analysis to the network edge [19]. Real-time location-based services and applications with mobility assistance are feasible because of the physical proximity of users and a high-speed internet connection to the cloud. To promote fog computing, load balancing approaches are utilized which may be done in two forms, static load balancing and dynamic load balancing [19].

Because most WSNs operate in unattended locations where human access and monitoring are nearly impossible, lifetime improvement has always been a critical concern. Clustering is one of the most effective approaches for organizing system operations in a coordinated manner to improve network scalability, reduce energy consumption, and extend network lifetime. During cluster creation, however, most of the prior techniques overload the cluster leader. To address this issue, various academics devised the concept of fuzzy logic, which is used in WSN decision-making [20]. The clustering hierarchy technique is another approach for data transfer in WSNs. This algorithm is one of the most potent ways for increasing the energy efficiency of WSNs and for maximizing the lifetime of WSNs. WSNs conserve energy by using hierarchical protocols centered on clustering hierarchy. Data might be collected and transmitted to a base station by nodes having more remaining energy. Nevertheless, earlier clustering hierarchy approaches [21] did not account for duplicated data acquired by nearby nodes or nodes that overlapped. Currently used clustering strategies include selecting cluster heads with higher leftover energy and rotating cluster heads on a regular basis to spread energy consumption across nodes in each cluster and lengthen network lifetime. Most earlier algorithms, on the other hand, did not take into account the predicted residual energy, which is then used to predict the remaining energy for selecting a cluster head and performing a round [22].

This study is aimed at working in a computational network of fog with a set of inhomogeneous wireless devices. The objective is to provide a computational distribution

method that reduces energy consumption in the nodes and satisfies the limitations of the edge delay. These nodes are dynamic and can both examine nodes and measure their communication links. The network can be used in smart cities or intelligent buildings that take information from sensors such as (traffic density or temperature) from the environment and use fog computing. Other works include network clustering and routing to reduce energy consumption and extend network life. For this purpose, fuzzy MADM algorithms are used to select optimal CHs. Therefore, the main aims are summarized as the following points:

(i) Provide computational distribution method in Dynamic WSN

(ii) Network routing using fuzzy MADM algorithms

The paper includes Introduction to present the main problem statement, and all need to be satisfied based on fog computing and routing WSN. Related Works provides a brief description of the literature review regarding fog computing. Methods and Materials represents the presented model and governing equations for both computation distribution and routing protocol. In Results and Discussion, the final findings and analysis results are illustrated. Finally, Conclusion summarizes the results and provides the future scope and direction of the study.

2. Related Works

Various methods have been proposed to analyze the spatial and temporal density of routing data; for example, the NMAST method [23] uses the ability of neighboring dynamics to measure the spatial and temporal density of data. k pathways can be utilized for visual investigation in applications such as traffic monitoring, public transit planning, and location selection as dynamic networks, unlike typical clustering algorithms that need several data-dependent hyperparameters [24]. Research of fog computing defined a new generation of support of WSN used in any aspect of smart cities, for example, using the system in the emergency system of fireman [25], traffic light control [26], agricultural system [27], and health monitoring system [28]. One of the challenges of WSN is data privacy. Gathering information and transferring to the base station needs some proper affords like designing security systems. For overcoming this challenge, the fog system is one efficient framework. In this case, an aggregator may be disconnected from the fog server and unable to send data directly. It can, nevertheless, share the encrypted data with an adjacent aggregator in order to send data to the fog server by adding its current collected data to the encrypted data. The relevant data values may be extracted by the fog server and saved in a local repository, which may then be updated in cloud repositories [29]. Storage, communication, transmission ratio, energy consumption, and resilience are all improved by the fog system [29]. As a result, job allocation and secure deduplication are two of the system's tasks. It detects data and protects against security risks. Sharma and Saini [30] proposed a Multi-Objective based Whale Optimization method for the modeling of a fog layer system for safe data deduplication. Average latency, customer happiness, network longevity, energy usage, and security strength all increased as a result of their work [30]. Szynkiewicz et al. [31] developed an energy-aware, secure sensing and computing system centered on static and dynamic clusters and edge and fog computing paradigms. The aggregated data stored at edges were transferred to the base station to analyze by gateways. The results of the implementation enhanced security and offload of data analysis [31].

The following are some examples of effective fuzzy algorithm applications in WSN. To model noisy power uncertainty, Mohammadi et al. [18] employed the fuzzy hypothesis test (FHT). Furthermore, using the Neyman–Pearson lemma on the FHT, they presented an optimum censoring strategy. It is demonstrated that the best censoring strategy may be found by comparing the energy of observed data to a threshold. The threshold would be determined by the local communication limitation and the noise uncertainty limitation, according to the findings [18]. Mohammadi et al. [32] looked at a decentralized detection problem for a WSN and utilized FHT to characterize the noise power uncertainty from a Bayesian perspective. The suggested method was assessed in terms of detection and false alarm probability. In the presence of noisy power uncertainty, simulations indicate that the suggested detector outperforms both the Anderson–Darling approach and the standard energy detector. Nayak and Vathasavai [20] looked into the pros and cons of a variety of clustering techniques. These algorithms are focused on CH efficiency, which might be adaptable, adaptable, and intelligent enough to transfer load across sensor nodes, extending the network lifetime. Menaria et al. [33] introduced an FT technique in WSN to manage faults that happen during data transmission from the sensor to the sink or base station due to link or node failure. An enhanced quadratic minimum spanning tree technique was used in the model. To increase fault tolerance in WSN, the revised technique introduced a unique approach to discover an alternate edge in the spanning tree in place of the broken or failed edge.

In a chapter, Kaur et al. [17] discussed the various aspects of cloud and fog computing platforms. In addition, both platforms' full architectures were provided, along with a comparison study. All application management techniques were examined, including resource coordination, distributed application deployment, and distributed data flow. Different load balancing algorithms were described by Singh et al. [19]. In fog computing settings, round robin load balancing is the simplest and most straightforward load balancing solution. The Source IP Hash load-balancing technique has a critical flaw in that each change might redirect to anybody with a different server, making it unsuitable for fog networks [19]. El Alami and Najid [21] developed an improved clustering hierarchy methodology for overlapping and nearby nodes based on the sleeping-waking process. As a result, data redundancy was reduced to a minimum, and network lifespan was increased. Unlike earlier hierarchical routing algorithms, which needed all nodes to gather and send data, the suggested technique just needed the waking nodes to do so. They use the method in both homogeneous and

Table 1: Summary of literature review of the paper.

Author	Year	Subject	Method	Application	Results
Giang et al. [34]	2020	Large scale, dynamic fog computing in WSN	Distributed node-RED	Base study	Reduce costs
Hossan & Nower [26]	2020	Fog-based WSN dynamic	Neighboring impact factor	Efficient dynamic traffic light control algorithm for multiple intersections	Reduces wait time, lowers fuel consumption, and boosts system throughput
Sharma & Saini [30]	2020	Task allocation and secure deduplication using fog computing	Hybrid Multiplier. Multi-Objective based Whale Optimization algorithm	Base study	Enhancement in average latency, user satisfaction, network lifetime, energy consumption, and security strength
Tsipis et al. [27]	2020	Latency-Adjustable Cloud/Fog Computing Architecture for Time-Sensitive monitoring	Cloud/fog computing paradigm	Environmental Monitoring agricultural activity	Improve efficiency, flexibility, and scalability of the approach in terms of latency
Zeng et al. [35]	2020	Energy powered Cyber-Physical Fog Systems	Mixed-integer linear programming	Cyber-Physical application	The high energy efficiency of our algorithm
Rani & Saini [28]	2020	Secure data collection of fog computing in WSN	The combination of fog and cloud can handle extensive data collection.	Health monitoring	Reduce the cost of data transportation and storage
Bellavista et al. [36]	2020	SDN-based multi-layer routing in fog environments	Multi-Layer Advanced Networking Environment	Smart city	Determines the most suitable path and configures the proper MLR forwarding mechanism
Jain & Goel [37]	2020	Energy efficient fuzzy routing protocol	Fuzzy C-means	Wireless sensor network	High performance, low energy consumption
Tortonesi et al. [38]	2019	Innovative information-centric service model for fog computing	Fog-as-a-service	Smart city environments	An effective platform for running fog services on heterogeneous devices
Sun et al. [39]	2019	Presenting an energy-efficient clustering method for fog computing in WSNs	Cross-layer-sensing clustering method and particles swarm optimization	Base study	Optimize the data aggregation efficiency and improve the network performance
Maatoug et al. [40]	2019	Fog computing framework for energy management	Fog computing framework	Smart building	Decreases latency and improves energy-saving and the efficiency of services among things with different capabilities
Sahith et al. [41]	2019	Face identification in fog computing framework for WSN	Radio communication module XBee, ZigBee protocol	Face identification	Data collection and the functionality of the system are good.
Mihai et al. [42]	2018	Intelligent Data Processing in fog system and WSN	Fog and mist computing approaches	Base study	Improve the information to noise ratio
Bhargava et al. [43]	2017	Fog-enabled WSN system for animal behavior analysis	Edge mining concept	Animal behavior analysis	Accuracy and suitability of the methods

heterogeneous networks. Lee and Cheng [22] suggested a fuzzy-logic-based clustering methodology with an energy prediction extension to extend the lifetime of WSNs. The suggested methodology was found to be more efficient than previous distributed algorithms in simulations. Because edge devices have restricted computing and energy resources, efficient sensor deployment and power management are critical design concerns that must be addressed in order to carry out a significant amount of computation and extend the lifespan of a sensing system to guarantee high-quality monitoring. One of the challenges of the edge-based system is data volume in edge devices. Regarding the exponential increment of data in edges, reducing this congestion can extend the WSN lifespan and improve power consumption. For overcoming this problem, Deng et al. [44] presented a compression method base of fog computing approaches. Their autoregressive analysis method reduced data congestion significantly in conjunction with improvement in power consumption. In some ways, mobile sinks work as fog nodes to connect WSNs and cloud systems. Data are received from sensor nodes and sent to the cloud system through fog nodes(sinks) [45]. Summary of some methods and research about the use of fog computing in WSN are provided in Table 1.

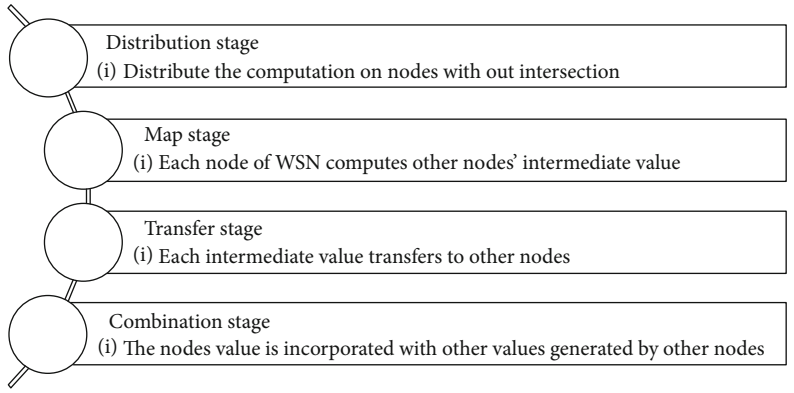


Figure 1: The flow chart of the presented distribution method.

3. Methods and Materials

3.1. The Fog Computing. The devices in fog computing are known as nodes. Each device with a network connection, computing, and storage can be a node that can be placed anywhere with a network connection. A variety of devices, from controllers to switches, routers, and cameras, can act as a fog in WSN. These nodes can be used in target areas such as the office or a vehicle. Each node in WSN is designated as a fog node and computed the primary services.

3.2. Architecture and Working. The conceptual diagram of the proposed fog computing for WSN is illustrated in Figures 1 and 2. In the presented fog computing, WSN is related to four stages of computation. Before starting the computation, data should be distributed to each node without any intersection (distribution stage). Then, in the second stage, each node of WSN computes other nodes' intermediate value from data received from the access point (map stage). In the third stage, each value transfers to other nodes (transfer stage). The node's value is incorporated with other values generated by other nodes in the combination stage. In the presented diagram, red arrow shows the first stage of the computation. At this level, data is divided into node computation memory. The nodes now act as fog computation systems. Then, each node calculated all other nodes' intermediate values at the same time. This stage (map) is depicted in green, blue, and yellow colors. Now each value should be exchanged by other node-related results. In order to achieve this goal, data is exchanged among nodes by the access point (transfer stage). In the final stage of combination, the values of the nodes are combined to reach their exact value.

The presented method with the main four steps is called Distribution-Map-Transfer-Combination (DMTC) method. The presented techniques are based on fog computing, and the main aim is to decrease the computational complexity and dead nodes as well as increase energy efficiency.

There are several reasons or objectives for using fog computing in WSN. These reasons ultimately increase organizational productivity. First is the reduction of latency in the WSN. One of the most significant benefits of fog computing is reducing latency. It is no longer necessary to send data for processing to cloud data centers or base stations, and eliminating this problem makes data analysis and processing much better and more efficient [46]. The second is increasing performance. Not sending data to cloud computing data centers and saving time can also reduce the amount of bandwidth required to do so. In contrast, this amount of bandwidth can be used to communicate with sensors and data centers or base stations [47]. Third is extensive geographical distribution. The use of fog computing with the network's decentralization allows for wider geographical distribution than traditional networking or cloud computing. It will lead to better quality service for the end user [48–51]. Fourth is analysis instantaneously. In many environments, the ability to analyze data immediately is essential. Eliminating inefficiencies and delays in cloud services means that the user can have an accurate and instantaneous data analysis [49, 52].

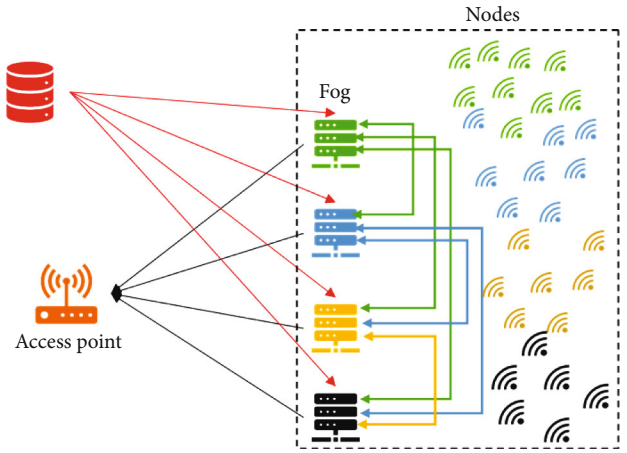


Figure 2: The conceptual diagram of the proposed model.

3.3. Governing Formulation. In this paper, linear mathematical programming is used to optimize energy consumption in the presented DMTC system. The objective function for computing is presented as Eq. (1):

$$M = E_m + E_t + E_c, \tag{1}$$

where E_m, E_t, and E_c are energy consumption for mapping, transfer, and combination stages, respectively. Also, n

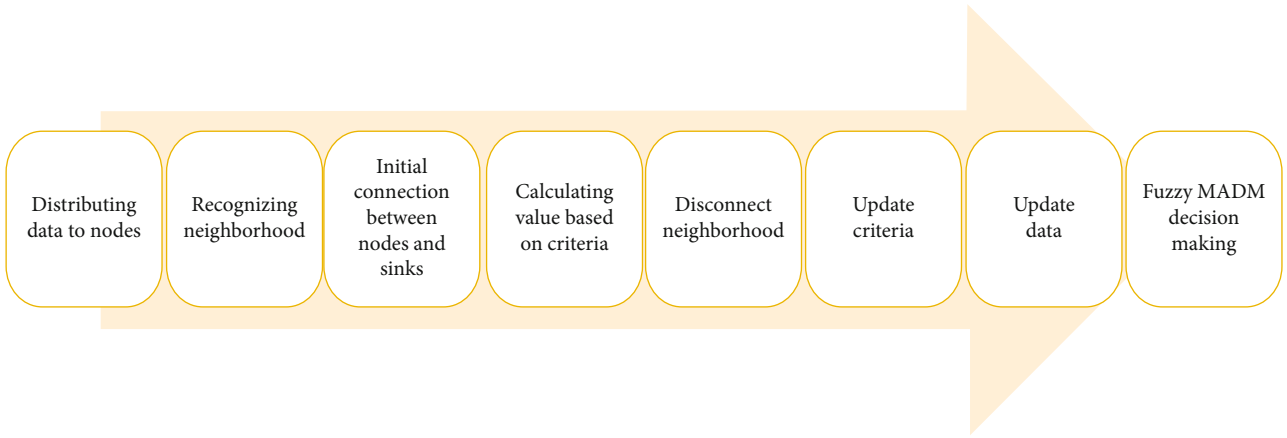


Figure 3: The flowchart of CH detection for routing.

is the node number. Energy consumed in mapping level is defined as follows:

$$E_m = C_n P_n (ND + l_n), \tag{2}$$

such that C_n is the number of CPU cores for processing single bit and P_n is the energy required for the process. Therefore, $C_n P_n$ is the amount of energy for processing a single node. N is the number of nodes, D is distributed data, and l_n is the size of the distributed file.

Moreover, in transfer level, the energy consumption is equal by Eq. (3).

$$E_t = C_n P_n. \tag{3}$$

In this equation, T is the number of bits for computation. E_s shows shuffle level energy consumption is the WSN. It equals by Eq. (4):

$$E_c = \frac{p_n (N-1) T l_n}{\mathrm{LB}.\log_2\left(1 + p_n |h_n|^2 / \Gamma \sigma^2\right)}, \tag{4}$$

where p_n, h_n, B, Γ, and σ^2 are the power of radiofrequency of node n, wireless channel, bandwidth, SNP gap, and noise power, respectively. The following constraints are exerted to the computation:

$$\begin{aligned} &\sum_{n=1}^{N} l_n = L, l_n \geq 0,\ N \geq 1, T \geq 0, \quad B \geq 1, L \geq 1, \\ &\frac{l_n C_n}{F_n} + \frac{E_t}{p_n} \leq \tau_n, \end{aligned} \tag{5}$$

where F_n is the number of CPU process per second in node n and τ_n is defined as the latency of node n. According to mathematical programming, we should obtain the minimum value of energy consumption in the WSN system.

3.4. Clustering and Routing Protocol. In WSN routing, only a small number of nodes must be connected to the base station to increase network lifespan and decrease energy consumption. These nodes are cluster heads (CHs). Because the nodes are dynamic, the most appropriate nodes should be selected as the CHs. In this section of the study, the Fuzzy Multiple Attribute Decision-Making (MADM) method was used to select the CHs. The Fuzzy MADM method uses three criteria: concentration, the energy level in each node, and the node's centrality. The properties of the network are as follows:

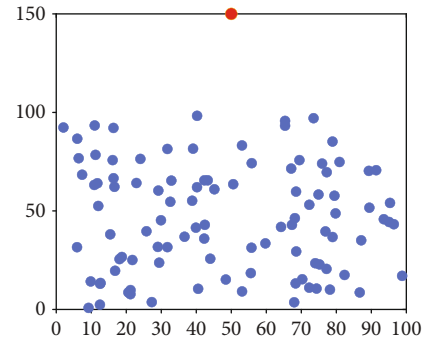


Figure 4: Initial wireless sensor network for the involving problem with 100 nodes and an access point.

(i) The base station must be away from the sensor nodes and immobile

(ii) All network nodes are heterogeneous and have energy limitations

(iii) Nodes have spatial information sent to the base station with the corresponding energy level in the phase adjustment phase

(iv) Nodes are dynamic

In this research, however, routing is based on clustering. However, the choice of CHs based on a method depends on multiple parameters. Therefore, in this study, unlike previous methods where the selection of CHs was mainly based on one criterion or a one-sided approach, in the proposed method of selection protocols, the CHs are chosen based on multicriteria.

According to the flowchart of Figure 3, first, the data is randomly distributed between the nodes. Then, the initial

Table 2: Parameter's value used in the presented DMTC algorithm.

Parameter	Letter	Range	Unit
Energy free space	E_{fs}	10	pJ/bit/m^2
Energy of multipath fading	E_{mp}	0.0013	pJ/bit/m^2
Energy dissipated per bit	E_{elec}	5.5	nJ/bit
Energy consumed in aggregating one-bit data	E_{DA}	5.5	nJ/bit
Number of CPU	C	500-1500	Cycle/bit
Wireless channel	h	10^{-3}	-
CPU process per second	F	0.1-1	GHz
Bandwidth	B	15	kHz
Power of process	P	10-200	pJ/cycle
Noise power	σ^2	1	nW
SNP gap	1	-	

connection between the nodes and the sinks is established to load the data of each node in the system. The criteria should be identified using the existing constraints, and the values based on them should be calculated. In the next step, in order to update the data, the connection between the user and the sink is disconnected, and the criteria will be updated and measured in the new phase so that the final selection can be made based on the Fuzzy MADM method by modifying the existing data and taken from the nodes. In general, the space considered in the flowchart can be described in the following sections.

(i) In the first stage, the establishment of nodes in the field begins so that the mechanism of neighbor detection to discover the general network and create an initial routing tree begins

(ii) In the second stage, the best route from the relay node to the sink is identified

(iii) The final step involves the operation stage, these criteria are monitored, and the value is dynamically changed in response to changes in the status of the network

This method optimizes the lifespan and reduces the error in the network by presenting new constraints and different assumptions. Adding node power consumption as a new constraint can have different challenges in the simulation of the proposed model.

```
Forall(nodes)
   Find(neighborhood);
   Connect (nodes and Access point);
   Distribute (data);
   Em=Eq.(2)
Forall (node i:N)
   I_ii=Calculate (int_value(i,…,N))
   Et=Eq.(3)
For (node i:N)
   I_i=I_i1+I_i2+…I_iN
   Ec=Eq.(4)
   E=Eq.(1)
For all (nodes)
   Calculating (Concentration, Centricity, Energy level)
   Disconnect (neighborhoods)
   Update (criteria)
   Update (data)
   Fuzzy MADM decision making
   Find (CHs)
   Calculate (energy, death node, pack size)
End
```

Algorithm 1: Pseudocode of the presented approach.

4. Results and Discussion

4.1. Architectural Properties of the WSN. As shown in Figure 4, the first sensor network is a square network with dimensions of 100 × 100 m, with the base station (BS) placed away from the sensors. In addition, all sensor nodes are provided 0.1 J of starting energy. As a result, the network's total starting energy is 10 J. The energy parameters E_{fs} and E_{mp} are 10 pJ/bit/m^2 and 0.0013 pJ/bit/m^2, respectively. E_{elec} and E_{DA} parameters have values of 5.5 nJ/bit and 5.5 nJ/bit, respectively. Simulation tests for 100 WSN installations were conducted to guarantee the correctness of the results. To offer a comparative description of the procedures, the average of the collected findings was employed. Experiment has numerous N clusters ranging from ten to twenty to find the ideal value of a cluster. For each value of N, the average energy consumption per cycle is determined. Moreover, the efficiency of optimum computing is studied through mathematical operations. The presented DMTC computing system includes regularly sharing w among the N nodes, without considering the nodes' computing capacities and the power of channel to access point. The used parameters for the simulation are illustrated in Table 2.

4.2. Results of Presented Distribution Analysis. Regarding Figure 4, the presented problem in the initial condition

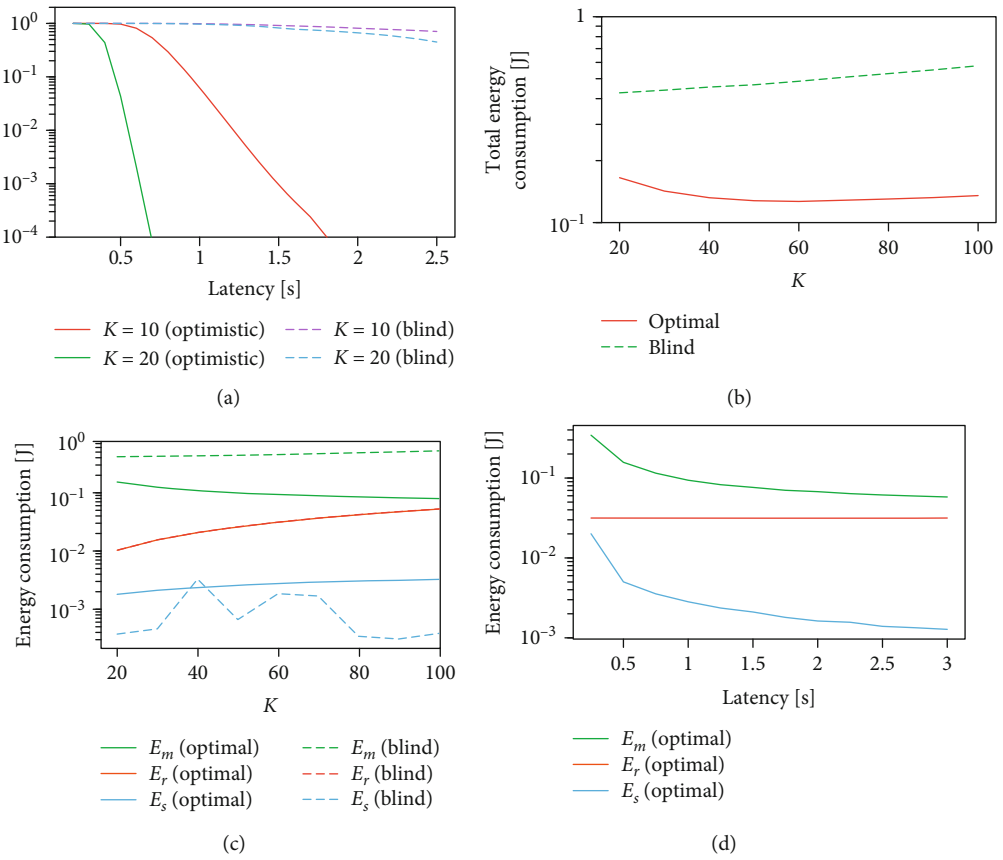


FIGURE 5: Results of the presented DMTC distributed computing method. (a) Interruption probability vs. latency, (b) total energy of fog nodes vs. number of nodes, (c) disruption of energy consumption vs. number of nodes, and (d) energy consumption vs. latency.

consists of 100 dynamic nodes of fogs with an access point. The solution area is 100×100 m the access point is located 50 m upper than the problem area. Before processing the network, the process is equally divided by each node based on the architecture of Figure 1. The presented method is implemented on the different number of nodes N. Two methods of computing are considered as optimistic and blind schemes.

The highest point of computational load for both of the schemes is calculated as the following equation.

$$L_{\text{optimistic}} = \tau_1 \frac{F_1}{C_1} + \cdots + \tau_K \frac{F_N}{C_N}, \tag{6}$$

$$L_{\text{blind}} = N \min \left\{ \tau_1 \frac{F_1}{C_1}, \cdots, \tau_K \frac{F_N}{C_N} \right\}. \tag{7}$$

If we consider each node's capacity as random values, computational load also is random. In this condition, interruption probability is shown in Figure 5(a) in different latency values in 10 and 20 nodes. Findings show that the optimistic distribution among nodes has a lower interruption in comparison with the blind model. In the optimistic method, the computing load is calculated as the sum of the process of each node. However, in the blind method, the load value is equal to N times the minimum value of the nodes process. The results of the distribution method in Figure 5(a) show that the rising number of the total system interruption is decreased that is one of the advantages of this method. Another advantage is the remarkedly low energy consumption of the optimistic approach shown in Figure 5(b) compared to the blind one. The process is done for 100 number nodes with one-second latency.

Regarding Eq. (1), total energy consumption in the presented system is constructed by three E_m, E_t, and E_c as energy consumption for mapping, transfer, and combination stages, respectively. The results of total energy decomposition on the three factors of Figure 5(b) are depicted in Figure 5(c). Based on the results, high percentage of energy belongs to E_m and E_c for mapping and composition, respectively. With an increasing number of nodes, mapping energy decreased. Regarding Figure 5(d), with rising latency, energy consumption for mapping stage is reduced. Another advantage of the presented method is that a slower process leads to reduced energy used. We used the Fuzzy MADM method for routing the wireless sensor network based on the presented distribution algorithm in the other parts of the paper.

4.3. Results of the Clustering Process. Regarding Figure 4, the initial network consists of 100 fog nodes and one access point for connection. The computing load is randomly distributed between nodes based on the methods mentioned above.

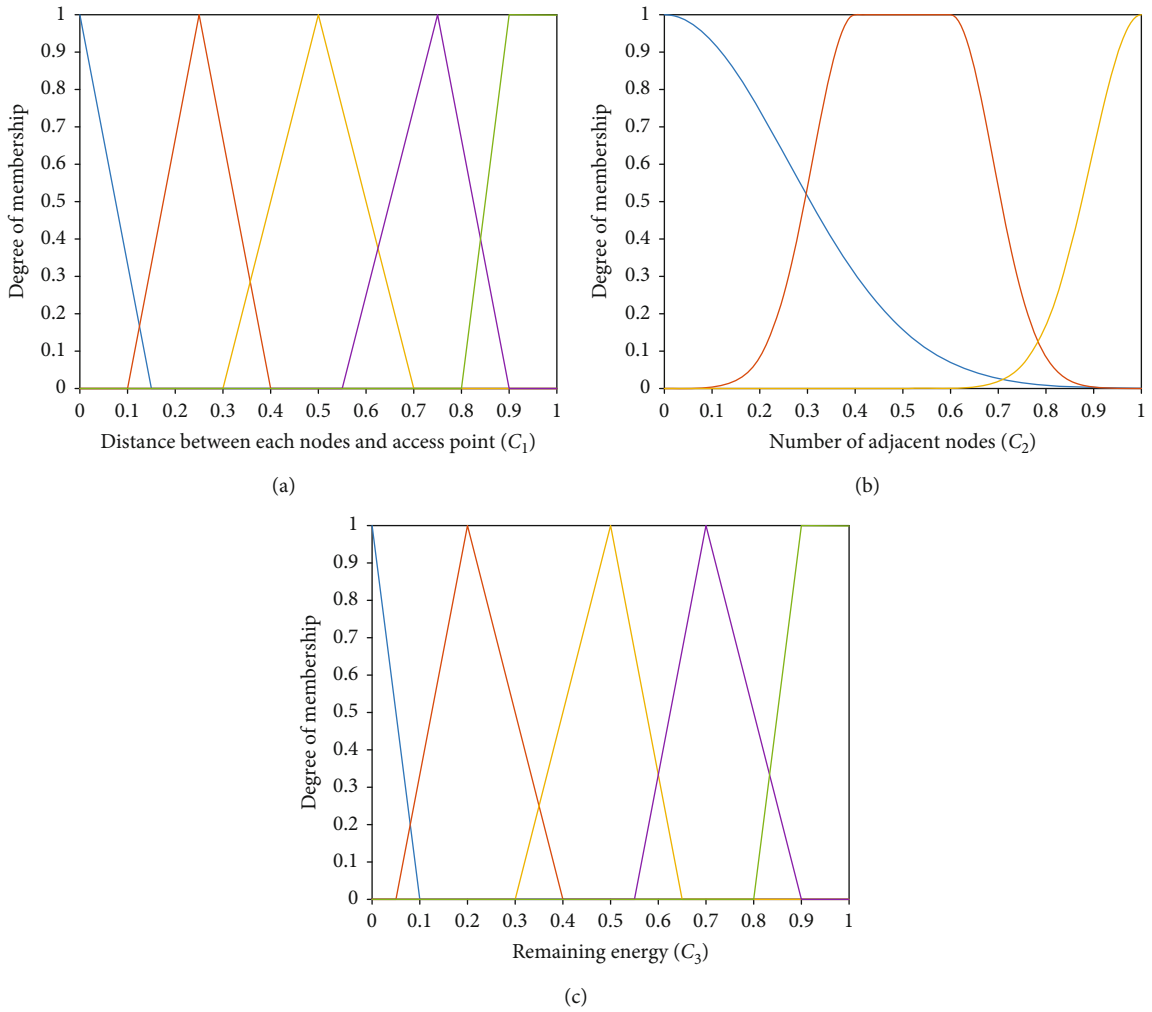


FIGURE 6: Plots of membership functions for the three components of the Fuzzy MADM method.

These processes are performed in any iteration of the presented routing method.

(i) Checking the equipment remaining energy and determining to inactivate equipment due to the depletion of energy

(ii) Determining cluster heads (CHs) based on the Fuzzy MADM methods

(iii) Clustering of the remaining nodes (except CHs) according to the shortest distance to one of the CHs

(iv) Transfer of information from nodes to CHs and then to access points based on radio transmission relationships that lead to energy consumption in nodes

(v) If termination is not done, return the mentioned loop, i.e., check the remaining energy in the nodes and determine the inactive nodes

For determining CHs based on Fuzzy MADM, first, a decision matrix is constructed. The number of rows of the matrix is the number of nodes N, and columns are equal to three numerical criteria of decision as follows:

C_1: distance between each node and access point.
C_2: number of nodes in the adjacency of nodes.
C_3: remaining energy of each node.

In the next step, the matrix is standardized to be ranged between 0 and 1. We used five values of very low, low, media, high, and very high for fuzzification of the matrix based on the adaptive Neuro fuzzy system. The fuzzified standardized criteria of C_1, C_3 using triangular and C_2 using second-order Gaussian function are depicted in Figure 6. The equations for the energy required to transmit information on WSNs comply with wireless communication laws as follows:

$$E_{TX} = \begin{cases} K\left(E_{\text{elec}} + d^2 E_{\text{fs}}\right) d \le d_0, \\ K\left(E_{\text{elec}} + d^4 E_{\text{mp}}\right) d \ge d_{0.} \end{cases} \tag{8}$$

The last steps are determining the CHs, calculating the energy consumption to send information from nodes to CHs and from the CHs to the access point, and implementing the node allocation to the clusters based on the minimum node's distance to the CHs. Total energy consumption is calculated as the amount of energy consumed in data

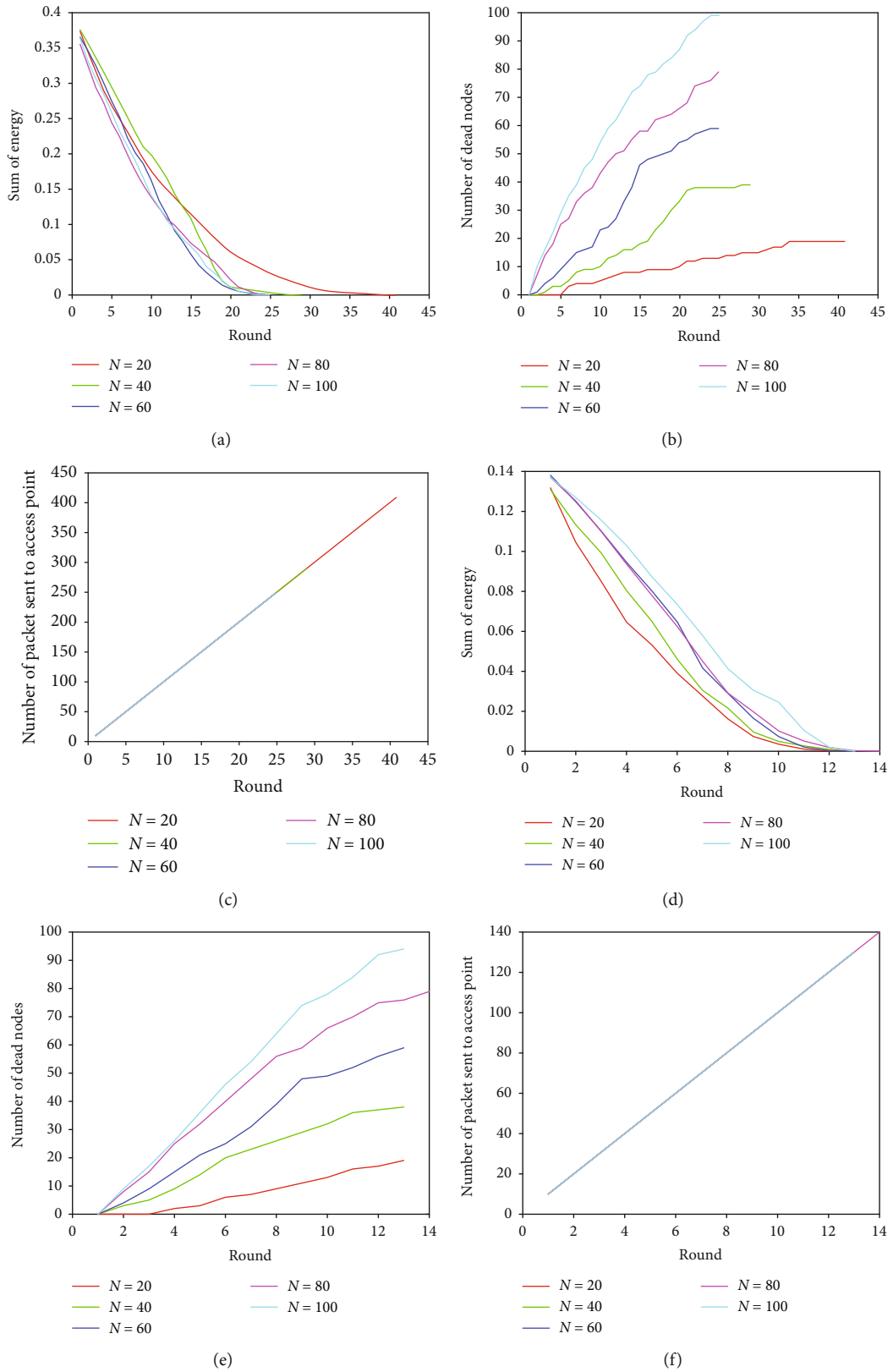


Figure 7: Results of routing for the presented method with Fuzzy MADM approach. (a) Energy consumption vs. round for optimistic model, (b) number of dead nodes vs. round for optimistic model, (c) number of the packet sent to the access point vs. round for optimistic model, (d) energy consumption vs. round for blind model, (e) number of dead nodes vs. round for blind model, and (f) number of the packet sent to the access point vs. round for blind model.

Table 3: The comparison between used protocol and literature review.

Protocol	Network size	Nodes	Death rate 1%	Death rate 50%	Death rate 100%	Computational complexity (s)
SPIN	100×100	100	251	283	379	630
OCM-FCM	100×100	100	400	980	1000	593
FD-LEACH	100×100	100	500	960	1000	591
PEGASIS	100×100	100	762	1216	1270	690
MH-EECDA	100×100	100	630	650	980	503
M-GEAR	100×100	100	951	998	1015	610
Fuzzy MADM (opt)	100×100	100	960	1200	1400	453
Fuzzy MADM (blind)	100×100	100	530	630	700	330

transmission, mapping, and composition by each node with network execution.

The routing is done using two methods of optimistic and blind, and the results are illustrated in Figure 7. Figure 7(a) shows total energy consumption for the network with the optimistic scheme for computational load distribution for the number of nodes $N = 20$, 40, 60, 80, and 100. For all the process of optimistic scheme, computational load $L_{\text{optimistic}}$ (see Eq. (6)) is identical. Energy consumption until all the nodes are dead shows that a network with many nodes is lower energy consumption. However, a network with 20 nodes used a higher value of energy.

Due to the optimal formation of clusters using fuzzy logic and selection of fuzzy CHs, long-distance transmission in the network is further reduced, which CHs show low energy consumption in each sensor node. It is one of the advantages of the optimistic method that has been aforementioned in the previous process. When the remaining energy of a sensor node in a network hits zero, it is called dead. The operational capacity of the network diminishes as the number of dead nodes in the network grows. As a result, sensor node mortality has a direct impact on network operation. Therefore, optimistic method endurance is reduced with the increasing number of nodes based on Figure 7(b). While 40, 60, 80, and 100 nodes are dead, only 20% of the node of 20 cases is dead. Also, the number of packed sent to access point is decreased Figure 7(c) with the increasing number of nodes. On the other hand, in the blind method, according to Figure 7(d), the increasing number of node energy consumption also has risen. Moreover, maximum energy consumption is belonging to the 100 node cases.

In this case, nodes' death is completed in almost the same round and with an identical percentage. Also, sent packages are the same approximately. The comparison between previous research is shown in Table 3. Based on the results, the mortality rate of the Fuzzy MADM method is higher than the other methods, which means that the percentage of death has lately occurred. It shows the reliability of networks based on the high lifespan of the network between the presented distributions. The optimistic scheme enhanced network endurance and led it to be competitive with other protocols. Based on the results of complexity analysis, the presented method is processed in a lower time than the other methods. Considering the SPIN method as a baseline, the presented optimal method and blind methods were 28% and 48%, respectively.

5. Conclusion

Practical engineering in data distribution between nodes in a wireless sensor network can meet the time lost in the irregular information channel. These wireless sensor systems can be made inside the structure of figuring appropriating the computational fog between a few nodes successfully. This investigation is aimed at working in the field of a computational system of fog with a lot of inhomogeneous wireless nodes. The goal is to give a computational distribution strategy that outcomes in diminished energy utilization in the network and fulfills the constraints of the edge latency. The nodes are dynamic and can both look at nodes and measure their correspondence joins. In this paper, we presented the DMTC distribution method for a dynamic wireless sensor network. In this system, one access point plays the base station roles in the system, and nodes are considered fog computing subsystems. The computational load is divided by fog nodes with two optimistic and blind models in the distribution methods. In the optimistic scheme, the computing load is distributed randomly on each node, and the total load is the sum of each node process. On the other hand, in the blind model, the load value equals N times the minimum value of fog node computation. Findings show that the optimistic distribution among nodes has a lower interruption in comparison with the blind model. Also, with the number of nodes, the total system interruption is dropped, which is one of the benefits of the presented approach. Another efficiency is the low energy consumption of the optimistic method. In addition, the high contribution of energy belongs to the mapping and composition stages of energy. Also, with the rising of fog nodes, mapping energy reduced. Moreover, with the growth of latency, energy consumption for the mapping stage is dropped and a slower process consumes a low value of energy. In the next step, the distribution system was implemented on a routing and clustering technique using Fuzzy MADM. Choosing suitable cluster heads can also significantly reduce energy consumption and increase the lifespan of the WSN. The implementation of the routing

method on optimistic and blind schemes revealed that large networks consume lower energy in an optimistic approach than small ones. Also, energy consumption is dropped with clustering and choosing cluster heads. Because nodes' mortality rate influences WSN efficiency, increasing nodes' number network endurance is decremented; however, in the blind method, the efficiency of the network with an increasing number of nodes reduced. To be concluded, the optimistic scheme is proper for an extensive network. However, the blind method is better for a small network.

Fog node resources may be virtualized and distributed to several users. Multitenant support in fog resources and scheduling compute jobs based on their QoS needs have not been thoroughly addressed in the available literature. Future study can be directed toward addressing this gap in the literature. The development of a real-world testbed for testing the operation of fog-based rules is typically quite highly priced and not scalable in many circumstances. As a result, many academics are looking for an effective toolbox for fog simulation to conduct preliminary evaluations of fog computing systems. Nevertheless, there are just a few fog simulators on the market right now. Future research might include the construction of a more efficient simulator for fog computing.

Data Availability

In this paper, a random dataset is used. And the parameters' values are extracted from articles.

Disclosure

The funding sources had no involvement in the study design, collection, analysis, or interpretation of data, writing of the manuscript, or in the decision to submit the manuscript for publication.

Conflicts of Interest

We declare no conflict of interest.

References

[1] N. Jeyakkannan and B. Nagaraj, "Online monitoring of geological methane storage and leakage based on wireless sensor networks," *Asian Journal of Chemistry*, vol. 26, Supplement 2, pp. S23–S26, 2014.

[2] P. Saini and A. K. Sharma, "E-DEEC - enhanced distributed energy efficient clustering scheme for heterogeneous WSN," in *2010 1st International Conference on Parallel, Distributed and Grid Computing, PDGC-2010*, pp. 205–210, Solan, India, 2010.

[3] S. H. Thimmaiah and G. Mahadevan, "Analysis of improved DV-distance algorithm for distributed localization in Wsns," *International Journal of Engineering Research and Development*, vol. 14, no. 2, pp. 16–20, 2018.

[4] A. Gallegos, T. Noguchi, T. Izumi, and Y. Nakatani, "Zone-based energy aware data collection protocol for WSNs," *IEICE Transactions on Communications*, vol. E101.B, no. 3, pp. 750–762, 2018.

[5] A. Petitti, D. di Paola, A. Milella et al., "A network of stationary sensors and mobile robots for distributed ambient intelligence," *IEEE Intelligent Systems*, vol. 31, no. 6, pp. 28–34, 2016.

[6] E. Pietrosemoli, "Long-distance, low-cost wireless data transmission," *URSI Radio Science Bulletin*, vol. 339, no. 339, pp. 23–31, 2011.

[7] B. Song, W. Xiao, and Z. Zhang, "Multi-step sensor scheduling for energy-efficient high-accuracy collaborative target tracking in wireless sensor networks," in *Proceedings -2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, GreenCom-iThings-CPSCom 2013*, pp. 1341–1345, Beijing, China, 2013.

[8] G. P. Sunitha, S. M. D. Kumar, and B. P. V. Kumar, "Energy efficient hierarchical multipath routing protocol to alleviate congestion in WSN," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 32, no. 1, pp. 59–73, 2019.

[9] M. Azharuddin and P. K. Jana, "PSO-based approach for energy-efficient and energy-balanced routing and clustering in wireless sensor networks," *Soft Computing*, vol. 21, no. 22, pp. 6825–6839, 2017.

[10] L. Chan, K. Gomez Chavez, H. Rudolph, and A. Hourani, "Hierarchical routing protocols for wireless sensor network: a compressive survey," *Wireless Networks*, vol. 26, no. 5, pp. 3291–3314, 2020.

[11] P. S. Mann and S. Singh, "Energy-efficient hierarchical routing for wireless sensor networks: a swarm intelligence approach," *Wireless Personal Communications*, vol. 92, no. 2, pp. 785–805, 2017.

[12] M. Jamuna Rani and C. Vasanthanayaki, "Network condition based multi-level image compression and transmission in WSN," *Computer Communications*, vol. 150, pp. 317–324, 2020.

[13] A. Lipare and D. R. Edla, "Cluster head selection and cluster construction using fuzzy logic in WSNs," in *2019 IEEE 16th India Council International Conference, INDICON 2019 - Symposium Proceedings*, pp. 1–4, Rajkot, India, 2019.

[14] M. A. Siddiqi, A. A. Mugheri, and M. Khoso, "Analysis on security methods of wireless sensor network," *Sukkur IBA Journal of Computing and Mathematical Sciences*, vol. 2, no. 1, pp. 52–60, 2018.

[15] G. Zhang and R. Li, "Fog computing architecture-based data acquisition for WSN applications," *China Communications*, vol. 14, no. 11, pp. 69–81, 2017.

[16] A. Nayyar and R. Singh, "A comprehensive review of simulation tools for wireless sensor networks (WSNs)," *Journal of Wireless Networking and Communications*, vol. 5, no. 1, pp. 19–47, 2015.

[17] A. Kaur, P. Singh, and A. Nayyar, "Fog Computing: Building a Road to IoT with Fog Analytics," in *Fog Data Analytics for IoT Applications*, pp. 59–78, Springer, 2020.

[18] A. Mohammadi, S. H. Javadi, D. Ciuonzo, V. Persico, and A. Pescapé, "Distributed detection with fuzzy censoring sensors in the presence of noise uncertainty," *Neurocomputing*, vol. 351, pp. 196–204, 2019.

[19] S. P. Singh, R. Kumar, A. Sharma, and A. Nayyar, "Leveraging energy-efficient load balancing algorithms in fog computing," *Concurrency and Computation: Practice and Experience*, p. e5913, 2020.

[20] P. Nayak and B. Vathasavai, "Energy efficient clustering algorithm for multi-hop wireless sensor network using type-2 fuzzy logic," *IEEE Sensors Journal*, vol. 17, no. 14, pp. 4492–4499, 2017.

[21] H. El Alami and A. Najid, "ECH: an enhanced clustering hierarchy approach to maximize lifetime of wireless sensor networks," *IEEE Access*, vol. 7, pp. 107142–107153, 2019.

[22] J. S. Lee and W. L. Cheng, "Fuzzy-logic-based clustering approach for wireless sensor networks using energy predication," *IEEE Sensors Journal*, vol. 12, no. 9, pp. 2891–2897, 2012.

[23] Y. Yang, J. Cai, H. Yang, J. Zhang, and X. Zhao, "TAD: a trajectory clustering algorithm based on spatial-temporal density analysis," *Expert Systems with Applications*, vol. 139, p. 112846, 2020.

[24] S. Wang, Z. Bao, J. S. Culpepper, T. Sellis, and X. Qin, "Fast large-scale trajectory clustering," *Proceedings of the VLDB Endowment*, vol. 13, no. 1, pp. 29–42, 2019.

[25] L. Campanile, M. Gribaudo, M. Iacono, and M. Mastroianni, "Performance evaluation of a fog WSN infrastructure for emergency management," *Simulation Modelling Practice and Theory*, vol. 104, article 102120, 2020.

[26] S. Hossan and N. Nower, "Fog-based dynamic traffic light control system for improving public transport," *Public Transport*, vol. 12, no. 2, pp. 431–454, 2020.

[27] A. Tsipis, A. Papamichail, G. Koufoudakis, G. Tsoumanis, S. E. Polykalas, and K. Oikonomou, "Latency-adjustable cloud/fog computing architecture for time-sensitive environmental monitoring in olive groves," *AgriEngineering*, vol. 2, no. 1, pp. 175–205, 2020.

[28] S. Rani and P. Saini, "Fog Computing: Applications and Secure Data Aggregation," in *Handbook of Computer Networks and Cyber Security: Principles and Paradigms*, pp. 475–492, Springer, 2019.

[29] A. Ullah, G. Said, M. Sher, and H. Ning, "Fog-assisted secure healthcare data aggregation scheme in IoT-enabled WSN," *Peer-to-Peer Networking and Applications*, vol. 13, no. 1, pp. 163–174, 2020.

[30] S. Sharma and H. Saini, "Fog assisted task allocation and secure deduplication using 2FBO2 and MoWo in cluster-based industrial IoT (IIoT)," *Computer Communications*, vol. 152, pp. 187–199, 2020.

[31] E. Niewiadomska-Szynkiewicz, A. Sikora, J. Kołodziej, and P. Szynkiewicz, "Modelling and simulation of secure energy aware fog sensing systems," *Simulation Modelling Practice and Theory*, vol. 101, article 102011, 2020.

[32] A. Mohammadi, S. H. Javadi, and D. Ciuonzo, "Bayesian fuzzy hypothesis test in wireless sensor networks with noise uncertainty," *Applied Soft Computing*, vol. 77, pp. 218–224, 2019.

[33] V. K. Menaria, S. C. Jain, N. Raju, R. Kumari, A. Nayyar, and E. Hosain, "NLFFT: a novel fault tolerance model using artificial intelligence to improve performance in wireless sensor networks," *IEEE Access*, vol. 8, pp. 149231–149254, 2020.

[34] N. K. Giang, R. Lea, and V. C. M. Leung, "Developing applications in large scale, dynamic fog computing: a case study," *Software: Practice and Experience*, vol. 50, no. 5, pp. 519–532, 2020.

[35] D. Zeng, L. Gu, and H. Yao, "Towards energy efficient service composition in green energy powered Cyber–Physical Fog Systems," *Future Generation Computer Systems*, vol. 105, pp. 757–765, 2020.

[36] P. Bellavista, C. Giannelli, and D. D. P. Montenero, "A reference model and prototype implementation for SDN-based multi layer routing in fog environments," *IEEE Transactions on Network and Service Management*, vol. 17, no. 3, pp. 1460–1473, 2020.

[37] A. Jain and A. K. Goel, "Energy efficient fuzzy routing protocol for wireless sensor networks," *Wireless Personal Communications*, vol. 110, no. 3, pp. 1459–1474, 2020.

[38] M. Tortonesi, M. Govoni, A. Morelli, G. Riberto, C. Stefanelli, and N. Suri, "Taming the IoT data deluge: an innovative information-centric service model for fog computing applications," *Future Generation Computer Systems*, vol. 93, pp. 888–902, 2019.

[39] Z. Sun, L. Wei, C. Xu et al., "An energy-efficient cross-layer-sensing clustering method based on intelligent fog computing in WSNs," *IEEE Access*, vol. 7, pp. 144165–144177, 2019.

[40] A. Maatoug, G. Belalem, and S. Mahmoudi, "Fog computing framework for location-based energy management in smart buildings," *Multiagent and Grid Systems*, vol. 15, no. 1, pp. 39–56, 2019.

[41] S. R. Sahith, S. R. Rudraraju, A. Negi, and N. K. Suryadevara, "Mesh WSN data aggregation and face identification in fog computing framework," in *Proceedings of the International Conference on Sensing Technology, ICST*, pp. 1–6, Sydney, NSW, Australia, 2019.

[42] V. Mihai, C. E. Hanganu, G. Stamatescu, and D. Popescu, "WSN and fog computing integration for intelligent data processing," in *Proceedings of the 10th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2018*, pp. 1–4, Iasi, Romania, 2019.

[43] K. Bhargava, S. Ivanov, C. Kulatunga, and W. Donnelly, "Fog-enabled WSN system for animal behavior analysis in precision dairy," in *2017 International Conference on Computing, Networking and Communications, ICNC 2017*, pp. 504–510, Silicon Valley, CA, USA, 2017.

[44] H. Deng, Z. Guo, R. Lin, and H. Zou, "Fog computing architecture-based data reduction scheme for WSN," in *1st International Conference on Industrial Artificial Intelligence, IAI 2019*, pp. 1–6, Shenyang, China, 2019.

[45] T. Wang, J. Zeng, Y. Lai et al., "Data collection from WSNs to the cloud based on mobile Fog elements," *Future Generation Computer Systems*, vol. 105, pp. 864–872, 2020.

[46] Y. Y. Shih, W. H. Chung, A. C. Pang, T. C. Chiu, and H. Y. Wei, "Enabling low-latency applications in fog-radio access networks," *IEEE Network*, vol. 31, no. 1, pp. 52–58, 2017.

[47] K. Bhargava and S. Ivanov, "A fog computing approach for localization in WSN," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, pp. 1–7, Montreal, QC, Canada, 2018.

[48] A. H. Gandomi, A. H. Alavi, A. Asghari, H. Niroomand, and A. M. Nazar, "An innovative approach for modeling of hysteretic energy demand in steel moment resisting frames," *Neural Computing and Applications*, vol. 24, no. 6, pp. 1285–1291, 2014.

[49] A. Matin Nazar, P. Jiao, Q. Zhang, K. J. I. Egbe, and A. H. Alavi, "A new structural health monitoring approach based on smartphone measurements of magnetic field intensity," *IEEE Instrumentation & Measurement Magazine*, vol. 24, pp. 49-50, 2021.

[50] M. H. Syed, E. B. Fernandez, and M. Ilyas, "A pattern for fog computing," in *ACM International Conference Proceeding Series*, pp. 1–10, New York, NY, USA, 2016.

[51] A. M. Nazar, K.-J. I. Egbe, P. Jiao, and A. H. Alavi, "A novel multi-mode magnetic triboelectric nanogenerator energy harvesting system," in *Behavior and Mechanics of Multifunctional Materials XV*, 2021.

[52] K. Intharawijitr, K. Iida, and H. Koga, "Analysis of fog model considering computing and communication latency in 5G cellular networks," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops*, pp. 1–4, Sydney, NSW, Australia, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5580179, 14 pages
https://doi.org/10.1155/2021/5580179

Research Article

IoT with BlockChain: A Futuristic Approach in Agriculture and Food Supply Chain

Sabir Awan,[1] Sheeraz Ahmed,[1] Fasee Ullah,[2] Asif Nawaz,[3] Atif Khan,[4] M. Irfan Uddin,[5] Abdullah Alharbi,[6] Wael Alosaimi,[6] and Hashem Alyami[7]

[1]*IQRA National University, Peshawar, Pakistan*
[2]*Department of Computer and Information Science, University of Macau, Macau*
[3]*ETS, Higher Colleges of Technology Dubai Women College, UAE*
[4]*Department of Computer Science, Islamia College Peshawar, Peshawar 25120, Pakistan*
[5]*Institute of Computing, Kohat University of Science and Technology, Kohat 26000, Pakistan*
[6]*Department of Information Technology, College of Computers and Information Technology, Taif University, P.O.Box 11099, Taif 21944, Saudi Arabia*
[7]*Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to M. Irfan Uddin; mirfanud@gmail.com

Received 20 January 2021; Revised 15 March 2021; Accepted 2 April 2021; Published 23 June 2021

Academic Editor: Suleman Khan

Agricultural food production is projected to be 70% higher by 2050 than it is today, with the world population rising to more than 9 billion, 34% higher than it is now. The farmers have been forced to produce more with the same resources. This pressure means that optimizing productivity is one of the main objectives of the producers but also in a sustainable way. Not only does agriculture face a decline in production, but it has also had to face limitations in data collection, storing, securing, and sharing, climate change, increases in input prices, traditional food supply chain systems where there is no direct connection between the farmer and the buyer, and limitations on energy use. Existing IoT-based agriculture systems have a centralized format and operate in isolation, leaving room for unresolved issues and major concerns, including data security, manipulation, and single failure points. This paper proposes a futuristic IoT with a blockchain model to meet these challenges. Further, this paper also proposes and novel energy-efficient clustering IoT-based agriculture protocol for lower energy consumption and network stability and compares its results with its counterpart low-energy adoptive clustering hierarchy (LEACH) protocol. The simulation results show that the proposed protocol network stability is 23% higher as compared to LEACH as first node of LEACH dies at 168 rounds while IoT-based agriculture first node dies after 463 rounds. Similarly, IoT-based agriculture protocol energy consumption is 68% lower than that of LEACH. The proposed protocol also extends the network life to more rounds and demonstrates an increase of 112%.

1. Introduction

The United Nations Population Division expects the world population, currently 7.8 billion (2020), to reach 10.9 billion after the end of the 21st century [1]. Due to rapid population growth, there is a high pressure on agriculture to increase food production sustainably. Food distribution and consumption simultaneously promote human well-being and preserves scarce natural resources. As a result, policymakers, development agencies, civil society organizations, and private enterprises have shown interest in investigating the role of the food and farm markets in supporting sustainable development for people and the world [2]. Agriculture is experiencing drastic changes and is facing numerous environmental and social problems. Many farmers are still depending on traditional farming practices and having no direct access to the market, and it has proven difficult to balance the demands on limited natural resources such as land

and water. Both changes in diets and consumer preferences and the fact of climate change contribute to the complexity of providing high-quality food to the end-user. Sustainable food and agricultural production cannot be accomplished by the conventional agriculture systems that have led to substantial deforestation, water scarcity, or soil erosion. Thus, advanced systems must be used which conserve and reinforce the basis of natural resources and increase production. A phase of transition to "holistic" methods is like smart agriculture [3]. Smart agriculture is an approach that guides agricultural field supervision in the era of climate change. The idea of smart agriculture was first introduced in 2009 and revived through the involvement and input of many stakeholders involved in design and implementation. In response to the debate on agricultural policy on climate change and sustainable development work, the main features of the smart agriculture approach have been developed. [4]. Smart agriculture aims to provide universally applicable systems for climate change management and agricultural food security. Smart agriculture uses advanced technologies such as the Internet of Things (IoT), big data, and cloud computing to monitor the field environment, analyze crop growth, and provide information to the farmer for decision-making. An information-based management cycle for smart agriculture is presented in Figure 1. [5].

1.1. Internet of Things (IoT). The Internet of Things (IoT) is a cutting-edge computing and networking technology that reflects the future. Smart computing will be focused on the Internet of Things in the future. IoT is now playing a significant role in the transformation of existing technologies from the home to the workplace [6]. IoT technology is used to link computers or nodes to networks for the sharing of information and communication [7]. The Internet of Things (IoT) has the potential to link billions of devices (smart objects). These smart objects can gather data and interact with other systems over the internet. IoT has evolved into a next generation technology with several agricultural applications [8].

1.2. Internet of Things in Smart Agriculture. In the field of smart agriculture IoT provides a wide range of applications such as soil and plant tracking, crop growth observation, and selection, assistance for irrigation assessment, and monitoring of the agriculture environment. In smart agriculture Internet of Things (IoT) technology is applied in diagnostics and control. To optimize agriculture, the implementation of IoT in the field has increased the productivity and effectiveness of farmers. It may help to determine field variables such as soil quality and plant biomass. It can also be used to test and monitor variables including temperature, soil moisture, and crop diseases. Besides, IoT can be used to track crop growth and yield influencing factors. Farmers can also figure out which crops are most suited for which conditions and can rotate crop accordingly [9]. IoT applications support farmers during crop planting, irrigation, crop processing, harvesting and postharvest, crop storage and transportation, and many other benefits in agricultural IoT systems. Soil moisture sensors, humidity sensors, leaf moisture sensors, solar radiation sensors, infrared light sensors, and rainfall predictors are among the field sensors used in IoT-based systems. In IoT scenarios, sensors can be installed in a variety of locations, including greenhouses, seed banks, cold rooms, agricultural machinery, transportation systems, and livestock, and the data collected can be processed in the cloud for monitoring and control [10].

1.3. Food Supply Chain Management. Food supply chain management is a process that explains how food from an agricultural field ends up on our tables. Supply chain management deals with production, refining, delivery, selling consumption, and disposal [11]. The process of the supply chain is summarized in Figure 2. In developing countries, the food supply chain faces several challenges, such as the need for confidence among stakeholders which often correlated with their credibility and traceability required by the end-users, and the difficulty of managing risks, delays, or disruptions is often occurred due to insufficient or lacking information [12]. Blockchain technology is one of the best ways to meet these challenges [13].

1.4. Blockchain Technology. Blockchain is a distributed database where data can be recorded and shared via a decentralized computing network while also providing security and privacy. And if the data is spread, only the owner who has the private key can make transactions. The other machines or computers on the network serve as validators. It safely records transactions between nodes in a public ledger without the need for a trusted third party [14]. In a centralized cloud approach where an asset or object is owned, it is either held under the control of the owner or through a trusted intermediary or a centralized authority such as a bank. A centralized computing system always has a single server for several clients, the server has higher resources than its client systems, and all processes are made through a single server which now has become a traditional approach [15]. Blockchain records transactions in block units, and every block includes the hash of the block history, hash of the current block, date, other details, and transactions for the block. When a sender node makes a transaction, it distributes it to the other nodes on the network. The receiving nodes verify the transaction and have proof of work. The node that succeeds in the proof of work will broadcast it to all the other nodes and connect the block to the chain. The transaction shall contain the public key of the recipient and shall be signed by the sender. Therefore, any other node will verify the validity of the transaction. Each block includes a hash of the previous block, which means that each block is connected to each other [16]. The key characteristics of blockchain are shown in Figure 3 [17].

Because of its unique characteristics, blockchain may be an evolutionary next step in the food supply chain also to add accountability to the supply chain system by exchanging accurate data between supply chain stakeholders. After the integration of IoT with blockchain, the overall visibility of food products across the supply chain will become a reality. The major advantages of blockchain and IoT in the food supply chain are real-time monitoring and sensing of original food items from the origin which identify major bottlenecks.

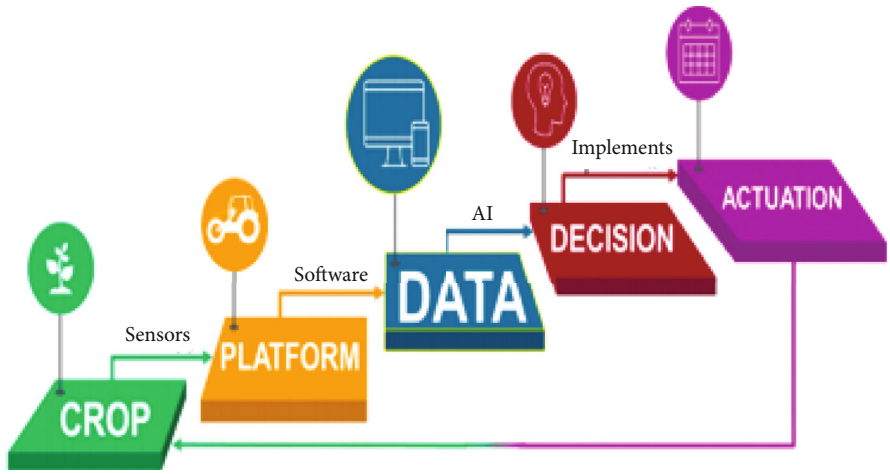


Figure 1: Information-based management cycle for smart agriculture.

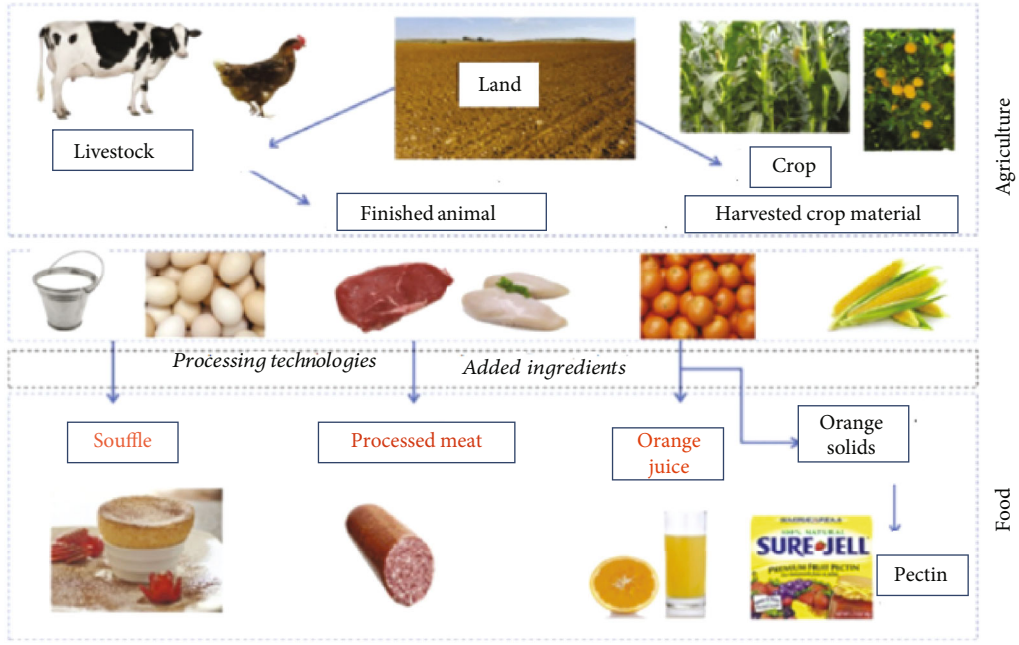


Figure 2: Agricultural food supply chain process.

As in the agricultural sector, farmers use chemical fertilizers, insecticides, and other materials to increase crop production, especially vegetable growers often use harmful pesticides to spray vegetables to improve their profitability, and consumers buy the same vegetables to eat which is a major health hazard. Secondly, food frauds and data alternation are also big challenges for supply chain stakeholders.

1.5. Blockchain in Food Supply Chain. Blockchain is an emerging technology for the agriculture sector, applying blockchain to the agricultural supply chain provides a digital database that monitors, tracks records, and processes digital and physical resources. Blockchain allows greater traceability and higher quality transactions. This technology can manage and integrate all procedures and dealings in real-time across the agricultural supply chain. Every transaction treated in a distributed manner may contain product-specific attributes and transactions that will be added by players in the supply chain. The supply chain players can recognize and investigate the association of products along with all stages of the supply chain and zoological technology practices (fertilizers, feed, and veterinary services [18]. The blockchain stores invariant records that are transparent and digitally accessible to all users in the supply chain; this technology has the potential to produce massive efficiency gains for each actor. The blockchain offers a forum for traceability in the agricultural supply chain, allowing end-users to monitor product from origins and ensure their authenticity. Figure 4 illustrates the benefits of using blockchain in the food supply chain. The ability to monitor a

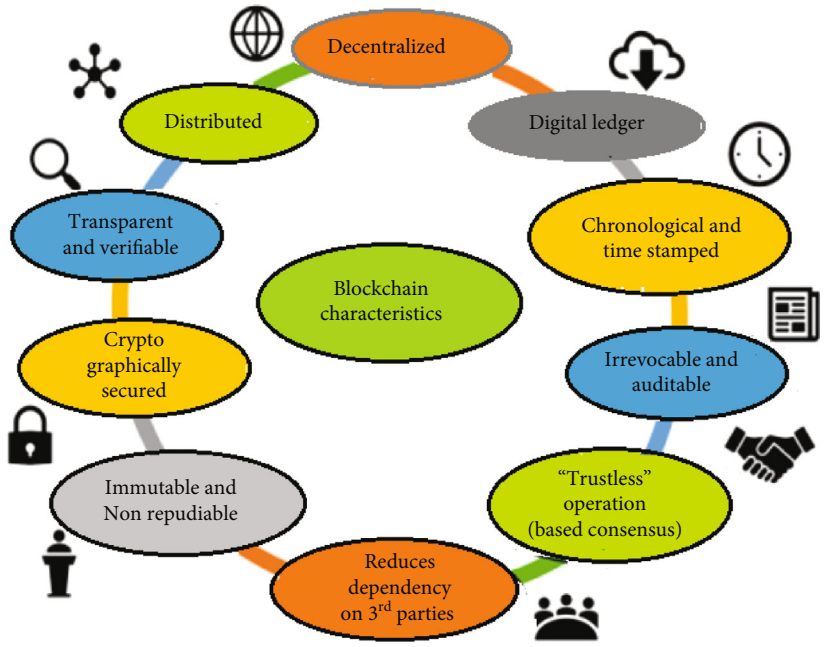


Figure 3: Key Characteristics of Blockchain Technology.

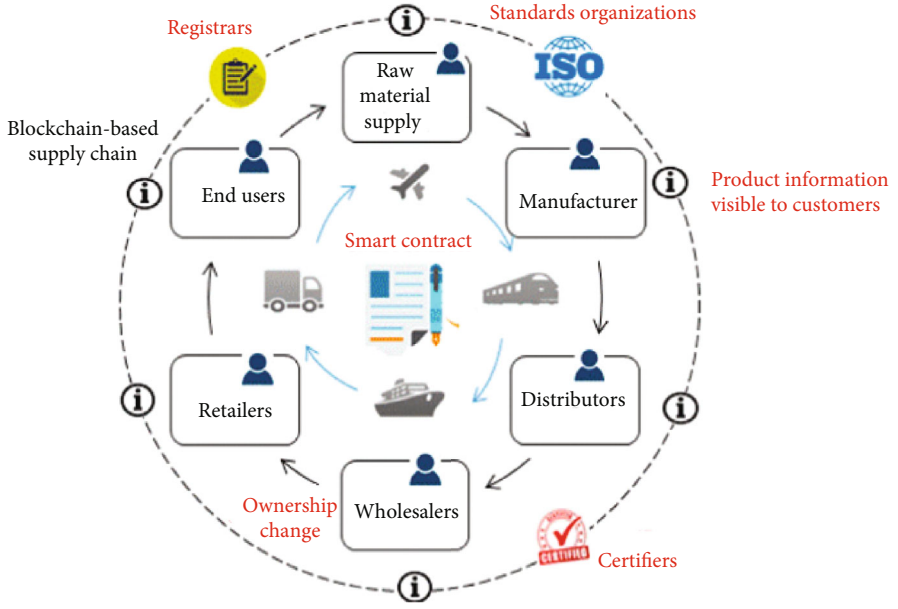


Figure 4: Advantages of Blockchain Technology in Food Supply Chain.

product's movements allows for legal liability for infringement on the product's authenticity, making regulation easier to regulate. The blockchain keeps the mechanism and the data connected at all times [19].

1.6. IoT Integration with the Blockchain. Industry 4.0 has shown that the Internet of Things is a fundamental technology and a key player in the digital revolution. In different network topologies, IoT is expected to depend more on sensing devices, a wide range of data, and more connected devices. As a result, it must be referred to as the Internet of Things 2.0. The Internet of Things 2.0 is now transitioning away from sensors and data technology and toward actionable intelligence technology [20]. IoT becomes more efficient when it is combined with cloud computing, artificial intelligence, and machine learning [21]. Integration of blockchain technology with IoT is another significant contribution to the digital transformation of various domains. By integrating with IoT, blockchain is projected to add $176 billion to the global economy by 2025 and $3 trillion by 2030 [22]. Figure 5 depicts the most important blockchain and IoT predictions for 2030.

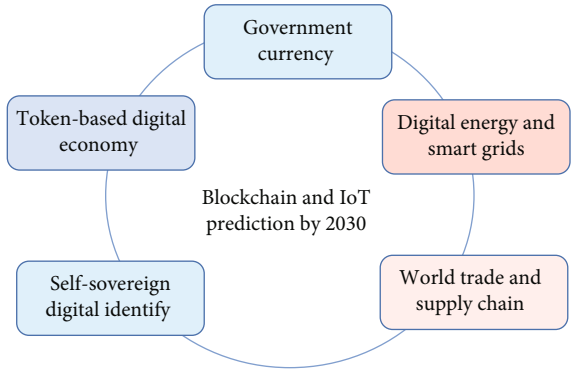


Figure 5: Blockchain and IoT prediction by 2030.

In the present world, the problem of health and food safety is deteriorating day by day because a proper food traceability system is not available. The biggest reason for foodborne disease is contamination which is hard to track in the conventional food supply chain because the mechanism is not visible [23]. However, IoT with a blockchain system will fix the problem by creating a distributed system and eliminating the third party participating in the system at the same time. Consumers have questions about the nature of the food products they purchase and the sources of the product they buy in the conventional supply chain. The key point here is to create customer interest, which is difficult to do as the system is not open. By building customer confidence, IoT and blockchain technology could eradicate these issues. Each product has a digital identity in this framework, which stores information from the point of origin up to the end of the retailer [24]. IoT with blockchain can provide a more transparent and connected system.

1.7. Communication Protocols. Communication protocols have been commonly used in various types of networks to improve communication efficiency. The LoRaWAN protocol was created to migrate data from sensor nodes and the IoT network platform that introduced the LoRaWAN backend service to the cloud for data sharing. In terms of integrating with other IoT systems and introducing new services, the planned framework was scalable and versatile. It is also horizontally scalable, which means it can improve efficiency by replicating new server occurrences [25]. Heizelmann created the low-energy adaptive wireless sensor network routing LEACH protocol, which was the first hierarchical cluster-based routing protocol for a wireless sensor network. The protocol nodes had been distributed across clusters, and each cluster had its head (CH). Member nodes of cluster sensory data share them via their selected CHs with the base station. Smart agriculture has been able to ensure that it not only meets the general needs of crops but also anticipates their specific needs, taking into account a variety of environmental factors. The challenge is to accurately predict potential crop losses to take appropriate measures. IoT nodes enable efficient routing protocols to obtain data and share information with farmers [26].

2. Literature Review

This section provides a thorough review of the related literature.

2.1. Review on IoT in Smart Agriculture. In paper [26], the authors proposed a smart agriculture system to provide information on soil, water supply, and the general state of the field to farmers. The objective of the proposed system was to make agriculture and irrigation more efficient and to make it easier for farmers to take appropriate decisions to increase agricultural production. The use of IoT networks to monitor environmental factors and the combination of this information with a user-specific web service enables farmers to make effective use of their knowledge to obtain the best results from their agriculture. In paper [27], researchers have developed an IoT-based agriculture data analysis system using IoT devices such as environmental field real-time monitoring sensors. The field observation was collected and transferred to the farmers through a web-based application with the help of a node and a Wi-Fi module. The system consists of sensors such as a soil moisture sensor and an ultrasonic sensor. The system has a low cost but uses high energy. The authors of the paper [28] used IoT to build a smart agriculture system for Indian farmers. Farmers in India have begun to commit suicide as a result of low agricultural output. To address this problem, an IoT-based smart agriculture system was created that uses a variety of sensors to monitor pests and soil moisture. Crop distribution companies were also involved in providing high-quality seeds to farmers to boost yields. The system architecture, on the other hand, is extremely durable, but it comes at a high price.

Researchers explored the potential benefits of using smart agriculture technologies such as smart soil and air sensors in paper [29], to assist farmers by covering the initial installation costs and providing smart farming advice through the integrated use of heterogeneous information sources. The work helped farmers in growing their ecological footprint by providing opportunities for innovation targeting and climate change adaptation options. The authors of the paper [30] used sophisticated Internet of Thing (IoT) agricultural applications to make accurate measurements, using sensor data from 60 scientific papers published between 2016 and 2018. Previous research has shown the importance of smart agriculture in water, grain, livestock, and irrigation management. When it came to sensor data collection, the best results were obtained when calculating temperature and humidity. Sensors for soil moisture and soil reaction are also available.

2.2. Review on Blockchain in the Supply Chain. Researchers in the paper [31] proposed a theoretical model to investigate the impact of blockchain on operational supply chain transparency and collaboration by establishing rapid trust between key partners involved in disaster relief supply chains. Collaboration enables stakeholders to work together, leads to productive and efficient real-time information exchange between partners, and provides many benefits such as increased transparency, flexibility, and reduced lead–time, as well as helping to improve supply chain resilience. The authors of the paper

[32] suggested blockchain adoption as a way to improve the entire supply chain, from safe, accurate, and open transactions to increasing supply chain participants' trust and efficiency by exchanging all transactions and related information across the entire network. The authors of [14] focused on weakening supply chain resilience in the era of information asymmetry, which occurs when one of the supply chain's members has significantly more product information than the others. They admitted that data inaccuracies or discrepancies might lead to incorrect decisions and an increase in technological risks including cybercrime, hacking, and theft, both of which could harm the company's credibility. One of these solutions, according to the writers, is establishing "confidence" among supply chain participants using a decentralized information sharing framework based on blockchain technology.

2.3. Motivation. In terms of technology, although the literature in smart agriculture using IoT is very rich, there is still a gap for improvement, whereas the use of blockchain technology in the food supply system literature is very limited. Besides, no solid research papers are published yet using a combination of IoT and blockchain technology for agriculture and food supply systems for addressing the information reliability issues. Most of the solutions are ad hoc-based or work separately. Limitations in the existing literature motivated us to develop a futuristic smart model based on smart technologies to meet the requirements of all stakeholders involved in agriculture and food supply chain by using IoT with blockchain technology through which traceability management can be applied through the food supply chain and addresses the main centralized format challenge that is asymmetric, opaque, and monopolistic. These limitations have driven us to build a distributed, efficient, and safe system to enhance crop production, ensure the end-to-end traceability of food products, and stop food fraud in the supply chain. The goal would be to develop a distributed and automatic agriculture food supply chain traceability model that enables various quality levels to be implemented based on consumer needs. The model that not only uses transaction data but also smart contract and enables end-to-end traceability and data security.

3. Materials and Methods

This section presents the smart model architecture and smart model protocol design.

3.1. Smart Model Architecture. As shown in Figure 6, a futuristic smart model for agricultural environmental monitoring and food supply chain consists of three layers: physical data layer, logical data layer, and web interface layer. This type of layered approach allows for scalable, extendable, and efficient framework implementation. In the physical data layer, a variety of IoT nodes are used to track the farm environment and crop growth. IoT nodes collect data from the cluster farm and send it to the base station via IoT gateway and wireless router, which then sends it to the database. A GPRS router is integrated into a single central board device that serves as a remote Radio Frequency Gateway (RFG) for wireless telemetry. To achieve effective control management and synchronization of two relatively unrelated data sources, data collection through IoT nodes for soil parameters and IoT crop monitoring information, the RFG gateway acts as a coordinator between two separate data streams and supports remote access, allowing for complete remote control of devices in the cluster farm. The logical data layer in an SQL database server stores the data obtained from the cluster farm. The SQL server, as a more intermediate layer, covers the complexity of multiple physical layer devices and allows database server data validation. Raw data is stored in the SQL database, which is then extracted to a local file system.

Energy is the core part of every network, and energy conservation is one of the key objectives of this research. To reduce energy consumption, this research proposed an IoT-based agriculture protocol that can be divided into three phases: IoT nodes, IoT cluster head, and sink. IoT nodes gathered data and transferred it to CH in the first phase, CH broadcasted information to sink in the second phase, and sink transferred data to the base station in the third phase. This research work implements an RFG server using a robust ultra-low-power Single Board Chip to seamlessly combine a variety of devices in the cluster farm (SBC TS-7260). The monitoring system switches between active and sleep mode automatically. The optional battery backup module supports the SBC sleep mode, which also operates with a built-in uninterrupted power supply (UPS). Different devices in the cluster farm are normally equipped with an RS232 serial port. Wireless routers are switched off during the system's sleep cycle to conserve electricity. SQL server obtains data by either pull or pushes operations. In a pull method, SQL server connects to the data source regularly and extracts the data. In a push operation, this server opens a port through which data can be passed through the data source. This design allows the server to adapt to different data source types with ease. The data was sent to the SQL server. The data manager may also require the RFG server to recollect data and retransfer data packets that were omitted. For each new data source, a new data manager is included with limited modifications to the database and web visualization layers, improving extensibility. Specific IoT identification codes link the nodes and observations. As a consequence, once they are linked to the database, the web application layer will recognize them right away. This design allows for flexible model development and web interface design. It also allows for data conversion to IoT networks, allowing for data sharing and interoperability between IoT devices through web services.

3.2. BlockChain Integration with IoT in the Smart Model. Exclusive blockchain characteristics will combine agricultural and food supply chain processes into a single smart system to ensure that consumers receive healthy food. Figure 7 shows a functional overview of the blockchain.

The role of stakeholders in the overall system is also discussed. The research used blockchain smart contracts to exchange data between mining nodes in the system. All business transactions are recorded in the shared ledger by mining nodes, and smart contracts receive all transactions in the blockchain in the form of function calls and generate

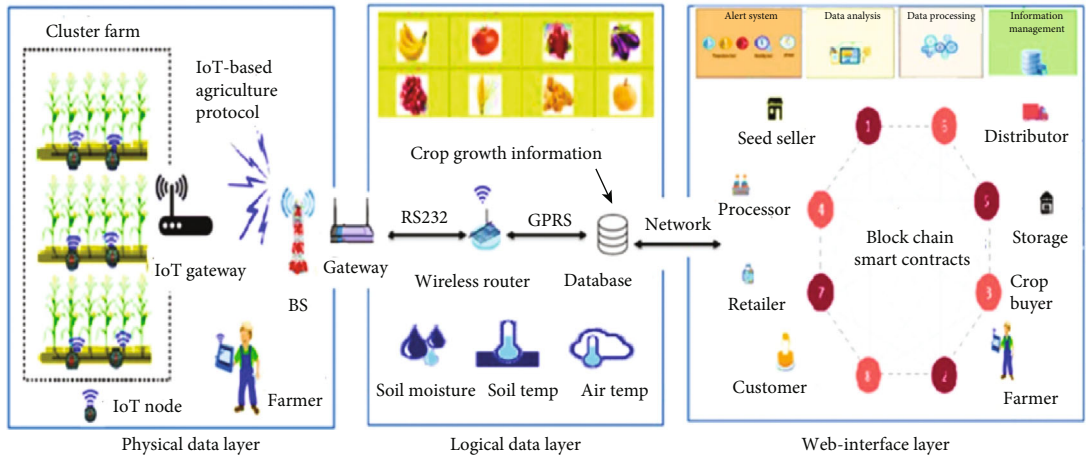


Figure 6: IoT and blockchain smart model.

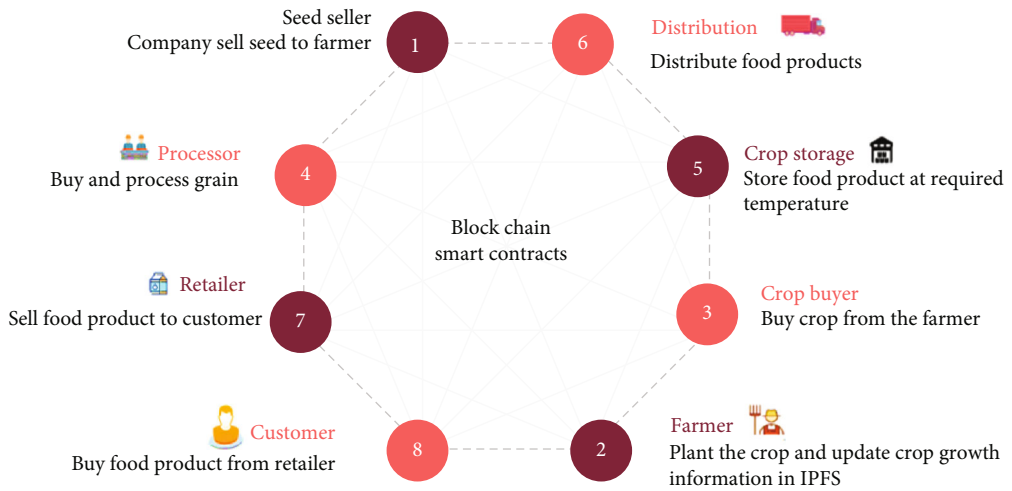


Figure 7: Blockchain-based food supply chain.

activities, as well as providing access to parties involved in the transaction to exchange control track and receive alerts in the event of a violation. Finally, smart contracts help to maintain the best conditions and respond to food supply chain misappropriations.

3.2.1. Seed Seller. Seed sellers, also known as seed vendors, organize and produce a wide range of seeds before selling them to farmers. They are obligated to sell high-quality seeds to growers to produce a high-quality crop. The seed seller uses blockchain to store information about seed germination and chemical composition, which all stakeholders can access and share. Seeds are stamped with a unique number, similar to serialized Global Trade Identification Numbers, that allows all parties to be electronically identified and track products in the seed trade.

3.2.2. Farmer. A farmer is someone who buys high-quality seeds and plants them. The farmer is responsible for recording information about fertilizers use and plant growth data in the form of MPEG files in the IPFS, where different nodes keep the data with high reliability throughout the entire process from planting to harvesting. Farmers can also use blockchain to communicate with crop buyers and sell their crops without relying on middlemen, who often get profit by mediating between sellers and buyers.

3.2.3. Crop Buyer. Crop buyer is a company that buys crop or a grain from a farmer and sells it to a processor.

3.2.4. Processor. The processor buys grain that does a moisture analysis to convert the raw grain to the final product. The processor is responsible for all operations in the blockchain where other players in the supply chain have access to this information.

3.2.5. Crop Storage. The storing of crops or grains is an important and unavoidable part of the crop production process and supply chain process. In warehouse, grains or crops are stored, and information is recorded in the blockchain where supply chain stakeholders can view and exchange this information.

3.2.6. Distributor. Distributors have a business arrangement with the suppliers or producers. Many distributors maintain exclusive purchasing agreements which limit or allow distributors to cover a certain area through the number of participants. The distributor is the producer's exclusive point of contact with prospective purchasers of those products. Distributors, though, seldom sell the goods of the manufacturer directly to the retailer. Distributors tend to work with wholesale agents who may buy large quantities of one commodity regardless of the very large quantity of one commodity they have on hand or can purchase from the manufacturer. The distributor is responsible for registering all food product information in the blockchain.

3.2.7. Retailer. Retailers are profit-making companies that directly sell products to consumers. Retailers search for products that suit their market objectives and find vendors at the most advantageous prices to make a profit. Generally, small quantities of an item can be ordered from a distributor by a retailer. In this setup, the retailer orders goods in batches of traceable identifiers which make it easy for the customer to verify product life cycle records.

3.2.8. Customer. The customer is the final user of the food product who purchases the food product from the retailer. This research smart model provides full access to the customer to trace production information from the origin to the sale point. In this setup, all supply chain stakeholders are bound to digitally sign blockchain smart contracts and to record information fairly and in case of dishonest blockchain automatically configure the incident into smart contracts to fine the concerned party. Another option is to install cameras in the field to automatically capture and transfer images to the blockchain for preaudit. The advantage of using product traceability technology in the food supply chain is the availability of reliable and real-time information to all stakeholders.

3.3. IoT-Based Agriculture Protocol for the Smart Model. IoT nodes are ideal for cluster farms because they consume less energy than WSN and can be further reduced through an efficient clustering protocol. Therefore, this research proposed a new clustering protocol IoT-based agriculture, as shown in Figure 8, based on the LEACH protocol, to reduce energy consumption and extend network life.

3.3.1. Assumption for Simulation. Assumptions for simulation are given below.

(i) IoT nodes are installed on a cluster farm randomly

(ii) IoT nodes send hello messages to the base station with local information

(iii) The initial number of clusters is calculated by taking the optimal values to vary with the node density before the node starts to expire, and the smaller clusters become larger clusters

(iv) The base station and sink are installed outside the cluster farm as presented in Figure 9

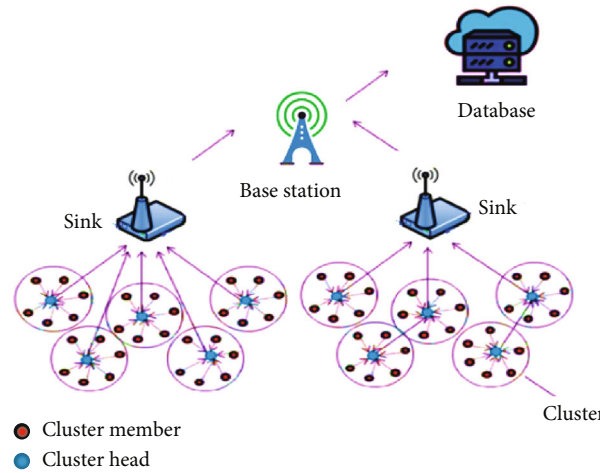


Figure 8: IoT-based agriculture clustering protocol.

3.3.2. Initialization Phase. A cluster farm with a total area of $500 \times 500\,\mathrm{m}^2$ was chosen and distributed into clusters, with 100 IoT nodes placed in each cluster at random. The IoT nodes were chosen based on their capabilities, such as soil moisture monitoring, temperature monitoring, and crop disease tracking. Nodes in clusters are distributed in such a way that one cluster node does not communicate with other cluster nodes, but only with the cluster heads of their respective clusters. Each cluster's CH is in charge of sharing sensed data with the BS via the sink node. To achieve a satisfactory SNR, a widely used LEACH "First Order Radio Model" is used to transmit a small message over a distance of d.

3.3.3. LEACH First Order Radio Model. As shown in Figure 10 and Table 1, the first order radio model dissipates $E_{\mathrm{elec}} = 50\,\mathrm{nJ/bit}$ to run the transmitter or receiver circuit system and $E_{\mathrm{amp}} = 100\,\mathrm{pJ/bit/m}^2$ for the transmit amplifier to achieve a satisfactory signal-to-noise ratio. These parameters are slightly better than the existing radio system second assumption that was an energy loss due to channel transmission. To transmit an m-bit message over a distance of d using a radio model, the radio dissipates as follows:

$$E_{T_x}(m, d) = E_{T_{x-\mathrm{elec}}}(m) + E_{T_{x-\mathrm{amp}}}(m, d), \tag{1}$$

$$\mathrm{E}_{T_x}(m, d) = E_{\mathrm{elec}} \times \mathrm{m} + \epsilon_{\mathrm{amp}} \times m \times d^2, \tag{2}$$

and to get this message, the radio expands

$$E_{R_x}(m) = E_{\mathrm{elec}} \times m. \tag{3}$$

With those parameters, transferring or receiving a packet is not a low-cost method; so, protocols must strive to reduce not only transmission distances but also the number of transmission and receiving operations per message.

3.3.4. Clustering Mechanism. IoT nodes are divided into groups in a cluster farm, and these groups are known as clusters; each cluster is represented by G and has a head node. Each cluster's nodes sense data and send it to a single head node; they do not communicate with other head nodes.

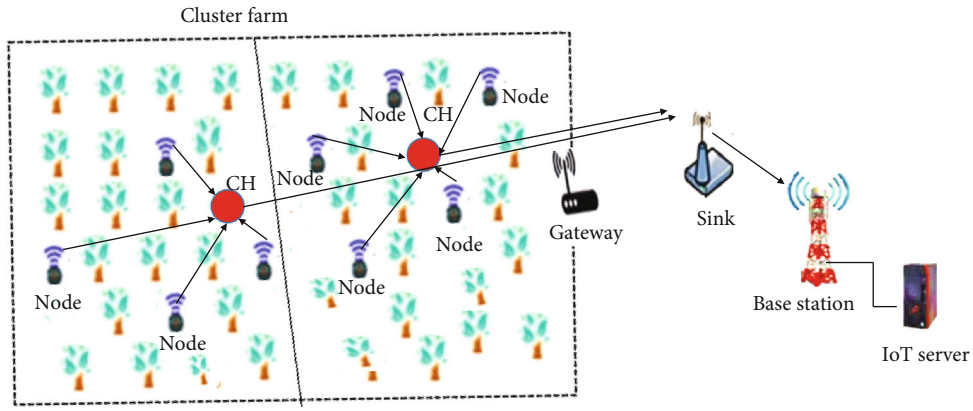


Figure 9: IoT-based agriculture cluster farm and data transmission.

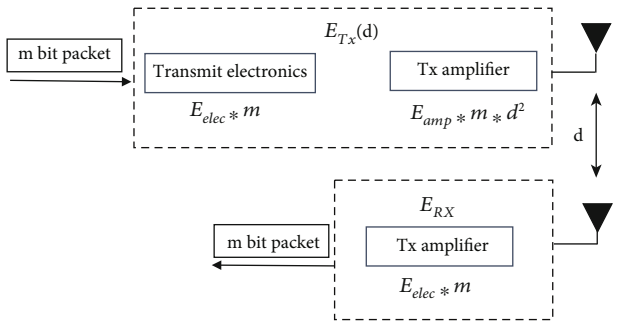


Figure 10: First order radio model.

Table 1: Radio characteristics.

Operations	Energy dissipated
Transmitter electronics $(E_{T_{x-elec}})$	50 nJ/bit
Receiver electronics $(E_{R_{x-elec}})$ $(E_{T_{x-elec}} = E_{R_{x-elec}} = E_{elec})$	100 pJ/bit/m^2
Transmit amplifier E_{amp}	

(1) Cluster Head (CH) Selection. The CH is responsible for collecting data from member nodes and transmitting aggregated data to the base station, it requires a lot of energy to do so, and the transmission process must be boosted with high power amplification. Two parameters are considered in the CH selection process: the history of nodes acting as CH and the optimal percentage of a node. The generation of a random number is used to make each node decision (between 0 and 1). The node will be selected as CH if the generated random number is threshold (T_n) for that round.

The following formula is used to calculate the threshold (T_n):

$$T_{(n)} = \begin{cases} \dfrac{P}{1 - Px(r \bmod (1/P))} & n \in G \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

In equation (4), p represents the percentage of CH, r represents the number of rounds, and G represents the set of nodes that were not selected as CH in the previous $1/p$ rounds. Every node in the cluster has a $1/p$ chance of being a CH in each round in this situation. When a node is chosen as CH, it sends an advertisement message to its nearby nodes, inviting them to be CH. The advertisement message is accepted by the member nodes, and they enter the CH.

(2) Data Transmission. Following CH selection, the data transmission schedule kicks off, with member nodes sending data to their assigned CH during their designated transmission time. Low-energy transmission was required for this type of transmission. Before transmission time is allocated, the member node can be turned off to save energy. The CH must keep the receiver to receive the full data and then combine all of the data into a single signal before transmitting it to the base station via a sink node.

(3) Routing Phase. The way to pick a traffic route in a network is routing. The IoT agricultural network is comprised of numerous modes of communication and protocols in the long and short range, which form the backbone for these networks to gather information and exchange this data with the farmers for review and decision-making. The IoT-based agriculture protocol has three levels of data routing in the first stage member node sense data and transfer it to their CHs, while in the second stage, the CHs accumulate and transfer data to the nearest sink, and in the third and last stage, sink sends this data into the base station. In such data communication, the function of the sink minimizes CH's energy consumption and reduces packet losses, especially for long haul communication. Figure 8 demonstrates the full data routing method.

(4) Setup Phase. For the first round, CHs are generated using the standard LEACH algorithm, and the cluster heads are chosen using equation (4). After transferring data, each node in the network consumes a specific amount of energy, and each node has consumed a different amount of energy. The distance between the sending and receiving nodes, denoted

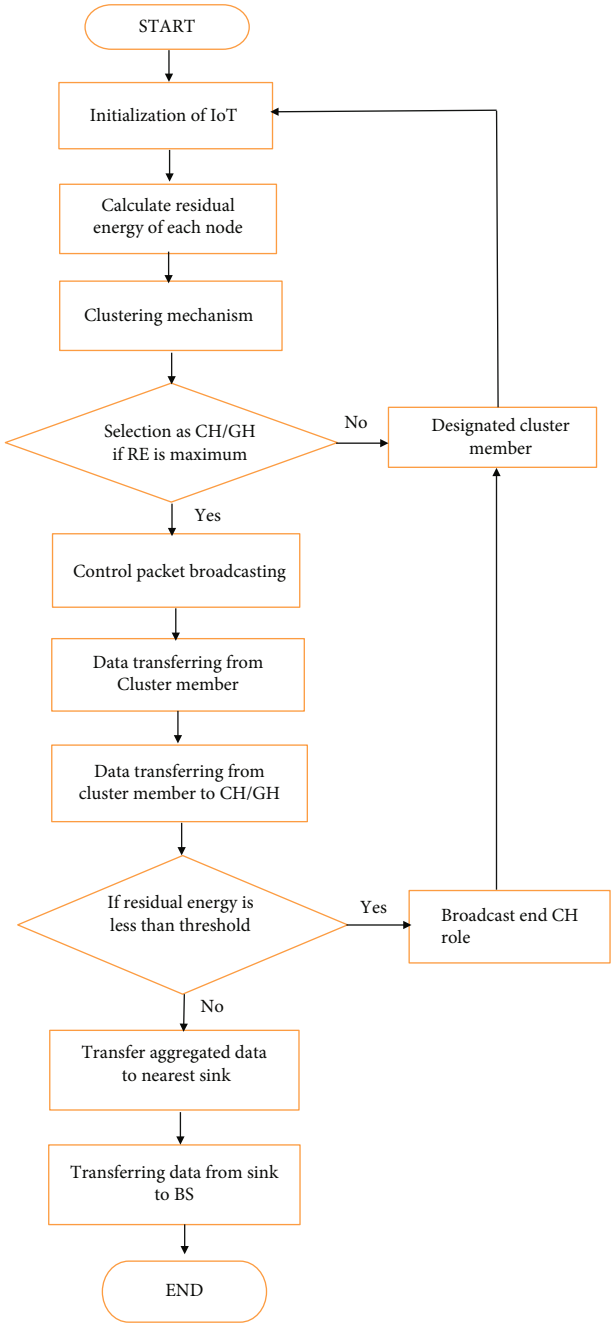


Figure 11: Workflow chart of CH mechanism.

by "d," is an important factor in power consumption. As a result, the CH is chosen for the next round using an improved equation, as shown below.

$$T_{(n)} = f_{(x)} = \left\{ \frac{P}{1 - P(r \bmod (1/P))} \times \frac{E_{\text{residual}}}{E_{\text{initial}}} m_{\text{opt}} \text{forall} \in \forall G. \right. \tag{5}$$

In above equation (5), E_{residual} represents the node remaining energy, and E_{initial} is the initial energy assigned. Thus, m_{opt} could be written as optimal number of clusters.

1. Initialization of IoT
2. Calculate residual energy (R.E) of each node ($n_i \ldots n_n$)
3. Clustering Mechanism
4. If (R.E n_i == Max (R.E)) Then
   CH = n_i
   Else
   Cluster_Member= n_i
   GoTo Step 2
5. Control Packet broadcasting
6. Data transferring from Cluster_Member
7. Data transferring from Cluster_Member to CH/GH
8. If (R.E n_i <threshold) Then
   n_i !=CH
   Cluster_Member=n_i
   GoTo Step 2
   Else
   Transfer aggregated data to nearest sink S_i
9. Transferring data from S_i to BS
10. END

Algorithm 1: The proposed scheme.

$$m_{\text{opt}} = \sqrt{\frac{E_{fs}}{E_{\text{amp}} d^4 (2m - 1) E_0 - m E_{DA}}}. \tag{6}$$

In the above equation (6), the m_{opt} is the optimal number of clusters, X is the diameter of the network, whereas E_0 is the node initial energy. For data transmission, this research extended the first order radio model specified in equation (2) to compute the energy consumption

$$E_{TX} = f(x) = \begin{cases} m \times \left(E_{\text{elec}} + E_{fs} \times d^2\right) d < d_o \\ m \times \left(E_{\text{elec}+E_{\text{amp}}} \times d^4\right) d \geq d_o \end{cases}, \tag{7}$$

where d_o presents distance threshold while E_{elec} and E_{fs} are denoted as energy dissipation with 50 nJ/bit and 10 pJ/bit/m^2 accordingly. m represents the packet size, and E_{amp} is the multipath model of transmitter amplifier with 0.0013 pJ/bit/m^4. Thus, E_{RX} can be determined as it represents the receiving energy

$$E_{RX} = m \times E_{\text{elec}}. \tag{8}$$

Because all nodes in the farm were deployed at random with the same initial energy, the sink node is deployed without energy constraints outside of the cluster farm, and nodes are informed of its location. In the first round, nodes report their location to the sink based on their signal strength. Those nodes that were chosen as grid heads (GH) begin performing their GH role and cease their previous role of data sensing during this process. Cluster heads are chosen as nodes with a high residual energy and a short Euclidean distance to sink and GH. The Euclidean distance between any two nodes a and b in the next two dimensions is calculated as follows:

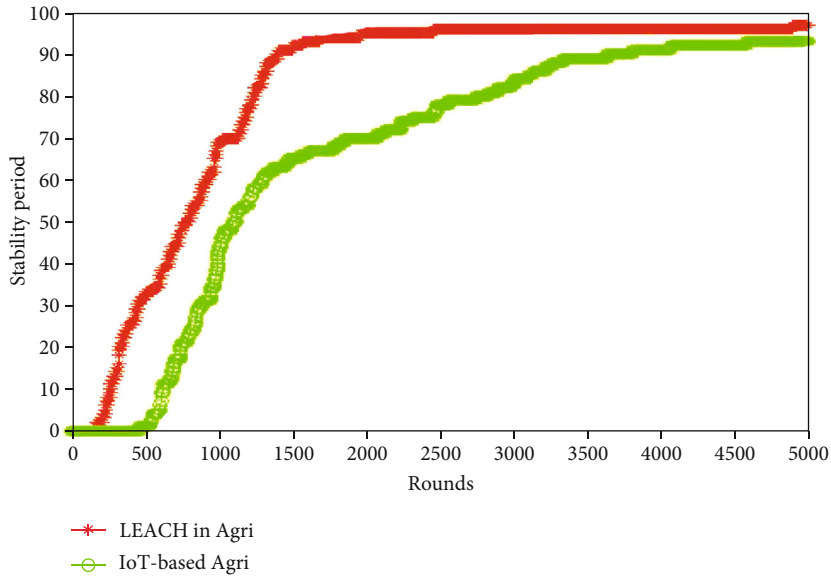


Figure 12: Network stability period of LEACH and IoT-based agriculture protocol.

$$d(a,b) = \sqrt{(x_2 - x_1) + (y_2 - y_1)},$$
$$\text{Min} \sum_{r=1}^{r=\max} \text{WE}_{\text{consumed}}(r) \forall_r \epsilon R. \tag{9}$$

$\text{WE}_{\text{consumed}}$ is calculated using equation by (6).

In the above equation, $\text{WE}_{\text{consumed}}$ is the amount of energy consumed in each round, and d_o denotes the distance between two nodes. The consumption of energy is assumed as E_{TX} and E_{RX}, and their parameters can be determined by using (2) and (3). To meet the objective mentioned in (5), this research divides the total node energy into different equivalent parts and presented as "EL" and can be calculated from (7)

$$\text{EL} = \frac{\text{Eo}}{\text{TL}}. \tag{10}$$

In the equation (10), EL stands for energy level, *Eo* stands for initial static node, and TL stands for total energy, which is dependent on the amount of energy consumed by each node as well as network density and packet size. The relationship between EL and TL is inverse. When the TL value is low, the EL value is high. The node in the cluster farm performs the role of CH until the values of EL do not reduce the residual energy, at which point, a termination of CH selection message is sent to all nodes, and a new process for CH selection begins. Only those nodes that are supposed to transfer data remain active in this mechanism to reduce energy consumption. When member nodes finish their data sensing process, CH begins to collect data from all member nodes, aggregate the data to remove any duplication, combine data into a single signal to save bandwidth, and transmit data in a single hop to the sink, where it is then transferred to the base station. The flowchart is shown in Figure 11, and the algorithm for the entire process is shown below.

4. Simulation Results and Discussion

To test the efficiency of our proposed IoT-based agriculture protocol, the MATLAB simulation is compared with the current LEACH protocol. IoT nodes were installed randomly in each simulated technique in the simulation of 5000 rounds. In the farm of $500 \times 500\text{m}^2$, all active IoT nodes transfer data to their respective cluster heads, and cluster heads transfer data to the base station via sink node. The simulation results are presented below.

4.1. Network Stability Period. The time it takes for the first node to die is known as the network stability period. Figure 12 depicts the LEACH network stability period and the proposed IoT-based agriculture protocol. LEACH has a shorter stability period than the proposed protocol. The reason for the improvement is that data is only transferred in the proposed protocol when there is a difference between previously taken data and current data. The results show that the LEACH first node dies after 168 rounds, while the proposed protocol first node dies after 463 rounds, indicating that the IoT-based agriculture protocol is 23% more efficient in terms of network stability.

4.2. Energy Consumption. The energy consumed by the nodes during data transmission is referred to as energy consumption. Figure 13 shows the simulation results for both protocols. The energy consumption of the nodes in the IoT-based agriculture protocol is 68 percent lower than that of LEACH, which can extend network life.

4.3. Network Life. The network's life is determined by the time it takes to finish the energy at the first node. Figure 14

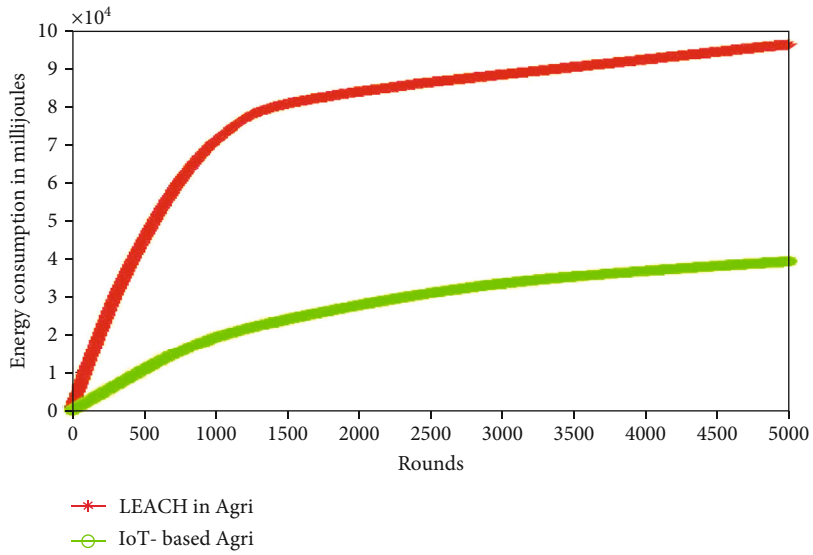


Figure 13: Energy consumption of LEACH and IoT-based agriculture.

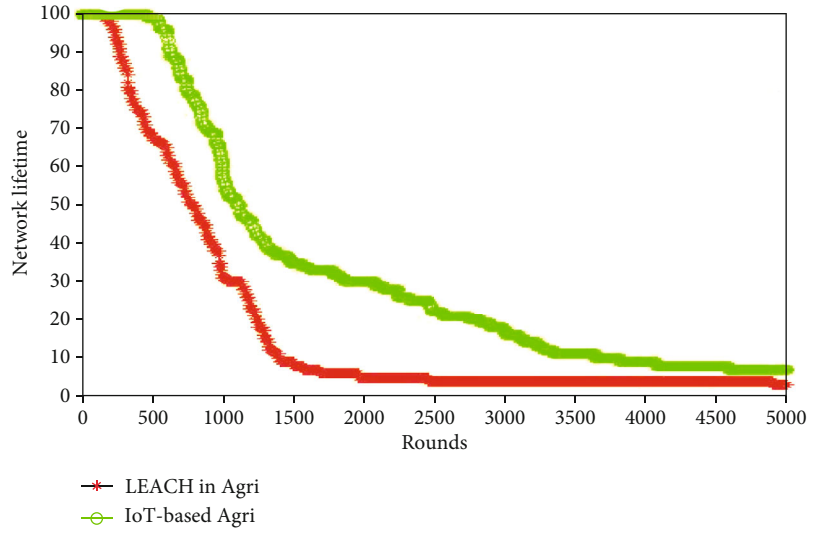


Figure 14: Network life of LEACH and IoT-based agriculture.

depicts the efficiency of LEACH's network life and IoT-based agriculture. LEACH assumes that CHs dissipate the same amount of energy for each round, resulting in unsuccessful CH selection and a reduction in network life, whereas the IoT-based protocol selects CHs based on the residual energy of the nodes and the optimal number of clusters, resulting in a 112 percent increase in network life.

5. Conclusions

The use of blockchain in the food supply chain has yielded various benefits, including the potential to expand and step towards the distributed network, as well as the ability to build a trustless environment for all processes. Despite the trustless existence of blockchain, it is difficult to completely establish trust between the agricultural product seller and buyer. Current blockchain-based food supply chain systems focus solely on the supply chain, with no choice to track food products back to their source to ensure product quality, while existing IoT-based agriculture systems are often isolated, with roles restricted to agriculture environment monitoring and farmers unable to directly communicate with market buyers. In this paper, we have provided complete solution for agriculture and food supply chain by integrating IoT with the blockchain. In addition, we have developed an energy efficient routing protocol for the proposed system to extend system life by reducing energy consumption. A novel approach has been applied in this research and developed a futuristic smart model for agriculture and the food supply chain, which offers an innovative way for farmers to acquire information on crops. IoT can provide farmers with information on crop yields, soil temperature, pest infestation, and soil quality that is essential for high crop production and provides precise data that can be used to improve farming techniques. Crop

tracking can be performed effectively to track crop growth and record growth information. Another unique aspect of the smart model is blockchain, which offers real-time updates on the safety status of food items to all actors of the supply chain, significantly eliminates the vulnerability of centralized information networks, and makes them secure, more widely available, more accessible, and more interactive. The smart model will greatly boost the efficiency and reliability of the food supply chain, which will inevitably increase food safety and regain customer trust in the food industry.

Data Availability

Data is available upon requesting the corresponding author.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was supported by Taif University Researchers Supporting Project number (TURSP-2020/254), Taif University, Taif, Saudi Arabia.

References

[1] V. Matyushok, V. Krasavina, A. Berezin, and J. S. García, "The global economy in technological transformation conditions: A review of modern trends," *Economic Research-Ekonomska Istraživanja*, vol. 2021, pp. 1–41, 2021.

[2] L. Movilla-Pateiro, X. M. Mahou-Lago, M. I. Doval, and J. Simal-Gandara, "Toward a sustainable metric and indicators for the goal of sustainability in agricultural and food production," *Critical Reviews in Food Science and Nutrition*, vol. 61, no. 7, pp. 1108–1129, 2021.

[3] H. C. Morang and D. Laskar, *Problems and prospects of agricultural development in the tribal areas of Golaghat and Sivasagar Districts of Assam*, 2020.

[4] L. Lipper and D. Zilberman, "A short history of the evolution of the climate smart agriculture approach and its links to climate change and sustainable agriculture debates," in *Climate Smart Agriculture*, pp. 13–30, Springer, Cham, 2018.

[5] V. Saiz-Rubio and F. Rovira-Más, "From smart farming towards agriculture 5.0: a review on crop data management," *Agronomy*, vol. 10, no. 2, p. 207, 2020.

[6] M. A. Bouras, F. Farha, and H. Ning, "Convergence of computing, communication, and caching in Internet of Things," *Intelligent and Converged Networks*, vol. 1, no. 1, pp. 18–36, 2020.

[7] S. R. J. Ramson, S. Vishnu, and M. Shanmugam, "Applications of Internet of Things (IoT)–an overview," in *2020 5th International Conference on Devices, Circuits and Systems (ICDCS)*, pp. 92–95, Coimbatore, India, March 2020.

[8] S. Nižetić, P. Šolić, D. López-de-Ipiña González-de-Artaza, and L. Patrono, "Internet of Things (IoT): opportunities, issues and challenges towards a smart and sustainable future," *Journal of Cleaner Production*, vol. 274, p. 122877, 2020.

[9] A. U. Mentsiev, A. R. Isaev, K. S. Supaeva, S. M. Yunaeva, and U. A. Khatuev, "Advancement of mechanical automation in the agriculture sector and overview of IoT," *Journal of Physics: Conference Series*, vol. 1399, no. 4, p. 044042, 2019.

[10] J. Mahalakshmi, K. Kuppusamy, C. Kaleeswari, and P. Maheswari, "IoT sensor-based smart agricultural system," in *Emerging Technologies for Agriculture and Environment*, pp. 39–52, Springer, Singapore, 2020.

[11] S. S. Kamble, A. Gunasekaran, and S. A. Gawankar, "Achieving sustainable performance in a data-driven agriculture supply chain: a review for research and applications," *International Journal of Production Economics*, vol. 219, pp. 179–194, 2020.

[12] P. Kittipanya-Ngam and K. H. Tan, "A framework for food supply chain digitalization: lessons from Thailand," *Production Planning & Control*, vol. 31, no. 2-3, pp. 158–172, 2020.

[13] X. Niu and Z. Li, "Research on supply chain management based on blockchain technology," *Journal of Physics: Conference Series*, vol. 1176, no. 4, p. 042039, 2019.

[14] D. Pavithran, K. Shaalan, J. N. Al-Karaki, and A. Gawanmeh, "Towards building a blockchain framework for IoT," *Cluster Computing*, vol. 23, no. 3, pp. 2089–2103, 2020.

[15] A. Shahid, A. Almogren, N. Javaid, F. A. Al-Zahrani, M. Zuair, and M. Alam, "Blockchain-based agri-food supply chain: a complete solution," *IEEE Access*, vol. 8, pp. 69230–69243, 2020.

[16] J. Yadav, M. Misra, and S. Goundar, "An overview of food supply chain virtualisation and granular traceability using blockchain technology," *International Journal of Blockchains and Cryptocurrencies*, vol. 1, no. 2, pp. 154–178, 2020.

[17] D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and C. Yang, "The blockchain as a decentralized security framework [future directions]," *IEEE Consumer Electronics Magazine*, vol. 7, no. 2, pp. 18–21, 2018.

[18] K. A. Clauson, E. A. Breeden, C. Davidson, and T. K. Mackey, "Leveraging blockchain technology to enhance supply chain management in healthcare: an exploration of challenges and opportunities in the health supply chain," *Blockchain in Healthcare Today*, vol. 1, no. 3, 2018.

[19] S. Saberi, M. Kouhizadeh, J. Sarkis, and L. Shen, "Blockchain technology and its relationships to sustainable supply chain management," *International Journal of Production Research*, vol. 57, no. 7, pp. 2117–2135, 2019.

[20] E. Oztemel and S. Gursev, "Literature review of industry 4.0 and related technologies," *Journal of Intelligent Manufacturing*, vol. 31, no. 1, pp. 127–182, 2020.

[21] M. Nazari Jahantigh, A. Masoud Rahmani, N. Jafari Navimipour, and A. Rezaee, "Integration of internet of things and cloud computing: a systematic survey," *IET Communications*, vol. 14, no. 2, pp. 165–176, 2020.

[22] A. Panarello, N. Tapas, G. Merlino, F. Longo, and A. Puliafito, "Blockchain and IoT integration: a systematic survey," *Sensors*, vol. 18, no. 8, p. 2575, 2018.

[23] R. Haas, D. Imami, I. Miftari, P. Ymeri, K. Grunert, and O. Meixner, "Consumer perception of food quality and safety in Western Balkan Countries: evidence from Albania and Kosovo," *Foods*, vol. 10, no. 1, p. 160, 2021.

[24] K. Pal and A.-U.-H. Yasar, "Internet of things and blockchain technology in apparel manufacturing supply chain data management," *Procedia Computer Science*, vol. 170, pp. 450–457, 2020.

[25] E. M. Torroglosa-Garcia, J. M. A. Calero, J. B. Bernabe, and A. Skarmeta, "Enabling roaming across heterogeneous IoT wireless networks: LoRaWAN MEETS 5G," *IEEE Access*, vol. 8, pp. 103164–103180, 2020.

[26] A. I. Badran and M. Y. Kashmoola, "Smart agriculture using Internet of Things: A Survey," in *Proceedings of the Proceedings of the 1st International Multi-Disciplinary Conference Theme: Sustainable Development and Smart Planning, IMDC-SDSP 2020, Cyperspace, 28-30 June 2020*, p. 10, 2020.

[27] T.-h. Kim, V. S. Solanki, H. J. Baraiya, A. Mitra, H. Shah, and S. Roy, "A smart, sensible agriculture system using the exponential moving average model," *Symmetry*, vol. 12, no. 3, p. 457, 2020.

[28] H. S. Anupama, A. Durga Bhavani, and A. B. A. Z. Fayaz, "Smart farming: IoT based water managing system," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 4, pp. 2383–2385, 2020.

[29] M. R. Ramli, P. T. Daely, D.-S. Kim, and J. M. Lee, "IoT-based adaptive network mechanism for reliable smart farm system," *Computers and Electronics in Agriculture*, vol. 170, p. 105287, 2020.

[30] K. Lova Raju and V. Vijayaraghavan, "IoT technologies in agricultural environment: a survey," *Wireless Personal Communications*, vol. 113, no. 4, pp. 2415–2446, 2020.

[31] N. Etemadi, Y. Borbon-Galvez, F. Strozzi, and T. Etemadi, "Supply chain disruption risk management with blockchain: a dynamic literature review," *Information*, vol. 12, no. 2, p. 70, 2021.

[32] S. Corrado, C. Caldeira, M. Eriksson et al., "Food waste accounting methodologies: challenges, opportunities, and further advancements," *Global Food Security*, vol. 20, pp. 93–100, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 9773402, 12 pages
https://doi.org/10.1155/2021/9773402

Research Article

SS-Drop: A Novel Message Drop Policy to Enhance Buffer Management in Delay Tolerant Networks

Obaid ur Rehman,[1] **Irshad Ahmed Abbasi**,[2] **Hythem Hashem**,[2] **Khalid Saeed**,[3] **Muhammad Faran Majeed**,[3] **and Sikandar Ali**[4,5]

[1]*Department of Computer Science, National University of Modern Languages, Islamabad, Pakistan*
[2]*Department of Computer Science, Faculty of Science and Arts at Belgarn, University of Bisha, Sabt Al-Alaya 61985, Saudi Arabia*
[3]*Department of Computer Science, Shaheed Benazir Bhutto University, Sheringal, Dir (U), Khyber Pakhtunkhwa, Pakistan*
[4]*Department of Computer Science & Technology, China University of Petroleum-Beijing, Beijing 102249, China*
[5]*Beijing Key Lab of Petroleum Data Mining, China University of Petroleum-Beijing, Beijing 102249, China*

Correspondence should be addressed to Irshad Ahmed Abbasi; aabasy@ub.edu.sa and Sikandar Ali; sikandar@cup.edu.cn

Received 15 April 2021; Accepted 21 May 2021; Published 9 June 2021

Academic Editor: Suleman Khan

A challenged network is one where traditional hypotheses such as reduced data transfer error rates, end-to-end connectivity, or short transmissions have not gained much significance. A wide range of application scenarios are associated with such networks. Delay tolerant networking (DTN) is an approach that pursues to report the problems which reduce communication in disrupted networks. DTN works on store-carry and forward mechanism in such a way that a message may be stored by a node for a comparatively large amount of time and carry it until a proper forwarding opportunity appears. To store a message for long delays, a proper buffer management scheme is required to select a message for dropping upon buffer overflow. Every time dropping messages lead towards the wastage of valuable resources which the message has already consumed. The proposed solution is a size-based policy which determines an inception size for the selection of message for deletion as buffer becomes overflow. The basic theme behind this scheme is that by determining the exact buffer space requirement, one can easily select a message of an appropriate size to be discarded. By doing so, it can overcome unnecessary message drop and ignores biasness just before selection of specific sized message. The proposed scheme Spontaneous Size Drop (SS-Drop) implies a simple but intelligent mechanism to determine the inception size to drop a message upon overflow of the buffer. After simulation in ONE (Opportunistic Network Environment) simulator, the SS-Drop outperforms the opponent drop policies in terms of high delivery ratio by giving 66.3% delivery probability value and minimizes the overhead ratio up to 41.25%. SS-Drop also showed a prominent reduction in dropping of messages and buffer time average.

1. Introduction

Wireless technology has a great impact over our lives due to pervasive communication. With the passage of time, a paradigm shift has occurred in networking and its applications. Challenged network, i.e., Remote Area Networks [1], Military Battlefield Networks [2], Postdisaster Response Networks (PDRNs) [3], Interplanetary Networks (IPNs) [4], and Energy Constrained/Sparse.

Wireless Sensor Networks (WSNs) [5] is a network where traditional hypotheses such as reduced data transfer error rates, end-to-end connectivity, or short transmissions have not much significance [6]. A wide range of applications scenarios are associated with such networks [7]. Delay tolerant networking is an approach that pursues to report the problems that reduce communication in disrupted networks [6]. The Delay tolerant network is referred to that network in which end-to-end connectivity rarely exists [8]. Delay

tolerant networking (DTN) architecture was presented in 2003 in which the problem of credible message delivery is tried to address by adopting the notion of store-carry and forward [6]. In such networks, store-carry and forward paradigm is used to achieve a successful sending of messages to destinations [9]. Providing the prearrangement of buffer policies is problematic. Especially in resource-oblige delay tolerant networks (DTNs) where a high ratio for delivery of data and least overhead is supposed to be achieved in dense congested circumstances [10].

Recent research suggests that proper buffer management policies can enhance effectiveness of DTN routing protocols. Routing in DTN is also a challenging problem [8, 11]. In order to minimize the impact of buffer overflow in such congested and constrained environments, buffer management policies should carefully select the messages to be discarded instead of a blind dropping. Several buffer management schemes have been proposed which consider various message attributes in order to select the appropriate messages to be dropped [9, 12, 13, 14].

The latest essential issue mostly disregarded by the DTN researchers was the influence of dropping messages in terms of buffer overflow. This is the reason why traditional buffer management policies like drop front, drop last, and drop random do not have any efficient mechanism to select a message to be dropped. As an alternative such schemes depends upon the order which the messages adopt for arriving and residing in the buffer. Many schemes are fully dependent on the priority assigned to each message. Such schemes are not suitable for DTNs because their performance is better in frequently connected networks rather than frequently disconnected ones. To improve the delivery ratio, to reduce communication overhead, to avoid redundant message drop, and to enhance buffer management are the key points of this research.

The paper is summarized as Section 2 consists of details about buffer management in DTN. Section 3 contains the significant work done. Section 4 comprises a proposed model of the system. Section 5 consists of simulation settings. Section 6 is all about results and discussion. And Section 7 includes brief conclusion and discusses future work.

2. Buffer Management in DTN

Buffer management is an important mechanism to control the buffer resources. Buffer management is responsible for scheduling of messages. An effective buffer management is necessary to make decisions for dropping messages. Whenever the buffer becomes overflow, it needs to drop some messages in order to make room for new ones [15].

2.1. Buffer Management Architecture. Figure 1 describes basic buffer management architecture at a DTN router. Upon arrival, an incoming message is classified according to the criterion implemented by that router and then stored in its buffer. At a contact opportunity, the scheduler decides the order by which messages should be replicated or transmitted as contact durations are limited. In case of buffer overflow, messages which should be discarded are decided by the buffer manager. In other words, scheduling strategies govern the order in which the messages are passed in case of bandwidth and contact constraints.

2.2. Buffer Management Policies in DTN. One of the major issues in DTN is that it faces the problem of buffer overflow due to which dropping of messages is the intense need. For the said reason, several message drop policies were used, and these techniques got different updates and versions time to time. A comprehensive and efficient buffer management technique is the soul requirement of DTN. Hence, for a short review, the existing message drop policies of the DTN have been divided into two different categories: policies having/-keeping network knowledge and policies do not having/keeping network knowledge.

2.2.1. Policies Having No Network and Its Component Knowledge. DTN possesses many buffer management policies in which some requires or use the local information available for a message in network like when the message enter in to the buffer or node, how much is time-to-live (TTL) for the message, etc. These network policies which does not keeps the knowledge of remaining network are quite useful as it is independently working, and it does not require any specific routing scheme to be implemented. Some examples of such policies are as follows:

(1) Drop Front (DF). This policy has the capability or mechanism of message dropping from the front of memory as the buffer becomes overflow. Drop front works on the first-in first-out mechanism. The main problem with this policy is that maximum end-to-end connectivity is mandatory for it, which is quite impossible in DTNs [16].

(2) Drop Last (DL). This scheme works in contrary to drop front. It uses the LIFO algorithm where the last incoming message is selected for dropping upon buffer overflow. This policy increases the probability of message dropping which directly increases the availability of free buffer space for new messages. On the other hand, it leads towards biasness by selecting every time the last message to drop [16].

(3) Drop Random (DR). This policy works on random selection of messages for dropping as buffer becomes overflow. The random deletion property ensures unbiasedness towards message selection. In this technique, each and every message has an equal priority of deletion. In contrast, it may lead towards the wastage of valuable information by dropping such messages which has to travel a lot in future in the network [17].

(4) Drop Youngest (DY). In DTN, every message has its life time called tme-to-live (TTL). TTL of a message depends upon how much time it spends and for how much time it will travel in a network. Drop youngest buffer management policy drops a very recent and freshly arrived message, whose TTL is the highest amongst all the arrived messages. The reason behind this policy is to minimize the overhead ratio as it

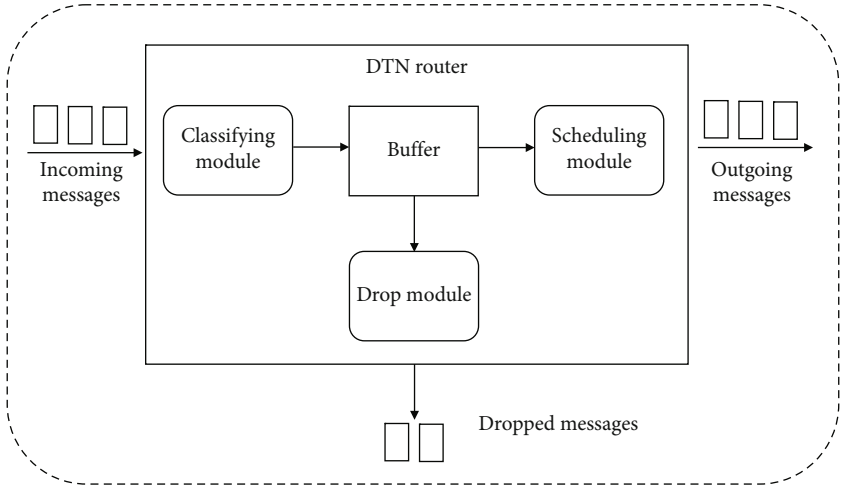


Figure 1: DTN buffer management architecture.

is very much clear that the message whose TTL is long has consumed very little resources [18].

(5) Drop Oldest (DOA). In disrupted networks, a message which traveled for more time in a network and left with very less TTL is considered to be the oldest message. Drop oldest is that type of policy, who's specific and targeted victims are the oldest messages. It always selects those messages to drop who are left with lowest TTL [16].

(6) Drop Largest (DLA). DLA selects size as a criterion for dropping of message as buffer becomes saturated [19]. It selects largest size messages in the buffer for dropping to free up space for the incoming messages. This scheme also ensures that messages having small size will be left for more time in the network, and they can live better with. Hence, they got more forwarding predictability as compared to large size messages [20].

2.2.2. Policies Having Network Knowledge. Other group of buffer management schemes keeps the information about network components such that the number of nodes in the network, number of copies of each message, gathering or meeting time, and ratio of the nodes. These policies take its decisions to drop messages on the basis of available information. Some examples of such schemes are as follows.

(1) Evict Most Forwarded First (MOFO). In DTN, messages move from sender to receiver through hop to hop count. MOFO drop policy considers the quantity of hop counts as a parameter to drop a message from buffer as it gets overflow. MOFO first sets a sending counter value for a message and then a message with the highest counter value that is selected to drop from buffer. This property of MOFO ensures that the messages with a lower counter value may have good chance to travel more within the network [21].

(2) RAPID's Utility-Based Drop. In 2007 [22], developed a new message drop policy which considers the information of network knowledge named as RAPID or resource allocation policy. This scheme intends to keep the information, based on locally optimizing variable which is the estimated delay in function per message used. In general, this can be stated as "the expected contribution of message to the given routing metric." This means that the variable or utility it considers is the expected delay of the message.

(3) Global Knowledge-Based Drop (GBD). Every message in the buffer will have some statistical attributes like message size, time-to-live (TTL), sending node information, and arrival time. In 2008, [23] presented a buffer management policy, that considers all such statistical information to originate every message value on the basis of contact history of a node in the network. The maximization of average delivery ratio and minimization of average delivery delay makes the GBD a very special dropping policy.

(4) History-Based Drop (HBD). Sustaining full and wide network knowledge is much difficult and impossible in real scenarios. That was the reason that global knowledge-based drop policy (GBD) failed in a real sense as it cannot be implemented easily. [22] introduced another message drop policy called history-based drop (HBD). In contrast to (GBD), (HBD) uses local information of the network to calculate its parameters and other useful information. HBD is similar to GBD in few aspects like it must have the information of deterministic nodemobility, and it also have the mandatory requirement of unlimited bandwidth and as well as the requirement of identical message size.

(5) Evict Most Probable First (MOPR). In DTN, a node may contain one or more than one messages at a time. Every message has its own local information like its (TTL) and size. (MOPR) is adding another variable to the forwarding message naming it the forwarding predictability (FP) value. This (FP) value starts from zero (0) and getting increments as far as the message travels in the network. This means that when buffer overflow occurs (MOPR), it will select this (FP) value as a parameter and will target a message for deletion having greater (FP) value amongst all. This mechanism produced

quite reasonable results but this much statistical calculations and values were much hard to sustain [23].

(6) Evict Least Probable First (LEPR). As discussed, MOPR used the high FP value for selection of message to drop. As opposed to MOPR, [24] introduced a new policy known as "Evict Least Probable First" (LEPR). It also first calculates the FP value like MOPR, but in contrast to MOPR, it drops the message with the least FP value [25].

3. Related Work

In DTN, congestion is occurred due to nonavailability of end-to-end connectivity, node mobility, energy constraints, and limited buffer space. Due to these reasons, the message is not delivered to destination on time. To avoid congestion, various drop policies have been adopted which drops the message from buffer to vacate a space for the incoming messages.

According to the scheme proposed in [26], the lengthiest message should be dropped. The purpose behind is the less possibility of messages to be forwarded which stay for a longer time inside the buffer, or it has already been delivered. Another advantage of this scheme is its ability to accommodate a large number of incoming messages with only fewer drops.

Evict Most Forwarded First (MOFO) is suggested by [24], as the name suggests attempts to increase the dissemination of messages over the network by dropping the message which has been sent the most number of times. This enables messages with a lower hop count to travel further within the network. MOFO which keeps track of numbers of times a message is forwarded and message of the highest number will drop and then it provides a good chance for the least forward message to be forwarded.

Drop largest (DLA) proposed in [19] is a popular drop policy which utilizes the size of message as an attribute for selection of messages to drop upon a full buffer. Another advantage of this scheme is its ability to accommodate a large number of incoming messages with only fewer drops. This gives small-sized messages more chance to be forwarded.

Another group of buffer management policies discussed by [27] requires some or complete knowledge about the network (e.g., amount of nodes in the network, quantity of replicas of message, and meeting ratio among two nodes), to make their message discarding decision.

RAPID or resource allocation problem in DTNs was proposed by [22], and it is used to implement a buffer management scheme that considers network wide information. It is an empirical approach which focuses on locally optimizing minimal utility, i.e., the expected delay in utility per message used.

Research conducted in [23], a decent buffer management strategy named global knowledge-based drop (GBD), proposed an idea which focuses on messages in the buffer of a node via overall information about the network. It uses some statistical knowledge to develop per message utility based on node's contact history. Furthermore, GBD works upon global knowledge about the network state. As global knowledge is problematic to acquire, hence GBD is difficult to implement. A deployable variant of GBD called the history-based drop (HBD) uses local information to evaluate global values. Being a variant of GBD, it too has unrealistic norms about the state of the network like deterministic node mobility, unlimited bandwidth, and same message size. This limits the practical implementation of this approach severely.

The authors in [28] have explained message scheduling policies, message dropping policies, and buffer management methods. In any node/router, if a channel is not available, then to ensure the right mobility of the message router/node is required to maintain some queue. Their strategies are FIFO which prefers to deliver messages on a first come first serve basis.

According to the research conducted in [18], an effective mechanism has been discussed to drop a message upon buffer overflow. This paper focuses on sized-based mechanisms. The idea given by the authors of this paper is named as weight-based drop (WBD) policy. Whenever the message arrives at a destination node and it has not found a free space in the buffer, so this mechanism will drop the one with the largest size among all in order to vacate a space for incoming messages.

Researchers in [29] elaborate the importance of dropping messages and buffer management in DTNs. Drop policies, iterative replication, and storing of messages produce congestion which is over hidden by dropping messages. The author has emphasized upon the fact that effective buffer management must be there to make decisions for dropping messages. Whenever the buffer becomes overflow, it needs to drop some messages in order to make room for new ones.

In [30], researchers have suggested a message drop policy with average latency, message delivery, and network overhead called max hop count (MHC). In this drop policy, a message stays for some time on each node in the network moving from source towards destination. Each message has a specific hop count which shows the number of nodes from which the message has been passed moving toward the destination. Maximum hop count means that a message has been passed from many nodes and has leaved replicas on each node, and so dropping such messages may not affect its delivery. Low hop count means that a message has passed from fewer amounts of nodes during transmission from source to destination and has leaved less number of message copies in the network, and dropping of such messages may lead poor delivery ratio. The movement model used for this purpose is random walk. The routing protocols for max hop count are Epidemic and PROPHER V2. The experimental results show that it works better than other message drop polices such as LIFO, FIFO, and MOFO in terms of delivery rate and overhead ratio with high TTL which is above 180 minutes and is much realistic than other polices, and the TTL of other polices in realistic environment may reach to days.

[31] proposed an effective buffer management policy which is called size-aware drop (SAD). SAD tries to improve the buffer consumption and reduce unnecessary message drop. It improves the data delivery ratio by determining the requirements based on discrepancy of recently arrived

message and existing free space. It picks the appropriate message for drop in order to minimize overhead. Due to appropriate message selection, the delivery probability of SAD is better than that of other polices. The authors have used two movement models, i.e., Random Way Point and Disaster. The performance of the SAD is better than DOA, DLA, and MOFO in terms of delivery rate, overhead ratio, and buffer time average. Its performance is better in the case if the messages available in the buffer have greater or equal size of an incoming message and in case of small messages through which process it selects more than one message to accommodate an incoming large message.

[32] suggested a new message drop policy for buffer management called novel buffer management scheme. It works on two parameter,s i.e., hop count and TTL. In this mechanism, a buffer is logically partitioned on hop count and threshold value. The messages with low hop count than threshold value are stored in the buffer according to their hop count, and messages with high hop count than threshold value are stored in the buffer according to their TTL values. The messages with low hop count in the buffer are to be transmitted first because these messages have been passed from fewer amounts of nodes and still need long movement to reach to the destination. Dropping of such messages may cause poor delivery ratio. The messages stored in the buffer with TTL valued are to be dropped first because it has traveled for long time in the network and have leaved many copied in the network. So, dropping of such messages enhances delivery rate. It takes intelligent decision about message delivery and message drop to manage the buffer and select the smart relay for routing. Experimental results show that it works better than other polices like drop least, drop most forwarded, and drop most recently received in terms of high delivery rate and low overhead ratio.

[33] proposed a buffer management policy called space time drop (ST Drop). The routing algorithms used for this work are epidemic-based, probabilistic, and social aware. In this policy, a message in the network with great space and time has high probability to be delivered, and that message from the buffer would be dropped that can accommodate the new arrived message. The space and time coverage is calculated by the formula $STc = Sc \times Tc$, where "Sc" is the coefficient of space and "Tc" is the coefficient of time. The routing strategies for this research work were Prophet and Bubble Rap. The results of the ST drop in terms of reducing overhead and increasing message delivery are high as compared to other message drop polices. It works better for improving the cost effectiveness of the opportunistic network.

[34] proposed a message drop policy under city-based environments for DTN which is called Priority Queue Based Reactive Buffer Management Policy (PQB-R). It classifies the buffer messages into source, relay, and destination lists. An individual drop metric has also been used to each list. The Drop Expire Message Module is activated by the buffer manager to drop a message from the buffer to create space for incoming message if it meets the algorithmic criteria. The new message will be accommodated if the available buffer space is larger than the message size and if not so, then the buffer will activate Reactive Message Drop Module to drop the messages from the buffer until the free buffer space becomes greater than message size, and the new message is accommodated. To drop a relayed message, it uses the metrics TTL, hop count (HC), message size (MS), and created time (CT), while to drop a source message, it uses the metrics TTL, CT, and MS. The shortest path map-based movement model has been used as the movement model, and the routing protocol is epidemic. It works better to reduce latency, message drop, and message transmission and to increase message delivery as compared to drop largest (DLA), LIFO, and MOFO.

[21] presented a message drop policy called Proposed Buffer Management Policy (PBMP) with number of forwards (NoF) and TTL. This policy drops those messages which have high number of forwarding and minimum value of TTL. The main theme of this proposed policy is to enhance the message delivery ratio and minimize the message drop. The proposed policy has been compared with MOFO under Epidemic Routing Protocol, and it has been observed that the proposed policy works better than MOFO. The delivery probability of proposed policy is 13.18% higher than MOFO. This delivery probability is high due to selection of two metrics, i.e., NoF and lowest TTL which delay the message in the network, and chances of delivery is increased, while in MOFO, only the number of forwards (NoF) is checked and drop those messages which has the highest NoF which yields low delivery probability.

In a summary, DTN is referred to such network in which end-to-end connectivity is rarely exists. In such networks, store-carry and forward paradigm is used to achieve a successful sending of messages to destinations. To maximize the delivery prospect in a DTN, it is critical to drop such messages upon a full buffer which are the least likely to be delivered to the final destination. Several buffer management schemes have been proposed and discussed in this section which considers various message attributes in order to select the appropriate messages to drop. A summary of common message drop policies is shown in Table 1. The proposed scheme SS-Drop implies a simple but intelligent mechanism to determine the inception size in order to select a message to be dropped as the buffer becomes overflow.

4. Proposed System Model

Recent studies state that whenever a new message arrives and its size is greater than the available free space in the receiving node, some messages need to be dropped in order to make room for the upcoming message. Every time dropping messages leads towards the wastage of valuable resources which the message has already consumed.

To overcome this problem, it is important to control excessive dropping of messages. For achieving this goal, different mechanisms have been proposed to determine which message has to be dropped. Like some researchers have agreed upon to drop the message with a short time-to-live (TTL). Other groups of researchers have suggested a scheme to drop such messages which traveled for a long time in the network. Later on studies have selected size as criteria for

Table 1: Summary of common message drop policies.

Message drop policy	Policy type	Major contributions
Drop front (DF) [16]	Policy without network knowledge	Drop front (DF) uses the first-in first-out (FIFO) strategy to drop messages in case of buffer overflow.
Drop last (DL) [16]	Policy without network knowledge	Unlike drop front (DF), drop last (DL) uses the last in first-out (LIFO) technique to drop.
Drop random (DR) [17]	Policy without network knowledge	Randomly select messages for drop to show unbiasedness.
Drop youngest (DY) [18]	Policy without network knowledge	Dropping of messages whose TTL is long.
Drop oldest (DOA) [16]	Policy without network knowledge	DOA selects the message with maximum residential time in the buffer for dropping.
Drop largest (DLA) [19] [20],	Policy without network knowledge	DLA policy drops the message with the largest size from the buffer as the buffer gets overflow.
Evict most forwarded first (MOFO) [21]	Policy with network knowledge	MOFO keeps the message sent counter for the sent messages, and the message with the highest sent counter value is dropped first.
RAPID's utility-based drop [22]	Policy with network knowledge	RAPID or resource allocation problem is a policy that considers network wide information. The schemes emphasize on a locally optimizing variable which is the expected delay in utility per message used.
Global knowledge-based drop (GBD) [23]	Policy with network knowledge	This policy focuses on the overall statistical information of messages. Moreover, GBD requires global knowledge of the network state.
History-based drop (HBD) [22]	Policy with network knowledge	The history-based message drop policy uses local information to calculate a comprehensive value for the selection of message to be dropped.
Evict most probable first (MOPR) [23]	Policy with network knowledge	A forwarding predictability (FP) value is assigned to every message. MOPR selects those messages for dropping that possesses the highest FP value.
Evict least probable first (LEPR) [24, 25]	Policy with network knowledge	LEPR selects a message to drop which has the minimum FP value.

dropping messages which can favor in avoiding excessive drops and also decreases undue retransmissions which contribute to small delivery probability and greater overhead due to increased resource consumption [35].

The proposed solution is a size-based policy which determines an inception size for the selection of messages to be deleted as buffer becomes overflow. The basic theme behind this scheme is that by determining the exact buffer space requirement, one can easily select a message of appropriate size to be discarded. By doing so, unnecessary message drop can be overcome and ignore biasness just before selection of specific sized messages.

4.1. Spontaneous Size Drop (SS-Drop). The proposed scheme Spontaneous Size Drop (SS-Drop) implies a simple but intelligent mechanism to determine this inception size. Figure 2 shows the model for estimation of inception size for message selection. Whenever the buffer receives new message, it starts analyzing the size of new message. If the arriving message's size is greater than the existing free space in the buffer, this scheme then follows a few simple steps. Firstly, it calculates the difference among the size of incoming messages and the available buffer space. By doing so, it estimates an inception value which it considers as the inception size. Now, any message residing in the buffer of the receiving node, which is equal to or greater than this inception size, is selected to drop.

$$\text{UBS} = \text{Total Buffer Size} - \text{Occupied Buffer}, \tag{1}$$

$$\text{Inception Size (IS)} = \text{Message Size} - \text{Unoccupied Buffer Size}. \tag{2}$$

Equation (1) calculated the value of unoccupied buffer size (UBS) by subtracting the value of the occupied buffer from that of total buffer size. In Equation (2), the value of an inception size will be calculated by subtracting the already calculated unoccupied buffer size from the size of incoming message. This calculation will help the DTN router to drop such messages from the saturated buffer whose size is equal to the inception size.

4.2. Working of SS-Drop. The flow chart in Figure 3 explains the working of a proposed policy that how it will calculate the inception size of a message. And then on the basis of this inception value, how it will drop the message upon buffer's overflow.

First of all, the proposed model will check the size of an incoming message and then calculate the unoccupied size of the buffer. On the basis of these calculations, if the incoming message size is less than that of unoccupied size, it will allow the message to buffer, but if this expression becomes false, the proposed model then calculates the inception size. After that, the inception message size will be checked with

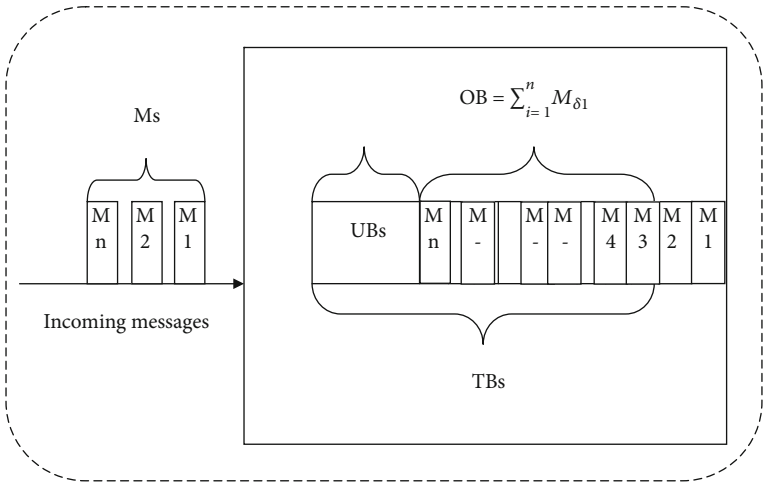


Figure 2: Proposed model for estimation of the inception size for message drop.

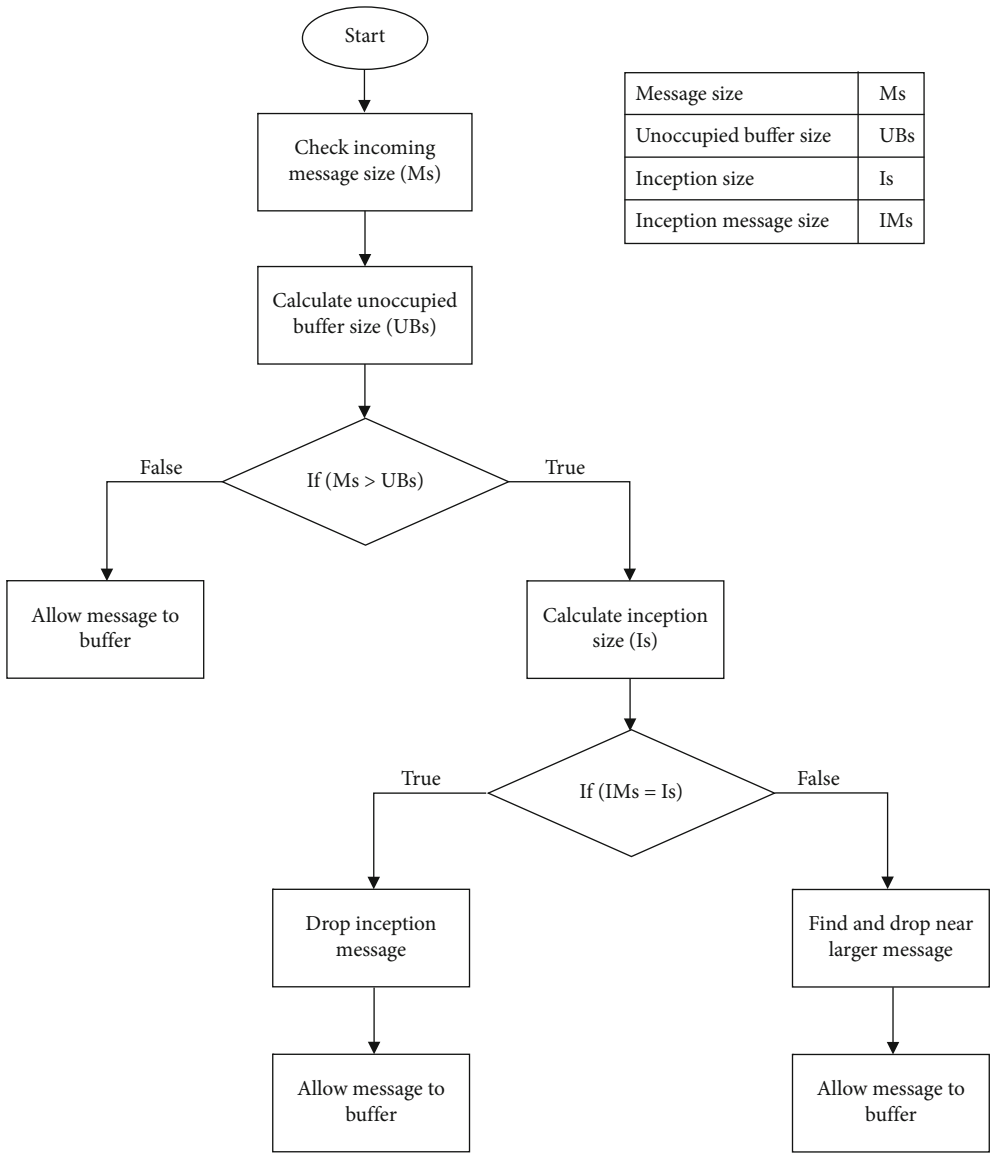


Figure 3: Flow chart of proposed policy.

```
    Input = Xi
    Output = Xj, >= j
1: Xi (Incoming message)
2: for i =1 to N do
3:        Compute S_i.X_i (check size of incoming message)
4: end for
5: for i =1 to k do
6:        Compute S_i.X_i (Check size of incoming message)
7:        OB_i = S.X_i + OB_i (Occupied buffer size)
8:        UB_i = TB_s − OB_i (Unoccupied Buffer Size)
9: end for
10: for j =1 to k do
11:       if X_i < UB_i then
12:          Enqueue X_i
13:       else
14:          Compute I_s = S_i.X_i − UB_i
15:          for I to n do
16:             if IS_i == S_i.X_i then
17:                Drop X_i
18:                Enqueue IS_i
19:             else
20:                Sub_proced(IS_i,S_i.X_i)
21:             end if
22:          end for
23:       end if
24: end for
```

Algorithm 1: Algorithm of SS-Drop policy.

```
    Input = IS_i.X_i
1: S((S_i.X_i)
2: if ((S_i +1X_i +1) > IS_i) then
3:       Drop X_i +1
4:       Enqueue X_i
5: end if
```

Algorithm 2: Algorithm of sub-NS.

the calculated inception size. If the message size and the inception size become equal, the mechanism will drop the inception message; otherwise, it will drop that message which is of near larger size than that of the inception size.

4.3. Pseudocode of SS-Drop. Algorithms 1 and 2 represent the algorithms of Spontaneous Size Drop.

5. Simulation Setting

In order to compare the efficiency of proposed novel message drop policy, i.e., Spontaneous Size Drop (SS-Drop) with other well-known traditional message dropping strategies (drop front, drop oldest, and drop largest), the simulation setup has been done by selecting Epidemic as a routing protocol due to its flooding nature of forwarding messages. The mobility model used in this research work is the shortest path map-based movement mobility model (SPMBM). FIFO has been used as a scheduling scheme throughout the simulation. Table 2 is summarizing the parameter list.

Table 2: Simulation parameters.

Parameters	Values		
	Set 1	Set 2	Set 3
No. of nodes	15	30	60
Size of buffer (MB)	3 MB	4 MB	5 MB
Simulator	Opportunistic network environment (ONE)		
Movement model	Shortest path map-based movement model		
Routing protocols	Epidemic routing protocol		
Group interface	Bluetooth		
Transmission speed	300 kbps		
Transmission range(m)	15 m		
Node speed (m/s)	1-3 (m/s)		
Simulation time	3600 sec		
Message creation interval	15 s-25 s		
Message sizes	100kB-1 MB		
Area of simulation	1600 m × 1600 m		

The parameters used in this simulation is the combination of three groups of nodes. First group consists of 15 nodes with 3 MB of buffer size, second group contains 30 nodes along with 4 MB buffer size, and the third group consists of 60 nodes with 5 MB buffer size. All the nodes have intragroup communication plus intergroup communication as well. Details of other parameters are as follows.

5.1. Epidemic Protocol. Epidemic is considered to be the pioneer protocol designed for communication in a dispersed network where end-to-end connectivity does not exist [36]. Epidemic is a flooding nature protocol. In this protocol, a unique ID is assigned to each and every message, and this ID is associated with the message and all its generated copies until the message has been dropped or successfully delivered to a particular destination. Epidemic sends and forwards the messages in the form of flood to all the nodes within the contact of the sender node, which increases the ratio of successful delivery, but on the other hand, it maximizes the overhead by consuming large amounts of valuable resources.

5.2. Shortest Path Map-Based Movement Mobility Model. Movement models dictate the nodes how to move in the simulation. Each model facilitates the nodes in terms of coordination, speed, and pause time. In this research work, the shortest path map-based movement mobility model is used as the mobility model. It is considered to be one of the mature shortest path models because it works on Dijkstra's shortest path algorithm in order to find out its way through the map [37].

6. Results and Analysis

To evaluate the effectiveness and performance of said messages, drop policies following metrics were used.

(1) Delivery probability

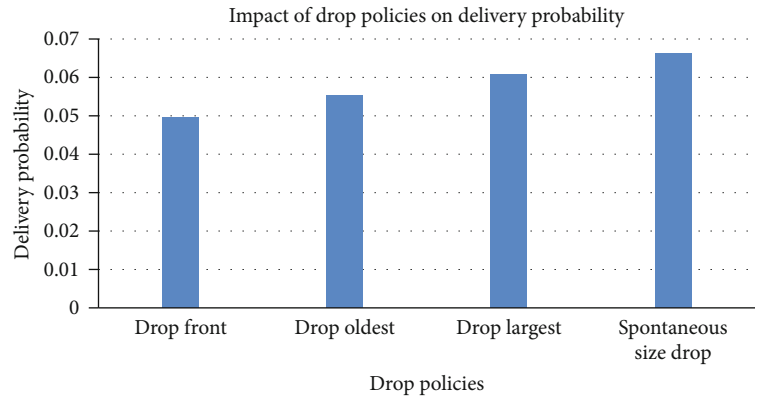


Figure 4: Comparison of message dropping srategies w.r.t delivery probability.

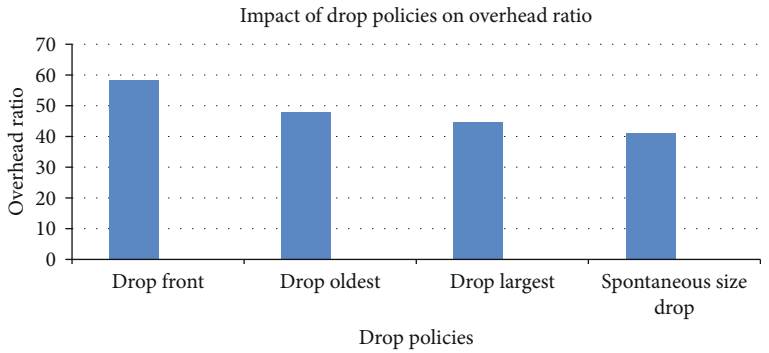


Figure 5: Comparison of message dropping strategies w.r.t overhead ratio.

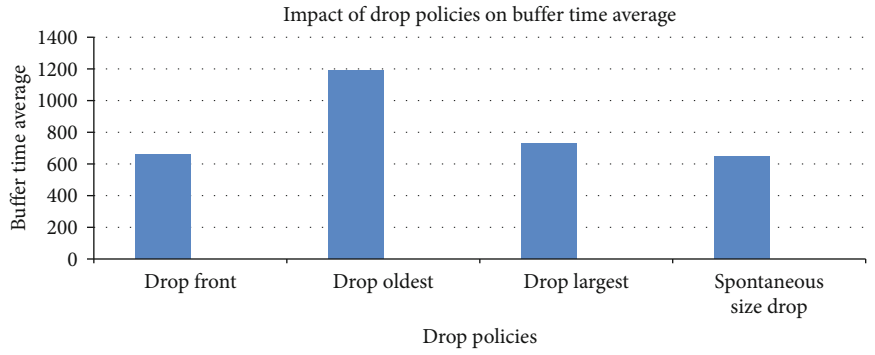


Figure 6: Comparison of message dropping strategies w.r.t buffer time average.

(2) Overhead ratio

(3) Buffer time average

(4) Dropped messages

(5) Messages delayed

6.1. Delivery Probability. Delivery probability can be explained as the percentage of the fully delivered messages within a given time period. Basically, it is the ratio of the number of messages delivered over the number of messages created. Delivery probability can be shown in Equation (3).

$$\text{Delivery Probability} = \frac{\text{Total Messages Delivered}}{\text{Total Messages Generated}}. \quad (3)$$

Figure 4 shows the effect of dropping strategies on delivery probability. It is seen that by applying a flooding nature Epidemic protocol, drop front policy gives a decent probability value of 0.496, while drop oldest and drop largest is slightly better and, respectively, gives the results 0.552 and 0.608, respectively. However, epidemics with spontaneous size drop due to its intelligent message dropping mechanism outperform all other dropping strategies by achieving the improved 0.663 delivery probability value.

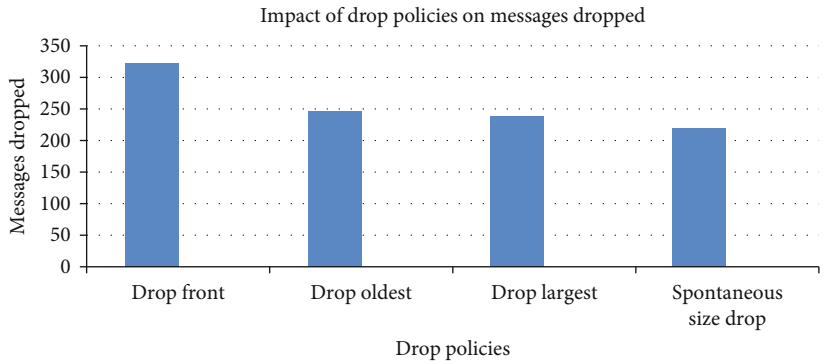


Figure 7: Comparison of message dropping strategies w.r.t messages dropped.

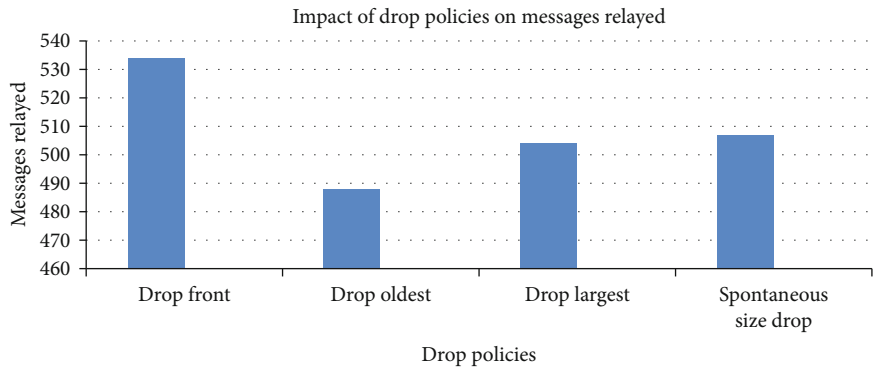


Figure 8: Comparison of message dropping strategies w.r.t messages relayed.

6.2. Overhead Ratio (OR). The word overhead ratio determines the percentage of total messages communicated (TMC) to the total messages delivered (TMD). Mathematically, it can be represented as in Equation (4).

$$\text{OR} = \frac{\text{TMC} - \text{TMD}}{\text{TMD}}. \tag{4}$$

To minimize the overhead ratio is also a vital metric of this research work. The key objective is to maximize the delivery probability opposite to minimize the overhead ratio. Figure 5 shows that by applying intelligent drop policy, overhead ratio can be reduced. The graph depicts drop front that has a high overhead ratio of 58.33 as compared to values of drop oldest 47.80 and drop largest 44.82. Message redundancy can be overcome by intelligently selecting messages, which results in minimizing overhead. The spontaneous size drop shows by giving optimized 41.25 overhead ratio.

6.3. Buffer Time Average. Normally, time taken by the message in the buffer at every node is referred to as average buffer time. It can be calculated by using the formula shown in Equation (5).

$$\text{Buffer Time Average} = \frac{\sum ni = 1 \text{ Buffer Stay Time of Message}}{\text{Total Messages in the Buffer}}. \tag{5}$$

Figure 6 analyzed the effect of existing policies and proposed Spontaneous Size Drop (SS-Drop) with respect to buffer time average. Epidemic with drop oldest policy allows messages to stay for a longer period of time in the buffer in order to achieve greater buffer time average. Epidemic with drop front, drop largest, and proposed scheme Spontaneous Size Drop achieves relatively smaller buffer stay time average. The reason is that drop front iteratively drops messages while both size-based schemes drop messages based on the requirements of buffer space. None of these schemes take time into consideration due to which relatively newer messages can be dropped as well. Thus, as a result, the overall buffer time average of messages is relatively smaller as compared to drop oldest. However, in this scenario, against popular belief, buffer time average does not seem to have much effect on the delivery probability.

6.4. Messages Dropped. Redundant message drop is another cause of increased overhead while the prime objective of any buffer management scheme is to control the number of messages dropped. Figure 7 shows that drop front causes the maximum number of messages to be dropped thus increasing overhead. The rest of the schemes (i.e., drop oldest, drop largest, and Spontaneous Size Drop) reduces this message drop significantly leading to reduced overhead ratio. This is due to selective behavior rather than blind drop of messages. SS-Drop has the least message drop ratio among the other buffer management schemes. This could be due to the fact that it has the ability to select intelligently a message to be dropped.

6.5. Messages Relayed. Figure 8 shows the effect of each scheme on the number of messages relayed in order to achieve successful delivery. Drop front has the lowest delivery probability with the highest overhead ratio which may have resulted from its maximum number of message relaying. Drop oldest has the least amount of messages relayed due to longer buffer time average. Drop largest and Spontaneous Size Drop closely follow behind.

7. Conclusion and Future Work

This research proposed an efficient buffer management policy Spontaneous Size Drop (SS-Drop) which estimates the difference of the available buffer size and the size of the incoming message and then determines an inception size for message discard.

To evaluate its performance, an Epidemic routing protocol is used along with the shortest path map-based movement model. In the shortest path map-based movement model scenario, the proposed policy SS-Drop showed complete superiority over the opponent drop policies drop front (DF), drop oldest (DOA), and drop largest (DLA) in terms of maximum probability and minimum overhead ratio. Spontaneous Size Drop also outperforms other drop policies in terms of dropping messages. Due to encouraging results obtained in this research, it is clear that the great potential lies in the inspection and exploration of this area. An interesting future direction is to evaluate the performance of the proposed scheme with other scheduling/forwarding policies. Designing a new scheduling scheme based on size, for further optimization, is an interesting extension to this work.

Data Availability

Data are available upon request to the corresponding authors.

Conflicts of Interest

The authors declare no potential conflict of interests.

Acknowledgments

This study was supported by the China University of Petroleum-Beijing and Fundamental Research Funds for Central Universities under grant no. 2462020YJRC001.

References

[1] R. Fulvio, *Application-gateway in a DTN environment*, 2021.

[2] K. M. Wook and C. Y. Won, "An efficient routing protocol with overload control for group mobility in delay tolerant networking," *Electronics*, vol. 10, no. 4, p. 521, 2021.

[3] Z. Jianguo, Z. Changjia, K. Yuqin, and T. Shenghui, "Integrated satellite-ground post-disaster emergency communication networking technology," *Natural Hazards Research*, vol. 1, pp. 4–10, 2021.

[4] S. Sanhua, "An effective congestion control scheme based on early offload for space delay/disruption tolerant network," *International Journal of Security and Networks*, vol. 16, no. 1, pp. 28–36, 2021.

[5] R. David, J. Arshad, and M. Luca, "Protocol transformation for transiently powered wireless sensor networks," in *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, pp. 1112–1121, Virtual Event Republic of Korea, 2021.

[6] F. Kevin, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 27–34, Karlsruhe, Germany, 2003.

[7] R. D. Satya, S. Koushik, M. Nandini, and P. Sinha Bhabani, "Delay and disruption tolerant networks: a brief survey," in *Intelligent and Cloud Computing*, M. Debahuti, B. Rajkumar, M. Prasant, and P. Srikanta, Eds., pp. 297–305, Springer Singapore, Singapore, 2021.

[8] D. K. Lobiyal, "Location based contact time energy efficient routing (LCTEE) approach for delay tolerant networks," *Wireless Personal Communications*, vol. 108, no. 4, pp. 2639–2662, 2019.

[9] M. G. Douglas and M. S. Kumar, "Catora: congestion avoidance through transmission ordering and resource awareness in delay tolerant networks," *Wireless Networks*, vol. 26, no. 8, pp. 5919–5937, 2020.

[10] Q. Ayub, S. Rashid, and M. S. Zahid, "Buffer scheduling policy for opportunistic networks," *International Journal of Scientific & Engineering Research*, vol. 2, no. 7, 2011.

[11] S. Nidhi, P. Sudhakar, and K. Sanjay, "Friendship and location-based routing in delay tolerant networks," in *Evolutionary Computing and Mobile Sustainable Networks*, V. Suma, B. Noureddine, and W. Haoxiang, Eds., pp. 781–789, Springer Singapore, Singapore, 2021.

[12] J. Sweta and C. Meenu, "A fuzzy logic based buffer management scheme with traffic differentiation support for delay tolerant networks," *Telecommunication Systems*, vol. 68, no. 2, pp. 319–335, 2018.

[13] R. Sulma and A. Qaisar, "Integrated sized-based buffer management policy for resource-constrained delay tolerant network," *Wireless Personal Communications*, vol. 103, no. 2, pp. 1421–1441, 2018.

[14] R. Animesh, A. Tamaghna, and D. B. Sipra, "Fairness in message delivery in delay tolerant networks," *Wireless Networks*, vol. 25, no. 4, pp. 2129–2142, 2019.

[15] L. Yong, Q. Mengjiong, J. Depeng, S. Li, and Z. Lieguang, "Adaptive optimal buffer management policies for realistic DTN," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, pp. 1–5, Honolulu, HI, USA, 2009.

[16] R. Sulma, A. Qaisar, Z. M. Soperi Mohd, and A. A. Hanan, "E-drop: An effective drop buffer management policy for DTN routing protocols," *International Journal of Computer Applications*, vol. 13, pp. 118–121, 2011.

[17] J. Sweta and C. Meenu, "Survey of buffer management policies for delay tolerant networks," *The Journal of Engineering*, vol. 2014, no. 3, pp. 117–123, 2014.

[18] R. Sulma, A. Qaisar, and A. A. Hanan, "Reactive weight based buffer management policy for DTN routing protocols," *Wireless Personal Communications*, vol. 80, no. 3, pp. 993–1010, 2015.

[19] R. Sulma and A. Qaisar, "Efficient buffer management policy DLA for DTN routing protocols under congestion," *International Journal of computer and Network Security*, vol. 2, no. 9, pp. 118–121, 2010.

[20] S. Rashid, Q. Ayub, M. S. Zahid, and A. H. Abdullah, "Impact of mobility models on DLA (drop largest) optimized DTN epidemic routing protocol," *International Journal of Computer Applications*, vol. 18, no. 5, pp. 1–7, 2011.

[21] S. V. Kumar and G. Nisha, "Comparision of MOFO drop policy with new efficient buffer management policy," *International Journal of Innovative Research & Studies*, vol. 8, pp. 456–460, 2018.

[22] B. Aruna, L. Brian, and V. Arun, "DTN routing as a resource allocation problem," in *Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 373–384, New York, United States, 2007.

[23] K. Amir, B. Chadi, and S. Thrasyvoulos, "Optimal buffer management policies for delay tolerant networks," in *2008 5th annual IEEE communications society conference on sensor, mesh and ad hoc communications and networks*, pp. 260–268, San Francisco, CA, USA, 2008.

[24] L. Anders and S. Phanse Kaustubh, "Evaluation of queueing policies and forwarding strategies for routing in intermittently connected networks," in *2006 1st International Conference on Communication Systems Software & Middleware*, pp. 1–10, New Delhi, India, 2006.

[25] S. A. Pereira, B. Scott, H. C. Massaki, and O. Katia, "DTN congestion control unplugged: A comprehensive performance study," in *Proceedings of the 10th ACM MobiCom Workshop on Challenged Networks*, pp. 43–48, New York, United States, 2015.

[26] J. A. Davis, A. H. Fagg, and L. B. Neil, "Wearable computers as packet transport mechanisms in highlypartitioned ad-hoc networks," in *Proceedings Fifth International Symposium on Wearable Computers*, pp. 141–148, Zurich, Switzerland, 2001.

[27] K. Sathita and E. Hiroshi, "Independent DTNs message deletion mechanism for multi-copy routing scheme," in *Proceedings of the Sixth Asian Internet Engineering Conference*, pp. 48–55, New York, United States, 2010.

[28] I. Saeid, "A novel queue management policy for delay-tolerant networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016.

[29] G. Animesha, J. Shweta, and B. D. Singh, "An analysis optimal buffer management policy to improve QOS in DTN routing protocol," *International Journal of Computer Applications*, vol. 975, p. 888, 2012.

[30] H. Youssef and A. Abdelmounaim, "MaxHopCount: a new drop policy to optimize messages delivery rate in delay tolerant networks," *IJIMAI*, vol. 4, no. 1, pp. 37–41, 2016.

[31] M. Momina, H. Fazle, I. Muhammad, M. A. Ali, and V. Vasilakos Athanasios, "An adaptive and efficient buffer management scheme for resource-constrained delay tolerant networks," *Wireless networks*, vol. 22, no. 7, pp. 2189–2201, 2016.

[32] C. C. Sobin, "An efficient buffer management policy for DTN," *Procedia Computer Science*, vol. 93, pp. 309–314, 2016.

[33] M. D. Silva, I. O. Nunes, R. A. Mini, and A. A. Loureiro, "ST-Drop: a novel buffer management strategy for D2D opportunistic networks," in *2017 IEEE Symposium on Computers and Communications (ISCC)*, pp. 1300–1305, Heraklion, Greece, 2017.

[34] A. Qaisar, N. M. Asri, R. Sulma, and H. H. Adnan, "Correction: priority queue based reactive buffer management policy for delay tolerant network under city based environments," *PloS one*, vol. 14, no. 10, article e0224826, 2019.

[35] W. John, C. Vania, and A. M. Dias, "Performance of opportunistic epidemic routing on edge-markovian dynamic graphs," *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1259–1263, 2011.

[36] A. Vahdat and D. Becker, *Epidemic routing for partially connected ad hoc networks*, 2000.

[37] E. Frans, K. Ari, K. Jouni, and O. Jörg, "Working day movement model," in *Proceedings of the 1st ACM SIGMOBILE workshop on Mobility models*, pp. 33–40, New York, United States, 2008.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6615767, 8 pages
https://doi.org/10.1155/2021/6615767

Research Article

Secure OFDM-Based NOMA for Machine-to-Machine Communication

Shafiq U. Rahman,[1] **Amber Sultan**,[1] **Roobaea Alroobaea**,[2] **Muhammad Talha**,[3] **Syed B. Hussain**,[1] **and Muhammad A. Raza**[4]

[1]*Department of Electronics and Electrical Systems, The University of Lahore, Pakistan*
[2]*Department of Computer Science, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia*
[3]*Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia*
[4]*Department of Information Technology, Bahauddin Zakariya University, Multan 60000, Pakistan*

Correspondence should be addressed to Syed B. Hussain; baqar.hussain@es.uol.edu.pk

Received 26 November 2020; Revised 13 December 2020; Accepted 6 May 2021; Published 19 May 2021

Academic Editor: Mohammed El-Hajjar

Machine-to-machine communication (M2M) has obtained increasing interest in recent years. However, its enhancement and broadcasting characteristics produced a new security challenge. We have suggested a novel dynamic Quadrature Amplitude Modulation (QAM) scheme for a totally elastic and dynamic mapping of user data by using chaos. This paper analyses physical layer security methods in Orthogonal Frequency Division Multiplexing-based Nonorthogonal Multiple Access (OFDM-NOMA) and introduces a secure data transmission mechanism created by dynamic QAM. The security robustness given by the suggested encryption scheme is assessed, where an overall keyspace of $\sim 10^{163}$ is achieved, which is sufficient to provide security against exhaustive attacks. The result of the scheme is verified through MATLAB simulation, where the bit error rate performance of our proposed scheme is compared with an unencrypted OFDM signal, and the performance of our proposed scheme is analyzed for an illegal user. The suggested dynamic mapping fulfills the fundamental obligations of cryptography for data security. Moreover, it enhances the level of security in OFDM-NOMA.

1. Introduction

Among the top technological trends of the world, one of the most promising applications is M2M communications. The widespread application of M2M communication boosts their use in various fields. In M2M communication, various intelligent devices are associated with wired or wireless connections to implement Internet of Things (IoT) generation networks. These devices interface with one another without direct human intervention. IoT can be an initiative to open novel job opportunities, such as environmental monitoring, smart grids, health care, intelligent transport systems, building automation, and smart houses [1]. Since the data in M2M communication is intercommunicated through an open channel, the security of those data which contain sensitive information is a major concern. M2M communication is incredibly defenseless against attackers for a few reasons. Initially, its parts regularly invest a greater part of their energy unattended; what is more, in this way, it is easy to attack them physically. Secondly, most of the communications are done remotely, which makes eavesdropping very simple during downlink transmission [2]. In [3], an anonymous authentication hybrid encryption scheme is discussed and the Advanced Encryption Standard (AES) scheme is implemented for the confidentiality and authenticity of the multidomain of M2M communication. The scheme has achieved mutual authentication for inter- and intradomain communication in the absence of the identity of M2M devices. The presented scheme can protect small data such as text; however, is not suitable for large data encryption. Recently, work

has been done on corporate chaos with encryption. Chaos is a part of mathematics that reviews the behavior of a dynamic framework that is very sensitive to the initial condition, randomness, and unpredictability. In [4], the authors suggested a private nonorthogonal multiple-access visible-light communication system encryption based on nonlinear dynamical systems. The encryption scheme assures privacy among all legal users against attackers during communication. A recent scheme has deployed a two-level encryption mechanism that encrypts the data of multiusers using different keys [5]. A chaotic NOMA scheme for downlink transmission is given in [6]. In [7], a four-dimensional (4D) encryption is utilized to improve the secrecy of OFDM Passive Optic Network (OFDM-PON). 4D-hyperchaotic mapping is utilized to produce four masking factors to get ultra-high-confidentiality encryption in four various dimensions [8]. In [9], encryption techniques are used in the physical layer based on frequency induction for OFDM signals to enhance the security against any present attackers. In [10], the authors have proposed determining a real-time secured transmission system with a chaos-based encryption scheme deployed in the physical layer of the OFDM-PON. In the encryption procedure, field-programmable gate array boards are used at the optical line terminal (OLT) and optical network unit (ONU), utilizing hyperdigital chaos. The scheme provides optimum security and has a large keyspace. However, according to [11], the encryption algorithms based on a single round permutation diffusion are capable of resisting the chosen plain text attack. Thus, a multifold and efficient substitution-permutation-based encryption scheme is necessitated for the encryption application in the physical layer of the OFDM, which provides better security to the data communication.

Moreover, all mentioned schemes have considered a conventional mapping criteria for the underlying modulation scheme. For the traditional QAM mapping procedure, all OFDM information symbol points are static on the constellation diagram. According to this fixed mapping rule, the cipher information can only inhabit the constellation's exact position. Alone, scrambling the constellation points among those fixed locations cannot provide strong security defense. This scrambled data can be used later for statistical analysis, and thus, the security of the data can be compromised. In [12], the authors tried to hide the underlying constellation by chaotic insertion of pilots to increase the security of the OFDM-PON system based on a chaotic system. However, with blind channel estimation, the channel can be estimated and useful data can be extracted as the characteristics of the constellation would be revealed. Therefore, to assure user data security with statistical analysis, it is required to make symbol mapping dynamic.

This manuscript has introduced a novel encryption scheme for physical layer security in OFDM-NOMA. The proposed scheme is based on a dynamic system that is capable of multimedia data and text data encryption. The proposed scheme aims at mapping QAM symbols anywhere independently and dynamically onto the complex plane. The dynamic mapping of QAM does not let the mapping be static, and thus, the user data is not compromised by statistical analysis. The scheme executes the two-dimensional Zaslavsky map to generate chaotic sequences using the initial conditions and parameters. As the first step of encryption, the scheme uses the obtained chaotic sequences to permute the QAM symbols. XOR encryption is performed on the permuted QAM symbol to produce uniformness in the encrypted data that ensures the scheme's resistance against histogram attacks. The last step of our multifold encryption is distributing the encrypted QAM symbols on the complex plane independently. The independent mapping of QAM symbols consequently improves the scheme's resistance to correlation attacks. The proposed scheme achieves a keyspace of up to $\sim 10^{163}$, and as a result, it enhances the security level of the OFDM encryption scheme.

2. Methodology

In this section, we have discussed the proposed encryption scheme that is being deployed in the physical layer of OFDM-NOMA. In OFDM-NOMA, the proposed encryption scheme chaotically encrypts the data before the transmission and then sends the data through an insecure network. The scheme uses the chaotic output values to randomize user data and map it dynamically onto the dimensions of a higher QAM.

Figure 1 illustrates the proposed dynamic mapping of QAM in comparison with the conventional static QAM mapping. Figure 1(a) shows the 16-QAM static mapping, where every QAM symbol has a fixed position in the constellation plane. In Figure 1(b), we have demonstrated how the position of the QAM symbol will be changed chaotically, such that the QAM symbols will be mapped anywhere in the constellation plane. Figure 1(c) shows the resultant noisy constellation after incorporating chaos to map a QAM symbol dynamically.

In our suggested scheme, multifold chaotic encryption is achieved through the two-dimensional (2D) Zaslavsky chaos [13]:

$$\begin{aligned} x_n + 1 &= x_{n+}\nu(1 + \mu y_n) + \epsilon\nu\mu[\cos(2\pi x_n) \bmod 1], \\ y_n + 1 &= e^{-r}[y_n + \epsilon \cos(2\pi x_n)], \\ \mu &= \frac{1 - e^{-r}}{\tau}. \end{aligned} \tag{1}$$

In the above equation, x_n and y_n are the input samples to the chaotic map. The symbols ν and μ denote the controlling parameters that are used to control the chaotic behaviors, and for $n = 0$, the values of x_n and y_n are the initial conditions. The encryption scheme will be used for three operations: permutation, XOR operation, and constellation shifting [14]. Let D be user data of dimension $M \times N$. Then, the detailed procedure of the encryption process of data D is discussed in equations ((2)), equations ((3)), equations ((4)), equations ((5)), equations ((6)), and equations ((7)).

Figure 2 shows the block diagram of the proposed scheme. Since user data can be extremely correlated, therefore, initially, the proposed encryption scheme permutes the plain data. The scheme uses the chaotic map for the selection of the new position. At the beginning, the permutation step iterates the chaos map $2 \times M \times N$ times. Let x_i and y_i be the obtained iterated sequences, where the elements of

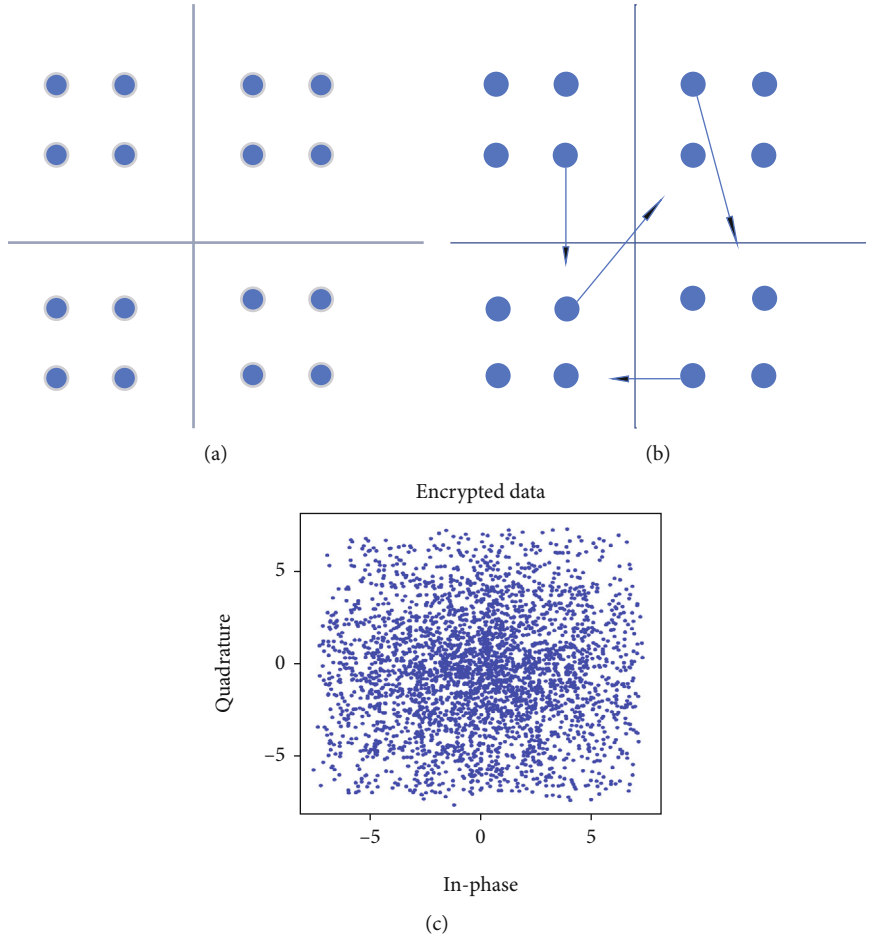


Figure 1: (a) Conventional 16 QAM. (b) Dynamic 16 QAM. (c) Resultant dynamically shifted 16 QAM.

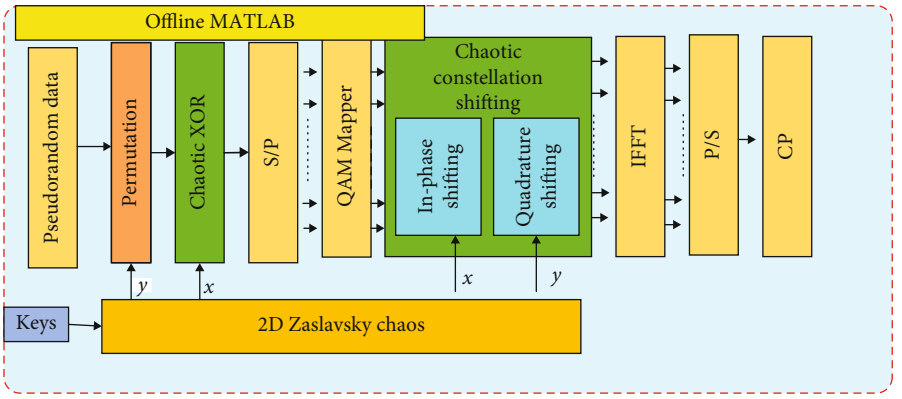


Figure 2: Block diagram of the proposed chaotic encryption scheme.

the sequences x_i and y_i belong to the close interval $[-2, 2]$. In the next step, we convert them into integer values using the following equations:

$$x_i' = \text{floor}\left(x_i \times 10^{15} \bmod M\right), \tag{2}$$

$$y_i' = \text{floor}\left(y_i \times 10^{15} \bmod N\right). \tag{3}$$

Equation (2) and equation (3) are used to transform the data from interval $[-2, 2]$ into the sets $\{0, 1, 2, 3, \cdots .M\}$ and $\{0, 1, 2, 3, \cdots, N\}$. Accordingly, in the new sequences x_i' and y_i', the elements occur randomly in the range between 0 and $M \times N$. Subsequently, we use the obtained random integer sequence and permute the data matrix D by using the equation given as follows:

$$D'(i, j) = D\left(x_i' + 1, y_i' + 1\right), \tag{4}$$

where $D(i, j)$ denote the pixel position of the original data matrix placed at the position ith row and jth column, and D' denote the permuted matrix. Consequently, one can get the permuted data.

As shown in Figure 2, the XOR operation is then used to increase the randomness in the ciphered data. The XOR operation scheme generates a sequence of random numbers through a chaotic map. Therefore, the order of the sequence is the same as the order of the data matrix. The obtained sequence is then XORed with the permuted data. The mathematical representation is given as follows:

$$C(i, j) = D'(i, j) \bigoplus S(i, j), \tag{5}$$

where $C(i, j)$ denote the data of the new data, and $S(i, j)$ denote the elements of the generated sequence. After the XOR operation, one can get the new data matrix C of the required ciphered data.

In Figure 2, the in-phase and quadrature shifting blocks show our proposed scheme. To permit an elastic mapping, we have utilized chaos for constellation shifting. The chaotic sequences x and y are used to shift the in-phase and quadrature parts of the QAM symbol as follows:

$$I_x = -2 + 4 * [\text{mod}\ (\text{abs}(y), \text{floor}(\text{abs}(y)))], \tag{6}$$

$$Q_y = -2 + 4 * [\text{mod}\ (\text{abs}(y), \text{floor}(\text{abs}(y)))], \tag{7}$$

where x and y are the chaos output value. The mod is used for the remainder, and abs is used to convert the negative value of x and y to a positive value. The floor function is utilized to round the number downward. Using equation (6) and equation (7), the in-phase I_x and quadrature-phase Q_y will give the value between $[-2, 2]$.

The encrypted data is converted into serial data after performing IFFT. This serial data is then appended with a cyclic prefix of length 1/16 of the OFDM symbol. Table 1 shows a detail of the parameters used during the offline MATLAB simulation.

As the proposed scheme is a symmetric encryption scheme, the decryption will be done by performing the reverse of all processes at the receiver side. However, for performing channel estimation with the help of pilots, a zero-forcing algorithm is used. The legal receiver would use the same initial keys to generate the chaotic sequence used by the sender and thus decrypt the received data. Moreover, due to the multifold and independent mapping of each QAM symbol, chaos reconstruction is not possible. The initial keys can be shared between the legal users wirelessly based on [15] or over an optical channel based on [16].

3. Results and Discussion

The possibility of the suggested encryption scheme is proven through simulation analysis. Figure 3 illustrates the simulation results after sending encrypted 16-QAM OFDM information. 64 subcarriers have contemplated conveying 16-QAM information. To get information from the channel, the pilots are added to the data. Then, the data is transmitted in parallel, and after performing IFFT, the ciphered information is passed from a parallel to a serial. A cyclic prefix of 1/16 length of the OFDM symbol is then attached. These ciphered signals are then transmitted over an AWGN channel with SNR ranging from 10 dB to 30 dB. The data encoding is carried out using MATLAB programs. Two types of users are contemplated at the receiving side in this simulation (i.e., User-1 and User-2). User-1 is considered the legitimate user with information about the preshared keys, while User-2 is considered the illegitimate user with no information about the preshared keys. The illegitimate user will get the same ciphered OFDM signal as that obtained by the legitimate user. However, the legitimate user will be able to demap the noisy constellation into a regular 16-QAM constellation with the help of the preshared keys. For an illegitimate user, even after blind channel estimation, the constellation will still reveal no information. In our simulation results, we have compared our proposed scheme with an unencrypted OFDM signal; the resulting BER curve is shown in Figure 3. Our proposed scheme maps the constellation points anywhere in the complex plane, i.e., the point can be located away from or near the origin, based on the output of the chaos. As long as the new location remains near the origin, there would be no power penalty; however, when the new QAM symbol is far from the origin of the complex plane, more power is required to transmit it. Therefore, when compared with an unencrypted OFDM signal, our proposed signal incurs a power penalty. As shown in Figure 3, the BER performance of our proposed scheme shows a power penalty of ~0.7 dB compared to an unencrypted OFDM signal at BER 10^{-3}. The power penalty can be reduced by allowing the dynamic mapping within the dimensions of the conventional 16 QAM. The BER curve for User-2 shows that an illegal user cannot receive any useful information at any SNR.

In the simulation results of Figure 4, we have compared encrypted signals with unencrypted OFDM signals. A decoded OFDM signal shows better execution when contrasted with encrypted signals. Figure 4 shows the simulation results after sending encrypted 64-QAM OFDM information. The power penalty at BER 10^{-4} is approximately the same for 16 QAM and 64 QAM. Therefore, our proposed scheme can be used

Table 1: Parameters used in simulation.

Parameters	Value
Subcarriers	64
Symbols	50
Pilots	4
Cyclic prefix	16
SNR	10-30
x_n	0.14
y_n	0.15
V	4
E	2.3
R	3
U	$1 - \exp(-r)/2$

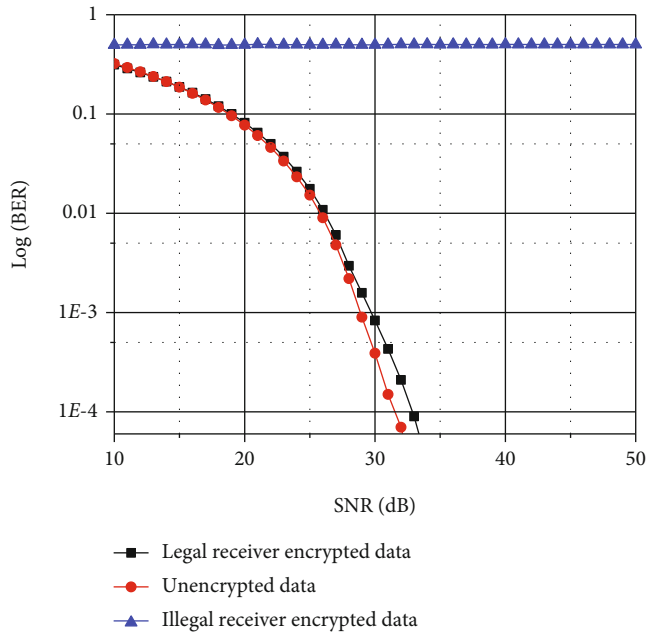


Figure 3: BER performance for encrypted 16 QAM.

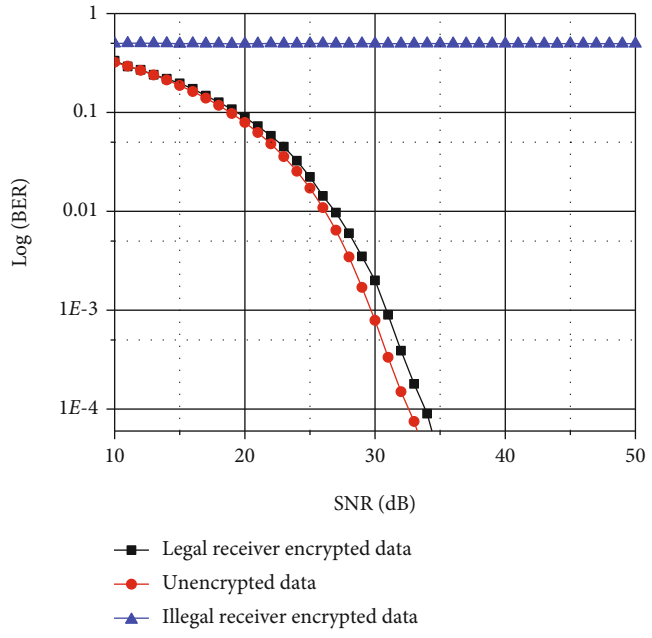


Figure 4: BER performance of encrypted 64 QAM.

for a higher dimension QAM as well. The BER is calculated after iterating the simulation model 50 times and taking the average of bit error values for all SNRs.

The sensitivity of the 2D chaotic system used in our suggested scheme is depicted in Figure 5. It is seen that with a little change, i.e., at the 10^{-15} position, in one input's initial value, the resultant output structure is completely different. Therefore, with the usage of 2D chaos to perform encryption, the confidentiality is improved. Moreover, this sensitivity of the 2D chaos system will alone provide a key space of 10^{-30}.

To examine the impact of the suggested chaotic constellation shifting plan, we decipher the obtained signals for all possible cases, where either the entire or different chaotic orders are unknown. Figure 6 illustrates the related BER analysis. We can see that in all these feasible cases, an illegal user is not able to recover the cipher data (i.e., BER ~ 0.5). The constellations, along with the encrypted data received by an illegitimate user with one identified shifting parameter, are exhibited as insets in Figure 6. With one known chaotic order, the suggested encryption scheme still hides the data

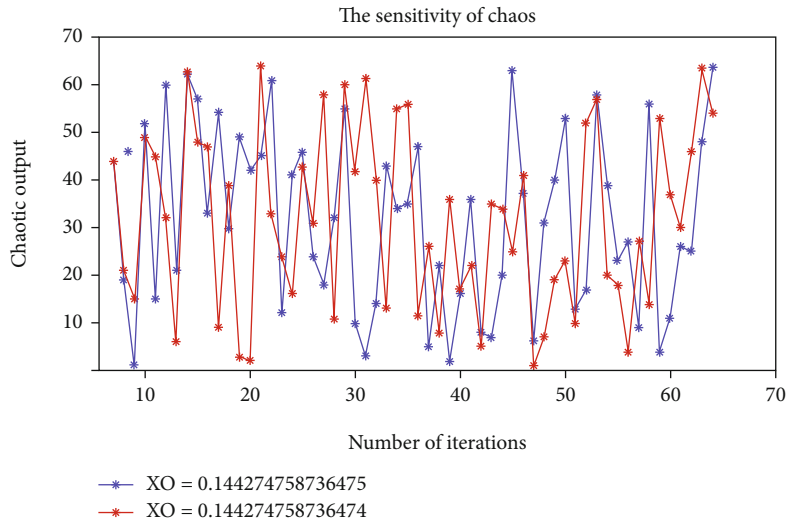


Figure 5: 2D chaos sensitivity diagram.

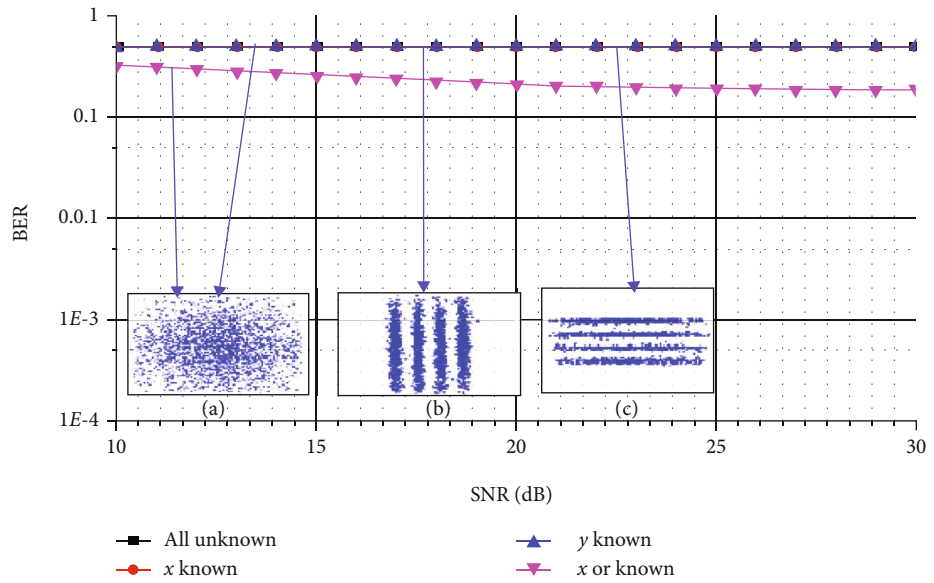


Figure 6: BER for illegal users.

successfully from an illegitimate user. As a result, it can be expressed that the suggested chaotic encryption scheme does not bargain the physical layer confidentiality of the information under all potential causes. Furthermore, finding the privacy key by relating unencrypted text with the cipher text will be challenging due to the multifold and autonomous QAM constellation shifting. Figure 6(a) shows that if the illegal user does not know about the chaotic sequence of mapping or only knows about the chaotic sequence of XOR operations, in both situations, they will receive a noisy constellation. If the illegal user knows only about the x-axis chaotic sequence, they will receive the constellation shown in Figure 6(b). If the illegal user knows about the chaotic y-axis sequence, they will only be able to demap the symbol mapping along the y-axis and thus receive the constellation as shown in Figure 6(c). Thus, it can be concluded that our proposed scheme provides efficient security even if the illegal user knows any one of the initial values.

To further evaluate our proposed scheme, we have transmitted an image. Figures 7(a) and 7(b) show the unencrypted image and its histogram, respectively, whereas Figures 7(c) and 7(d) show the encrypted image and its histogram, respectively. It can be seen that the encrypted image does not reveal any information about the real image. Moreover, the histogram of the encrypted image is almost uniform. Therefore, any attack on the histogram to reveal information about the image will not be successful. Therefore, the proposed scheme provides good image encryption as well.

The strength of the suggested chaotic encryption scheme can be assessed by determining the set of all keys used during encryption and decryption. The keyspace of the suggested scheme is quantitatively evaluated as follows. The chaotic

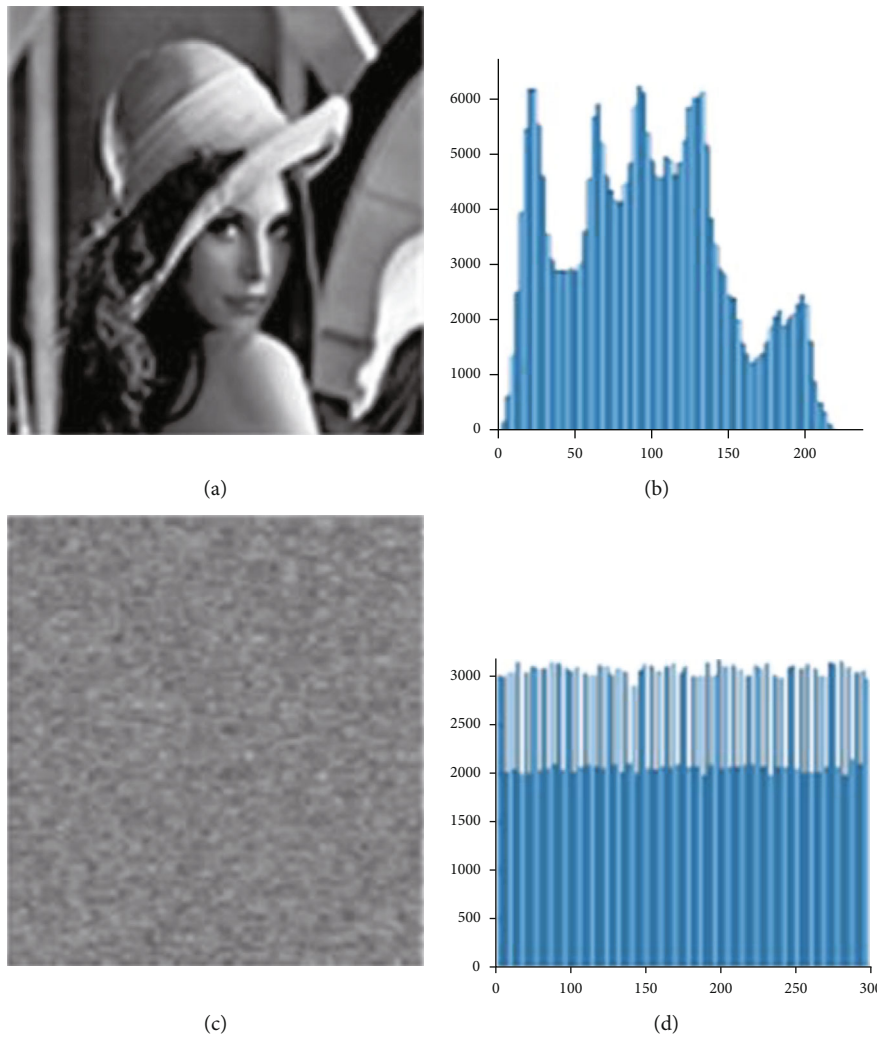


Figure 7: (a) Unencrypted transmitted image; (b) histogram of unencrypted transmitted image; (c) encrypted transmitted image; (d) histogram of encrypted transmitted image.

XOR gives a 2^{mN} keyspace, where m symbolizes the order of the QAM mapper used in the scheme, and N indicates the length of the OFDM symbol. The quadrature shifting and in-phase parameters generate $r \times 10^k \times 10^k$ keyspaces, where r is a binary number and k is the decimal position considered by (3) and (4), respectively. Consequently, the general formula to evaluate the approximate number of keyspaces is equal to $r \times 10^k \times 10^k \times 2^{mN}$. Since a total of 64 subcarriers are utilized in the suggested encryption scheme, a 16-QAM modulation is used and every single subcarrier is characterized by 4 bits. Therefore, when the chaotic XOR sequence is complete, this generates a keyspace approximately equal to 2^{256}. The chaotic parameters I and Q comprise two values before the fractional element and the digit number during constellation shifting considered on the tiny part. For clarity, we measured up to 4 positions of the decimal point; therefore, it will produce an extra keyspace of $10^4 \times 10^4 \times 2$. Similarly, when the chaotic permutation sequence is completed, this also generates a keyspace of 2^{256}. Besides, the constellation is affected because of the expansion and deduction of the moving parameters, so it boosts the order of the keyspace up to $10^4 \times 10^4 \times 2^{256} \times 2^{256} \times 4 \times 2$ ($\sim 10^{163}$). The resultant values guarantee the strength of the proposed scheme against the brute force attack.

4. Conclusion

In this paper, we exhibit a novel scheme for OFDM-based NOMA physical layer security, where we apply the confused constellation moving to multifold OFDM information encryption. Simulation results reveal that, because of the execution of constellation shifting, an elastic constellation with dynamic in-phase and quadrature shifted dimensions is accomplished. The corresponding noisy constellation successfully encodes the OFDM information with chaotic scrambling. The security robustness given by the suggested encryption scheme is assessed, where an overall keyspace of $\sim 10^{163}$ is achieved. The proposed scheme can be used to provide security to both uplink and downlink data transmission. The possibility of the scheme is certified through the

simulation, where the bit error rate performance of our proposed scheme shows a power penalty of ~0.7 dB in comparison with an unencrypted OFDM signal. However, in the presence of high noise, the power penalty would increase. The simulation results illustrate that the suggested scheme can be a good candidate for the upcoming secure OFDM-NOMA.

Data Availability

Data can be taken any time through email correspondence.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful to the Taif University Researchers Supporting Project (number TURSP-2020/36), Taif University, Taif, Saudi Arabia. The authors are also grateful to the Deanship of Scientific Research and King Saud University for funding this research work.

References

[1] W. Zhan and L. Dai, "Massive random access of machine-to-machine communications in LTE networks: throughput optimization with a finite data transmission rate," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5749–5763, 2019.

[2] F. Hussain, L. Ferdouse, A. Anpalagan, L. Karim, and I. Woungang, "Security threats in M2M networks: a survey with case study," *Computer Systems Science and Engineering*, vol. 270, 2016.

[3] Y. Qiu, M. Ma, and S. Chen, "An anonymous authentication scheme for multi-domain machine-to-machine communication in cyber-physical systems," *Computer Networks*, vol. 129, pp. 306–318, 2017.

[4] Y. Yang, C. Chen, W. Zhang et al., "Secure and private NOMA VLC using OFDM with two-level chaotic encryption," *Optics Express*, vol. 26, no. 26, pp. 34031–34042, 2018.

[5] J. S. Khan, W. Boulila, J. Ahmad et al., "DNA and plaintext dependent chaotic visual selective image encryption," *IEEE Access*, vol. 8, pp. 159732–159744, 2020.

[6] N. Horiike, E. Okamoto, and T. Yamamoto, "A downlink non-orthogonal multiple access scheme having physical layer security," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, 2018.

[7] J. Zhao, B. Liu, Y. Mao et al., "High security OFDM-PON with a physical layer encryption based on 4D-hyperchaos and dimension coordination optimization," *Optics Express*, vol. 28, no. 14, pp. 21236–21246, 2020.

[8] F. M. Almansour, R. Alroobaea, and A. S. Ghiduk, "An empirical comparison of the efficiency and effectiveness of genetic algorithms and adaptive random techniques in data-flow testing," *IEEE Access*, vol. 8, pp. 12884–12896, 2020.

[9] M. Jacovic, K. Juretus, N. Kandasamy, I. Savidis, and K. R. Dandekar, "Physical layer encryption for wireless OFDM communication systems," *Journal of Hardware and Systems Security*, vol. 4, no. 3, pp. 230–245, 2020.

[10] J. Zong, A. A. Hajomer, L. Zhang, W. Hu, and X. Yang, "Real-time secure optical OFDM transmission with chaotic data encryption," *Optics Communications*, vol. 473, 2020.

[11] J. Liu, Y. Ma, S. Li, J. Lian, and X. Zhang, "A new simple chaotic system and its application in medical image encryption," *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 22787–22808, 2018.

[12] Q. Chen, M. Bi, X. Fu et al., "Security scheme in IMDD-OFDM-PON system with the chaotic pilot interval and scrambling," *Optics Communications*, vol. 407, pp. 285–289, 2018.

[13] R. Hamza and F. Titouna, "A novel sensitive image encryption algorithm based on the Zaslavsky chaotic map," *Information Security Journal: A Global Perspective*, vol. 25, no. 4-6, pp. 162–179, 2016.

[14] A. Alsufyani, R. Alroobaea, and A. Ahmed, "Detection of single-trial EEG of the neural correlates of familiar faces recognition using machine-learning algorithms," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 2855–2860, 2019.

[15] X. Yu, X. Zhou, C. Xu, L. Wang, D. Shen, and H. Zhou, "A NOMA-based quantum key distribution system over Poisson atmospheric channels," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Waikoloa, HI, USA, 2019.

[16] Y. Wu, Y. Yu, Y. Hu, Y. Sun, T. Wang, and Q. Zhang, "Channel-based dynamic key generation for physical layer security in OFDM-PON systems," *IEEE Photonics Journal*, vol. 13, no. 2, pp. 1–9, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6672297, 12 pages
https://doi.org/10.1155/2021/6672297

Research Article

Mobility-Aware Routing Algorithm for Mobile Ad Hoc Networks

Chalew Zeynu Sirmollo[1] **and Mekuanint Agegnehu Bitew**[2]

[1]*School of Computing and Informatics, Mizan-Tepi University, Tepi, Ethiopia*
[2]*Faculty of Computing, Bahir Dar Institute of Technology, Bahir Dar University, Bahir Dar, Ethiopia*

Correspondence should be addressed to Mekuanint Agegnehu Bitew; memekuanint@gmail.com

Received 3 November 2020; Revised 13 February 2021; Accepted 3 May 2021; Published 13 May 2021

Academic Editor: Ihsan Ali

Mobile ad hoc network (MANET) is a group of wireless mobile nodes that create a temporary network without the help of any central administration or standard support services. Mobility of nodes determines the overall performance of MANET networks. High mobility of nodes causes frequent changes in the network topology, and this leads to link breakage and increases reinitiating of the route discovery process. MANETs commonly use broadcasting techniques for route discovery process. However, it can cause redundant rebroadcasts, packet collisions, and channel contention. The main objective of this paper is to design and develop the mobility-aware routing algorithm (MARA) to enhance the performance of the routing protocol in MANETs. The proposed scheme allows mobile nodes to rebroadcast or discard received broadcasted messages. The decision is based on the combination of node speed, distance between nodes, and residual energy of nodes. These parameters are considered both in route request and route reply process to reduce the chance of link breakage and broadcast storm problems. The proposed algorithm has been evaluated based on the performance metrics: packet delivery ratio, average end-to-end delay, throughput, and routing overhead. We have used network simulator NS-2 V-2.35. The simulation results revealed that MARA outperforms ad hoc on-demand distance vector (AODV), mobility and direction aware (MAD), and mobility and energy-aware (MAE) routing protocols.

1. Introduction

MANET is an autonomous system of mobile nodes with routing capabilities connected by wireless links [1]. Every mobile node can act as a router, and it can communicate directly with another node in its physical neighborhood. MANET is a collection of mobile nodes that can dynamically change locations to form a network to exchange information. The main features of MANETs are automatic self-configuring, self-maintenance, inexpensive deployment, and lack of the need for fixed network infrastructures or centralized administration [2]. Due to its flexible nature, MANETs are widely applicable in battlefield communications, disaster relief, emergency operation, and educational, commercial, rescue and search operations [3]. One of the basic challenges in MANET is the designing of dynamic routing protocols that can be mobility aware and efficient to determine the routes between the communicating nodes with better performance and less overhead [1]. Routing protocols in MANETs can be categorized into three groups based on how routing information is acquired and maintained by mobile nodes [1, 4, 5]. The first category belongs to proactive (table-driven) routing protocols. In these routing protocols, mobile nodes calculate routes to all reachable nodes a priori and maintain consistent and up-to-date routing information using a periodic route update process [1]. Examples of proactive routing protocols are destination-sequenced distance-vector routing (DSDV) and optimized link state routing protocol (OLSR). The second category of routing protocols is named as reactive (on-demand) routing protocols. In these protocols, when any node wants to communicate to the other node, it applies on-demand route discovery mechanism for creating connections. Examples of reactive routing protocols are dynamic source routing (DSR) and AODV. Hybrid routing protocols are the third category of MANET routing protocols. Hybrid routing protocols combine the best practices of both proactive and reactive routing protocols. The zone routing protocol (ZRP) is an example of hybrid routing protocols.

Reactive routing protocols such as AODV [5] do not adapt well in high-mobility environments. Reactive protocols select the shortest path between the source and destination nodes. However, the shortest path may not be always reliable or active due to mobility of nodes. The frequent breakage of the established path degrades the performance of MANETs [6]. When an active route between the source and the destination node breaks, the routing protocol executes route maintenance procedures. However, this consumes network resources and eventually influences negatively on the performance of the network [6].

Broadcasting is an information distribution mechanism for sending a packet from a source to all nodes within a network [3]. Several routing protocols broadcast route request to seek multihop route to the destination. For example, AODV [5] typically uses broadcasting in the route discovery process. In MANETs, broadcasting occurs frequently for finding a route to a specific host, paging a specific host, and other network operations [7]. Broadcasting through flooding causes contention when many adjacent nodes broadcast concurrently [3]. Additionally, node mobility creates continuously changing network topology, in which routing paths are broken and new routes are formed dynamically. Therefore, broadcasting is one of the challenging problems in MANETs. Several approaches have been proposed to solve the issue as we have discussed in Related Works. However, they did not combine multiple parameter metrics simultaneously with decision methods such as the speed, direction, and residual energy of mobile nodes to aware dynamic change of the network topology.

In this paper, we proposed new mobility-aware routing scheme for MANETs, which can adapt dynamic changes of the network topology. The proposed algorithm considers node mobility and a broadcasting decision mechanism to reduce the chance of link breakage and broadcast storm problems. We used node speed, node direction, and residual energy of nodes.

2. Related Works

Broadcasting in MANETs can be classified as simple (blind) flooding, location-based, distance-based, neighbor-knowledge-based, counter-based, and probabilistic-based [8–12]. In [12]; the authors have proposed a dynamic probabilistic-based routing scheme. In this scheme, packets forwarded to the neighbor node with dynamically computed probability, which is forwarding probability. The probability function is calculated based on the density of the local neighbors and cumulative amount of neighbor mobile nodes. The proposed approach showed improvement on performance as compared to AODV and fixed probability AODV (FP-AODV) routing protocols. However, the throughput is very low when the region of the network is spare. This is because of poor connectivity ratio among the mobile nodes in the sparse region and the end-to-end delay is high. Hence, the failure of a route request packet is high. The authors in [13] aimed to establish more stable data routes over vehicular ad hoc networks. However, they did not consider residual energy during route path selection. The authors of [14] proposed a mobility and load aware routing scheme which has a stable route and load balancing among several routes in high-mobility and high traffic load situation. This scheme uses speed and traffic load of intermediate nodes to determine the reliable route during the route discovery process. The Markovian decision process tool was used to rebroadcast or discard when each node receives a request packet. Simulation results showed that this scheme reduced the effects of broadcast storm problem, increased throughput, and reduced routing overhead as compared to the AODV protocol. However, the authors did not consider residual energy and direction of nodes. In [15], the authors proposed personalized ad hoc on-demand distance vector algorithm that deals with link breakage during exchange of data in MANETs. In this algorithm, selection of routing path based on delay and bandwidth metrics helped to improve the quality of service. It did not consider residual energy and speed of mobile nodes. In [16], the authors designed and implemented a mobility adaptive ad hoc on demand distance vector routing protocol which extends the AODV routing protocol using a hello message. This protocol has a capability of predicting mobility of nodes. However, it increases overhead due to maintenance of an active neighbor list to provide an alternative link during mobility. In [17], the authors proposed an MBMA-OLSR routing scheme. The paper extends and enhances the conventional MP-OLSRv2 routing protocol. Control messages (hello and TC) gather all the network topology information to construct the network graph. Through the network graph, MBMA-OLSA can avoid nodes with higher speeds and lower residual energy and choose more stable links. MBMA-OLSR is significantly superior to MP-OLSRv2 in throughput, end-to-end latency, and packet delivery rates in the range of 5-30 m/s. In [18], the authors presented the new energy and mobility-aware multipoint relay (EMA-MPR) selection mechanism which is an extension of the conventional MP-OLSRv2 protocol. Energy and mobility aware parameters decide the willingness of the node to become a multipoint relay node in the OLSR routing protocol. The proposed scheme concentrated on the fact that considering nodes speed alongside the residual battery of the node results in stable paths to the destination. As the authors stated, the scheme provides its effectiveness for high-speed scenarios with heavy traffic. However, it did not consider the distance between the communicating nodes through movement. In [19], the authors proposed a multicriteria-based hybrid multipath routing protocol. This study computed the multicriteria node rank metric by combining several parameters that are associated with energy and quality of service (QoS) for reducing the control overhead. Based on the computed metric, the multipoint relay nodes are selected by considering the energy and QoS metric. This work focuses on control routing overhead. In [20], the author has studied the performance of the proposed multipath routing protocols for energy-efficient and QoS awareness depending on the node's mobility in MANETs. This works focus on evaluating the performance of the multipath battery and queue aware routing (MBQA-OLSRv2), multipath battery, and mobility aware routing (MBMA-OLSRv2). The obtained results show that the MBMA-OLSRv2 scheme outperformed MBQA-OLSRv2

routing, the scheme in terms of several metrics like throughput, total packets dropped, energy cost per packet, consumed energy, and delay particularly in high-mobility scenarios. Mobility and direction aware (MDA) proposed in [21] is aimed at establishing stable routes and reducing the chance for link breakages based on node speed and direction with respect to the source and destination nodes. Compared to AODV, MDA improved delivery ratio, end-to-end delay, routing overhead, and energy consumption. However, it did not consider the number of the redundant rebroadcasting packets generated by nodes. The selected most reliable path may contain low-power mobile nodes, which leads to link breakage. In addition, MDA considered the speed of nodes' movement only in the route request packet (RREQ) broadcasting process. It did not take into account the direction of the node in the RREQ process. The mobility and energy-aware routing algorithm (MEA) proposed in [20] used relative mobility and residual energy of nodes. However, the remaining energy and relative mobility were considered only in the route replay phase; it did not considered these parameters in the route discovery process. In addition, the scheme has a broadcast storm problem because nodes broadcast RREQ packet into its neighbor without any consideration.

In general, many researchers tried to enhance the performance of MANET routing algorithms by considering different metrics. However, there are limitations in combining multiple parameters with decision-making techniques to exchange information through the network and to reduce mobility and broadcasting problems. To the best of our knowledge, there is no work, which considers node speed, distance calculation, and residual energy simultaneously to establish a stable and reliable route between the source and destination nodes in route request and route reply phases. In this paper, we have considered node speed, direction value, and the remaining energy of a mobile node as a factor and proposed the MARA scheme.

3. Proposed Routing Algorithm

Nodes in MANETs change position quite frequently, and this has a great impact on the performance of the routing algorithm. To reduce this problem, the node's mobility should be considered when designing a routing protocol for MANETs. Our proposed algorithm is capable of adapting to frequently changing network topology because we considered the node's mobility. The knowledge of the position is significant for successful routing of packets. We assumed that nodes are equipped with global position system (GPS) and connect to each other using an omnidirectional antenna. Each node sends a RREQ message to the neighbor nodes to get their position at the current time. When a node broadcasts RREQ message, neighbors will get the required parameters and store in its neighbor fields in the routing table. Based on this information, mobile nodes can compute their speed and distance between neighbor nodes. Hence, nodes are aware of neighboring node information.

Table 1: Modified RREQ message format [5].

....	X_{position}	Y_{position}	Speed (m/s)

3.1. Description of Route Discovery and Reply Phases

3.1.1. Route Discovery Phase. The proposed algorithm is an on-demand routing protocol. When a source node wants to transmit data to a destination node and it has no routing entry for this destination, then the route discovery process of the proposed algorithm is initiated. The initiation of a route discovery process is done by broadcasting new RREQ packets to all neighbor nodes. This new RREQ packet is an extension of the AODV RREQ packet, Table 1. Figure 1 shows the flow chart of route discovery and reply phases of the proposed algorithm, MARA. When a RREQ packet is received by a neighbor node, it checks whether the RREQ packet is a duplicate or not. If it is a duplicate, the node discards or drops the RREQ packet immediately. If it is not, it searches for the reverse route towards the source node. If a route exits, it updates the existing route otherwise create a reverse route. If the receiving neighbor node is not the destination node and there is no valid route, it checks whether its speed is below a predefined threshold speed. Then, the neighbor or intermediate node checks whether the two nodes' direction value is one. In addition, check whether the neighbor node residual energy value is greater than the threshold value of energy. Finally, if the condition is satisfied, the node updates the route to the originator, increments the hop count by one, and broadcast the RREQ. If it is not, the node discards or drops the RREQ packet. Each node will follow this process until the RREQ packet reaches the destination or an intermediate node that has a valid route to the destination.

3.1.2. Route Reply Phase. In the above section, we have described how the RREQ packet reaches the destination or intermediate node that has a valid route to the destination. The destination node needs to verify the speed and direction of its neighbors through which the request was received. This process has two different conditions based on where this node is: the destination itself or an intermediate node that has an active route to the destination. If the node is an intermediate node which has an active route to the destination, the node updates its route and initiates RREP packet to the previous-hop from which it has received the RREQ packet only the previous hop, and if its intermediate node is moving with the speed below the average speed, all neighbor nodes and the direction value is one. If the node is the destination, it updates its route and initiates RREP packet to the previous-hop from which it has received the RREQ packet; if only the previous hop is moving with the speed below average speed, all neighbor nodes and the node direction value is one, and if not, discard or drop the RREQ packet. If the number of previous-hop direction values is greater than equal to one, generate RREPs and select the path which contains nodes with having minimum relative speed values in order to achieve better link stability. If not, it discards or drops the

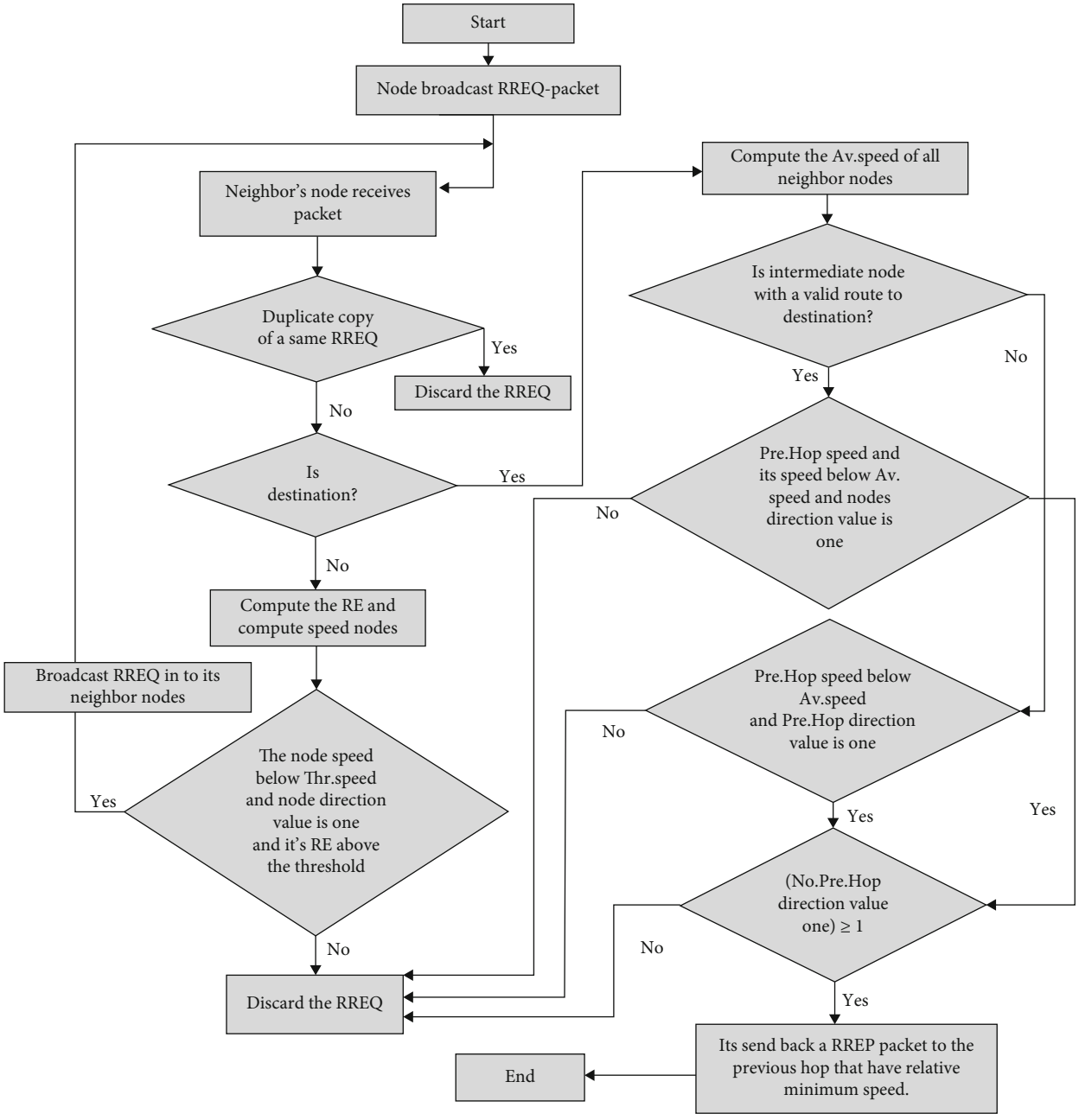


Figure 1: Flowchart of MARA in the route discovery and reply phases.

RREQ packet. Each intermediate node that received the RREP packet until the RREP packet reaches the original source node will follow this process.

3.2. Distance Calculation between Mobile Nodes. Location, speed, and direction of mobile nodes change dynamically in MANETs. We computed the Euclidean distance between two neighbor nodes to determine whether the nodes are moving towards each other or away from each other. Suppose that there are two mobile nodes n_1 and n_2 with transmission range of r, as shown in Figure 2. Nodes n_1 and n_2 are moving at the speed of V_{n1} and V_{n2}, respectively.

The distance between n_1 and n_2, $D(n1, n2)(t)$ at a time t can be calculated as

$$D(n1, n2)(t) = \sqrt{\left(X_{n2(t)} - X_{n1(t)}\right)^2 + \left(Y_{n2(t)} - Y_{n1(t)}\right)^2}, \quad (1)$$

where $(X_{n1(t)}, Y_{n1(t)})$ and$(X_{n2(t)}, Y_{n2(t)})$ are the locations of $n1$ and $n2$ at a time t, respectively.

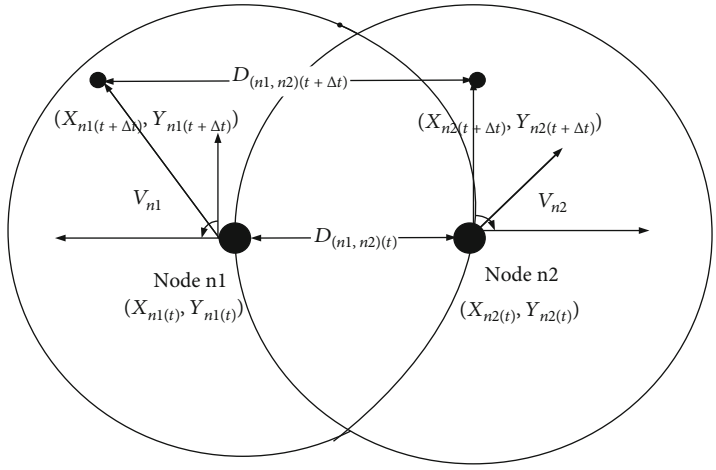


Figure 2: Position estimation for mobile nodes [22].

At $(t + \Delta t)$ time, the distance between n_1 and n_2, $D(n1, n2)(t+\Delta t)$ can be calculated as

$$D(n1, n2)(t+\Delta t) = \sqrt{\left(X_{n2(t+\Delta t)} - X_{n1(t+\Delta t)}\right)^2 + \left(Y_{n2(t+\Delta t)} - Y_{n1(t+\Delta t)}\right)^2}, \quad (2)$$

where $(X_{n1(t+\Delta t)}, Y_{n1(t+\Delta t)})$ and $(X_{n2(t+\Delta t)}, Y_{n2(t+\Delta t)})$ are the locations of $n1$ and $n2$ at a time of $(t + \Delta t)$, respectively.

As mobility discussed in [23], mobility is the average change in distance between all nodes over that period of time. It is a function of speed and movement pattern. Speed of a node at a time t can be calculated as follows:

$$V_{(t,t+\Delta t)} = \frac{/(X_2 - X_1) + (Y_2 - Y_1)/}{(t+\Delta t) - t}, \quad (3)$$

where $V_{(t,t+\Delta t)}$ is the speed field values that are discussed below in the structure of the route request and routing table. Hence, each node can compute its distance from its neighbors as well as its speed at any time t.

After determining the distance between n_1 and its neighbors at t and $(t+\Delta t)$, we can determine whether n_1 is joining or separating from its neighbors. If $D(n1, n2)(t) > D(n1, n2)(t+\Delta t)$, then nodes n_1 and n_2 are closer to each other within the time interval between t and $(t + \Delta t)$. Hence, the two nodes are joining each other for this time interval. On the other hand, if $D(n1, n2)(t) < D(n1, n2)(t+\Delta t)$, the nodes are moving away from each other and the nodes have a high probability to be disconnected.

3.3. Route Request Message Format. We modified the RREQ message format used in the AODV protocol because the route discovery process of MARA is based on speed, distance between nodes, and the remaining energy of nodes. The AODV protocol did not consider these parameters in its RREQ message format. The entries of the RREQ message format of the AODV protocol extended by adding three new fields to its structure. The newly added fields are

Table 2: Modified routing table Format [5].

.....	Speed (m/s)	Direction values

(i) X_{position}: this field contain the X coordinate of the mobile node

(ii) Y_{position}: this field contain the Y coordinate of the mobile node

(iii) Speed: this field indicates the speed of the mobile node

The dots in Table 1 indicate the RREQ message fields of the well-known AODV protocol [5]. When a node broadcasts a RREQ message, neighbor nodes get the required parameters from the broadcasted RREQ message and store these parameters in their neighbor fields. In this regard, every node gets updated location information from its neighbors at a certain interval of time.

3.4. Routing Table Format. A routing table is a set of rules, which contains information necessary to forward a packet about its origin in addition, along the best path towards its destination. It stores information in the form of a table. Based on equations (1) and (2), the routing table of the proposed algorithm contains two new fields, as shown in Table 2. These are the following:

(i) Speed of the neighbor node

(ii) Direction values: it contains the value of node movement based on the difference of distance between communicating mobile nodes as we discussed in Section 3.1

The direction in the routing table is used to set whether neighbors are joining or separating. In order to set the direction value, we compare the distance between the mobile nodes using equations (1) and (2). If the nodes are joining

```
Notations:
1. SA: Source or initiator address
2.DA: Destination address
3.IN: Intermediate node that has a valid route to the desired destination node
4.EDi: Euclidian distance between the initiator node and a destination node in the initial time
5.EDf: Euclidian distance between initiator node and destination node of in final time
6.RT: Route table
7.RRM: Route request message
8.RREP: Route request reply
9.Received pair: Source address and RREQ_ID
10.REn: Residual energy of node
11.SIn: Speed of source or initiator nodes.
12.SNn []: Speed of all neighbor nodes in a recipient routing table
13.SPn: Previous hop speed
14.PHn: Number of previous hop
15.SN: Source node
16.DM: Direction of movement if EDi > EDf set to 1 and EDi < EDf set to 0 in recipient routing table
17.BEGIN
18.SN Broadcast RRM into all neighbor nodes, then find match received pair to do this:
19.   Boolean found = False
20.      For (a = 1 to Rout_Table. Length)
21.      If (Received_pair == RT [a]) then
22.         found = True;
23.         End if
24.      if (found) then
25.      Drop RRM
26.      Else
27.   If (SA! = DA) then
28.      If (REn > Threshold_Energy && SIn < Threshold_Speed && Its DM==1) then
29.      Update RT
30.      Update RRM
31.      Broadcast RRM into its neighbor nodes
32.                  Else
33.                        Drop RRM
34.                        Loop Break
35.                     End else
36.            End if
37.      Else
38.   Switch (Node category)
39.         Case: "IN" then
40.               Count =0
41.                  For (b = 1 to neighbors. Length)
42.                     Count = Count + SNn [b]
43.                  End for
44.      Average_Speed = count/neighbors. Length
45.      If (SPn & SIn < Average_Speed && both DM==1) then
46.                  Fun()
47.                        Counter =0
48.For (i = 1 to Rout Table. Length)
49.            if (PHn value of DM =1) then
50.                  Counter ++51.                  End if
52.               End for
53.            if (Counter > = 1) then
54.            Update Routing Table
55.            It send back an RREP packet to the previous hop that has minimum speed.
56.                  Drop Received RRM
57.                     Loop Break
58.                     Else
59.                           Drop RREP packet
60.                     Loop Break
```

Algorithm 1: Continued.

```
61.                    End Fun
62.                End else
63.            End If
64.        End If
65.            End Case
66.      Case: "Final destination" then
67.                Count = 0
68.      For (c = 1 to neighbors. Length)
69.                Count = count + SNn[c]
70.                   End for
71.      Average_Speed = Count/neighbors. Length
72.      If (SPn < Average_Speed && DM==1) then
73.                Call to fun()
74.             Update Routing Table
75.   Its send back a RREP packet to the previous hop
76.                Drop Received RRM
77.                Loop Break
78.                     Else
79.                Drop Received RRM
80.                   Loop Break
81.                End else
82.             End If
83.         End Case
84.   End If
85. End for
86.End BEGIN
```

Algorithm 1: Pseudocode of the proposed routing algorithm (MARA).

each other and the distance between the two nodes is constant, the direction value is set to one. Otherwise, the direction value is set to zero. Each mobile node maintains its routing table based on this way.

3.5. Residual Energy of Mobile Nodes. To improve the lifetime of MANETs, we considered the residual energy of nodes at the time of transmission. Because the mobile nodes run using battery, its energy is restricted. A mobile node loses a certain amount of energy when it transmits/receives packets. The proposed algorithm used a generic radio energy model [24] to estimate initial energy (IE), residual energy (RE), and consumption energy (CE) at a time t. The following equation is used to compute the RE of nodes:

$$\mathrm{RE}_{(t)} = \mathrm{IE} - \mathrm{EC}_{(t)}, \tag{4}$$

where $\mathrm{EC}_{(t)}$ is the energy consumed by a node after time t . To estimate the energy consumed for transmitting N number packets, we used the following equation:

$$\mathrm{TE}_{(t)} = N * \mathrm{PT}_{(t)}, \tag{5}$$

where $\mathrm{TE}_{(t)}$ refers to the transmission energy of a node and $\mathrm{PT}_{(t)}$ refers to the power spent through the transmission of N number of data packets and at a certain period of time the neighbor nodes exchange routing information. The estimated energy consumption in a receiving N numbers of packets can be calculated by

$$\mathrm{RE}_{(t)} = N * \mathrm{PR}_{(t)}, \tag{6}$$

Table 3: Simulation parameters of all scenarios.

Parameters	Value
Routing protocol	MARA, AODV, MAD, and MAE
Propagation model	Two-ray ground
Simulation area	1000 m × 1000 m
MAC type	IEEE 802.11
Antenna type	Omnidirectional
Interface type	Phy/WirelessPhy
Packet rate	10 packets/sec
Traffic type	Constant bit rate (CBR)
Min speed	0
Max speed	10, 20, 30, 40, and 50 m/s
Number of nodes	10, 20, 30, 40, 50, 60, and 70
CBR flows	15
Mobility model	RWP and RPGM
Packet size	512 bytes
Transmission range	250 m
Queue length	50
Pause time	10 s
Simulation time	120 s

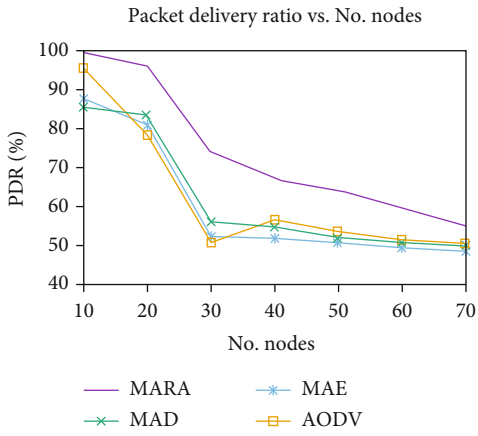


Figure 3: PDR vs. number of nodes using RWP.

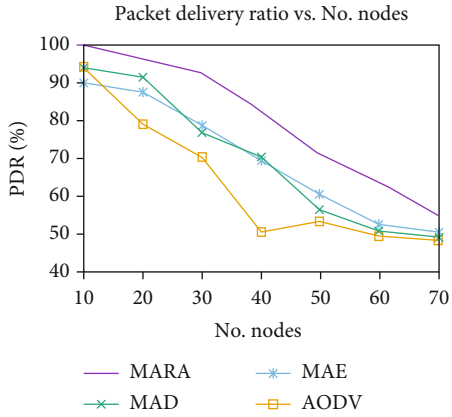


Figure 4: PDR vs. number of nodes using RPGM.

where $RE_{(t)}$ refers to the receiving energy of a node and $PR_{(t)}$ refers to the power spent through receiving N number of data packets and at certain period of time the neighbor nodes exchange routing information. Therefore, the total energy consumption of a node at the time t for transmission and reception is given by

$$EC_{(t)} = TE_{(t)} + RE_{(t)}. \quad (7)$$

Based on this, our proposed algorithm considered the residual energy in the broadcasting decision in order to reduce the chance of a broadcast storm problem. The following pseudocode shows the detailed explanation of the route request and route reply process of the proposed algorithm.

3.6. Simulation Scenarios. To evaluate the proposed scheme, we used network simulator two (NS2) [25] under the random waypoint (RWP) model [26] and reference point group mobility model (RPGM) [27]. The deployment considers 10, 20, 30, 40, 50, 60, and 70 mobile nodes with a simulation area of 1000 m × 1000 m. We have evaluated the performance of the proposed algorithm by varying node density and node mobility. Each node uses an IEEE 802.11 MAC layer protocol to send and receive messages. The connection

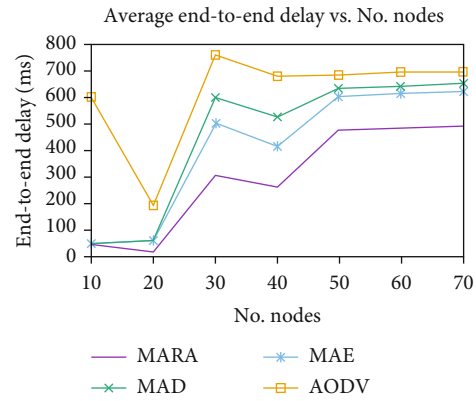


Figure 5: AETED delay vs. no. of nodes using RWP.

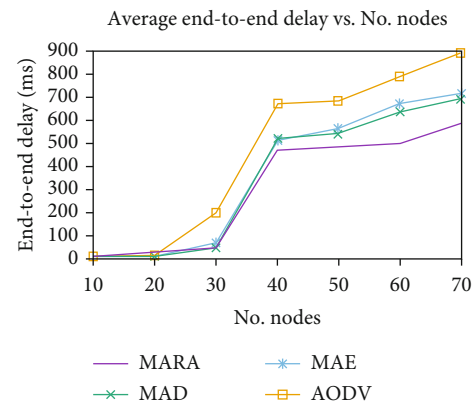


Figure 6: AETED delay vs. no. of nodes using RPGM.

between the source and destination node spread randomly over the network. The network bandwidth is 2 Mbps. We used a two-ray ground model for radio propagation [28]. We have considered two simulation scenarios: impact of node density and node mobility. Table 3 shows the details of the simulation parameters.

3.7. Performance Evaluation Metrics. In order to evaluate the performance analysis of the proposed algorithm, we considered the packet delivery ratio, average end-to-end delay, throughput, and routing overhead which are widely used performance metrics MANET [29, 30]. The description of these performance metrics is given below:

(a) Packet delivery ratio (PDR) refers to the ratio between the number of packets sent by constant bit rate sources and the number of packets received by the CBR sink at the destination. It is the measure of reliability of a protocol

(b) Average end-to-end delay (AETED) refers to the average time data packets spent to reach to the desired destinations. It includes delays caused by buffering throughout the route discovery process, processing in intermediate nodes, queuing at the interface queue, and retransmission delays

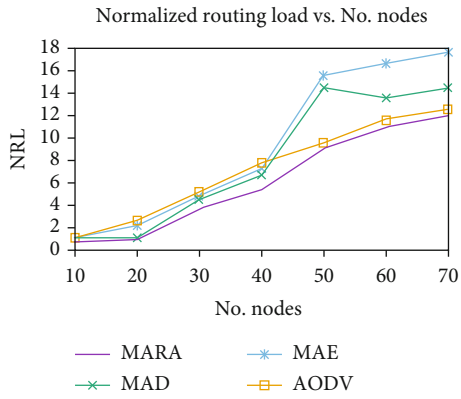


Figure 7: NRL vs. no. of nodes using RWP.

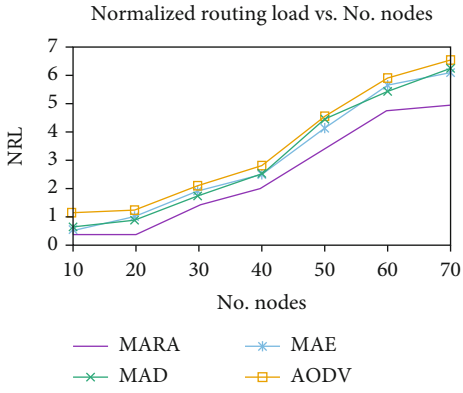


Figure 8: NRL vs. no. of nodes using RPGM.

(c) Throughput is defined as the total number of packets received by the destination in a given time period and typically measured in bits per second (bps)

(d) Normalized routing load (NRL) is defined as the number of routing packets transmitted per data packet delivered at the destination. When a packet sent over several hops, each transmission of the packet counts as one transmission

4. Simulation Results and Discussion

In this section, we present the simulation results and discussions of the proposed algorithm. The performance of proposed routing algorithm, MARA, is compared with AODV [5], MAD [21], and MAE [31]. We used the quantitative performance metrics used in [22, 23].

4.1. Impact of Node Density. In order to check the impact of node density, the number of nodes deployed has been changed over a 1000 m × 1000 m area by keeping all other parameters unchanged. Based on the mobility models RWP and RPGM, the packet delivery ratio of all protocols is decreasing as the number of nodes increases (Figures 3 and 4). Increasing the density causes link breakage and congestion which leads to packet drop. In order to check the impact of node density, the number of nodes deployed has been changed over a 1000 m × 1000 m area by keeping all other parameters

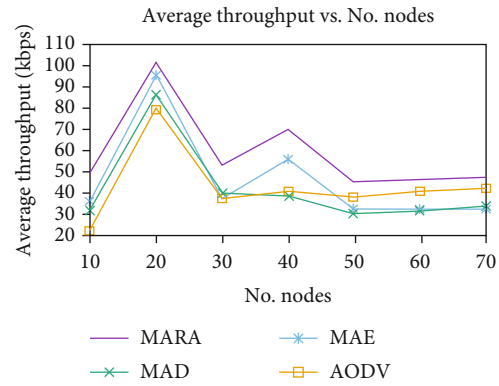


Figure 9: Throughput vs. no. of nodes using RWP.

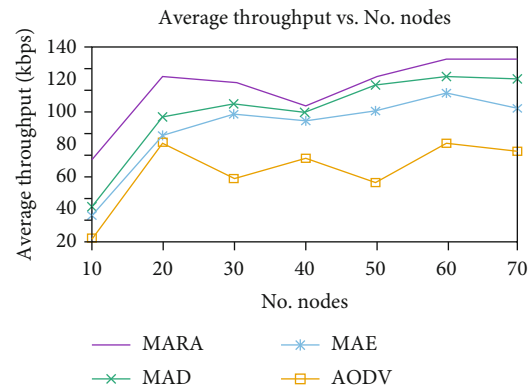


Figure 10: Throughput vs. no. of nodes using RPGM.

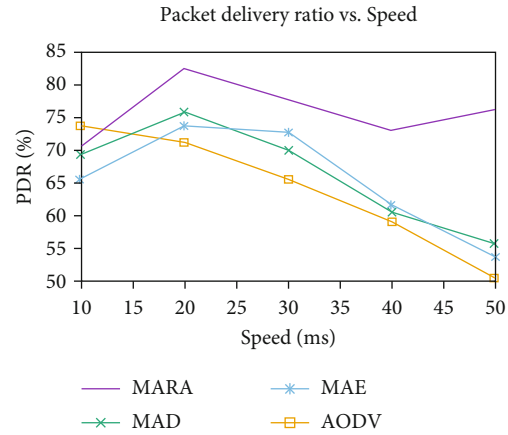


Figure 11: PDR vs. speed using RWP.

unchanged. The packet delivery ratio of all protocols is decreasing as the number of nodes increases. Increasing the density causes link breakage and congestion which leads to packet drop. The packet delivery ratio (PDR) of the proposed algorithm, MARA, is higher than AODV, MAD, and MAE. MARA reduces the chance of several link breakages, which reduces the amount of packet loss. It also decreases the number of rebroadcasting packets, which decreases the possibility of collision and contention. Since we considered speed, distance between communicating nodes, and low energy in RREQ and RREP phases, the established path between the

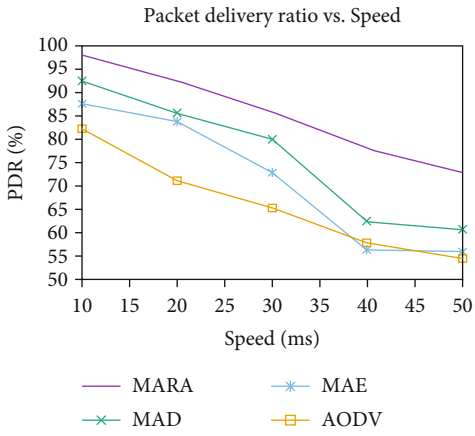


Figure 12: PDR vs. speed using RPGM.

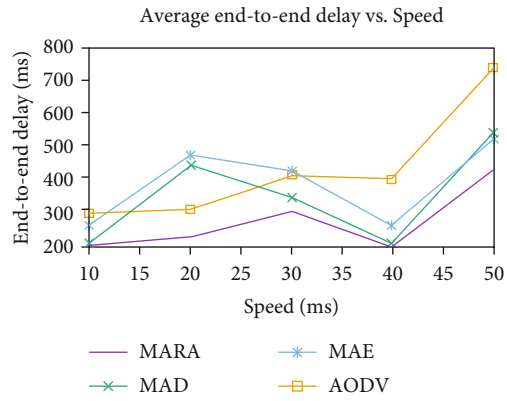


Figure 14: AETED vs. speed using RPGM.

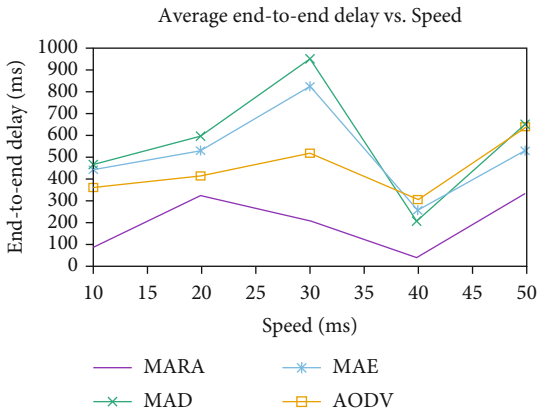


Figure 13: AETED vs. speed using RWP.

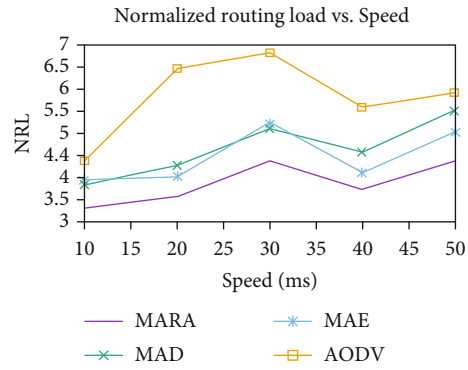


Figure 15: NRL vs. speed using RWP.

nodes has less chance of breakage. Increasing the density of nodes increases end-to-end delay (Figures 5 and 6). Increasing the network density increases the possibility of packet collision and contention. This leads to frequent rebroadcasting of packets and increases end-to-end delay. MARA achieved a significant improvement as compared to AODV, MAD, and MAE. MARA reduces the possibility of reinitiating RREQ packets.

Increasing the number of nodes increases the routing load gradually in Figures 7 and 8. The routing load of MARA is less than AODV, MAD, and MAE in both mobility models. Because of the reduction of redundant rebroadcast of the RREQ packet, there is less chance of packet collisions and less link breakage. Reducing the link breakage reduces the reinitiation of route discovery and maintenance process. The obtained result in Figures 9 and 10 describes the throughput of the proposed algorithm as a function of the maximum number of nodes. The results demonstrate that MARA scheme outperforms AODV, MAD, and MAE in all cases. MARA considers node mobility to select the best route, which has less probability of link break. However, the AODV protocol selects the shortest route and does not consider mobility. Selecting a route that may cause the frequent link breakage affects the overall throughput of the network.

4.2. Impact of Node Mobility. In order to evaluate and compare, the performance of MARA, AODV, MAD, and MAE simulations have been conducted by changing the speed of mobile nodes. 50 mobile nodes have been deployed on a 1000 m × 1000 m area. We have analyzed the impact of node mobility on the PDR as shown in Figures 11 and 12. Increasing the node speed decreases the PDR of all protocol. This is because a valid route begins to break as the speed of node increases that causes to initiate RREQ retransmitting, which leads to consume more bandwidth and increase rebroadcast. However, MARA outperforms AODV, MAD, and MAE in all cases. The packet delivery ratio in MARA has been improved due to rebroadcasting and chance of link break. As we have considered high speed, direction, and low energy of the mobile nodes, the established path has less chance of breakage. Therefore, fewer rebroadcast of the routing message causes smaller bandwidth consumption and reduces collisions and contentions, which affected the significance of the network.

Figures 13 and 14 results indicate the impact of node mobility on the performance of MARA, AODV, MAD, and MAE. MARA has achieved less end-to-end delay as compared to the other schemes in all cases. This is because the route discovery process of MARA is less than AODV, MAD, and MAE due to the link breakage possibility being less. Figures 15 and 16 show the impact of node mobility

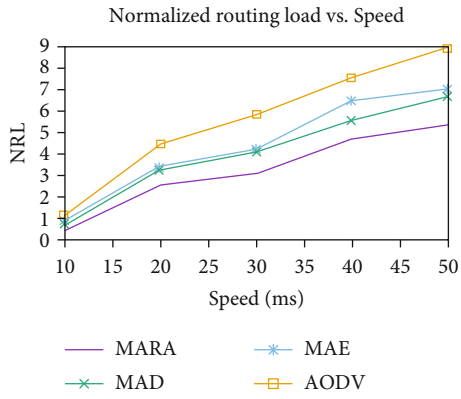


Figure 16: NRL vs. speed using RPGM.

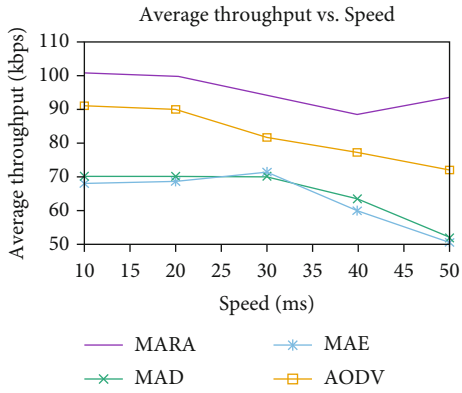


Figure 17: Throughput vs. speed using RWP.

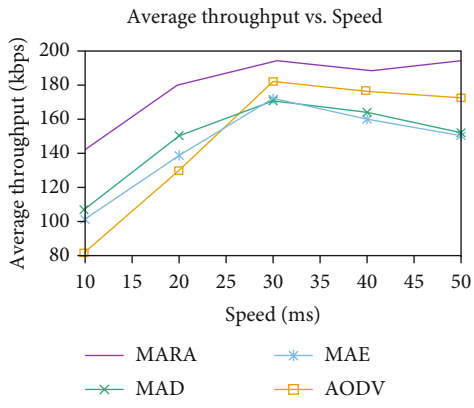


Figure 18: Throughput vs. speed using RPGM.

on the performance of MARA, AODV, MAD, and MAE protocol in terms of network routing load. The overhead has increased when we increased the speed of mobile nodes. The reason is the paths between the sources and destination node repeatedly breaks and reestablishes due to mobility. Moreover, the RREQ packets may not reach the desired destinations. Hence, this leads to reinitiate a route discovery process that eventually increases the network routing load or overhead. MARA scheme has achieved better performance in terms of routing load in all cases, because it reduces the unnecessary rebroadcast of the RREQ packet and selects the reliable route at the destination with reducing the redundant rebroadcast of the RREQ packet.

The obtained result in Figures 17 and 18 shows that increasing the speed of nodes decreases the throughput. Due to mobility, the valid route between the source and destination node breaks causes to reinitiate RREQ packet. This leads to more rebroadcast and greater bandwidth consumption. Therefore, the throughput decreased when the node speed is increased. The proposed algorithm reduces link breakage and number of rebroadcast RREQ packets that makes it to minimize bandwidth consumption and reduces the chance of collisions and contentions. Therefore, it has a higher throughput.

5. Conclusions and Future Work

MANET is a collection of mobile nodes that can dynamically change locations to form a network to exchange information. Mobility causes link breakage and increases the reinitiating of the route discovery process. In this study, we designed a MANET routing algorithm that overcomes the limitation of existing routing protocols. We combined multiple parameter metrics such as speed, direction, and residual energy of mobile nodes for decision-making in route discovery and route reply processes. Unlike the previous works, the proposed routing algorithm is based on node speed, direction, and residual energy to select more stable routes, among the intermediate nodes located in the path of the source and destination nodes. It has shown through extensive simulations that the proposed schemes in several operating conditions and scenarios. The proposed algorithm was tested and evaluated through simulation by varying node density and node speeds in different mobility models. MARA outperforms AODV, MAD, and MAE in terms of the widely used performance metrics: packet delivery ratio, average end-to-end delay, throughput, and normalized routing load.

In our future works, we will consider other decision-making techniques like wireless link quality and routing load. In addition, the network topology in MANETs is very dynamic due to the mobility of nodes. Hence, the wireless link is vulnerable to attacks. Therefore, it needs a secure solution to the dynamic behavior of the network.

Data Availability

The data used in this research is generated using NS2 and available upon requesting the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. K. Sarkar, T. G. Basavaraju, and C. Puttamadappa, *Ad Hoc Mobile Wireless Networks: Principles, Protocols, and Applications*, CRC Press, Second edition, 2013.

[2] C. S. Ram and M. B. S. Murthy, *Ad Hoc Wireless Networks: Architectures and Protocols*, Pearson education, India, 2004.

[3] S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, *Mobile Ad Hoc Networking: Cutting Edge Directions*, Wiley-IEEE Press, Second edition, 2013.

[4] B. Bhushan, S. Gupta, and C. K. Nagpal, "Comparison of on demand routing protocols," *International Journal of Information Technology and Computer Science*, vol. 5, no. 3, pp. 61–68, 2013.

[5] P. C. D. SR and E. M. Belding-Royer, "Ad hoc on-demand distance vector (AODV) routing," *Second IEEE Work. Mob. Comput. Syst. Appl.*, pp. 1–37, RFC Editor, United States, 2003.

[6] G. Singh, D. Saini, R. Rishi, and H. Rohil, "Role of link expiration time to make reliable link between the nodes in MANETs," *International Journal of Applied Engineering Research*, vol. 11, no. 7, pp. 5321–5325, 2016.

[7] Y. Tseng and Y. Chen, "The broadcast storm problem in a mobile ad hoc network," *Proceeding of ACM/IEEE MobiCom*, vol. 8, no. 2–3, pp. 153–167, 2002.

[8] B. Williams and T. Camp, "Comparison of broadcasting techniques for mobile ad hoc network," in *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing - MobiHoc '02*, pp. 194–205, Lausanne, Switzerland, 2002.

[9] V. Sharma and A. Vij, "Broadcasting methods in mobile adhoc networks," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 582–587, Greater Noida, India, May 2017.

[10] J. Sharma, V. Bhatia, and G. Kaur, "A review on broadcasting strategy for mobile ad-hoc wireless network," *International Journal of Advanced Computer Engineering*, vol. 7, no. 8, pp. 692–699, 2018.

[11] M. Bakhouya, "Broadcasting approaches for mobile ad hoc networks," in *2013 International Conference on High Performance Computing & Simulation (HPCS)*, pp. 705–707, Helsinki, Finland, July 2013.

[12] K. Shanmugam, K. Subburathinam, and A. V. Palanisamy, "A dynamic probabilistic based broadcasting scheme for MANETs," *Scientific World Journal*, vol. 2016, article 1832026, pp. 1–8, 2016.

[13] K. A. Darabkh, M. S. A. Judeh, H. Bany, and S. Althunibat, "Mobility aware and dual phase AODV protocol with adaptive hello messages over vehicular ad hoc networks," *AEU - International Journal of Electronics and Communications*, vol. 94, pp. 277–292, 2018.

[14] Y. Khamayseh, G. Obiedat, and M. B. Yassin, "Mobility and load aware routing protocol for ad hoc networks," *Journal of King Saud University-Computer and Information Sciences*, vol. 23, no. 2, pp. 105–113, 2011.

[15] P. Kamalakkannan, T. Salem, and R. S. Kumar, "Personalized RAODV algorithm for reduce link break in mobile ad hoc networks," in *2012 Fourth International Conference on Advanced Computing (ICoAC)*, pp. 1–6, Chennai, India, December 2012.

[16] T. A. Murshedi and X. Wang, "Mobility adaptive ad-hoc on demand distance vector routing protocol in MANET," *International Journal of Future Generation Communication and Networking*, vol. 8, no. 6, pp. 71–82, 2015.

[17] W. A. Jabbar, M. Ismail, and R. Nordin, "Energy and mobility conscious multipath routing scheme for route stability and load balancing in MANETs," *Simulation Modelling Practice and Theory*, vol. 77, pp. 245–271, 2017.

[18] W. A. Jabbar, M. Ismail, R. Nordin, and R. M. Ramli, "EMA-MPR: energy and mobility-aware multi-point relay selection mechanism for multipath OLSRv2," in *2017 IEEE 13th Malaysia International Conference on Communications (MICC)*, pp. 1–6, Johor Bahru, Malaysia, November 2017.

[19] W. A. Jabbar, W. K. Saad, and M. Ismail, "MEQSA-OLSRv2: a multicriteria-based hybrid multipath protocol for energy-efficient and QoS-aware data routing in MANET-WSN convergence scenarios of IoT," *IEEE Access*, vol. 6, pp. 76546–76572, 2018.

[20] W. A. Jabbar, "Mobility-based performance comparison of MBQA-OLSRv2 and MBMA-OLSRv2 routing protocols," in *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, pp. 281–286, Phuket, Thailand, October-November 2019.

[21] A. Swidan, H. B. Abdelghany, and R. Saifan, "Mobility and direction aware ad-hoc on demand distance vector routing protocol," *Procedia Computer Science*, vol. 94, pp. 49–56, 2016.

[22] T. K. Vu and S. Kwon, "On-demand routing algorithm with mobility prediction in the mobile ad-hoc networks," Sch. Electr. Eng. Univ. Ulsan, 2016.

[23] T. Larson and N. Hedman, *Routing protocols in wireless ad-hoc networks - a simulation study*, Master's thesis Luleå Univ. Technol., 1998.

[24] T. D. Nguyen, J. Y. Khan, and D. T. Ngo, "A distributed energy-harvesting-aware routing algorithm for heterogeneous IoT networks," *IEEE Transactions on Green Communications and Networking*, vol. 24, no. 14, pp. 1–10, 2018.

[25] K. Fall and E. K. Varadhan, *The ns Manual (formerly ns Notes and Documentation)*, A Collab. between Res. UC Berkeley, LBL, USC/ISI, Xerox PARC, no. 3, 2011.

[26] A. Sharma, Gurpreet, and J. Singh, "Mobility models for MANETs : mathematical perspective," *International Journal of Advanced Research in Engineering and Applied Sciences*, vol. 2, no. 5, pp. 59–68, 2013.

[27] X. Hong, M. Gerla, G. Pei, and C. C. Chiang, "A group mobility model for ad hoc wireless networks," in *Proceedings of the 2nd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems - MSWiM '99*, pp. 53–60, 1999.

[28] M. Khan, M. F. Majeed, M. F. Majeed, and J. Lloret, "The impact of mobility speed over varying radio propagation models using routing protocol in MANET," in *Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). AI2SD 2019. Lecture Notes in Networks and Systems, vol 92*, M. Ezziyyani, Ed., pp. 277–288, Springer, Cham, 2019.

[29] D. Bhatia and D. P. Sharma, "A comparative analysis of proactive , reactive and hybrid routing protocols over open source network simulator in mobile ad hoc network," *International Journal of Applied Engineering Research*, vol. 11, no. 6, pp. 3885–3896, 2016.

[30] A. Adlakha and V. Arora, "Performance evaluation of AODV and DSR routing protocols under constrained situation," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, no. 7, pp. 189–191, 2015.

[31] U. Rashid, O. Waqar, and A. K. Kiani, "Mobility and energy aware routing algorithm for mobile ad-hoc networks," in *2017 International Conference on Electrical Engineering (ICEE)*, pp. 1–5, Lahore, Pakistan, March 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5513349, 12 pages
https://doi.org/10.1155/2021/5513349

WILEY | Hindawi

Research Article

A Method of Optimizing Network Topology Structure Combining Viterbi Algorithm and Bayesian Algorithm

Xiaoxiao Shi

Hubei University of Technology, WuHan, Hubei 430068, China

Correspondence should be addressed to Xiaoxiao Shi; suixing3693@163.com

Received 16 February 2021; Revised 23 March 2021; Accepted 11 April 2021; Published 10 May 2021

Academic Editor: Rahim Khan

With Internet entering all walks of life, development of internet and usage expansion demand better performance, especially the application of 5G network that adopts NAS networking mode. Some of the network bandwidth cannot fully support the current network demand, which causes network fluctuations and other concerns. In this paper, a method for optimizing the topological structure of the bottom layer of the communication network is proposed that has outage performance close to optimal routing scheme. In specific, path in areas with poor network conditions is first optimized using Viterbi algorithm. Then, network element nodes on the path are optimized using Bayes recommendation algorithm for reasonable flow distribution. Dual planning of improved Viterbi algorithm is used to realize the main and standby path planning, and then, Bayesian recommendation algorithm based on the average value is used to optimize the network elements. Therefore, it is very efficient to realize overall performance optimization.

1. Introduction

Network topology refers to the physical layout of interconnecting various devices with transmission media, that is, how to connect computers and other devices in the network. The topology diagram shows the network configuration and mutual connection of network servers and workstations. Its structure mainly includes star structure, ring structure, bus structure, distributed structure, tree structure, mesh structure, and honeycomb structure. In the structural hierarchy, most networks adopt a three-layer network architecture: access layer, convergence layer, and core layer. The stable operation of the overall network is achieved through the rational distribution of the three-tier structure.

Based on the physical layout, there are also some logical structure designs. The logical structure design makes the data transmission more efficient and stable. There are some problems with the current logical network structure, such as instability, the design of the logical structure does not consider the whole, and less consideration for scalability. And due to the development of 5G networks, there are higher requirements on the current network conditions for 5G networks that adopt nonindependent networking (NSA) mode [1]. Since it is not suitable for redesigning the entire network topology in nonindependent networking, partial gradual optimization is required to make the entire network system more stable and efficient.

In the overall network structure, the topological structure of the communication network consists of two parts, namely, the network element node, which is actually a single transit base station, which receives, processes, and forwards all communication data. Among the forwarded network elements, the board is involved. With ports, there are multiple boards in the same network element, and there are multiple ports to choose from on the same board. There will be a data table inside the board to store and forward the corresponding port information. By changing the port in this way, we can change the link of business data. We can make the network condition better by changing the data link and make the usage rate of network elements reach a better level, so as to better respond to the needs of the network. The other component is optical fiber. The optical fiber connects the connection between network element nodes. As a bridge between network element nodes, there may be many direct paths between two nodes

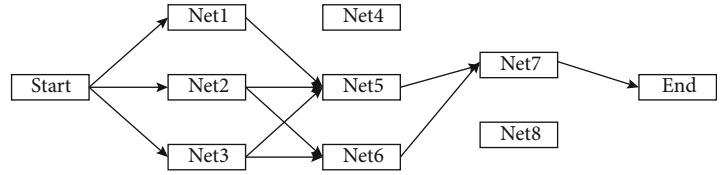


Figure 1: State transition expansion diagram.

to choose from. We need to follow multiple constraints. Choose the most suitable path from them to meet our network needs, thereby improving the overall network conditions. Only through the comprehensive optimization of the above two points can the overall optimization of the network topology be achieved, and our ever-growing network needs can be better. In the current development, research related to the optimization of network structure has been continuously advancing. From the structural design of the power grid to the structural design of the Internet, the realization of the bottom layer uses the research methods related to topological networks. Genetic algorithm is commonly used for the optimal routing scheme. This method can be applied to the optimization of multiple goals [2]. This method can be applied to the optimization of a single line. In the same situation, research on electric power, such as Jiang Hang's optimal structural design considering power loss and other factors, optimizes the parts in stages to improve the overall effect. For optimizing the communication delay in the network structure, the multivolume distributed subgradient optimization algorithm is used to transform the delay by expanding the dimension. Find the optimal communication situation through the adjacency matrix to achieve a better convergence effect, thereby reducing network delay [3]. It is also possible to generate a topology network algorithm through machine learning that includes a topology generation algorithm based on Lasso and its supplementary rules and the dynamic topology reconstruction algorithm by LSTM prediction [4].

The above algorithms have improved the network status to a certain extent, but in the application process, most operators still tend to use manual practical experience. This also leads to the fact that the overall network status has not been obtained. This research starts from practical applications and improves the performance of the network from three perspectives, namely, (1) starting from the situation where the network element is not restricted by the forwarding, the path planning method is proposed to the maximum; (2) individual optimization of network elements that are easily overlooked also ensures the forwarding efficiency of the network; and (3) the comprehensive consideration and the addition of manual evaluation indicators can better cope with some special scenarios.

2. Related Work

2.1. Path Planning Algorithm. The Viterbi algorithm is a very commonly used path planning algorithm. In the application process, it is mainly used in the hidden Markov model. In the model, the hidden state of the observation path is analyzed to find the maximum posterior probability of the Viterbi path [5]. In this way, the Markov chain of the optimal solution is found in the hidden area of the joint finite state and the observable process. In the Markov hypothesis, each state behind an observed phenomenon has a probability value, and the optimal solution is found in the process; we only need to find the state with the largest probability value is the optimal solution [6, 7]. It is precisely because this state is the result probability of the superposition of the previous states; it is the state reached after multiple observations so that the fence network can be used to reflect the model well. In an ideal fence network, I use a 3×3 network structure for research. Let T be the probability of different states in a stage, so there is an observation state of $T = (t1, t2, t3)$, and the hidden state is defined as $R = (r1, r2, r3)$; through this situation, we can have the solution formula (1):

$$(r1, r2, r3) = \arg\max P(r1, r2, r3 \mid t1, t2, t3). \quad (1)$$

The equivalent formula (2) can be obtained by extension:

$$\operatorname{argmax} \prod_{i=1}^{N} P(r_t \mid r_i) \cdot P(r_i \mid r_{i-1}). \quad (2)$$

The state analysis of the model is expanded as shown in Figure 1:

In the conventional calculation process, the number of combinations calculated according to this 3×3 model is 9, that is, 3N combinations. In the case of multiple rounds of superimposed combinations, this method requires more calculations. When using the Viterbi algorithm to solve the model, its time complexity is proportional to the length of the stage, and the complexity is $O(N * D\text{^}2)$, where N is the length of the stage and D is the width. For the application of the Viterbi algorithm, the advantages can be explained in the relevant literature: (1) If the path with the highest probability passes through a certain point of the fence network, the subpath from the starting point to this point must also be the path with the highest probability from the beginning to the point. (2) Assuming that there are k states at the i-th moment, there are k shortest paths from the beginning to the k states at time i, and the final shortest path must pass through one of them. (3) According to the above properties, when calculating the shortest path of the $i + 1$th state, we only need to consider the shortest path from the beginning to the current k state values and the shortest path from the current state value to the $i + 1$th state value, such as finding the shortest path when $t = 3$ is equal to finding the shortest path of all state nodes when $t = 2$ plus the shortest path of each node from $t = 2$ to $t = 3$.

In order to verify this property, we record the intermediate variable of the shortest path at the t stage as formula (3):

$$\delta_t(i) = \max P(i_t = i, i_{t-1}, \cdots, i_1, o_t, \cdots, o_1 \mid \lambda), i = 1, 2, \cdots, N. \tag{3}$$

Among them, i_t is the maximum probability path at stage t, o_t is the observed value, and λ is the parameter. This formula represents the intermediate variable of a certain state, and the recursive formula (4) can be calculated by further reasoning:

$$\delta_{t+1}(i) = \max \left[\delta_t(j)a_{ji}\right] b_i(o_{t+1}). \tag{4}$$

It represents the maximum probability of the observation of o_{t+1} from the transition to state i at the $t + 1$ stage, so we define o_{t+1} as the target point, which is the E node. In this process, the path nodes traversed can be expressed by formula (5):

$$\psi_t(i) = \underset{1 \le j \le N}{\operatorname{argmax}} \left[\delta_{t-1}(j)a_{ji}\right]. \tag{5}$$

Among them, $\psi_t(i)$ is the node through which the path with the maximum probability of state i passes when stage t . Through this formula, we can record all the nodes passed in the path of maximum probability, thereby saving the path.

The application of genetic algorithm to multitarget path search is one of the more commonly used algorithms. The main implementation process of the algorithm is divided into four steps, including initialization, selection, crossover, and mutation. In terms of implementation, given m selection targets, when applied to network path search, all possible paths need to be traversed. The time complexity is $O(N * D\text{^}2)$. According to the experiment, first choice, initial population, random choose a path with a certain node, form a chromosome set from node S to node E, select the fitness function according to different nodes, and perform genetic operations. The fitness function is based on formula (6):

$$F = \sum_{i=1}^{n} f_i(l_i). \tag{6}$$

Among them, $f_i(l_i)$ represents the i-th objective fitness function, and the optimal selection path of the i-th objective depends on the fitness function of l_i and the corresponding constraint information, so the value of f_i depends on l_i. In the result analysis, the lower the value of the fitness function F, the better the path. The specific experimental steps are as follows. First, for the m sets of data, we retain some excellent chromosomes and then perform cross-mutation according to the fitness functions of the respective nodes and perform next-generation operations on the new populations generated into new components and test the effectiveness of chromosomes. Finally, the solutions are sorted by the adaptability function to find the optimal path solution.

2.2. System Recommendation Algorithm. Regarding system recommendation, the recommendation system is more and more widely used at this stage, most of which are used to recommend content related to personal preferences. For users, recommendation is a specific type of intelligent information that determines the user's preferences by analyzing the user's behavioral data history and then recommends content that may be of interest to the user [8]. The main recommendation systems can be divided into two categories, content-based recommendation methods and collaborative filtering-based recommendation methods. Collaborative filtering recommendations mainly depend on users expressing their personality through ratings of movies, music, etc., which are more widely used. In this kind of recommendation, the result of recommendation is generally determined by displayed feedback and implicit feedback. The data set is formed by classifying the ratings of users and items. This type of data is displayed feedback, which is easier to collect, and its application range is also very limited [9]. Another kind of feedback is achieved through the user's preferences for items. This type of feedback is implicit feedback. In this process, implicit feedback is only negative and positive, so analysis is more difficult. In general, for implicit feedback, we use a pairwise ranking method for ranking. In the pairwise comparison, we can know the preference order relationship between different items. In the application of pairwise ranking, the optimal ranking algorithm is a method of generating personalized recommendations based on implicit feedback data, which is Bayesian personalized recommendation (BPR) [10]. In this recommendation method, it mainly treats all unobserved feedback as negative feedback, treats all observed items as positive examples, and assumes that the user's preferences are independent; it assumes that users prefer positive feedback items, not items without positive feedback. BPR groups the data according to the association and then ranks the group's data into positive or negative effects. Its purpose is to maximize the distance between observed and unobserved interactions. Formally, the objective function of BPR is formula (7):

$$L_{\text{BPR}} = \sum_{(u,i,j) \in D} - \ln \sigma\left(\widehat{x}_{u,i} - \widehat{x}_{u,j}\right) + \lambda \|\Theta\|^2. \tag{7}$$

Among them, $\widehat{x}_{u,i}$ represents u's preference for item i in a set of data, $\sigma(x)$ is a sigmoid function, λ is a model-specific regularization parameter to prevent overfitting, and Θ is a collection of multiple training pairs in the formula. Thus, the recommendation list result is obtained according to $\widehat{x}_{u,i}$. However, in the rating or ranking, the projects in which users do not participate account for the vast majority of projects. We can find that it is flawed in the application process. In order to make this research more effective in optimizing the network, we introduce a Mean Bayesian Personalized Ranking Algorithm (MBPR), which converts the displayed feedback data into implicit feedback data.

For the data processing of the MBPR method, the method defines the observed interactions above the average score as positive feedback and marks it as +1, and the observed

interactions below the average score as negative feedback and marks it as 0, mark the unobserved interaction as a random value between (0,1). We give a parameter matrix formula about W and H according to Bayesian posterior and Bayesian formula. W and H decompose all item rating matrices into preference matrix W and item rating matrix H.

$$p(\Theta \mid >_u) \propto p(>_u \mid \Theta)p(\Theta). \tag{8}$$

Among them, $>_u$ represents the total sorting scheme of the items selected by the user. We assume that the item selection scheme of each project is completely independent, and we can get formula (9):

$$\prod_{u \in U} p(>_u \Theta) = \prod_{(u,i,j) \in D_S} p(i >_u j \mid \Theta). \tag{9}$$

Among them, i and j, respectively, represent the two option values that are compared pair by pair in the item selection, that is, the comparison result of i and j when simulating the user's selection of the item. From this, we can see that the order of j is higher than the order of i (10):

$$p(i >_u j \mid \Theta) = \sigma\left(\widehat{x}_{u,i,j}(\Theta)\right). \tag{10}$$

Among them,

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \tag{11}$$

$\sigma(x)$ is the sigmoid function. We use this formula as the loss function, and in this way, the item selection of different network elements is mapped to hidden factors, and recommendations are formed through these factors. In this process, we also need to determine the maximum posterior probability of the parameters W and H. In citing the MBPR method, its objective function formula (12):

$$L_{\mathrm{MBPR}} = \ln p(\Theta \mid >_u) \propto \ln p(>_u \mid \Theta)p(\Theta). \tag{12}$$

As a result, we began to use the method to apply the data and train the model to determine the W and H parameters. Use the SGD method in the literature to learn parameters, according to formula (13):

$$\frac{\partial L_{\mathrm{MBPR}}}{\partial \Theta} = \sum_{(u,i,j) \in D_S} \frac{\partial}{\partial \Theta} \ln \sigma\left(\widehat{x}_{u,i,j}\right) + \lambda \frac{\partial}{\partial \Theta} \|\Theta\|^2 \propto \sum_{(u,i,j) \in D_S} \frac{-e^{\widehat{x}_{u,i,j}}}{1 + e^{\widehat{x}_{u,i,j}}} \times \frac{\partial}{\partial \Theta} \widehat{x}_{u,i,j} + \lambda \Theta, \tag{13}$$

where

$$\frac{\partial}{\partial \Theta} \widehat{x}_{u,i,j} = \begin{cases} h_{i,f} - h_{j,f} & \text{if } \Theta = w_{u,f}, \\ w_{u,f} & \text{if } \Theta = h_{i,f}, \\ -w_{u,f} & \text{f } \Theta = h_{j,f}, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

After constant iteration, the W and H parameters become more accurate, and the fit is better. After the parameters are determined, we have established a complete MBPR recommendation model.

3. Application

3.1. Path Planning Realization. This research is based on improving the overall structure of the network topology to meet the growing demand. The specific implementation logic diagram is shown in Figure 2.

In this method, the optimization of the overall network structure is mainly achieved by optimizing routes and network element nodes, and the optimization of the network is evaluated through evaluation tools. Applying the Viterbi algorithm to the optimization of the network topology, we select the network conditions of a certain city for experiments and perform local optimization for the lines that reach the critical value.

In terms of path optimization, the network conditions are optimized mainly from the following starting points:

(1) The main and standby routing rates are the same. In order to ensure a more stable service path, the main and backup paths of the service path are optimized to make them more highly available

(2) The huge ring node ratio ensures that the service path is a shorter path and avoids the increase in data transmission delay

(3) Link CIR bandwidth occupancy rate, the occupied bandwidth of a single line traffic should not be too large

(4) Through the optimization of the above four indicators, from the network element service carrying rate, due to the limited number of services carried by a single network element node, in order to ensure the safety of the network element node, make it in a relatively suitable utilization rate range

Through the optimization of the above four indicators, the system structure of the entire topology network is improved, so as to better meet the needs of future network development for the network structure. The demand analysis shows that the research has important significance and practical value.

When planning the path, use the model of the simulated fence network to establish a similar network connection diagram. The network element node is the junction point of the fence network, and the connection between the network

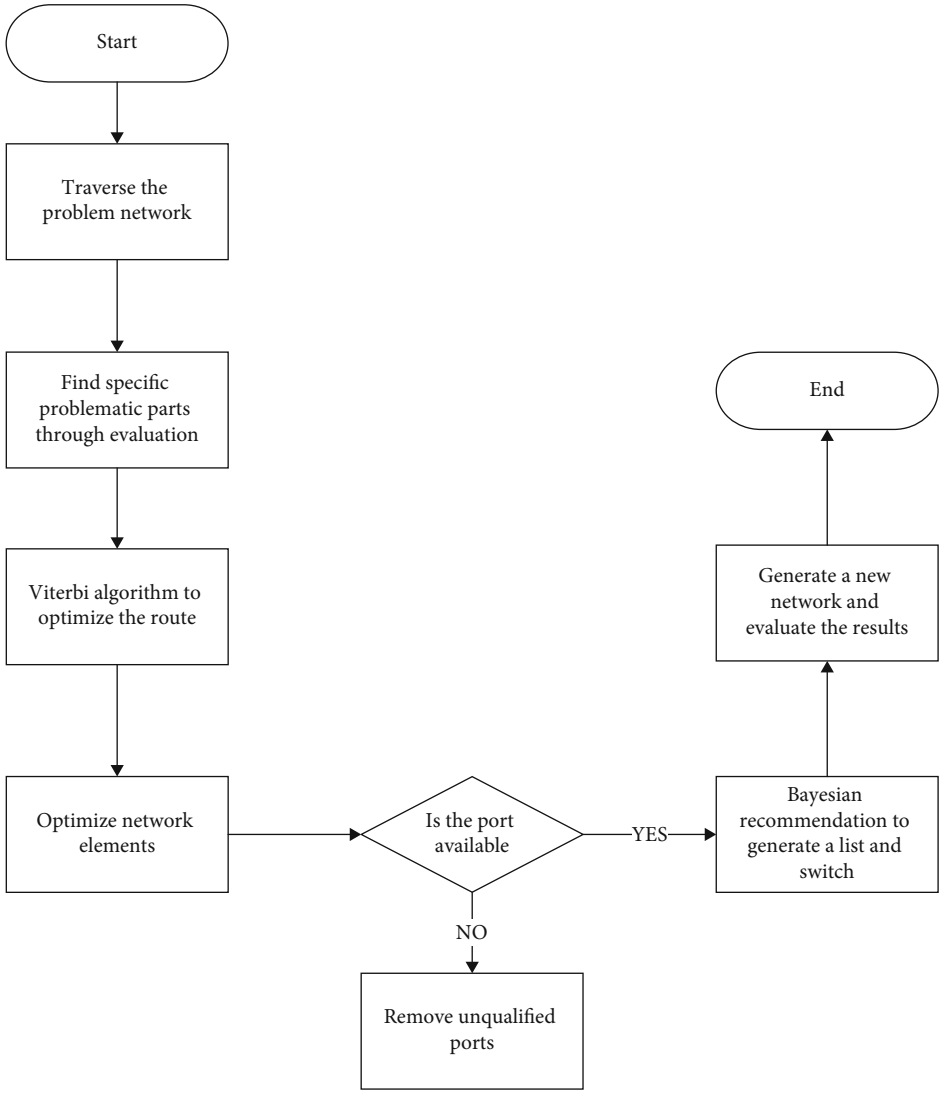


Figure 2: Implementation plan.

element and the network element is routed according to the connection of the optical fiber.

The second thing that needs to be considered is the evaluation tool based on network conditions. The main considerations for a single line are the problem of active and standby routes, network element service carrying rate, giant ring node ratio, and CIR broadband occupancy rate. For all reference factors and other issues, the route is dynamically scored in the form of weighting and penalty [11]. For the problem of active and standby routes, if the primary path and the backup path overlap, the network element node will be punished, and the rated traffic of the network element should not be exceeded. No processing is performed under 80% of the cases, and the load exceeding 80% of the rated capacity will be punished. Juhuan's judgment is that the number of network elements on a single line is calculated. Therefore, with the number of network elements node on a single line calculated, the more network elements node, the lower the score. The CIR broadband occupancy rate mainly considers the service data carried by a single fiber. The higher the occupancy rate, the lower the score. Establish a path evaluation model through this form and then select the optimal path through the Viterbi algorithm.

Based on the application of the Viterbi algorithm in the network research system, the nodes starting from a single convergent network element are not balanced in the same hidden state at the same stage in the path process, and some special nodes cannot be used too much. For optical fibers and some other unusable network elements, we adopt a negative feedback mode to make the selected path more reasonable. Through the above weighted parameters and related algorithms, we have established a path planning model.

In the optimization applied to the topological network structure, firstly, the network data is preprocessed, and the optical fibers of the entire network structure are first traversed to establish a complete regional topological network diagram. As the area may be large, we select some parts for display, as shown in Figure 3:

In a certain area, the approximate network element distribution is shown in the above figure. According to the

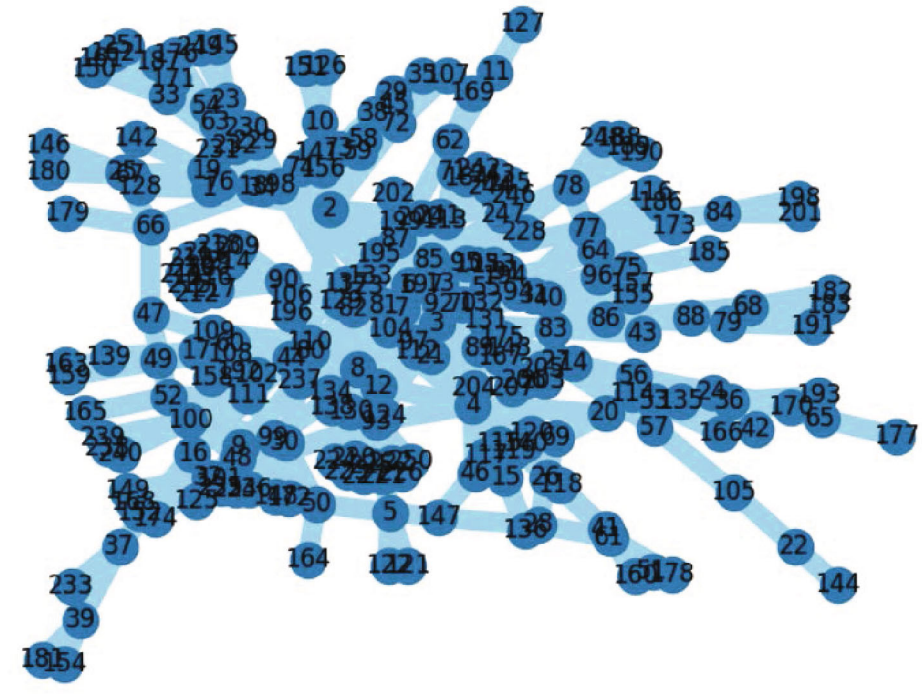


Figure 3: Partial visualization of topological network.

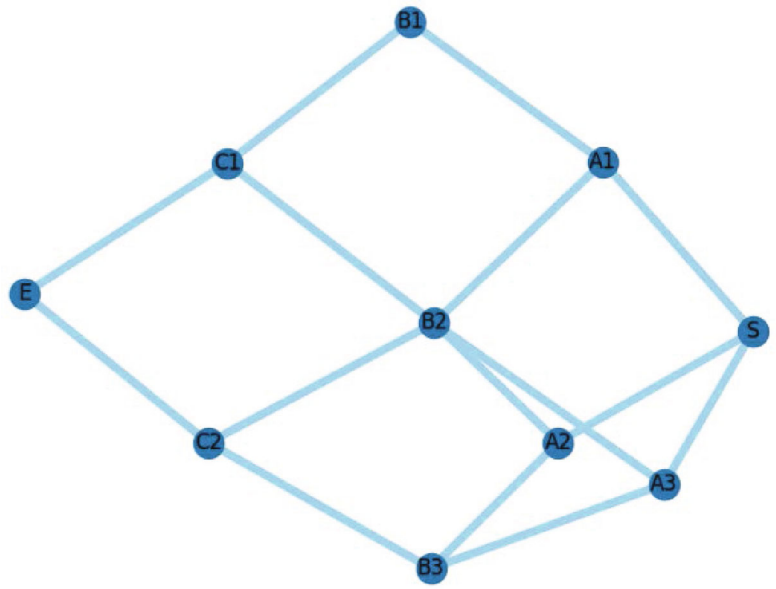


Figure 4: Typical structure diagram.

evaluation report of the entire network, we find out the network elements that need to be improved. We select the more typical cases to analyze and the application of the model. In this study, we mainly optimize the weak parts of the entire topology network. There may be many different situations, but because in the topology network architecture, we mainly optimize the connection between the convergent network elements, so we choose a more typical network structure as an example for experimental verification.

The actual circuit diagram model related to the construction of optical fiber is shown in Figure 4:

In this scenario, the two network elements S and E are the starting and ending network element nodes, and the other network element nodes are nodes on the path. Since only a part of the entire topological network is selected for analysis, the network needs to be initialized. Each node in the topological network may exist in other paths. In order not to damage other local network conditions, we need to the initial status of

a network element is initialized and weighted, and the score is weighted according to the current business volume and CIR occupancy of the network element. For special dedicated line network element nodes, they are removed.

In the application process, the main path is mainly maintained, and the backup path is an optional item. Therefore, the backup path that is not interfered by the main path is selected under the condition that the main path is optimal. In the course of the experiment, we need to consider the three characteristics of network element service carrying rate, giant ring node ratio, and CIR broadband occupancy rate to select the optimal path. In order to ensure the reliability of the experiment, genetic algorithm is introduced for experiment comparison [12]. The excellent primary path is selected, and the backup path is sorted in the same way to select the better path, thereby determining the primary path and the backup path.

For using the Viterbi algorithm, find the optimal path as the main path [13–17]. In the process of finding the main path, that is, starting from the network element of the sink node S, find the optimal path to the network element of the sink node E and find an optimal solution for the main and backup paths. Before optimizing, we start from node S, and there are multiple nodes in the node connecting S. We use the number of nodes to reach S node as the intermediate state. According to the formula, the node extending from node S will reach one at each stage; there will be a corresponding optimal solution with the maximum probability, and the optimal path from S to E is found by the continuous recursion of the model.

After the optimal path is generated, the corresponding path will be temporarily recorded, and the backup path will be searched according to the record, and the network element will be initialized again based on the main path search. In order to avoid the same network element line as much as possible, for the selected network element, the value is passed into the model in the form of negative feedback, and then, the alternate path selection is performed, and the optimal path selection is performed in a manner similar to the main path selection. When the temporarily recorded path appears, the corresponding penalty, thereby reducing the score, has reached the final priority path.

After achieving path optimization, we also need to plan specific network element nodes to ensure the normal and efficient operation of the nodes to improve overall performance. In this process, the most important thing is the selection of the access and output ports of the network element. After judgment, this is more suitable for the recommended algorithm.

3.2. Network Element Optimization. In this research, network elements undergo a reasonable distribution of transaction for a balanced traffic to prevent single node downtime. Two Bayes recommendation algorithms are introduced for experimental comparison, and intelligent decision-making is performed to allocate transaction. In practice, feedback data in Bayes recommendation algorithm can be mainly divided into three categories: At the same time, two Bayesian recommendation algorithms are introduced for experimental comparison, and intelligent decisions are made to allocate some business conversions. In the application process, in the Bayesian recommendation algorithm, the feedback data can be mainly divided into three categories: (1) the rating of the local network by the network status scoring tool, (2) resource usage in specific network elements, and (3) the topological network engineer's rating of the port allocation and operation of a single network element.

For using BPR, the unobservable data is negative feedback, and the observable data is positive feedback. In this case, the second type is explicit data, which is positive feedback, and the first and third types are implicit data, which are understood as negative feedback [17, 18]. Because the three types of data are more closely related, they are more suitable for data feedback using the converted MBPR method. The data processing of this method is well adapted to this application [19]. In this application, it is mainly to recommend an ideal port selection plan from a network element. We can assume that the selected port is an item. The user makes a decision on the port selection through three types of data feedback. Therefore, our training data screened based on the best performing data among the three types of feedback and conducted training with manual selection of the results. After screening, we obtained the selection of ports and boards among 1,000 network elements. The status of the network elements as well as the information related to the boards and ports is graded and put into the training model, respectively.

After establishing the model, we began to conduct application experiments. According to the application analysis of using Bayesian recommendations in network element optimization, we used the graph process to conduct experiments, as shown in Figure 5.

First, we grab the relevant data of the network element from the running network of a certain city and then preprocess the data and collect the relevant data according to the way of dividing the data into three categories. For the first type of data, we use a general network condition assessment tool. Because it is an overall scoring system, only parts can be rated. Then, we need to rate a single network element node, so we assume that the local rating represents the average rating of all network elements in the area. To ensure the integrity of the data, we set the local area rating as the rating of all the network elements in the area for evaluation. For the second type of data, since there are many indicators related to the status of the network element, it mainly includes port resource utilization, the same board rate of the active and standby network elements, traffic balance, link bandwidth occupancy rate, network element service load rate, network metarate, etc.; for the above indicators, we use the form of clustering to reduce the dimensionality of multiple indicator data. In the process of dimensionality reduction, it is found that three of the indicators account for a relatively high proportion. These are also important indicators that we need to solve.

(1) *Port Resource Utilization.* For the current network element, there are multiple ports, and the port selection of the service should be allocated reasonably to make the network more secure

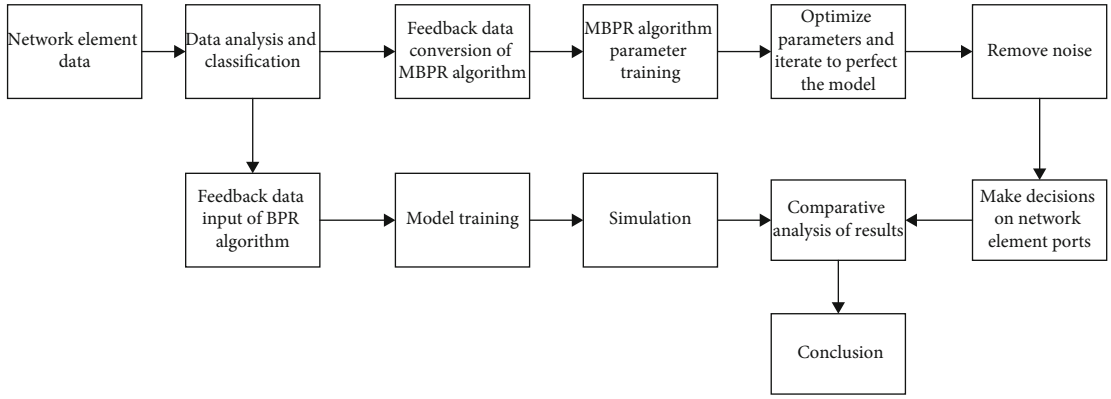


Figure 5: System recommendation algorithm flow design.

(2) The main and backup network elements are on the same board. In the case that the main and backup paths cannot be the same path and when the two paths pass through the same network element at the same time, the connection cannot appear on the same board. Avoid data confusion and instability due to accidents

(3) *Flow Balance Degree*. For each board in the network element, it is necessary to achieve a balanced flow state as much as possible. If the local flow is too large, it will also cause the failure of the network element node

Therefore, we analyze the need to pay attention to three indicators in the network element.

In practical applications, ports in the same board can forward related services within the same board according to the built-in data table. We parameterize the preprocessed data to form the second type of data parameters. For the third type of data, he is mainly based on the evaluation of the relevant network elements by engineers and experts. These evaluations involve the performance of the network elements in the actual application process. In this evaluation system, the network elements are mainly divided into two. The categories are qualified and to be improved. Through the comprehensive statistics of the three types of data, we have formed three types of feedback information on the source of the network element evaluation information. Then, analyze the feedback information according to Bayesian' basic formula (15):

$$P(B_i \mid A) = \frac{P(B_i)P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j)P(A \mid B_j)}. \tag{15}$$

For the application of the BPR algorithm, the data is grouped based on maximizing the separation of observed feedback and unobserved feedback, the second type of data is added to the model training as positive feedback, and the first and third types of data are added to the model as negative feedback training.

Thanks to the independence of preference behaviors between items, the partial order of different ports of the same item is independent of each other. We treat each piece of data as an independent part to improve the effectiveness of the feedback data. Put the processed sample data into the model for training and follow the code design drawing for training.

According to the objective function of BPR, after gradient descent, we calculate the corresponding loss function and adjust the corresponding weights of positive feedback and negative feedback in the iterative process, and when the loss function is minimized, we input test data and get the corresponding result.

For the application of the MBPR algorithm, according to the maximum posterior of the Bayesian algorithm, it is mainly to determine the matrix of W and H parameters. We map the second type of data to the factors of the port by means of mapping, and then according to the posterior, analyze the first and third types of data. Through the training of sample data, iterate continuously according to the formula. In the iterative process, the parameters are initialized many times to avoid the phenomenon of parameter overfitting. We introduce parameter iteration formula (16); after constant initialization of parameters, the parameters are better.

$$\Theta \longleftarrow \Theta + \eta\left(\frac{-e^{\hat{x}_{u,i,j}}}{1+e^{\hat{x}_{u,i,j}}} \times \frac{\partial}{\partial\Theta}\hat{x}_{u,i,j} - \lambda \times \Theta\right). \tag{16}$$

And to avoid some disturbance factors, in this process, the noise is continuously filtered out. The main source and port value are partially exceeded. When the $\sigma(x)$ loss function is the smallest, we obtain the corresponding parameters and then use the test data for correlation experiment of.

4. Analysis of Experimental Results

4.1. Path Planning Experiment Analysis. In order to verify the experimental results, we select a local network for route planning experiments, use genetic algorithms for experimental comparison, analyze the calculation time, and compare the coincidence rate of the path. The result is shown in Table 1:

Through the comparison of the two experimental methods, we found that the sample data shows that the coincidence node rate of the primary and backup paths using the genetic algorithm is significantly higher than that of the improved Viterbi algorithm. In order to ensure the stability

Input Training samples, number of iterations n, regularization parameters
Output model parameters
Initialization parameters
While i<n do
according to $P(B_i \mid A) = P(B_i)P(A \mid B_i)/\sum_{j=1}^{n} P(B_j)P(A \mid B_j)$ Maximum a posteriori P(W,H|>u)P(W,H|>u) to solve the model parameters W, H, use the sigmoid function to set the threshold
End
Return Recommended list

Algorithm 1: BPR designing process.

Table 1: Algorithm performance comparison.

Algorithm	Calculation time	AVG network element overlap
GA	0.251	5
Viterbi	0.011	2

Table 2: Comparison of experimental data generated by BPR and MBPR.

Method	Data	Port	Sparsity	Levels
BPR	1000	26241	85.40%	3-5
MBPR	1000	26241	92.66%	4-5

of the network and avoid the network failure due to a single line problem, therefore, it is better to use a scheme with a low number of network elements, so the Viterbi algorithm is better. In terms of calculation time, the Viterbi algorithm effectively reduces the time complexity due to its core idea of dynamic programming. Therefore, the Viterbi algorithm performs better in large-scale application scenarios. Comprehensive comparison Viterbi is more suitable for network path planning than genetic algorithm.

4.2. Network Element Optimization Experiment Analysis. In order to ensure the effectiveness of network element optimization, we select 1000 sets of data for experiments to compare the efficiency of the Bayesian recommendation algorithm. After comparing the experimental data, see Table 2.

Through the analysis of sample data, the results produced by the two schemes can be used for reference. In terms of the posterior nature of the joint evaluation tool scoring and the engineer's rating, the port score recommended by the MBPR method is better. During the experiment, MBPR has added some parameters to prevent data disturbance and has undergone denoising processing, so the time complexity is lower than that of BPR. When using the BPR method for recommendation, the method is mainly to obtain the recommendation list result according to $\widehat{x}\_(u, i)$. When the user does not participate in the project, the effect will not be ideal. In this application, more attention is paid to the first category compared with the posterior data of the third type of data; the weight of the second type of data is low, and it is cited in negative feedback, so there are some unsatisfactory conditions. And when its effectiveness is judged by the AUC index, its comparative effect is shown in Figure 6.

Through the analysis of the sorting accuracy in the sample data, we can make a selection according to different scenarios. In large-scale application scenarios and when computing resources are lacking, we can use the BPR method to improve efficiency. In order to pursue better performance, it is more recommended to use the MBPR method for recommendation ranking.

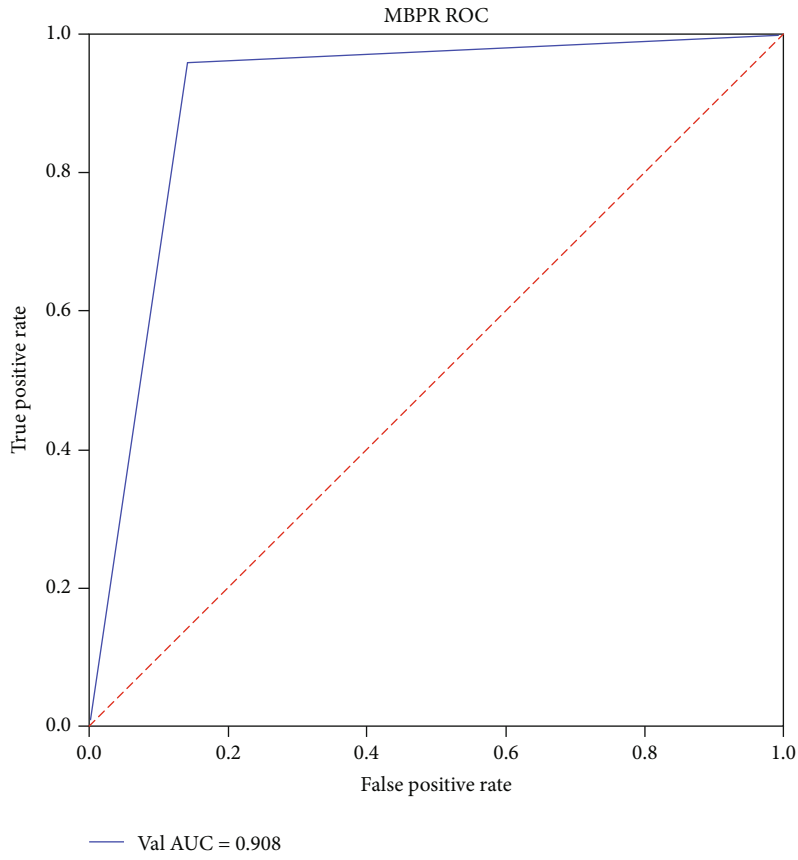
In evaluating the effect of the recommendation algorithm, we introduce a model evaluation program. When optimizing the port, it is important to establish a reliable recommendation list by selecting and recommending the board and port forwarded by the network. For the model training of the selected 1000 sets of data, each network element will have a corresponding initial value of the network element condition. Since the initial value of different network elements may have large differences, it is likely that serious defects will occur when the model is overfitted. Therefore, when collecting data, choose as large a difference as possible sample and guarantee the difference of the sample. In order to verify the effectiveness of the recommendation algorithm, we use a general model evaluation method to verify the port ranking through accuracy. This evaluation method has been widely used in the system to recommend evaluation indicators. Specific data refer to formula (17):

$$\text{Precision} = \frac{1}{\text{net}} \sum_{u=1}^{\text{net}} \frac{\text{utilizatation} * 0.4 - \text{samerout} * 0.4 + \text{flow} * 0.2}{\text{origin}}. \tag{17}$$

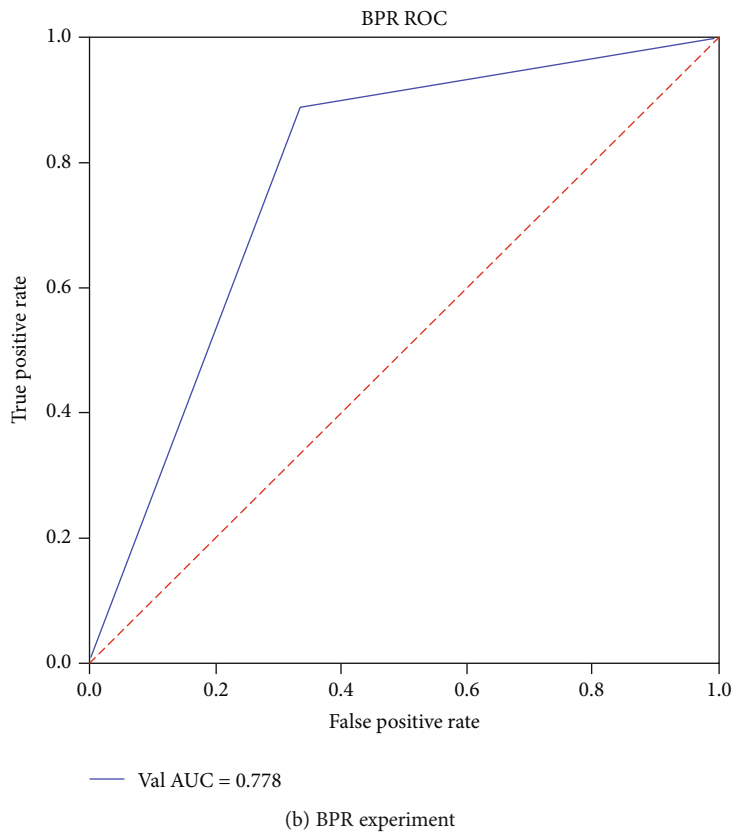
Among them, port_{new} represents the recommended port score, $\text{port}_{\text{origin}}$ represents the original port score, and origin represents the original data. The accuracy of the original data and the optimized data are compared to evaluate the effect of the model. The corresponding results of the two algorithms were verified by the prior data indicators, and the results are shown in Figure 7.

Through the change of the relevant parameter value, we get the change of accuracy as shown in Table 3.

During the experiment, in order to ensure that the data can reach the convergence effect, we uniformly set the number of iterations to 10000, and in the same 1000 samples, set 300 sets of data as training samples and 700 sets as test samples. After testing, the recommendation accuracy rate based on the MBPR method is higher, and then, different types of



(a) MBPR experiment



(b) BPR experiment

Figure 6: AUC curve.

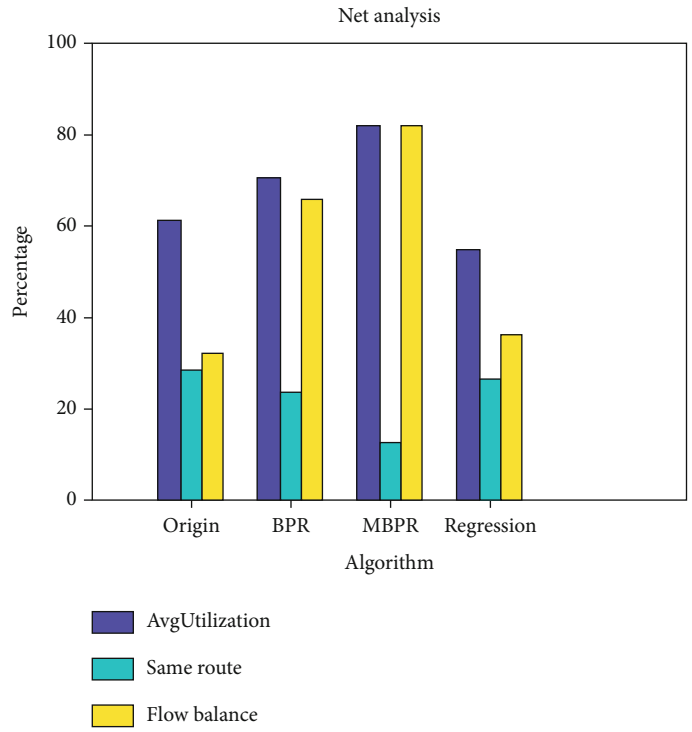


Figure 7: Comparison of changes in the three main indicators of optimization.

Table 3: Model evaluation parameters.

Algorithm	Precision
BPR	76.20%
MBPR	85.03%
Regression	62.66%

feedback data are analyzed, and the prior data is compared with the posterior data. In the baseline comparison result of the overall data, the prior data weight is higher. In the posterior feedback data, the impact coefficient of the data set based on the evaluation tool is higher than that of the manually rated data. The result obtained by using the regression model and the prediction method is very small, and there is a decline in some indicators. After studying the literature related to this article, add relevant theoretical and experimental analysis [20–24] .Comprehensive consideration, the use of MBPR performs better.

5. Conclusion

In this paper, we propose a near optimal routing scheme between two network elements by different weighting forms and dual Viterbi algorithm. Recommended value was selected, and optimization results were obtained by average Bayes algorithm. A set of solutions that meet requirement of efficient and balanced network elements were achieved. It is shown that the proposed routing scheme achieves a significant performance improvement in despite of accuracy of the algorithm that still needs to be improved. The proposed scheme can be implemented with low computational cost and is efficient for large-scale topology networks.

Data Availability

The source of the basic data and the network data of a certain city is inconvenient to disclose due to confidentiality.

Conflicts of Interest

The author declares that he/she has no conflicts of interest.

Acknowledgments

This work is funded by the National Natural Science Foundation of China under Grant No. 61772180.

References

[1] L. I. Guo, Z. U. O. Xue-qi, and W. E. I. Yu, "The selection strategies of 5G NSA anchor," *Digital Technology & Application*, vol. 38, no. 12, pp. 26–28, 2020.

[2] X. C. TianLu and L. J. SunJianwei, "Topology optimization design of time TDMA inter-satellite link based on heuristic genetic algorithm," *Computer Measurement & Control*, vol. 28, no. 12, pp. 155–160+171, 2020.

[3] L.-Q. Liujun, "Distributed subgradient method for multi-agent optimization with communication delays," *Journal of Hefei University of Technology*, vol. 36, no. 5, pp. 559–565, 2013.

[4] L. Xiaoyu, *Research on Topology Generation and Reconfiguration Optimization of Distribution Network Based on Machine Learning*, Beijing University of Posts and Telecommunications, 2019.

[5] A. Caliebe, "Properties of the maximum a posteriori path estimator in hidden Markov models," *IEEE transactions on information theory*, vol. 52, no. 1, pp. 41–51, 2006.

[6] O. Cappé, E. Moulines, and T. Rydén, *Inference in hidden Markov models*, Springer, 2005.

[7] S. Derrode and W. Pieczynski, "Unsupervised data classification using pairwise Markov chains with automatic copulas selection," *Computational Statistics & Data Analysis*, vol. 63, pp. 81–98, 2013.

[8] I. T. Christou, E. Amolochitis, and Z.-H. Tan, "AMORE: design and implementation of a commercial-strength parallel hybrid movie recommendation engine," *Knowledge and Information Systems*, vol. 47, no. 3, pp. 671–696, 2015.

[9] W. Zhou, J. Li, Y. Zhou, and M. H. Memon, "Bayesian pairwise learning to rank via one-class collaborative filtering," *Neurocomputing*, vol. 367, pp. 176–187, 2019.

[10] J. Ding, G. Yu, X. He, F. Feng, Y. Li, and D. Jin, "Sampler design for bayesian personalized ranking by leveraging view data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 2, pp. 667–681, 2021.

[11] J. Jo, H.-G. Kim, I.-C. Park, B. C. Jung, and H. Yoo, "Modified viterbi scoring for HMM-based speech recognition," *Intelligent Automation & Soft Computing*, vol. 25, no. 2, pp. 351–358, 2019.

[12] H. B. Su, Y. L. Shi, and Z. Z. Hou, "Multiobjective and multipath optimization selection methods based on genetic algorithms," *Microelectronics & Computer*, vol. 23, no. 10, pp. 41–43, 2006.

[13] J. Lember and J. Sova, "Existence of infinite Viterbi path for pairwise Markov models," *Stochastic Processes and their Applications*, vol. 130, no. 3, pp. 1388–1425, 2020.

[14] J. Lember, K. Kuljus, and A. Koloydenko, "Theory of segmentation," in *Hidden Markov Models, Theory and Applications*, P. Dymarski, Ed., pp. 51–84, InTech, 2011.

[15] J. Hamilton, *Regime Switching Models, in: Macroeconometrics and Time Series Analysis*, Springer, 2010.

[16] K. Kuljus and J. Lember, "On the accuracy of the MAP inference in HMMs," *Methodology and Computing in Applied Probability*, vol. 18, no. 3, pp. 597–627, 2015.

[17] P. Lanchantin, J. Lapuyade-Lahorgue, and W. Pieczynski, "Unsupervised segmentation of randomly switching data hidden with non-Gaussian correlated noise," in *Signal Processing*, vol. 91, no. 2pp. 163–175, Elsevier, 2011.

[18] W. Pan, H. Zhong, C. Xu, and Z. Ming, "Adaptive Bayesian personalized ranking for heterogeneous implicit feedbacks," *Knowledge-Based Systems*, vol. 73, pp. 173–180, 2015.

[19] J. F. Wang and P. Han, "Adversarial training-based mean Bayesian personalized ranking for recommender system," *IEEE ACCESS*, vol. 8, pp. 7958–7968, 2020.

[20] Y. Zhai, D. Liu, C. Wu, and R. She, "A recommendation approach based on Bayesian networks for clone refactor," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1999–2012, 2020.

[21] C. Cai, H. Xu, J. Wan, B. Zhou, and X. Xie, "An attention-based friend recommendation model in social network," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 2475–2488, 2020.

[22] S. Bin, G. Sun, N. Cao et al., "Collaborative filtering recommendation algorithm based on multi-relationship social network," *Computers, Materials & Continua*, vol. 60, no. 2, pp. 659–674, 2019.

[23] W. Jiang, J. Chen, Y. Jiang et al., "A new time-aware collaborative filtering intelligent recommendation system," *Computers, Materials & Continua*, vol. 61, no. 2, pp. 849–859, 2019.

[24] Y. Shen, Y. Li, J. Sun et al., "Hashtag recommendation using LSTM networks with self-attention," *Computers, Materials & Continua*, vol. 61, no. 3, pp. 1261–1269, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5599907, 12 pages
https://doi.org/10.1155/2021/5599907

Review Article

Evaluation and Quality Assurance of Fog Computing-Based IoT for Health Monitoring System

QingQingChang[1], **Iftikhar Ahmad**[2], **Xiaoqun Liao**[3] and **Shah Nazir**[2]

[1]*School of Information Management, Shanghai Linxin University of Accounting and Finance, 995 Shangchuan Road, Pudong New District, Shanghai 201209, China*
[2]*Department of Computer Science, University of Swabi, Khyber Pakhtunkhwa, Pakistan*
[3]*Information and Network Center, Xi'an University of Science and DS Technology, Xi'an 710054, China*

Correspondence should be addressed to Xiaoqun Liao; liaoxiaoqun642@sina.com and Shah Nazir; snshahnzr@gmail.com

Received 25 January 2021; Revised 25 March 2021; Accepted 13 April 2021; Published 23 April 2021

Academic Editor: Ihsan Ali

Computation and data sensitivity are the metrics of the current Internet of Things (IoT). In cloud data centers, current analytics are often hosted and reported on suffering from high congestion, limited bandwidth, and security mechanisms. Various platforms are developed in the area of fog computing and thus implemented and assessed to run analytics on multiple devices, including IoT devices, in a distributed way. Fog computing advances the paradigm of cloud computing on the network edge, introducing a number of options and facilities. Fog computing enhances the processing, verdicts, and interventions to occur through IoT devices and spreads only the necessary details. The ideas of fog computing based on IoT in healthcare frameworks are exploited by shaping the disseminated delegate layer of insight between sensor hubs and the cloud. The cloud proposed a system adapted to overcome various challenges in omnipresent medical services frameworks, such as portability, energy efficiency, adaptability, and unwavering quality issues, by accepting the right to take care of certain weights of the sensor network and a distant medical service group. An overview of e-health monitoring system in the context of testing and quality assurance of fog computing is presented in this paper. Relevant papers were analyzed in a comprehensive way for the identification of relevant information. The study has compiled contributions of the existing methodologies, methods, and approaches in fog computing e-healthcare.

1. Introduction

Fog computing is an infrastructure located somewhere between the data source and the cloud in which information computing, storage, and applications are located to process the data and information. Fog computing, like edge computing, takes the cloud's benefits and power closer to where information is produced and operated. The words fog computing and edge computing are interchangeably used by many individuals as both require taking knowledge and computation adjacent to where the information is formed. It is mostly done to enhance reliability, but it may also be done for reasons of protection and adherence. The distributed approach to fog computing addresses IoT needs, and perhaps even the enormous volume of information produced by smart sensors and IoT devices, that would also be time-consuming and expensive to submit for analysis and processing to the cloud. Fog computing decreases the required bandwidth and decreases the connectivity between receptors and also the cloud that can have a detrimental impact on IoT results. Fog computing offers the server counterpart to the IoT to manage the information gathered on a daily basis. By exporting gigabytes of Internet traffic from the core network, it eliminates the need for expensive bandwidth additions [1, 2]. Many designed structures have been developed by researchers depending on the best and mechanized cycle with the hope that current patient consideration techniques can be strengthened and fresh limits have been generated considering the gigantic data upset that ensures the framework is clever. Therefore, a simple technique and a novel smart flow model for savvy mending emphasis are the systematic mechanism of need assessment for the brilliant

work cycle of the mending group, considering a few evocative methods used to get-together requirements. Moreover, this research measure offers a better solution than knowing the boggling mending emphasis coordination system consideration and flattens out requesting office work measure of the specialist. Recreation performance shows that the average Quick Flow Model will work better than the current work steps [3].

Modern healthcare approaches are challenging errands to gain more researcher insights. The application of Healthcare 4.0 technique will contribute to the penetration of medical care information where programmers can obtain complete admission to the email records, texts, and reports of patients. In reality, an assured modern healthcare strategy will provide all stakeholders with completion, counting patients, and parental figures. In addition, the research provides a broad written audit, investigating best in class guidelines for preserving security and safety in modern healthcare. It has also explored the blockchain-based response to two specialists and expert networks for offering experiences. Finally, in modern healthcare, current issues and potential protection and security exploration bearings are added [4].

The contribution of the proposed study is to present an overview of e-health monitoring system in the context of testing and quality assurance of fog computing. Several relevant papers associated with the proposed study were analyzed in a comprehensive way. The study has compiled the contributions of the existing methodologies, methods, and approaches in fog computing in e-healthcare.

The organization of the paper is as follows: Section 2 presents the literature study of the proposed research. Section 3 shows the approaches for evaluation and quality assurance of fog computing-based IoT for health monitoring. Section 4 represents statistics of the research done in the area. The paper concludes in Section 5.

2. Literature Study

Research in the area of healthcare and IoT has gained more attention for devising new algorithms, approaches, techniques, and mechanisms for solving different problems. The integrity of IoT in medical care medicine is discussed by incorporating a comprehensive literature due to the lack of and less convincing medical care administrations to meet the rising demands of a growing population with persistent diseases. It is recommended that this involves a move from facility-driven care to quiet-driven medical services where each specialist is regularly aligned with each other, for example, medical unit, patients, and administration. This IoT e-health biological patient-driven model includes a multilayer infrastructure facility. Various case instances of administration and applications that are updated on certain layers adopt this mist-driven IoT engineering. These models range from portable well-being, assisted living, e-medication, inserts, and structures for early admonition to population management in savvy urban communities. At that point, it has finally got IoT e-healthcare challenges, such as executive data, adaptability, guidance, interoperability, gadget network-human interfaces, security, and safety [5]. Hartmann et al. [6] presented a report describing the existing and evolving edge processing systems and processes for medical care applications, to differentiate system preconditions and difficulties for various use cases. The application for connected devices focuses particularly on the grouping of well-being information, including critical sign monitoring and fall recognition. Other low-dormancy applications conduct explicit side effect scans for illnesses, such as walking irregularities in patients with Parkinson's infection. In addition, it presents a detailed audit of eager figuring information tasks that include transition, encryption, validation, characterization, decrease, and forecasting. Indeed, edge figuring has some related problems, even with these focal points, including prerequisites for refined protection and data reduction techniques to allow their cloud-based partners to perform equivalently, but with smaller capacity. It has been acknowledged that potential analysis headings in edge figures for medical facilities give consumers a wider spread of life whenever they tend to achieve. All information is collected in the concept of the information lake, regardless of its length, its abundance, and its pace. It may be a test to put away all this data regardless of whether the invention provides a few arrangements, for example, on reason, on the cloud or half-breed clouds, as well as the foundation and atmosphere. The Internet of Things has modified the concept of securing information in the atmosphere of the information lake, and the volume cut-off points could be reached earlier rather than later for certain information lakes. As of late, a novel concept, called mist registering, has been introduced. The exchange of information intake steps between the sensor that provides knowledge and the information lake that burns through information is a fundamental feature of haze figuring. Initially, this section discusses the principle of mist registration and the associated difficulties and then explores the alternative options to be considered when managing a knowledge lake [7].

Jaimes et al. [8] presented a study in which a crowd detecting measure is illustrated and evaluated that involves effective collaboration in brilliant contexts between crowd sensing participants, using a simple mist that registers the empowered Internet of Things. A haze figuring IoT model involves a layer of figuring hubs that reside closer to the detecting gadgets, with this layer of mist hubs lying in the organization and the cloud in the center of portable and detecting gadgets. This encourages us to propose a model in brilliant circumstances for crowd sensing that involves both competition and cooperation between members of the edge organization who are close to crowd sensing. To test the show of the specific proposal, recreations are added. The work demonstrates desirable attributes regarding the number of dynamic participants, the number of tests obtained, and inclusion within a given investment plan, considering the limited involvement of crowd detecting members on the edge layer that can serve various atmosphere applications. One of the new research areas is investigating the critical hypothesis, challenging framework, and innovation of continuous inquiry over streaming data for cloud processing. This review describes the related innovation of the investigation depending on random hash, finding out how to hash and summarize, investigating the problems and difficulties of the

ongoing question in the climate of asset-restricted mist processing, ultimately analyzing in detail the vital methodology and techniques for the issue, even decreasing the estimation, encoding techniques depending on figuring out how the development of systematic reviews strategy for inquiry over web-based Internet of Thing details, and the related research question structure study bearings and others. In addition, a Hybrid Dynamic Quantization approach for finding out how to hash has been proposed; studies show that other quantization methods are beaten by DAQ [9].

Kelati et al. [10] have discussed recent advances in metered energy usage knowledge in locally formed administrations. It also studies and analyzes interference, reliable existing, and effective force strategies that demonstrate stable load. This study readily retrieves either nonmeddle or judgmental approaches. This study demonstrates that engineering utilizes advances in the strategy of the savvy instrument and haze registering worldview for planning crude oil data. The framework is experiencing a change in perception to increase the need for everyday comfort of metropolitan networks and to provide healthcare administrations that are practical and competent. Patients with intellectual disabilities can be tested and illustrated by analyzing the power usage of home devices. After this, the article describes the execution stage based on replication to create unique models of family devices and check the AI measurement for the identification operation. Kumari et al. [11] presented an approach which addressed basic nature and difficulty of investigating mist data. The FDA's point-by-point scientific categorization is concerned with the cycle model. We need efficient and persuasive arrangements to handle such big data, such as information mining, analysis, and reduction to be distributed on a cloud at the edge of haze gadgets. For the most part, the current creative work attempts focused around conducting big data investigations lack the challenge of supporting mist knowledge analysis. The proposed model tackles numerous exploration challenges, such as availability, adaptability, and interaction with mint nodes, nodal coordination, variability, efficiency, and the essence of administration needs. We present two contextual studies to view the proposed cycle model. Li et al. [12] offered the production processes for edge fog IoT phase beginning to be completed. These models are applied to a solid situation: the analysis of information streams provided by inserted cameras. The administrations rely on cloud capacity and computing resource systems, transforming their engineering into more dispersed one-dependent eager offices provided by Internet service providers. It is indistinct between the IoT equipment association and cloud system, which is the largest portion in terms of energy utilization. The approval consolidates predictions on a growing array of IoT gadgets on real proving grounds running application-focused and recreations with prominent test systems to discuss the scaling up. The outcomes for this case are indeed the portion of the cloud infrastructure that inserts the processing assets devouring multiple times more than the IoT part containing the IoT equipment and the remote passageway.

Liu et al. [13] presented a framework for half, and half protection saving clinical option emotionally supporting network in haze cloud services, called HPCS, is proposed in this paper. A fog worker uses a lightweight information mining technique in HPCS to gradually screen patients' disease safely. In an authentication manner, the recently found abnormal appearances can be further shipped away from the cloud worker for rising projection. In particular, the goal is to prepare another secure reassessed internal item convention for mist workers to achieve a healthy lightweight single-layer neural organization. In addition, the security safeguarding convention of piecewise polynomial estimation allows cloud workers to safely execute any initiation capabilities in different neural organization layers. Besides that, another framework called security safeguarding division estimate convention is planned to take care of the estimation flood issue. At that phase, we show that by changing the constant and exacting quality of recreations, the HPCS meets the goal of patient possibly the best status checking without preventive splashback with unpermitted parties. To deliver the level of comfort, capability, and digitalization for consumers, the current and impending IoT administrations are exceptionally promising. It takes high security, assurance, validation, and recovery from attacks to get the option to complete such an environment in a constantly creating manner. A stable IoT structure is important at present, joining the crucial reforms in IoT structures designed to achieve start to finish. A detailed analysis is combined in this exploration of security-related problems and threat wellsprings in IoT properties or applications. Precisely, when taking a gender at privacy concerns, recent progress in maintaining a serious level of confidence in IoT applications appears to be made. Four basic changes are assessed to extend the degree of IoT security, including cryptography, fog figuring, edge processing, and machine learning [14].

3. Approaches for Evaluation and Quality Assurance of Fog Computing-Based IoT for Health Monitoring

Numerous platforms, approaches, and techniques are established in the field of fog computing and thus implemented and evaluated to run analytics on multiple devices, such as IoT devices, in a distributed way. Fog computing improves the paradigm of cloud computing on the network edge, introducing a number of options and facilities. Manocha et al. [15] presented a novel scientific fog supported to upgrade an individual's living accomplishments by a deep learning-empowered real position-based inconsistency recognition structure. An effort was made to record predicted movement scores on the cloud to extend the efficacy of the proposed augmented reality treatment by pursuing the ceaseless time arrangement plan to include potential well-being references to an approved clinical expert. In addition, a shrewd risk profile age structure is proposed to gradually insinuate clinical subject matter experts and managers regarding an individual's actual real status. The age of the alert is straight forwardly relative to the anticipated actual abnormality and the size of well-being seriousness. The determined results legitimize the prevalence of the proposed examination checking arrangement over the traditional cloud-based observing

arrangements by accomplishing high movement expectation, precision, and less dormancy rate in dynamics. Mutlag et al. [16] offered a study with the purpose to implement a deliberately writing audit of cloud processing developments in the field of IoT frameworks for medical services and review the history. The implications of the scientific categorization have been isolated into three main classes; systems and models, frameworks, audit, and summary. For demanding applications, ongoing low inertness, and high reaction time, particularly in medical services applications, fog figuring is considered necessary. Separate activities with glare registration were established. Compared to distributed computing, cloud processing decreased inertness without doubt. Specialists show that extensions of reproduction and research ensure that a detailed image passivity are provided.

Fog figuring is still starting and needs strong preparation to obtain a successful, productive, and effectively deployable replacement for the now prevalent cloud as essentially achievable cost [17]. In this article, a new asset-productive framework is presented for a multidistrict haze processing worldview for disseminated video synopsis. The portals of the sensor field depend on the Raspberry Equity gadget. Validation tapes are distributed over different hubs, and a breakdown is provided over the structure of cloud, which is periodically pushed to the cloud to decrease the consumption of data transfer resources. To test the proposed system, a number of realistic remaining tasks are used as observation recordings. Trial results indicate that the proposed device has virtually nothing overhead with great adaptability over off-the-rack costly database arrangements, even by using an exceedingly restricted asset, a single board, accepting its adequacy for brilliant urban areas assisted by IoT [17]. Olakanmi and Odeyemi [18] represented a security conspiracy that provides executives with viable data, and safe admission to patient data in an e-health setting is supported. In addition, the methodology underpins the useful conveyance of medical services among carers through compelling automation for data sharing. It will help clinical emphasis on carers to function more effectively and for patients to receive better treatment. Receiving wearable clinical gadgets and distributed computing offers an immense amount of data for quick and momentary access. Nevertheless, it provides some details on the bottlenecks, security, and safety challenges of managers. Using the symmetric key and modified cipher text-strategy trait-based encryption, a two-layer security approach is obtained to provide fine-grained admission control, time-sensitive repudiation of land, and agreeable assignment of well-being management among caregivers.

3.1. E-Health Approaches in Pandemic. Otoom et al. [19] presented a study suggesting an ongoing system for COVID-19 discovery and checking. The proposed structure uses the IoT system to collect client constant manifestation information, to identify suspected cases of Covid19 early, to screen the care reaction of people who have just recovered from the infection, and to gather and analyze significant information to understand the concept of the infection. The platform consists of five main segments: Collection and Uploading of Symptom Data, Isolation Focus, Data Analysis Center (AI), Health Advisors, and Network Equipment. This study proposes eight Artificial Intelligence calculations, specifically Support Vector Machine, Naive, Reverse Nearest Neighbor, Linear Regression, State Diagram, and Proposed General. In contrast to the part of the relevant side effects, the analysis was aimed at testing these eight calculations on a real COVID embodiment dataset. The results indicate that five of these eight analyses achieved an efficiency of more than 90 percent. Parasuraman and Sangaiah [20] presented a study that explores the systematic needs of vast spaces and devoured massive amounts of power to needless electronic measures. The coordinated structure was to form dispersed structures with higher efficiency at the end of the ongoing years. The normal registration process turns out to be more expensive and inviolate to oversee in the current years as information requests and online customers are rapidly extended. Conventional processing is unacceptable for getting to the data wherever and whenever. Cloud calculation is a web-based figure with comprehensive running effects and unsurprising features across companies, partnerships, data innovation, architecture, programming, and data stockpiling, providing easy and updated planning tools and on-demand preparation of resources. In fact, vendors may assume that their customer information placed on their base is safe and, in addition, very much guaranteed, so the strongest security efforts need to be divided to deal with the difficulties of putting away data at an outsider data center.

In the light of compact IoT and cloud side administration, the authors created two overlay arrangement in this paper. ITaaS contains arrangements for (a) the IoT side to regularly support information assortment from IoT gadgets to a passage and (b) the cloud back-end side to help exchange stockpile and prepare information. ITaaS provides the vanguard of innovation to allow fast application arrangements in the space of interest. E-health and distant tracking are conspicuous and promising applications of this breakthrough. A distant patient observation framework as a proof of idea and the coordination of the proposed scheme uses a beat oximeter and devices for detecting pulse observation. Similarly, the spine system with high client concurrence and high information streams was stressed, and we show that the solicitations are performed at around 1 second, a number that means a good presentation by considering the number of solicitations, the organization inactivity, and the general (two GB RAM) [21].

3.2. Geo-Based Dissemination. The concept of fog registering in healthcare frameworks is exploited by shaping a geo-disseminated delegate layer of insight between sensor hubs and the cloud. The cloud proposed system will adapt to various challenges in omnipresent medical services frameworks, such as portability, energy efficiency, adaptability, and unwavering quality issues, by accepting the right to take care of certain weights of the sensor network and a distant medical service group. Particularly in clinical conditions, a prosperous use of weight associated gateways will empower enormous arrangements of pervasive observing frameworks. A model is presented for a smart e-health gateway known as UT-GATE, where a portion of the higher level highlights reviewed has been modified. In addition, an Internet of

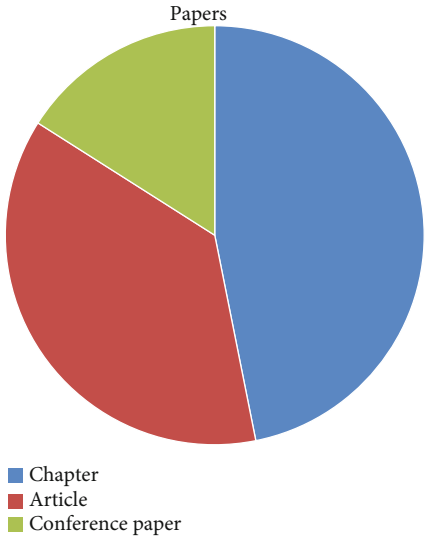


Figure 1: Paper types.

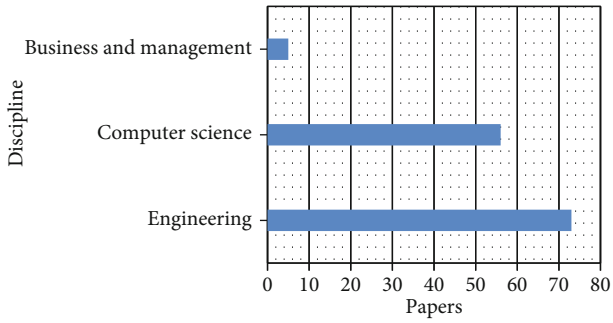


Figure 2: Disciplines in the area.

Things early warning score check was conducted to essentially demonstrate the efficacy and validity of our system for clinical contextual studies. The proof of concept configuration demonstrates an Internet of Things observing system with improved and broad knowledge of the platform, energy ability, accessibility, operation, connectivity, stability, and durability [22]. The study advocates the critical role of modern guidelines and edge authentication components for the diffusion of the largely expanded consumer experience in conjunction with presented collection management and surveys the modern insights that can gain from both the IoT and edge processing situation, discussing in depth about each of the taxonomic segments at that stage. Second, it presents two use cases executed for all intents and purposes that have as of late used the edge-IoT worldview together to fix metropolitan savvy living problems and, third, for e-medical services such as the proposed novel fog-based engineering and developed demo proving ground. The test results showed promising results in limiting emphasis on IoT cloud research or doorway. It concludes with discussions on various boundaries, such as engineering, prerequisite capacity, helpful problems, and determination rules, associated with the endurance of layer joining [23].

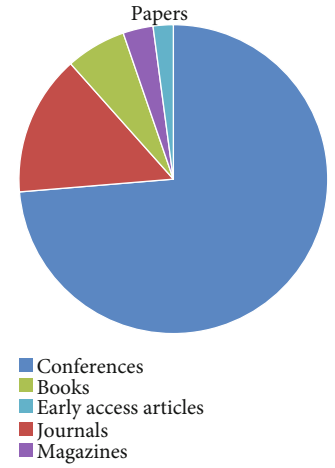


Figure 3: Paper types.

Rehman et al. [24] have completed genome datasets of different organisms readily available, and a lot more are being sequenced. In understanding the functioning of normal living beings, these genomic mechanisms are of utmost importance and have many applications in our everyday lives. It is a daunting job to control this gigantic measure of knowledge with conventional methods. Analysis of such data may take hours or days to produce results which have caused ideal models of current distributed computing to face various difficulties. Among the indicated qualities, fog processing is commonly used by specialists around the world for flexible asset distribution. Cloud registration uses the cloud at the back end, thus expanding the spectrum of cloud to things by taking resources close to the edge of gadgets, thus defeating various impediments to the worldview of distributed computing. In view of the interesting properties of haze, such as low jitter, low idleness, enhanced protection, and so on, it is argued that the philosophy of fog extraction has extraordinary potential for high embedded platforms for data and information. Sanchez-Gallegos et al. [25] presented a study on the plan creation, as well as implementation of an engineering model to build on request edge-mist cloud handling frameworks to deal consistently with enormous data and simultaneously execute NFR filling administration. Effective and calculated squares, revised as microservices and nanoservices, are recursively interconnected in this model to construct edge-haze cloud planning systems as a rationalist administrative framework. Coherence plans generate information through the cloud and edge structure squares and enable a model developed using this model to demonstrate the accomplishment of this model, which was tested in a situation study based on the handling of data to endorse a simple dynamic methodology in distant patient observation. This research examines situations in which end-clients and clinical staff received bits of information when planning electrocardiograms provided by sensors in remote IoT devices, much as doctors were accommodated and admonished when examining and identifying anomalies in the broken down ECG content on the web. It was also considered a situation in which associations deal with different concurrent edge-

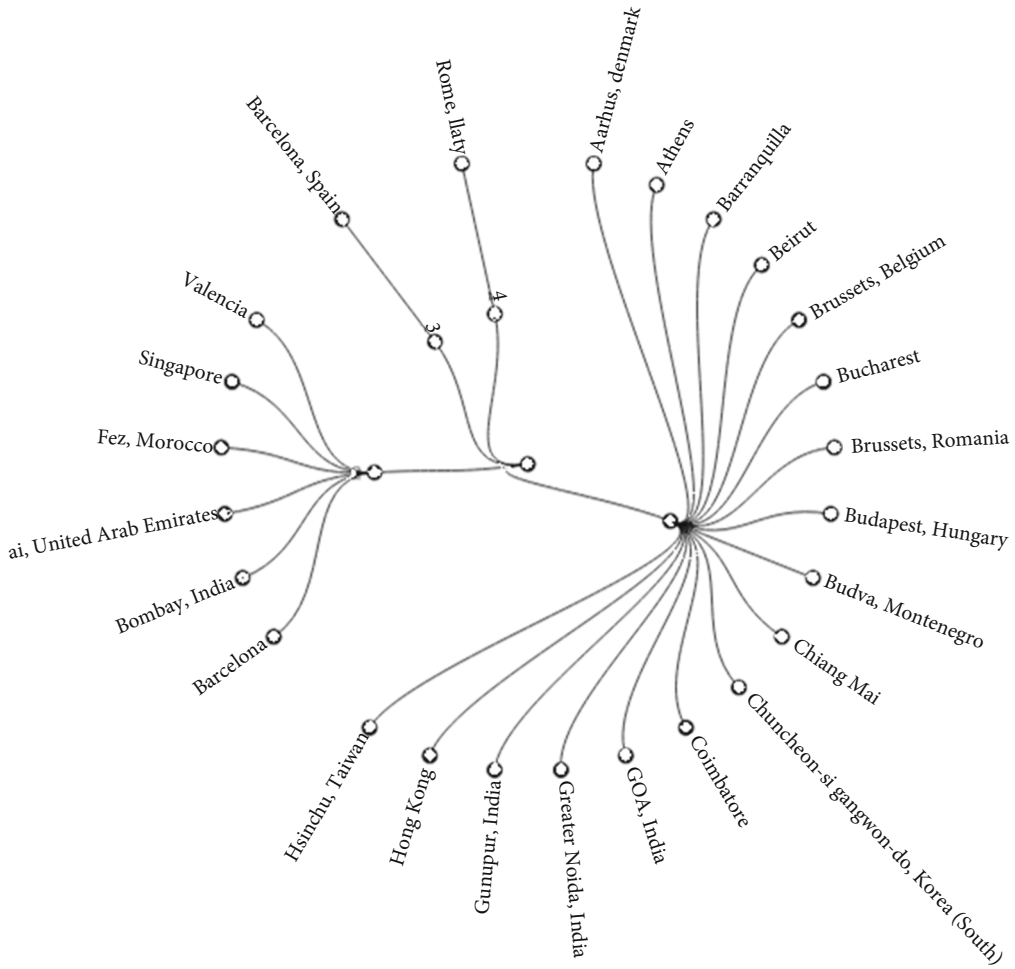


Figure 4: Conference locations.

mist cloud systems for the preparation of information and material transmitted to inside and outer workers.

3.3. Real-Time Mobility and Robust Streaming. García-Valls et al. [26] presented the plan and approval of a system that improves the administration season of the fog workers' chosen exercises; undoubtedly, most of those exercises are described by distant patients. It crosses the limits of current processors to parallelize explicit exercises that can be a sudden spike in demand for saved centers; what is more, it depends on the nature of administration, certification of information circulation stages to improve correspondence, and reaction times to versatile patients. A significant test of e-health administrations on the cloud, instead of various administrations running on shrewd large cities, is that they typically conduct various computational exercises conducting broad data handling on realistic information that should be protected. The overhaul of distant patient hubs can be enhanced by using the limits of current processors. The proposed approach is approved for a model execution of recreated computationally serious e-health collaborations, diminishing the reaction time by 4x when center reservation is enacted. In comparison to cloud space, the latest ideal models of edge and cloud figuring offer innovative arrangements by bringing assets closer to the customer and offering low idleness and energy efficient responses for knowledge planning. In any event, there are various limitations and spotlights on the latest mist models from restriction. It is suggested in this study that a new structure called health fog to integrate deep learning in edge registering gadgets and conveyed it for the genuine use of the fog-enabled cloud system programmed heart-disease inspection. Fog bus is used to convey and evaluate the presentation of the proposed monitoring. In various cloud calculation situations and for different customer needs, health fog is configurable for different operation modes that offer the best quality of service or forecast accuracy, as necessary [27]. To minimize the spread of the infection and protect the health of patients who need to stay in an emergency clinic, home hospitalization is a standout among other alternative arrangements. This paper proposes a system for home hospitalization based on IoT, fog, and cloud processing; these are among the key developments that have led in a big way to improving the field of medical services. These systems enable patients in their homes and among their families to recover and obtain care, where awareness and the ecological condition of the hospital stay room are observed, to encourage specialists to follow the hospital stay cycle and to make recommendations, through control units and flexible applications created for this

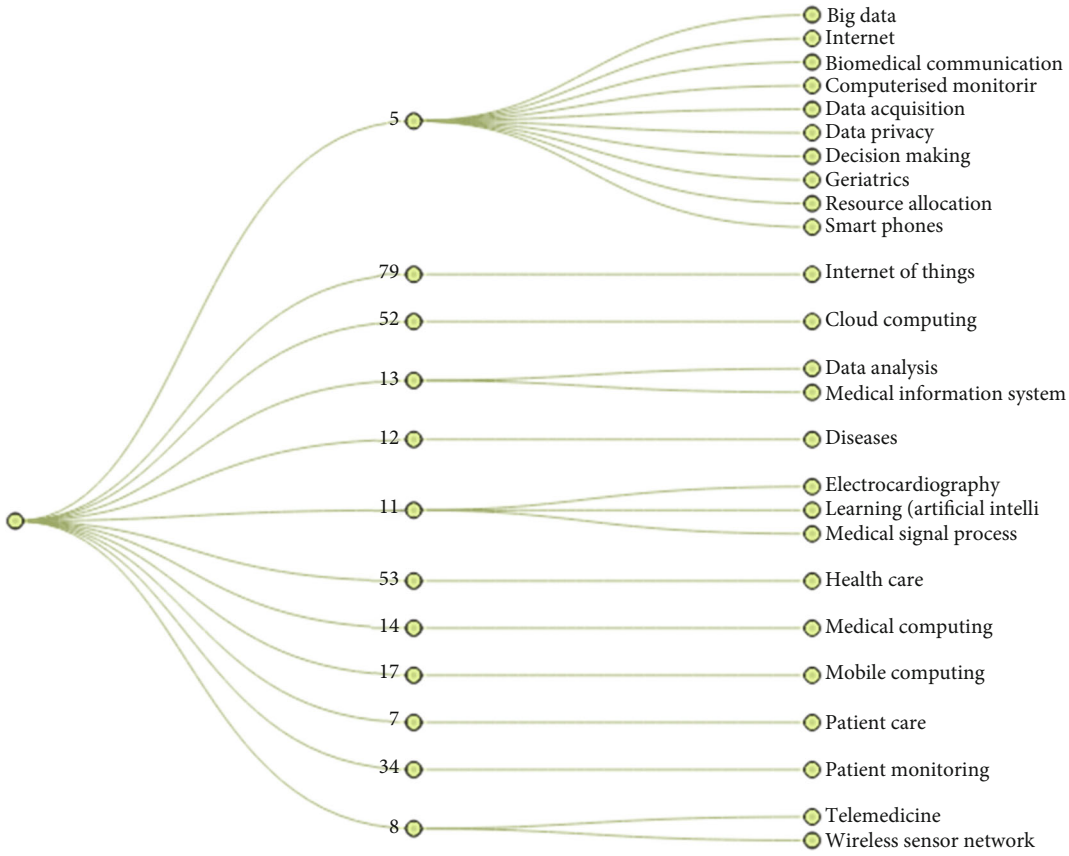


Figure 5: Publication topics.

purpose, for patients and their supervisors. The after effects of the test have shown a remarkable appreciation of this framework by patients and specialists alike [28].

The use of IoT gadgets for ML deduction saves the cloud disadvantage of high dormancy in the enterprise, unsuitable for delay-touch apps such as fall locators. The present fall recognition structures, however, require induction on the mist, and there is no evidence of it under real circumstances, nor documentation regarding the dynamic challenge of the structure. To collect tolerant observing data, a handheld trihub accelerometer is used. This study suggests a genius Open IoT engineering in the cloud to assist the far-off sending and the DL model board. Two DL models have been submitted to advance assets, and their exhibition and derivation time using virtualization are analyzed. The results show the adequacy of our fall system, which offers a more convenient and accurate solution than traditional fall finder frameworks, greater competence, 98.75 percent accuracy, lower deferral, and improvement in administration [29]. Farahani et al. [5] proposed a comprehensive AI-driven IoT e-health engineering focused on the concept of a collective machine learning method in which insight is transmitted through devices Despite the energizing advances in the shift from center-driven to understanding-driven medical care, the device enables medical service professionals to continuously screen the associated data of subjects anywhere anytime and has constant noteworthy interactions that ultimately strengthen the dynamic force. Using a comprehensive ECG-based arrhythmia position contextual analysis, the plausibility of such engineering is tested. From plan recommendations, for example, relating to overheads, energy usage, inertia, and implementation, to designing and conveying advanced AI strategies to such engineering, this illustrative model explores and discusses immeasurably important parts of the proposed engineering. Yacchirema et al. [30] introduced an innovative system based on distributed and cloud computing technologies that provides new opportunities to assemble novel and inventive administrations to support the rest of apnea and to resolve the current constraints in combination with IoT and large knowledge levels. In particular, the structure is focused on a few remote low-power organizations with brilliant heterogeneous gadgets. An edge center offers IoT association and interoperability in cloud computing and prehandling IoT information to continuously recognize occasions that can jeopardize the elderly and function similarly. In the cloud, for additional handling and investigation, a generic motivating agent background broker supervises, stores, and infuses information into the massive information analyzer. The presentation and emotional appropriateness of the system were evaluated separately using more than thirty GB size datasets and a poll satisfied by medical professionals educated. Results show that the system knowledge study enhances the dynamics of the experts to screen and direct rest apnea care, as well as improving the personal satisfaction of older people.

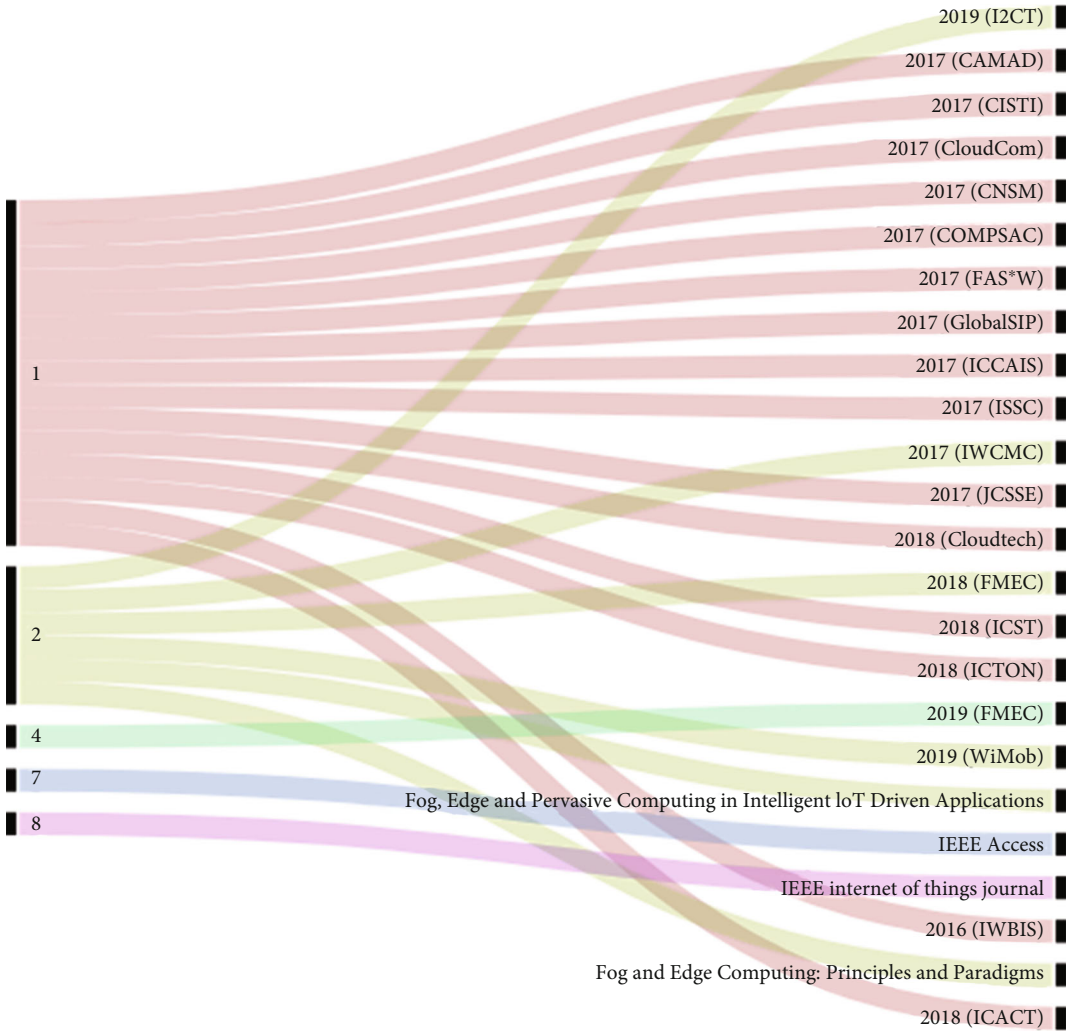


Figure 6: Publication title.

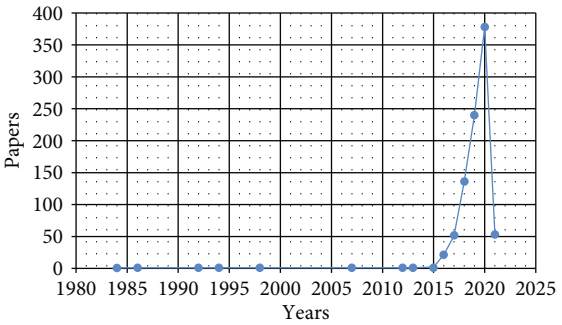


Figure 7: Year of publication.

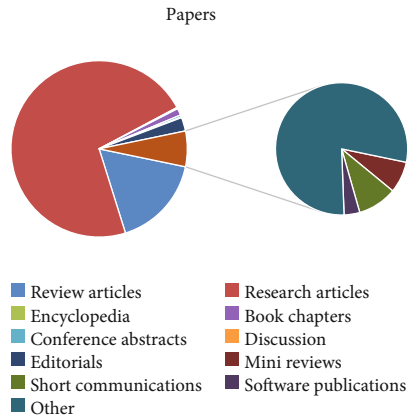


Figure 8: Publication type.

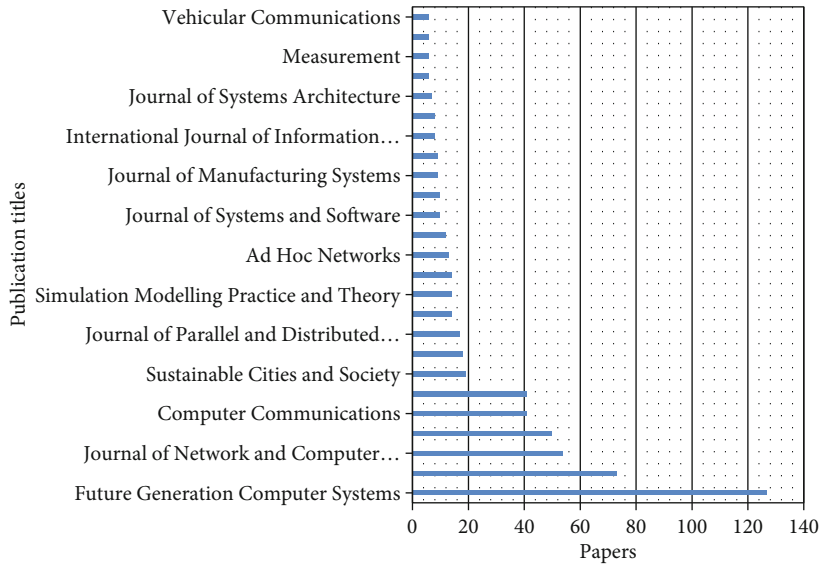


Figure 9: Publications titles.

4. Statistics of the Research in the Area

It is difficult to guarantee the security of sensitive information in an acceptable stored information in view of the fact that after the information is delivered to the data-driven entity in the type of piece, it is no longer limited by the information distributor. In addition, terminal clinical sensors are typically asset-driven in certain real-time health applications, limiting the immediate receipt of expensive cryptographic natives. To overcome these challenges, an asset-skilled secure information sharing strategy is proposed in the data-driven e-health system, the one which uses encryption based on the related literature trait and adapts it to the previously stated system regarding essential security needs. It likewise misuses the calculation assets of fog hubs and utilizes rethinking cryptography to boost framework productivity. The evaluation shows that the strategy can fundamentally reduce the overhead estimate of resource-restricted terminal clinical gadgets and can more effectively support ongoing e-health applications [31]. Aladwani [32] proposed to use fog registering between sensors and distributed computing to competently collect measurement information, reduce the measurement of information transferred between the cloud and the sensors, and increase the efficacy of the whole system. Remote sensor organizations that use health care observation in the territory send a large amount of companies of varying degrees of importance and length to fog registration all time. Eventually, estimation of the ability to reliably provide task needs and render the primary factor in the need for tasks is their importance, paying no attention to their duration. This study is aimed at enhancing the execution of static business booking calculations by using another technique called classification of tasks and categorization of virtual machines based on the significance of enterprises. IoT-characterized enterprises rely on their importance in three classes: high-significance errands, medium-significance enterprises, and low-significance errands that depend on the status of the patient. They will be added to the MAX-MIN booking equation to measure the exhibition achieved by these techniques.

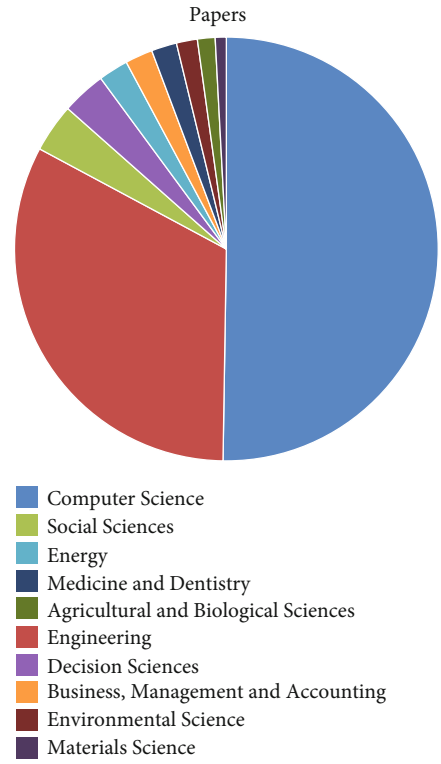


Figure 10: Subject area.

Karatas and Korpeoglu [33] proposed that a topographically circulated multiple leveled cloud in this paper, fog registration-based IoT architecture, and proposed procedures for setting IoT information in the sections of the proposed engineering. Information is considered in various kinds, and different applications can involve each kind of information.

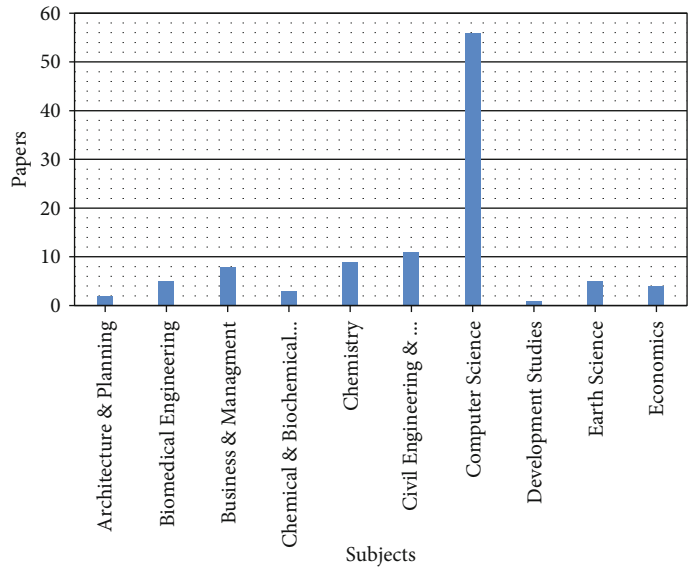


Figure 11: Subjects of the area.

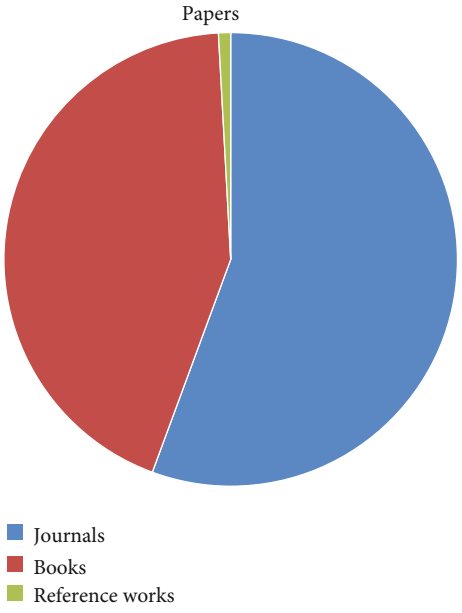


Figure 12: Publication types.

The model of the problem of information situation is a problem of improvement and proposes calculations for the effective, viable situation of information generated and devoured by IoT hubs that are topographically relevant. Data used for different applications is packed away in an environment that is essentially accessed by applications using that type of information for only a single period. To test the plan, comprehensive recreation trial is conducted and the results show that the design and situation techniques can productively position and store information while providing great execution to applications and organization's as far as access inertness and data transfer capability are devoted. The current gadgets that are used today are also becoming all more impressive in terms of highlights and skills, but they are still not equipped to perform shrewd, self-governing, and savvy orders, such as those often needed for shrewd medical services, concerning helped living, virtual reality, and increased reality; we need another substance to perform undertakings for emerging IoT and distributed computing applications; assignment offloading is desirable. Between IoT hubs, sensors and edge gadgets can happen. Offloading can be done based on different components that involve an application's computational needs, load change, board energy, executive inertness, etc. This review presents a scientific categorization of late discharge plans that have been suggested, such as cloud, distributed computing, and IoT, for space. It also discusses the middleware developments that enable offloading in a cloud-IoT scenario and the components that are critical for offloading in a particular scenario. Additionally, it presents an exploration preprint submitted to Future Generation Computer Systems on May 2, 2018, opening concerning offloading in edge and cloud processing [34].

The search process of the proposed research was carried out in various popular libraries including Springer, ScienceDirect, IEEE, and Wiley Online. The key reason of the search in these libraries was to identify the most associated materials for the process of analysis. The analysis was done from different perspectives such as to identify the publications on year-wise basis and to identify the type of publication, title of publication, topics of publication, location of publications, and so on. Figure 1 depicts the paper types in the library of Springer.

Figure 2 represents the disciplines of the area in the given library. More papers are published in the area of engineering.

Figure 3 shows the types of papers in the IEEE library. In this library, more articles were published as conference papers.

Figure 4 shows the conference location in the same library.

Figure 5 depicts the topics of publication in the library where more papers are published in the area of IoT.

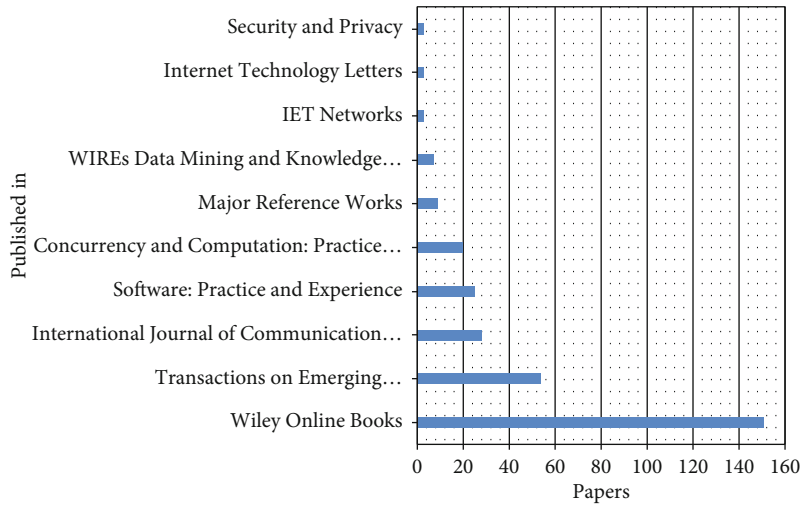


Figure 13: Papers published.

Figure 6 depicts the publication title.

Figure 7 graphically represents the number of publications done in a given year in the Library of ScienceDirect.

The publication types are given in Figure 8 for the given library.

The publication titles are presented in Figure 9. More publications regarding the area of research were done in "Future Generation Computer Systems."

The subject areas are presented in Figure 10. The figure shows that more publications are done in the field of "Computer Science."

The library of Wiley online was searched for identifying associated materials. Figure 11 depicts the subject areas of research in the library.

The publication types are mentioned in Figure 12. More publications are done as journal category.

Figure 13 graphically demonstrates the articles published.

5. Conclusion

Fog computing is a computing infrastructure located nearby data sources and the cloud, in which information computing, storage, and applications are positioned to process the data and information. Fog computing advances the paradigm of cloud computing on the network edge, introducing a number of options and facilities. Fog computing enhances the processing, verdicts, and interventions to occur through IoT devices and spreads only the necessary details. The ideas of fog computing based on IoT in healthcare frameworks are exploited by shaping the disseminated delegate layer of insight between sensor hubs and the cloud. An overview of e-health monitoring systems in the context of testing and quality assurance of fog computing is presented in the study under consideration. Relevant materials were searched and analyzed in a widespread manner. The study has compiled the contributions of the existing methodologies, methods, and approaches in fog computing in e-healthcare. This review will be an evidence for the researchers to devise new approaches and platforms for handling and managing various situations associated with researches in the area.

Data Availability

The data will be provided upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] S. Khan, S. Nazir, I. García-Magariño, and A. Hussain, "Deep learning-based urban big data fusion in smart cities: towards traffic monitoring and flow-preserving fusion," *Computers & Electrical Engineering*, vol. 89, article 106906, 2021.

[2] B. Wu, S. Nazir, and N. Mukhtar, "Identification of attack on data packets using rough set approach to secure end to end communication," *Complexity*, vol. 2020, Article ID 6690569, 12 pages, 2020.

[3] M. Rath and V. K. Solanki, "Performance improvement in contemporary health care using IoT allied with big data," in *Handbook of Data Science Approaches for Biomedical Engineering*, V. E. Balas, V. K. Solanki, R. Kumar, and M. Khari, Eds., pp. 103–119, Academic Press, 2020.

[4] J. J. Hathaliya and S. Tanwar, "An exhaustive survey on security and privacy issues in Healthcare 4.0," *Computer Communications*, vol. 153, pp. 311–335, 2020.

[5] B. Farahani, M. Barzegari, F. Shams Aliee, and K. A. Shaik, "Towards collaborative intelligent IoT eHealth: from device to fog, and cloud," *Microprocessors and Microsystems*, vol. 72, article 102938, 2020.

[6] M. Hartmann, U. S. Hashmi, and A. Imran, "Edge computing in smart health care systems: review, challenges, and research directions," *Transactions on Emerging Telecommunications Technologies*, no. article e3710, 2019.

[7] A. Laurent, D. Laurent, and C. Madera, *Book, Data Lakes, First Edition. Edited by © ISTE Ltd 2020. Published by ISTE Ltd and*

John Wiley & Sons, Inc., ISTE Ltd and John Wiley & Sons, Inc, 2020.

[8] L. G. Jaimes, A. Chakeri, and R. Steele, "Localized cooperation for crowdsensing in a fog computing-enabled internet-of-things," *Journal of Ambient Intelligence and Humanized Computing*, 2018.

[9] X. Jiang, P. Hu, Y. Li et al., "A survey of real-time approximate nearest neighbor query over streaming data for fog computing," *Journal of Parallel and Distributed Computing*, vol. 116, pp. 50–62, 2018.

[10] A. Kelati, I. B. Dhaou, A. Kondoro, D. Rwegasira, and H. Tenhunen, "IoT based appliances identification techniques with fog computing for e-health," in *2019 IST-Africa Week Conference (IST-Africa)*, pp. 1–11, Nairobi, Kenya, May 2019.

[11] A. Kumari, S. Tanwar, S. Tyagi, N. Kumar, R. M. Parizi, and K.-K. R. Choo, "Fog data analytics: a taxonomy and process model," *Journal of Network and Computer Applications*, vol. 128, pp. 90–104, 2019.

[12] Y. Li, A.-C. Orgerie, I. Rodero, B. L. Amersho, M. Parashar, and J.-M. Menaud, "End-to-end energy models for edge cloud-based IoT platforms: application to data stream analysis in IoT," *Future Generation Computer Systems*, vol. 87, pp. 667–678, 2018.

[13] X. Liu, R. H. Deng, Y. Yang, H. N. Tran, and S. Zhong, "Hybrid privacy-preserving clinical decision support system in fog-cloud computing," *Future Generation Computer Systems*, vol. 78, pp. 825–837, 2018.

[14] M. Mahbub, "Progressive researches on IoT security: an exhaustive analysis from the perspective of protocols, vulnerabilities, and preemptive architectonics," *Journal of Network and Computer Applications*, vol. 168, article 102761, 2020.

[15] A. Manocha, G. Kumar, M. Bhatia, and A. Sharma, "Video-assisted smart health monitoring for affliction determination based on fog analytics," *Journal of Biomedical Informatics*, vol. 109, article 103513, 2020.

[16] A. A. Mutlag, M. K. Abd Ghani, N. Arunkumar, M. A. Mohammed, and O. Mohd, "Enabling technologies for fog computing in healthcare IoT systems," *Future Generation Computer Systems*, vol. 90, pp. 62–78, 2019.

[17] M. Nasir, K. Muhammad, J. Lloret, A. K. Sangaiah, and M. Sajjad, "Fog computing enabled cost-effective distributed summarization of surveillance videos for smart cities," *Journal of Parallel and Distributed Computing*, vol. 126, pp. 161–170, 2019.

[18] O. Olakanmi and K. Odeyemi, "FEACS: a fog enhanced expressible access control scheme with secure services delegation among carers in E-health systems," *Internet of Things*, vol. 12, article 100278, 2020.

[19] M. Otoom, N. Otoum, M. A. Alzubaidi, Y. Etoom, and R. Banihani, "An IoT-based framework for early identification and monitoring of COVID-19 cases," *Biomedical Signal Processing and Control*, vol. 62, article 102149, 2020.

[20] S. Parasuraman and A. K. Sangaiah, "Fog - driven healthcare framework for security analysis," in *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*, pp. 253–270, elsevier, 2018.

[21] E. G. M. Petrakis, S. Sotiriadis, T. Soultanopoulos, P. T. Renta, R. Buyya, and N. Bessis, "Internet of Things as a Service (iTaaS): challenges and solutions for management of sensor data on the cloud and the fog," *Internet of Things*, vol. 3-4, pp. 156–174, 2018.

[22] A. M. Rahmani, T. N. Gia, B. Negash et al., "Exploiting smart e-health gateways at the edge of healthcare Internet-of-Things: a fog computing approach," *Future Generation Computer Systems*, vol. 78, pp. 641–658, 2018.

[23] P. P. Ray, D. Dash, and D. De, "Edge computing for Internet of Things: a survey, e-healthcare case study and future direction," *Journal of Network and Computer Applications*, vol. 140, pp. 1–22, 2019.

[24] H. U. Rehman, A. Khan, and U. Habib, "Fog computing for bioinformatics applications," in *Book Chapter*, pp. 529–545, elsevier, 2020.

[25] D. D. Sanchez-Gallegos, A. Galaviz-Mosqueda, J. L. Gonzalez-Compean et al., "On the continuous processing of health data in edge-fog-cloud computing by using micro/nanoservice composition," *IEEE Access*, vol. 8, pp. 120255–120281, 2020.

[26] M. García-Valls, C. Calva-Urrego, and A. García-Fornes, "Accelerating smart eHealth services execution at the fog computing infrastructure," *Future Generation Computer Systems*, vol. 108, pp. 882–893, 2020.

[27] S. Tuli, N. Basumatary, S. S. Gill et al., "HealthFog: an ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments," *Future Generation Computer Systems*, vol. 104, pp. 187–200, 2020.

[28] H. Ben Hassen, N. Ayari, and B. Hamdi, "A home hospitalization system based on the Internet of things, fog computing and cloud computing," *Informatics in Medicine Unlocked*, vol. 20, article 100368, 2020.

[29] D. Sarabia-Jácome, R. Usach, C. E. Palau, and M. Esteve, "Highly-efficient fog-based deep learning AAL fall detection system," *Internet of Things*, vol. 11, article 100185, 2020.

[30] D. Yacchirema, D. Sarabia-Jácome, C. E. Palau, and M. Esteve, "System for monitoring and supporting the treatment of sleep apnea using IoT and big data," *Pervasive and Mobile Computing*, vol. 50, pp. 25–40, 2018.

[31] L. Dang, M. Dong, K. Ota, J. Wu, J. Li, and G. Li, "Resource-efficient secure data sharing for information centric E-health system using fog computing," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kansas City, MO, USA, May 2018.

[32] T. Aladwani, "Scheduling IoT healthcare tasks in fog computing based on their importance," *Procedia Computer Science*, vol. 163, pp. 560–569, 2019.

[33] F. Karatas and I. Korpeoglu, "Fog-based data distribution service (F-DAD) for Internet of Things (IoT) applications," *Future Generation Computer Systems*, vol. 93, pp. 156–169, 2019.

[34] M. Aazam, S. Zeadally, and K. A. Harras, "Offloading in fog computing for IoT: review, enabling technologies, and research opportunities," *Future Generation Computer Systems*, vol. 87, pp. 278–289, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 5551497, 9 pages
https://doi.org/10.1155/2021/5551497

Research Article

Second-Order Delay Differential Equations to Deal the Experimentation of Internet of Industrial Things via Haar Wavelet Approach

Yongtao Xuan[1], **Rohul Amin**[2], **Fakhar Zaman**[2], **Zohaib Khan**[2], **Imad Ullah**[2] and **Shah Nazir**[3]

[1]*Guangdong University of Science and Technology, Dongguan City, 523083 Guangdong Province, China*
[2]*Department of Mathematics, University of Peshawar, Peshawar 25120, Pakistan*
[3]*Department of Computer Science, University of Swabi, Ambar 23430, Pakistan*

Correspondence should be addressed to Yongtao Xuan; xyt0512@126.com and Shah Nazir; shahnazir@uoswabi.edu.pk

Received 25 January 2021; Revised 24 February 2021; Accepted 20 March 2021; Published 7 April 2021

Academic Editor: Rahim Khan

In this article, an efficient numerical approach for the solution of second-order delay differential equations to deal with the experimentation of the Internet of Industrial Things (IIoT) is presented. With the help of the Haar wavelet technique, the considered problem is transformed into a system of algebraic equations which is then solved for the required results by using Gauss elimination algorithm. Some numerical examples for convergence of the proposed technique are taken from the literature. Maximum absolute and root mean square errors are calculated for various collocation points. The results show that the Haar wavelet method is an effective method for solving delay differential equations of second order. The convergence rate is also measured for various collocation points, which is almost equal to 2.

1. Introduction

Delay differential equations (DDEs) are type of DEs in which the solution of the unknown function is given in the previous time interval. A system whose performance does not depend directly on time is a time-invariant delay system. This systems can be defined by constant coefficients of the nth order ordinary DEs [1]. DDEs are used for modelling of various phenomenon such as modelling of systems with memory, electric circuits, and mechanical systems. The application of these systems in population dynamics [2] can be used in communication networks, mass transportation, remote controls, and biological systems. Many of the processes, both natural and artificial, in medicine, chemistry, engineering, economics, etc. involve time delays. The Internet of Things (IoT) contributes in facilitating the needs of daily life such as IoT for healthcare using effects of Mobile computing [3] and nonlinear delay integro-differential equations for wireless sensor network and IoT [4].

There are numerous approaches available in the literature for the solution DDEs of second order. Seong and Majid [1] developed the Adams Moulton technique to solve the second-order DDEs. Ibrahim [5], used 2h-step Spline method to solve the DDEs. Sedaghat [2], utilized the Chebyshev polynomials method to find the solution of DDEs. Ehigie et al. [6] implement a 3-point block technique to solve DDEs of second order. Chebyshev wavelet technique was developed by Ghasemi and Kajan [7] to solve the DDEs. Ahmad et al. [8] solved the DDEs by a block hybrid method. Multistep methods was used by Okunuga and Ehigie [9] to solve the DDEs. Brown [10] used a method of implicit multistep to solve the DDEs. Ismail et al. [11] found the solution of DDEs by 3-point block methods. Ehigie et al. [12] used a method of 2-step continuous linear multistep to solve the second-order DDEs. Some other well-known methods are the following: Runge-Kutta [13], Shift Walsh matrix method [14], Hermite interpolation method [15], method of retarded initial value problems [16], one-step collocation method [17],

coupled block technique [18], one-step block techniques [11], implicit block technique [19], Direct integration implicit variable method [19], predictor-corrector method [20], Taylor method [21], fuzzy mapping and control method [22], variational iteration method [23, 24], and Galerkin method [25]. A structure for the IIoT cloud-fog hybrid network for industry data processing was proposed by Liu et al. [26]. Sahal et al. [27] studied the strong point and flaws of open source technologies for big data. Khan et al. [28] offered the idea of IIoT in a novel manner for supporting readers to comprehend the IIoT. Gulati and Kaur [29] analysed the main opportunities assimilated from the idea of IoT into industry with suggesting reference architecture. The use of Haar wavelet have wide-ranging applications in scientific computing. The Haar Collocation Technique (HCT) is used for fractional Riccati type differential equations [30], Birthmark based identification [31], delay Fredholm-Volterra integral equations [32], delay integrodifferential equations [4], systems of fractional differential equations [33], and fractional integrodifferential equations [34] in recent literature. This article studies the solutions of second-order DDEs, that is, we develop numerical technique using Haar wavelet with constant delay.

In this paper, we discuss the solution of the second-order DDEs using a HCT to deal with the experimentation of IIoT, consider linear DDEs as

$$\begin{cases} w''(t) = a(t)w'(t) + b(t)w(t) + c(t)u(t) + d(t)w(t-\tau) + e(t)u(t-\tau), \\ w'(0) = \alpha_1, w(0) = \alpha_2, \\ w(t) = \phi(t), -\tau \le t < 0, \end{cases} \tag{1}$$

where $u(t)$ is a control function, $\phi(t)$ is the delay condition, and $w(t)$ is a state function.

The paper is organized as some basic results and notions are given in Section 2. Section 3 provides the HCT solution for linear DDEs of second order. In Section 4, the HCT validation is given. The results are discussed in Section 5, and the conclusion is given in the last part of the paper.

2. Haar Wavelet

Scaling function on $[\alpha_1, \alpha_2)$ is [35]

$$h_1(p) = \begin{cases} 1 & \text{for } p \in [\alpha_1, \alpha_2), \\ 0 & \text{elsewhere.} \end{cases} \tag{2}$$

Mother wavelet on $[\alpha_1, \alpha_2)$ is

$$h_2(p) = \begin{cases} 1 & \text{for } p \in \left[\alpha_1, \dfrac{\alpha_1 + \alpha_2}{2}\right), \\ -1 & \text{for } p \in \left[\dfrac{\alpha_1 + \alpha_2}{2}, \alpha_2\right), \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

The other terms can be written as

$$h_i(p) = \begin{cases} 1 & \text{for } t \in [\eta_1, \eta_2), \\ -1 & \text{for } t \in [\eta_2, \eta_3), \\ 0 & \text{elsewhere,} \end{cases} \tag{4}$$

where $\eta_1 = \alpha_1 + (\alpha_2 - \alpha_1)(\zeta/d), \eta_2 = \alpha_1 + (\alpha_2 - \alpha_1)(\zeta + 0.5/d), \eta_3 = \alpha_1 + (\alpha_2 - \alpha_1)(\zeta + 1/d)$, where $d = 2^r$, and $\zeta = 0, 1, \cdots, d-1$. The number i can be calculated as $i = d + \zeta + 1$. If we take interval $[0, 1]$, then values of η_1, η_2, and η_3 are: $\eta_1 = \zeta/d$, $\eta_2 = 1/2 + \zeta/d$, $\eta_3 = 1 + \zeta/d$. Any member in $L^2[0, 1)$, is $u(p) = \sum_{k=1}^{\infty} \lambda_k h_k(t)$, we truncate this series is $u(t) \approx \sum_{k=1}^{N} \lambda_k h_k(t)$. Let

$$p_{i,1}(t) = \int_0^t h_i(x)dx,$$
$$P_{i,1}(p) = \begin{cases} p - \eta_1 & \text{for } p \in [\eta_1, \eta_2), \\ \eta_3 - p & \text{for } p \in [\eta_2, \eta_3), \\ 0 & \text{elsewhere.} \end{cases} \tag{5}$$

Also, $p_{i,2}$ is

$$p_{i,2}(p) = \int_0^p p_{i,1}(r)dr. \tag{6}$$

We obtain

$$P_{i,2}(p) = \begin{cases} \dfrac{1}{2}(p - \eta_1)^2 & \text{if } p \in [\eta_1, \eta_2), \\ \dfrac{1}{4m^2} - \dfrac{1}{2}(\eta_3 - p)^2 & \text{if } p \in [\eta_2, \eta_3), \\ \dfrac{1}{4m^2} & \text{if } p \in [\eta_3, 1), \\ 0 & \text{elsewhere.} \end{cases} \tag{7}$$

Generally,

$$P_{i,n}(p) = \int_0^p P_{i,n-1}(r)dr. \tag{8}$$

Thus, [9],

$$P_{i,n}(t) = \begin{cases} 0 & \text{if } t \in [0, \eta_1), \\ \dfrac{(t-\eta_1)^n}{n!} & \text{if } t \in [\eta_1, \eta_2), \\ \dfrac{[(t-\eta_1)^n - 2(\eta_1-\eta_2)^n]}{n!} & \text{if } t \in [\eta_2, \eta_3), \\ \dfrac{1}{n!}[(t-\eta_1)^n - 2(\eta_1-\eta_2)^n + (t-\eta_3)^n], & \text{if } t \in [\eta_3, 1). \end{cases} \tag{9}$$

The $[a_1, a_2]$ interval for HCT is discretized as

$$p_i = a_1 + (a_2 - a_1)\frac{i - 0.5}{2M} \; i = 1, 2, 3, \cdots, 2M. \tag{10}$$

The points defined in the above Eq. (10) are called collocation points (CPs).

3. Numerical Method

Here, we describe the proposed method for second-order DDEs to deal with the experimentation of IIoT. Let $w''(t) \in L^2[0, 1)$, then

$$w''(t) = \sum_{i=1}^{N} \lambda_i h_i(t). \tag{11}$$

Integrating Eq. (11) from 0 to t,

$$w'(t) - w'(0) = \sum_{i=1}^{N} \lambda_i p_{i,1}(t), \tag{12}$$

from Eq. (1), $w'(0) = \alpha_1$, so we get

$$w'(t) = \alpha_1 + \sum_{i=1}^{N} \lambda_i p_{i,1}(t). \tag{13}$$

Now, integrating Eq. (13) from 0 to t, the following relation yields:

$$w(t) - w(0) = \alpha_1(t) + \sum_{i=1}^{N} \lambda_i p_{i,2}(t), \tag{14}$$

from Eq. (1), $w(0) = \alpha_2$, so we have

$$w(t) = \alpha_2 + \alpha_1(t) + \sum_{i=1}^{N} \lambda_i p_{i,2}(t). \tag{15}$$

By putting Eq. (11), Eq. (13), and Eq. (15) in Eq. (1), we get

$$\begin{aligned} &\sum_{i=1}^{N} \lambda_i h_i(t) - a(t)\sum_{i=1}^{N} \lambda_i p_{i,1}(t) - b(t)\sum_{i=1}^{N} \lambda_i p_{i,2}(t) \\ &= \begin{cases} a(t)\alpha_1 + b(t)(\alpha_2 + \alpha_1(t)) + c(t)u(t) + d(t)\phi(t-\tau) + e(t)u(t-\tau), \text{for} t < 0, \\ a(t)\alpha_1 + b(t)(\alpha_2 + \alpha_1(t)) + c(t)u(t) + d(t)\left(\alpha_2 + \alpha_1(t-\tau) + \sum_{i=1}^{N} \lambda_i p_{i,2}(t)(t-\tau)\right) + e(t)u(t-\tau), \text{for} t > 0. \end{cases} \end{aligned} \tag{16}$$

Discretizing this Eq. (16) at t_j CPs, we have

$$\begin{aligned} &\sum_{i=1}^{N} \lambda_i h_i(t_j) - a(t_j)\sum_{i=1}^{N} \lambda_i p_{i,1}(t_j) - b(t_j)\sum_{i=1}^{N} \lambda_i p_{i,2}(t_j) \\ &= \begin{cases} a(t_j)\alpha_1 + b(t_j)(\alpha_2 + \alpha_1(t_j)) + c(t_j)u(t_j) + d(t_j)\phi(t_j-\tau) + e(t_j)(u(t_j)-\tau), \text{for } t_j < 0 \\ a(t_j)\alpha_1 + b(t_j)(\alpha_2 + \alpha_1(t_j)) + c(t_j)u(t_j) + d(t_j)\left((\alpha_2 + \alpha_1(t_j-\tau)) + \sum_{i=1}^{N} \lambda_i p_{i,2}(t_j-\tau)\right) + e(t_j)u(t_j-\tau), \text{for } t_j > 0. \end{cases} \end{aligned} \tag{17}$$

The above system in matrix notations as given by

$$M\lambda = B, \tag{18}$$

where

$$M = \left[m_{ij}\right]_{N\times N}, \lambda = \left[\lambda_j\right]_{N\times 1}, B = \left[b_j\right]_{N\times 1},$$
$$b_j = \begin{cases} a(t_j)\alpha_1 + b(t_j)(\alpha_2 + \alpha_1 t_j) + c(t_j)u(t_j) + d(t_j)\phi(t_j - \tau) + e(t_j)u(t_j - \tau) \text{ for } t_j < 0, \\ a(t_j)\alpha_1 + b(t_j)(\alpha_2 + \alpha_1 t_j) + c(t_j)u(t_j) + d(t_j)(\alpha_2 + \alpha_1(t_j - \tau) + e(t_j)u(t_j - \tau) \text{ for } t_j > 0, \end{cases} \tag{19}$$
$$m_{ij} = \begin{cases} h_i(t_j) - a(t_j)p_{i,1}(t_j) - b(t_j)p_{i,2}(t_j) \text{ for } t_j < 0, \\ h_i(t_j) - a(t_j)p_{i,1}(t_j) - b(t_j)p_{i,2}(t_j) - d(t_j)p_{i,2}(t_j - \tau) \text{ for } t_j > 0, \end{cases}$$

Hence, λ_i is calculated as $\lambda = M^{-1}B$. This is a linear system of order $N \times N$, which is solved by the Gauss elimination technique. By putting these λ_i's in Eq. (15), we get the required solution of second-order DDEs defined in (1).

4. Numerical Examples

Let w_{ap} be approximate and w_{ex} is exact solution for CPs and GPs, then maximum absolute L_{cp} error is $L_{cp} = \max |w_{exc} - w_{apc}|$. The M_{cp} root mean square errors at CPs is $M_{cp} = \sqrt{(1/N)(\Theta|w_{exc} - w_{apc}|^2)}$. In CPs, convergence rate R_{cp} is $R_{cp} = \log [w_{apc}(N/2)/w_{apc}(N)]/\log 2$.

Example 1. Consider DDE of second order [8]

$$w''(t) = -\frac{1}{2}w(t) + \frac{1}{2}w(t - \pi), 0 \le t \le 8\pi, \tag{20}$$

with delay condition

$$\mathrm{w(t)} = \mathrm{sint}, -\pi \le \mathrm{t} \le 0, \tag{21}$$

and initial condition

$$\mathrm{w}(0) = 0, \mathrm{w}'(0) = 1. \tag{22}$$

The analytical solution is $w(t) = \sin t$.

Example 2. Consider the following second-order DDE [1]

$$w''(t) = -\frac{1}{2}w(t) + \frac{1}{2}w(t - \pi), t \in [0, \pi], \tag{23}$$

with delay condition

$$w(t) = 1 - \sin t, -\pi \le t \le 0, \tag{24}$$

and initial condition

$$w'(0) = -1, w(0) = 1. \tag{25}$$

The exact solution is $w(t) = 1 - \sin t$.

Example 3. Consider the following second-order DDE [1]

$$w''(t) = w(t - \pi), t \in [0, \pi], \tag{26}$$

with delay condition

$$w(t) = \sin t, -\pi \le t \le 0, \tag{27}$$

and initial condition

$$w'(0) = 1, w(0) = 0. \tag{28}$$

The exact solution is $w(t) = \sin t$.

Example 4. Consider DDE of second order as [36]

$$w''(t) = -w(t) + w(t - 1), 0 \le t \le 2, \tag{29}$$

with delay condition

$$w(t) = t^2 + 3t + 2, -1 \le t \le 0, \tag{30}$$

and initial condition

$$w(0) = 2, w'(0) = 0. \tag{31}$$

The exact solution is $w(t) = t^2 + t - 2 + 4\cos t - \sin t, 0 \le t \le 1$.

Example 5. Consider the following second-order DDE [8]

$$w''(t) = -\frac{\sin t}{2 - \sin t}w(t - \pi), 0 \le t \le 8\pi, \tag{32}$$

Table 1: L_{cp}, $R_c(N)$, M_{cp}, and CPU time (seconds) for Example 1.

J	$N = 2^{J+1}$	L_{cp}	$R_c(N)$	M_{cp}	$R_c(N)$	CPU time (seconds)
1	4	3.97951×10^{-03}		2.69531×10^{-03}		0.00158
2	8	1.04840×10^{-03}	1.9326	6.78403×10^{-04}	1.92440	0.00193
3	16	2.68212×10^{-04}	1.9563	1.69886×10^{-04}	1.96674	0.00461
4	32	6.77789×10^{-05}	1.9850	4.24894×10^{-05}	1.98446	0.01682
5	64	1.70330×10^{-05}	1.9936	1.06234×10^{-05}	1.99250	0.06525
6	128	4.26915×10^{-06}	1.9966	2.65593×10^{-06}	1.99631	0.20768
7	256	1.06864×10^{-06}	2.0068	6.63892×10^{-07}	1.99817	0.81604
8	512	2.67328×10^{-07}	1.9892	1.65997×10^{-07}	1.99909	3.27512
9	1024	7.68531×10^{-08}	1.7977	4.14994×10^{-08}	1.79843	13.0245

Table 2: L_{cp}, $R_c(N)$, M_{cp}, and CPU time (seconds) for Example 2.

J	$N = 2^{J+1}$	L_{cp}	$R_c(N)$	M_{cp}	$R_c(N)$	CPU time (seconds)
1	4	3.97951×10^{-03}		2.69531×10^{-03}		0.00158
2	8	1.04840×10^{-03}	1.9326	6.78403×10^{-04}	1.92440	0.00193
3	16	2.68212×10^{-04}	1.9563	1.69886×10^{-04}	1.96674	0.00461
4	32	6.77789×10^{-05}	1.9850	4.24894×10^{-05}	1.98446	0.01682
5	64	1.70330×10^{-05}	1.9936	1.06234×10^{-05}	1.99250	0.06525
6	128	4.26915×10^{-06}	1.9966	2.65593×10^{-06}	1.99631	0.20768
7	256	1.06864×10^{-06}	2.0068	6.63892×10^{-07}	1.99817	0.81604
8	512	2.67328×10^{-07}	1.9892	1.65997×10^{-07}	1.99909	3.27512
9	1024	7.68531×10^{-08}	1.7977	4.14994×10^{-08}	1.79843	13.0245

Table 3: L_{cp}, $R_c(N)$, M_{cp}, and CPU time (seconds) for Example 3.

J	$N = 2^{J+1}$	L_{cp}	$R_c(N)$	M_{cp}	$R_c(N)$	CPU time (seconds)
1	4	4.26909×10^{-03}		2.83925×10^{-03}		0.00140
2	8	1.13460×10^{-03}	1.9145	7.14201×10^{-04}	1.91174	0.00166
3	16	2.91783×10^{-04}	1.9572	1.78823×10^{-04}	1.95921	0.00334
4	32	7.39440×10^{-05}	1.9774	4.47228×10^{-05}	1.98039	0.01211
5	64	1.86096×10^{-05}	1.9903	1.11817×10^{-05}	1.99038	0.03886
6	128	4.66777×10^{-06}	1.9969	2.79551×10^{-06}	1.99524	0.14378
7	256	1.16886×10^{-06}	2.0062	6.98881×10^{-07}	1.99763	0.53690
8	512	2.92454×10^{-07}	1.9901	1.74720×10^{-07}	1.99882	2.59874
9	1024	7.31435×10^{-08}	1.9980	4.36801×10^{-08}	1.99940	11.0023

with delay condition

$$w(t) = 2 + \sin t - \pi \le t \le 0, \tag{33}$$

and initial condition

$$w'(0) = 1, w(0) = 2. \tag{34}$$

The exact solution is $w(t) = 2 + \sin t$.

5. Discussion

The second-order derivative in DDE is expressed as Haar function and the value of the first derivative is obtained by the process of integration. By applying the HCT, we obtain a system of linear equations by substituting CPs. The method of Gauss elimination is used for this system. Finally, by utilizing these coefficients, the solution at CPs is obtained. L_{cp} and M_{cp} errors for different number of CPs are given in Tables.

Table 4: L_{cp}, $R_c(N)$, M_{cp}, and CPU time (seconds) for Example 4.

J	$N = 2^{J+1}$	L_{cp}	$R_c(N)$	M_{cp}	$R_c(N)$	CPU time (seconds)
1	4	7.24881×10^{-03}		4.60001×10^{-03}		0.00197
2	8	1.95251×10^{-03}	1.8925	1.14978×10^{-03}	2.01260	0.00277
3	16	5.05791×10^{-04}	1.9491	2.87450×10^{-04}	1.98991	0.04996
4	32	1.28646×10^{-04}	1.9801	7.18631×10^{-05}	1.99899	0.01505
5	64	3.24351×10^{-05}	1.9821	1.79658×10^{-05}	2.00402	0.05394
6	128	8.14288×10^{-06}	1.9929	4.49145×10^{-06}	1.99517	0.19726
7	256	2.03997×10^{-06}	2.0035	1.12286×10^{-06}	2.00321	0.77262
8	512	5.10526×10^{-07}	1.9929	2.80716×10^{-07}	2.00000	3.07790
9	1024	1.27698×10^{-07}	2.0057	7.01790×10^{-08}	1.99794	12.3348

Table 5: L_{cp}, $R_c(N)$, M_{cp}, and CPU time (seconds) for Example 5.

J	$N = 2^{J+1}$	L_{cp}	$R_c(N)$	M_{cp}	$R_c(N)$	CPU time (seconds)
1	4	4.26909×10^{-03}		2.83925×10^{-03}		0.00196
2	8	1.13460×10^{-03}	1.9145	7.14201×10^{-04}	1.9868	0.00301
3	16	2.91783×10^{-04}	1.9572	1.78823×10^{-04}	2.0040	0.00947
4	32	7.39440×10^{-05}	1.9774	4.47228×10^{-05}	1.9935	0.02115
5	64	1.86096×10^{-05}	1.9903	1.11817×10^{-05}	2.0097	0.06661
6	128	4.66777×10^{-06}	1.9969	2.79551×10^{-06}	1.9922	0.25637
7	256	1.16886×10^{-06}	2.0062	6.98881×10^{-07}	1.9990	1.03078
8	512	2.92454×10^{-07}	1.9901	1.74720×10^{-07}	2.0041	4.09742
9	1024	7.31435×10^{-08}	1.9980	4.36801×10^{-08}	1.9967	16.4587

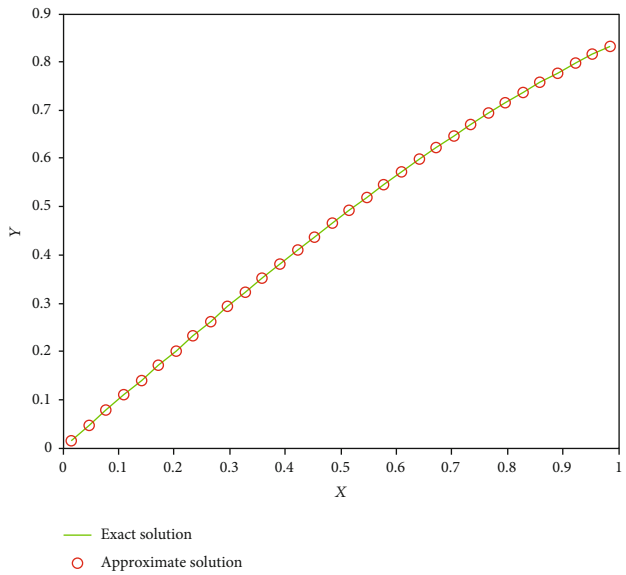


Figure 1: Comparison of both exact and approximate solution for 32 CPs for Example 1.

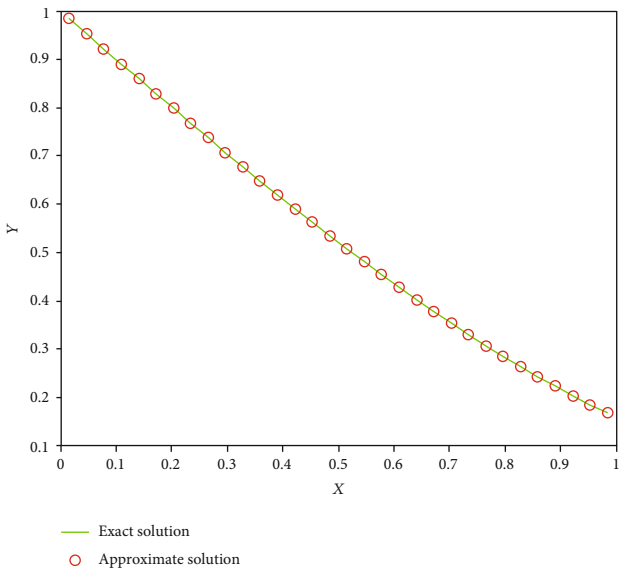


Figure 2: Comparison of both exact and approximate solution for 32 CPs for Example 2.

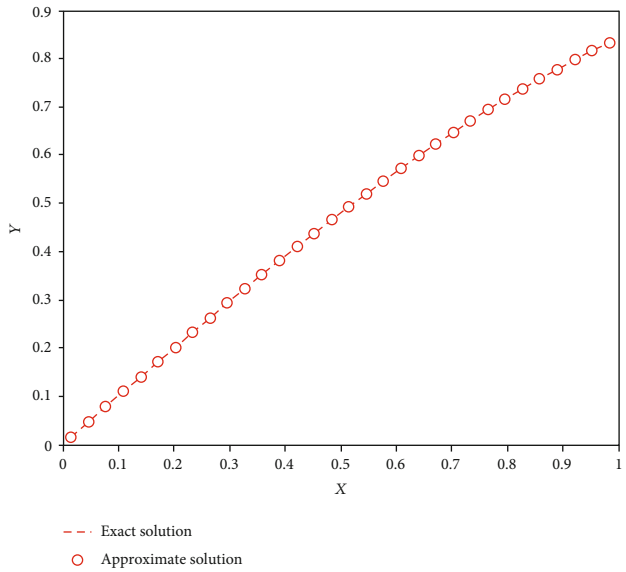


Figure 3: Comparison of both exact and approximate solution for 32 CPs for Example 3.

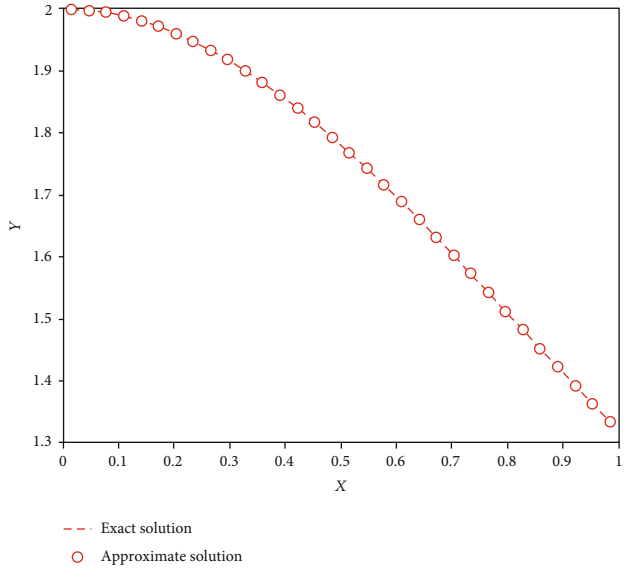


Figure 4: Comparison of both exact and approximate solution for 32 CPs for Example 4.

$R_c(N)$ and CPU time (seconds) are also reported in tables for each example. For Example 1, L_{cp} and M_{cp} errors for different number of CPs are shown in Table 1. Table 2 shows the errors for different number of CPs for Example 2, Table 3 represents errors for different number of CPs for Example 3, Table 4 shows the errors for different number of CPs for Example 4, and Table 5 shows the errors for different number of CPs for Example 5. All errors are decreased by taking more CPs. R_{cp} is determined which is nearly equal to 2, supporting the results shown in [37] by Majak et al. The comparison of both numerical and analytical solution at $N = 32$ CPs is also shown in figures. Figure 1 represents the comparison of approximate and exact solution for Example 1, Figure 2 represents the comparison of approximate and exact solution for

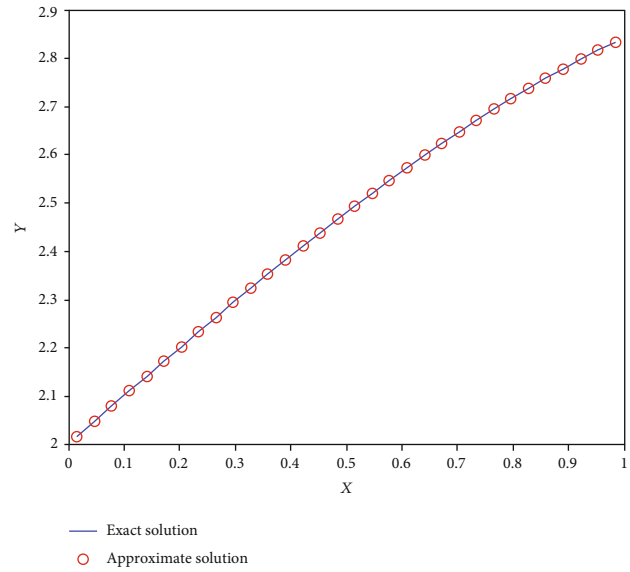


Figure 5: Comparison of both exact and approximate solution for 32 CPs for Example 5.

Example 2, Figure 3 represents the comparison of approximate and exact solution for Example 3, Figure 4 represents the comparison of approximate and exact solution for Example 4, and Figure 5 represents the comparison of exact and approximate solution for Example 5.

6. Conclusion

HCT scheme is devoted for the solution of second-order DDEs to deal with the experimentation of IIoT. The Haar technique is applied to linear DDEs for dealing with the experimentation of the Internet of Industrial Things. The Matlab software is utilized to experiment the Haar wavelet technique with the examples, and the numerical solution is compared with the exact solution. We compare the obtained solution with the exact solution and also we compute the L_{cp} and M_{cp} errors to show the accuracy of the Haar wavelet technique. We give some test problems for the illustration of our results. The experimental rate $R_c(N)$ of convergence for different number of CPs is also calculated which is approximately equal to 2. The results show that HCT is efficient for solving second-order DDEs. Our future work addresses to overcome the limitation of this study. Moreover, we will apply this technique to high order DDEs and system of DDEs.

Data Availability

No data is available.

Conflicts of Interest

There is no conflicting interest.

References

[1] H. Y. Seong and Z. A. Majid, "Solving second order delay differential equations using direct two-point block method," *Ain Shams Engineering Journal*, vol. 8, no. 1, pp. 59–66, 2017.

[2] S. Sedaghat, Y. Ordokhani, and M. Dehghan, "Numerical solution of the delay differential equations of pantograph type via Chebyshev polynomials," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 4815–4830, 2012.

[3] S. Nazir, Y. Ali, N. Ullah, and I. García-Magariño, "Internet of things for healthcare using effects of mobile computing: a systematic literature review," *Wireless Communications and Mobile Computing, page*, vol. 2019, article 5931315, pp. 1–20, 2019.

[4] R. Amin, S. Nazir, and I. García-Magariño, "A collocation method for numerical solution of nonlinear delay integro-differential equations for wireless sensor network and internet of things," *Sensor*, vol. 20, no. 7, article 1962, 2020.

[5] M. A. Ibrahim, A. El-Safty, and S. M. Abo-Hasha, "2*h*-Step spline method for the solution of delay differential equations," *Computers & Mathematics with Applications*, vol. 29, no. 8, pp. 1–6, 1995.

[6] J. O. Ehigie, S. A. Okunuga, and A. B. Sofoluwe, "3-point block methods for direct integration of general second-order ordinary differential equations," *Advances in Numerical Analysis*, vol. 2011, Article ID 513148, 14 pages, 2011.

[7] M. Ghasemi and M. Tavassoli Kajani, "Numerical solutions of time-varying delay system by chebyshev wavlets," *Applied Mathematical Modelling*, vol. 35, pp. 5235–5244, 2011.

[8] S. Z. Ahmad, F. Ismail, and N. Senu, "Solving oscillatory delay differential equations using block hybrid methods," *Journal of Mathematics*, vol. 2018, Article ID 2960237, 7 pages, 2018.

[9] S. A. Okunuga and J. Ehigie, "A new derivation of continuous collocation multistep methods using power series as basis function," *Journal of the Nigerian Association of Mathematical Physics*, vol. 3, pp. 43–50, 2009.

[10] R. L. Brown, "Some characteristics of implicit multistep multiderivative integration formulas," *SIAM Journal on Numerical Analysis*, vol. 14, pp. 982–993, 1977.

[11] F. Ismail, L. K. Yap, and O. Mohamad, "Explicit and implicit 3-point block methods for solving special second order ordinary differential equations directly," *International Journal of Mathematical Analysis*, vol. 3, no. 5, pp. 239–254, 2009.

[12] J. O. Ehigie, S. A. Okunuga, and A. B. Sofoluwe, "On generalized 2-step continuous linear multistep method of hybrid type for the integration of second order ordinary differential equations," *Archives of Applied Science Research*, vol. 2, pp. 362–372, 2010.

[13] J. B. Rosser, "A runge-kutta for all seasons," *SIAM Review*, vol. 9, no. 3, pp. 417–452, 1967.

[14] W. L. Chen and Y. P. Shi, "Shift walsh matrix and delay differential equations," *IEEE Transactions on Automatic Control*, vol. 23, pp. 265–280, 1978.

[15] H. J. Oberle and H. J. Pesh, "Numerical treatment of delay differential equations by hermite interpolation," *Numerische Mathematik*, vol. 37, no. 2, pp. 235–255, 1981.

[16] H. Arnt, "Numerical solution of retarded initial-value problems: Local and global error and stepsize control," *Numerische Mathematik*, vol. 43, no. 3, pp. 343–360, 1984.

[17] A. Bellen, "One-step collocation for delay differential equations," *Journal of Computational and Applied Mathematics*, vol. 10, no. 3, pp. 275–283, 1984.

[18] H. C. San, Z. A. Majid, and M. Othman, "Solving delay differential equations using coupled block method," in *2011 Fourth International Conference on Modeling, Simulation and Applied Optimization*, p. 11, Kuala Lumpur, Malaysia, April 2011.

[19] Z. A. Majid, N. A. Azmi, and M. Suleiman, "Solving second order ordinary differential equations using two point four step direct implicit block method," *European Journal of Scientific Research*, vol. 31, pp. 29–36, 2009.

[20] T. Allahviranloo, N. Ahmady, and E. Ahmady, "Numerical solution of fuzzy differential equations by predictor-corrector method," *Information Sciences*, vol. 177, no. 7, pp. 1633–1647, 2007.

[21] S. Abbasbandy and T. Allahviranloo, "Numerical solutions of fuzzy differential equations by taylor method," *Computational Methods in Applied Mathematics*, vol. 2, no. 2, pp. 113–124, 2002.

[22] S. L. Chang and L. A. Zadeh, "On fuzzy mapping and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 2, pp. 30–34, 1972.

[23] H. Jafari, M. Saeidy, and D. Baleanu, "The variational iteration method for solving n-th order fuzzy differential equations," *Central European Journal of Physics*, vol. 10, no. 1, pp. 76–85, 2012.

[24] Z. M. Odibat, "A study on the convergence of variational iteration method," *Mathematical and Computer Modelling*, vol. 51, no. 9-10, pp. 1181–1192, 2010.

[25] C. Hwang and M. Y. Chen, "Analysis of time-delay systems using the Galerkin method," *International Journal of Control*, vol. 44, no. 3, pp. 847–866, 1986.

[26] W. Liu, G. Huang, A. Zheng, and J. Liu, "Research on the optimization of iiot data processing latency," *Computer Communications*, vol. 151, pp. 290–298, 2020.

[27] R. Sahal, J. G. Breslin, and M. I. Ali, "Big data and stream processing platforms for industry 4.0 requirements mapping for a predictive maintenance use case," *Journal of Manufacturing Systems*, vol. 54, pp. 138–151, 2020.

[28] W. Z. Khan, M. H. Rehman, H. M. Zangoti, M. K. Afzal, N. Armi, and K. Salah, "Industrial internet of things: recent advances, enabling technologies and open challenges," *Computers and Electrical Engineering*, vol. 81, p. 106522, 2020.

[29] N. Gulati and P. D. Kaur, "Towards socially enabled internet of industrial things: architecture, semantic model and relationship management," *Ad Hoc Networks*, vol. 91, article 101869, 2019.

[30] M. M. Khashan, R. Amin, and M. I. Syam, "A new algorithm for fractional riccati type differential equations by using haar wavelet," *Mathematics*, vol. 7, no. 6, p. 545, 2019.

[31] S. Nazir, S. Shahzad, R. Wirza et al., "Birthmark based identification of software piracy using haar wavelet," *Mathematics and Computers in Simulation*, vol. 166, pp. 144–154, 2019.

[32] R. Amin, S. Nazir, and I. G. Magariño, "Efficient sustainable algorithm for numerical solution of nonlinear delay fredholm-volterra integral equations via haar wavelet for dense sensor networks in emerging telecommunications," *Transactions on Emerging Telecommunications Technologies*, no. article e3877, 2020.

[33] T. Abdeljawad, R. Amin, K. Shah, Q. Al-Mdallal, and F. Jarad, "Efficient sustainable algorithm for numerical solutions of

systems of fractional order differential equations by haar wavelet collocation method," *Alexandria Engineering Journal*, vol. 59, no. 4, pp. 2391–2400, 2020.

[34] R. Amin, K. Shah, M. Asif, I. Khan, and F. Ullah, "An efficient algorithm for numerical solution of fractional integro-differential equations via Haar wavelet," *Journal of Computational and Applied Mathematics*, vol. 381, p. 113028, 2021.

[35] I. Aziz and R. Amin, "Numerical solution of a class of delay differential and delay partial differential equations via haar wavelet," *Applied Mathematical Modelling*, vol. 40, no. 23-24, pp. 10286–10299, 2016.

[36] H. M. Radzi, Z. A. Majid, F. Ismail, and M. Suleiman, "Two and three point one-step block methods for solving delay differential equations," *Journal of Quality Measurement and Analysis*, vol. 8, pp. 29–41, 2012.

[37] J. Majak, B. S. Shvartsman, M. Kirs, M. Pohlak, and H. Herranen, "Convergence theorem for the Haar wavelet based discretization method," *Composite Structures*, vol. 126, pp. 227–232, 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6693316, 13 pages
https://doi.org/10.1155/2021/6693316

Research Article

Performance Improvements of AODV by Black Hole Attack Detection Using IDS and Digital Signature

Md Ibrahim Talukdar[1], **Rosilah Hassan**[2], **Md Sharif Hossen**[1], **Khaleel Ahmad**[3], **Faizan Qamar**[2] and **Amjed Sid Ahmed**[4]

[1]*Department of Information and Communication Technology (ICT), Comilla University, Cumilla, Bangladesh*
[2]*Centre for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), 43600 UKM Bangi, Selangor, Malaysia*
[3]*Department of Computer Science and Information Technology, Maulana Azad National Urdu University, Hyderabad, India*
[4]*Computing Department, Engineering Faculty, Global College of Engineering and Technology, Oman*

Correspondence should be addressed to Md Sharif Hossen; mshossen@cou.ac.bd, Khaleel Ahmad; khaleelahmad@manuu.edu.in, and Faizan Qamar; faizanqamar@ukm.edu.my

Received 24 November 2020; Revised 3 January 2021; Accepted 13 February 2021; Published 2 March 2021

Academic Editor: Ihsan Ali

In mobile ad hoc networks (MANETs), mobile devices connect with other devices wirelessly, where there is no central administration. They are prone to different types of attacks such as the black hole, insider, gray hole, wormhole, faulty node, and packet drop, which considerably interrupt to perform secure communication. This paper has implemented the denial-of-service attacks like black hole attacks on general-purpose ad hoc on-demand distance vector (AODV) protocol. It uses three approaches: normal AODV, black hole AODV (BH_AODV), and detected black hole AODV (D_BH_AODV), wherein we observe that black holes acutely degrade the performance of networks. We have detected the black hole attacks within the networks using two techniques: (1) intrusion detection system (IDS) and (2) encryption technique (digital signature) with the concept of prevention. Moreover, normal AODV, BH_AODV, and D_BH_AODV protocols are investigated for various quality of service (QoS) parameters, i.e., packet delivery ratio (PDR), delay, and overhead with varying the number of nodes, packet sizes, and simulation times. The NS2 software has been used as a simulation tool to simulate existing network topologies, but it does not contain any mechanism to simulate malicious protocols by itself; therefore, we have developed and implemented a D_BH_AODV routing protocol. The outcomes show that the proposed D_BH_AODV approach for the PDR value delivers around 40 to 50% for varying nodes and packets. In contrast, the delay decreases from 300 to 100 ms and 150 to 50 ms with an increase in the number of nodes and packets, respectively. Furthermore, the overhead changes from 1 to 3 for various nodes and packet values. The outcome of this research proves that the black hole attack degrades the overall performance of the network, while the D_BH_AODV enhances the QoS performance since it detects the black hole nodes and avoids them to establish the communication between nodes.

1. Introduction

Wireless sensor network (WSN) is an interesting research nowadays in the field of communication. The improvement of tiny-structured, resource-constraint, cost-effective sensors is getting simpler. Also, they seem to be able to perceive the parameters of the environment, accumulate relevant data from the area, and convey information to the users. The Internet of Things (IoT) represents a major and significant component for the 4.0 industrial revolution, and its implementation requires extensive research to ensure that it will operate appropriately [1]. Wireless networks are classified as infrastructure-based networks with a central access point and ad hoc with no access point. Mobile ad hoc network (MANET) is a dynamic network without fixed infrastructure due to its wireless nature that can be deployed as multihop

packet networks. It is a wireless network and has a dynamic topology due to its mobility nature [2]. Also, there is no fixed infrastructure, and each node can act as a source, a destination, or a bridge to forward information packets for the nodes that are out of the transmission range [3–5]. These nodes or devices can have different speeds, transmission ranges, data rates, and packet sizes. Some unique characteristics of MANET are autonomous, dynamic topology, multihop, etc. [6]. These networks are also constrained to transmission ranges, packet losses, security, QoS, etc. Moreover, routing is a fundamental requirement to establish a basic communication among various nodes.

MANET protocols can be described as reactive, proactive, and hybrid in general [7]. The primary function of routing in MANETs is to establish routes among different mobile nodes that satisfy QoS requirements such as bandwidth and end-to-end delay and can be able to operate within the limited energy constraints [8, 9]. There are various kinds of MANET protocols including AODV, dynamic source routing (DSR), destination sequenced distance vector (DSDV), reverse-AODV (RAODV), ad hoc on-demand multipath distance vector (AOMDV), and temporarily ordered routing algorithm (TORA) [10]. The general-purpose AODV is chosen for black hole simulation because it outperforms other reactive routing protocols under important QoS parameters [11, 12]. In fact, the AODV and DSR protocols are two of the most on-demand protocols used in MANETs [13]. Moreover, it combines both DSR and DSDV routing protocols and gets the advantages of both of them [14, 15]. The AODV [16] is a reactive routing protocol that follows route discovery and route maintenance mechanisms and guarantees a loop-free routing by using sequence numbers.

The infrastructureless architecture makes MANETs to numerous attacks [17] such as denial-of-service (DoS) attacks, which create the worst impact on energy consumption [18]. An approach in [19] was focusing on designing an energy-efficient cluster-based on queen-bee (QB) algorithm for wireless sensor networks. This algorithm's high rate results in premature convergence that improves the capability of finding the optimum value of the local minimum. It considers normal and strong mutation, so the diversity of children will be higher and premature divergence can be neglected. The outcome proves that the proposed QB algorithm delivers better results than the genetic algorithm (GA) in terms of energy efficiency that ultimately helps increase the network's lifetime. The authors in [20] proposed a new hierarchical clustering algorithm (HCAL) and corresponded protocol for large-scale MANETs (LMANET). The idea is to utilize the combined weight matric of both table-driven and on-demand routing in order to locate a dominant set of nodes. The interlink between the LMANET has been established by using the node's relative degree and link expiration time. The results have been evaluated in terms of delay, total rounds of cluster head, cluster head time, overhead, and PDR. The outcome of the proposed HCAL protocol outperforms. They are compared with various routing approaches such as dynamic Doppler velocity clustering, signal characteristic-based clustering, dynamic link duration clustering, and mobility-based clustering algorithms.

The black hole is one of the fatal attacks which acts like a hole that destroys all data packets by itself [21]. The malicious nodes also interrupt the route discovery that causes network packets to be absorbed by the attacker. In route discovery of AODV, the intermediate nodes are liable for finding a correct route to the destination by sending "hello" packets to the neighbors. Whereas in AODV, malicious nodes instantly respond to the source with a false route reply as if it has a correct route to the destination instead of forwarding discovery packets to neighboring nodes [22]. Consequently, the source node immediately forwards its data packets to the destination node through the malicious node, presuming it is an actual route. As a result, the network is affected by a black hole attack where malicious nodes are knowingly misbehaving and damaging the node interface. In general, nodes in the network will restlessly be trying to find a path for the destination, which makes the node consume its resources and lose packets [23].

This paper has implemented black hole attacks on general-purpose AODV protocol with three approaches: normal AODV, black hole AODV (BH_AODV), and detected black hole AODV (D_BH_AODV), wherein we observe that black holes acutely degrade the performance of networks. Hence, we have detected the attackers within the networks using two techniques, i.e., IDS and digital signature encryption technique with the concept of prevention. The IDS detects malicious nodes through the modification of AODV that requires a time stamp, and the digital signature detects malicious nodes through key comparisons. The results have been investigated for various QoS parameters, such as PDR, delay, and overhead with varying the number of nodes, packet sizes, and simulation times.

The rest of the paper is structured as follows. Section 2 includes the related works. In Section 3, we discuss the various issues related to MANET security. Section 4 illustrates the black hole attack and its implementation in AODV. Section 5 explains simulation tools and environment settings. Section 6 represents the black hole attack detection elaborately in AODV using IDS and digital signature. Finally, the conclusion and future remarks are shown in Sections 7 and 8, respectively.

2. Related Works

The ad hoc networks have various application areas in real-world wireless communication scenarios, including sensor networks, military fields, personal area network (PAN), and Bluetooth [24, 25]. Hence, MANET becomes an important research area to establish reliable communication among nodes in an adverse environment [26]. However, these networks fall into various security problems. In recent years, numerous methods but not limited to cryptographic techniques, modification of protocols, IDS, etc. [27, 28] have been suggested by many researchers to improve MANET security. More specifically, the authors in [29] proposed a neuro-fuzzy technique related to IDS for MANETs. The authors of [30] introduced the IDS to detect and identify attackers through the fuzzy technique. The authors in [31] proposed an enhanced trust detection algorithm to improve the detection

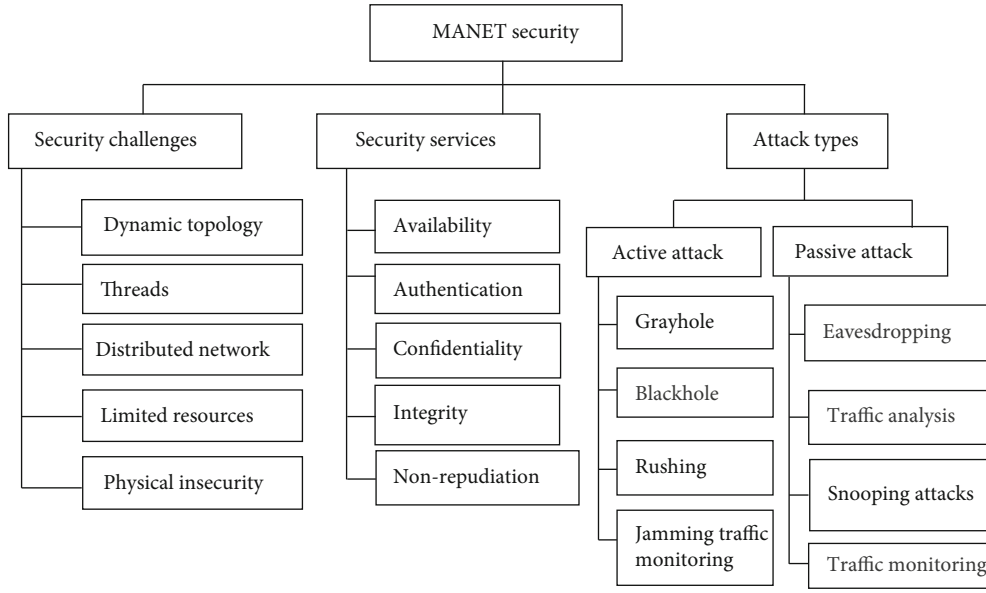


Figure 1: Security issues in MANET.

and prevention probability of black hole attackers in MANETs. This method skips the black hole nodes in MANETs, increases the network throughput, and reduces the packet loss as well as the power consumption in the presence of malicious black hole nodes. Another approach in [32] proposed an efficient detection approach that shows low overhead to the network. This approach enhances the delivery ratio by 45.6% for dense networks and 41% for sparse networks. Besides, it improves the dropped packet by 75% for dense networks and 63% for sparse networks. An approach in [33] proposed the honeypot-based security solution which uses cross-layer security to ensure better packet delivery with minimum packet dropped and decrease end to end to delay and network load. Similarly, the authors [34] proposed a new protocol based on a dynamic destination sequence number threshold value, which detects and prevents black hole nodes with a better performance than the black hole attack. Another research in [35] proposed lightweight mathematical-based concepts with less computational complexity to detect the hostile nodes and obstruct the black hole nodes in MANETs. Moreover, there are various approaches [36, 37] that improved MANET's security issues.

3. Security Issues in MANET

The security issues in MANETs are highly challenging due to no predefined boundary, adversary inside the networks, no centralized control, and limited energy resource. MANETs are affected by numerous types of threats and attacks. Various attacks but not limited to the black hole, impersonation, wormhole, eavesdropping, man-in-the-middle attack, gray hole, etc. badly interrupt routing mechanism and degrade the execution of ad hoc networks [38, 39]. These attackers are either active attacks or passive attacks [40]. Among these attackers, the black hole attack is one of the fatal attacks that has been considered in this research. Basically, there are two techniques to protect against attacks in MANETs, namely, proactive and reactive [41]. The proactive approach tries to prevent attackers from launching attacks in the initial stage through numerous cryptographic methods, whereas the reactive method follows the empirical process and responds accordingly to detect security threats. A complete security solution integrates both approaches and includes three sections, i.e., detection, prevention, and reaction. Various mitigation and prevention security approaches such as availability, confidentiality, authorization, authentication, integrity, nonrepudiation, and anonymity might be ensured to establish secure routing [42]. Figure 1 shows the security issues in MANET.

4. Black Hole Attack Implementation in AODV

The AODV protocol used in MANET suffers from a black hole attack wherein an attacker consumes the network traffic and fells all data packets [43, 44]. A black hole is an active attack wherein a malicious node awaits neighboring nodes to forward route request (RREQ) messages. When the malicious node accepts an RREQ message, it instantly sends the route reply (RREP) message of false copy to the sender with the maximum sequence number before other nodes send an actual true one. Therefore, the sender of RREQ presumes that route discovery is accomplished and begins to transmit packets to the malicious node. The black hole attack scenario is explained in Figure 2. Let nodes S, D, and B be the source, destination, and malicious node, respectively. Initially, source node S broadcasts the RREQ message for destination node D to establish a path for communication; however, the malicious node B instantly responds to source node S with a false RREP message exhibiting that it has the maximum sequence number of destination node D, though it is coming from destination node D. Presuming that the destination node D is just behind malicious node B with the single-hop count, source node S refuses the newly received RREP packets come from intermediary node N or M. The

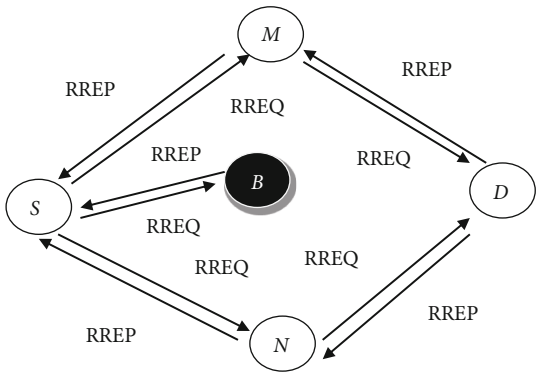


Figure 2: Black hole attack scenario.

source node *S* then begins to send out packets to the malicious node *B* and presumes that these packets will definitely reach destination node *D*; however, in actuality, malicious node *B* will fall packets and it stops forwarding any packet to any other nodes. The network operation is badly interrupted since the black hole malicious node *B* consumes all the packets.

The crucial question is that "Which node becomes a malicious node from the nodes?". One must design the attacking nodes since no node automatically acts as an attacker. In this research, we have designed black hole nodes by modifying the pair of *aodv.h* and *aodv.cc* files. The ns-2.35 simulator has been used in the study to design a network structure and to diminish the black hole node. According to the nodes' trust value, we design a node to identify that either is an attacker or not. A node may have high, medium, or low trust values (as Boolean values). We can make fewer trust value nodes as a malicious node. In this study, we have designed a black hole attack within AODV by using the five steps:

5. Simulation Tools and Environment Settings

This study utilized a discrete event simulator "NS2" to evaluate the MANET protocols. A tool named "cbrgen" under "~ns/indep-utils/cmu-scen-gen" is used to find random traffic among the nodes using transmission control protocol (TCP) or constant bit rate (CBR) connection. Moreover, "setdest" under "~ns/indep-utils/cmu-scen-gen/setdest/" is used for generating the traces of nodes by random movement with the velocity of the node to any location (not fixed) within the considered wireless region. Node's mobility is distributed in a random waypoint [45] fashion which can manually create traffic connections and node mobility for a small network. The wireless network environment is constructed using moving nodes. The CBR traffic patterns with specified simulation area, channel, time, etc. are used to design networks. Here, we have considered the random waypoint [46] as the mobility model, which adds the concept of pause time to the random walk model [47]. Table 1 shows the general simulation parameter. We have fixed the number of nodes equal to 60 when we vary the packet sizes, i.e., 512, 1000, 1800, and 2100 bytes. In contrast, we have fixed the packet size = 1000 when we change the number of mobile nodes, i.e., 20, 60, 80, and 100. Moreover, we have performed the simulations four times and then take the average results in order to calculate the packet delivery ratio, average delay or latency, and overhead ratio.

We choose AODV protocol with a specified simulation area, omnidirectional antenna, random waypoint mobility model, and CBR connection, and transmitted and received power to design the black hole attack. Then, a black hole is designed inside the class files and TCL scripts with few nodes. The TCL scripts are run with the commands "*ns blackhole_nodes.tcl*" where "blackhole_nodes" is the script's name and *.tcl* is an extension. These scripts will generate two files, namely, *nam* (*.nam*) and trace (*.tr*) files. It has then analyzed the trace files through AWK scripts, which will provide performance value such as PDR, delay, and overhead. Also, we plot the performance using xgraph. Here, *nam* is used for analyzing network simulation traces and practical packet traces. After running a *nam* file for 30 nodes, we can see the node positions and definitions according to declaration where node 16 is defined as an attacker, nodes 2, 8, 13, and 15 are defined as the source, and nodes 0, 6, 7, and 14 are defined as the destination. Here, the attacker is positioned in the middle of the network to succeed in a black hole attack. Multiple sources and destinations are linked through mesh topology while designing networks.

Figure 3 shows the network animator (*nam*) screen for the TCL of black holes. These TCL scripts generate trace (*.tr*) files that can be analyzed through AWK scripts or xgraph. After evaluating the trace files using AWK scripts, simulation results are collected and plotted into graphs. These graphs or xgraphs carefully exhibit the comparison among protocols.

The network's scalability means that with the growth of the number of nodes in the network, the algorithm maintains the same outcome for different network sizes. With the increase of the number of mobile nodes, the network size increases; hence, the proposed D_BH_AODV algorithm justifies the higher delivery, lower delay, and lower overhead. At present, only 100 nodes are used with 2100 bytes in the present scenario but can be extended for more number of nodes and packets in the future.

6. Black Hole Attack Detection and Simulation Analysis

This section has been divided into subheadings, where the black hole attack detection in AODV is using IDS and digital signature. It provides a concise and precise description of the experimental results and their interpretation, and the experimental conclusions are drawn.

According to the number of attacker nodes, the black hole attacks can be split into two types: (1) single black hole node and (2) cooperative black hole nodes. In a single black hole node, there is only one attacker exists in the network. On the contrary, a cooperative black hole node occupies multiple attackers in the network. In this paper, the black hole attacker has been identified in two ways: (1) IDS [48] or modification of the AODV protocol technique with a single attacker and (2) encryption technique such as digital

Step 1: Variable (attacker) declaration
We declare a variable malicious as Boolean within the code *aodv.cc*, and *aodv.h*, firstly modifying the code in *aodv.h* file as below:
Boolean malicious; // or BH
Step 2: Variable (attacker) initialization
We initialize the attacker variable as a false within the constructor of *aodv.cc*.
Step 3: The normal node is a black_hole (BH), what's happening to the malicious or attacker node value inside some block of code in *aodv.cc*
file command () function if (argc ==2)
add some lines of code and replace it as the below code
```
if (strcasecmp (argv[1], "black_hole") == 0)
{
attacker = true;
return TCL_OK;
}
```
Step 4: The attacker node is true what will be?
```
if (attacker == true) {
   printf ("Packets are dropped index of node and number of packets %d is as %d \n",
   index, t_count++);
   drop (p,DROP_RTR_ROUTE_LOOP); //dropped all packet based on this function
}
```
After this completion of work, open the command prompt and go to the ~ns-2.35/ then finally run the make command
$ make
If there are no mistakes in the above technique of packets dropped, your compilation and execution will be successful.
Step 5: Finally, we go running Tool Command Language (TCL) file with AODV protocol, with Attacker (BH) modified code and normal code, then comparing total experimental outcomes.
$ ns *AODV.tcl*

Algorithm 1

Table 1: General simulation parameters for black hole evaluation.

Parameters	Values
Protocol	AODV
Modified routing protocols	BH_AODV, D_BH_AODV
Mobility mode	Two-ray ground
Antenna	Omni antenna
Channel	Wireless channel
Simulation time	160 sec
Mobility model	Random waypoint
Simulation area	1100×750
Traffic	CBR
Packet size	1000 bytes
Variation of packets	512, 1000, 1800, 2100 bytes
MAC	MAC/802-11
Mobile nodes	60
Variation of nodes	20, 60, 80, 100
Mobility speed	6 m/s
Data rates	0.1 mbps
Performance metrics	PDR, delay, overhead
Simulator	NS 2.35

signature with cooperative black hole node. The following subsections discuss first detecting single black hole attackers using IDS and then detecting multiple attackers using a digital signature.

6.1. IDS and Digital Signature. Intrusion detection system (IDS) detects unwanted activities and security violations to systems [49]. The goal of IDS is to automate the intrusion detection that tries to interrupt the availability, integrity, or confidentiality. Different types of IDS, including signature-based IDS, anomaly-based IDS, and hybrid IDS, are introduced to improve MANET security [50]. At first, IDS detects the black hole nodes and tracks its route, and then it informs the sender node about the malicious node by transmitting a high sequence number so that the sender node does not use that path and does not send any message to that node and searches for a new route to establish a successful secure communication between two nodes [51].

A digital signature is an encryption technique, in which the nodes that are not verified properly are treated as a black hole and dropped. In our approach, we assign a short signature to all nodes and then each node should be verified to get a message from its neighboring nodes. If the signature is matched, then the routing table is updated; otherwise, all the updates are removed. This information is sent to all nodes in the network. In this research, it is used to verify the black hole attack within AODV. In AODV, a RREQ is forwarded

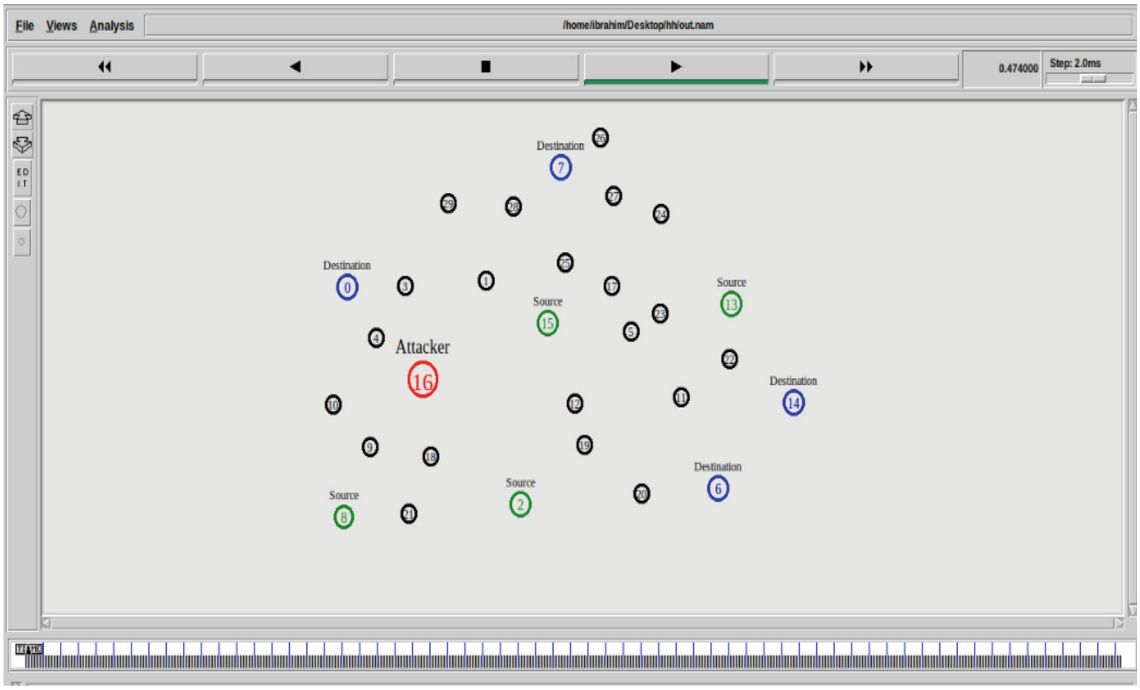


Figure 3: The *nam* scenario for black hole simulation.

to the neighboring nodes by the source node until the destination is found. The RREQ packet header retains all the visiting nodes' id while broadcasting RREQ packets to the destination. The destination node containing all of the nodes' id in its header unicasts the reply where each visiting node adds its digital signature. When the receiving node receives the packet compared to the digital signature of the previous node from its database and if the signature is matched, then that node is legitimate; otherwise, that node is considered as an attacker. Whenever an attacker node is detected, that information is broadcast to the neighbors. Hence, in this way, all packets are assigned with a digital signature to prevent the malicious attacks. Digital signature requires much more calculation overhead in signing/decrypting and verifying/encrypting for node activities. In this subsection, we have designed and detected multiple attackers through an encryption standard (digital signature) in which every node has its own key, and the packet transmission has been performed when the key is in a valid state. In this case, both source and destination nodes will exchange the keys before the packet transmission. If the node's key is found to be in a valid state, that node will be considered a trusted node to start a packet transmission. As the attacker node is not aware of the correct key for the transmission, it cannot get any packet from the source node, which ultimately enhances the overall network performance.

6.2. Black Hole Attack Detection Using IDS and Analysis. In this research, a black hole attacker is implemented inside the AODV protocol by modifying AODV using the node's trust value. To mitigate the black hole attacks, a trust-based mechanism has been used to analyze the packets dropped within the time stamp given on TCL code if the black hole is true. The various QoS parameters, such as PDR, delay, and overhead, have been analyzed through node and packet variations as shown in Figure 4 and Figure 5, respectively. The comparison has been shown among normal AODV, BH_AODV, and D_BH_AODV to investigate protocols' performance.

For node variations (20, 60, 80, and 100), PDR is increasing when the number of nodes is greater than 60 as shown in Figure 4. Moreover, for varying the packet sizes, i.e., 512, 1000, 1800, and 2100, D_BH_AODV outperforms BH_AODV and lags than normal AODV in the case of PDR as shown in Figure 5. In both cases, i.e., for varying the number of nodes and packet sizes, normal AODV shows good delivery. In this research, our consideration is the presence of black hole attacks through communication among nodes in the network. Hence, from both Figures 4 and 5, BH_AODV degrades the delivery of normal AODV for varying the nodes and packet sizes, whereas our proposed D_BH_AODV shows greater delivery than BH_AODV.

It is very often that the delay measurement shows irregularities in performance exhibition. Even small changes in parameters significantly affect the performances. For varying the number of mobile nodes and the packet sizes as shown in Figures 6 and 7, we can see that normal AODV exhibits higher average delay than BH_AODV and D_BH_AODV. However, for both cases, as shown in Figures 6 and 7, it is obvious that D_BH_AODV shows a lower average delay than normal AODV but exhibits a higher delay compared to BH_AODV.

In general, high overhead degrades network performance because the number of packet replication increases to send a successful packet to the destination. Hence, it increases the transmission cost also. So, low overhead is preferable. The black hole attack increases the rate of packet replication. For varying the number of nodes and packet sizes shown in

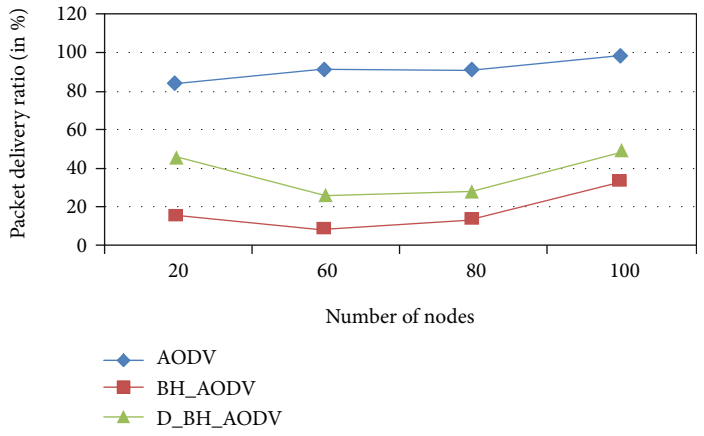


Figure 4: PDR with varying nodes.

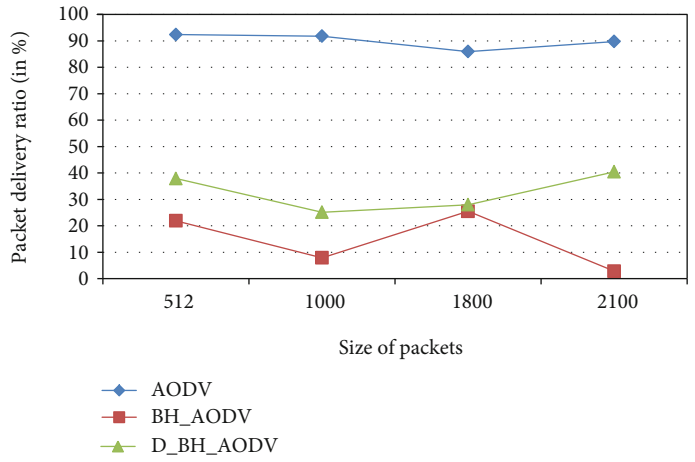


Figure 5: PDR with varying packets.

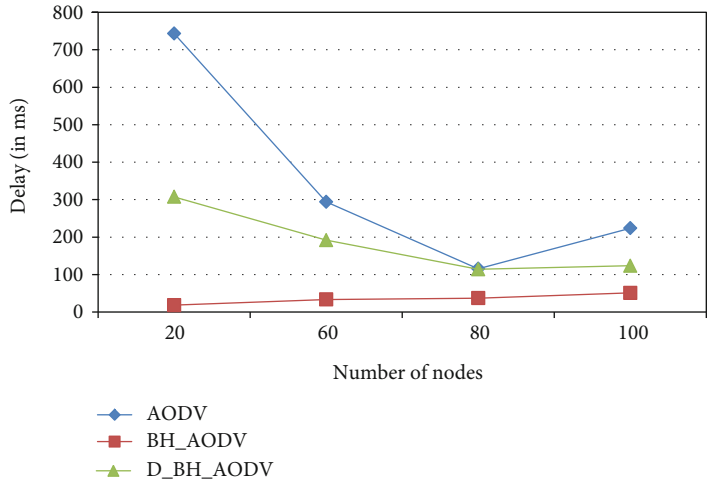


Figure 6: Delay with varying nodes.

Figures 8 and 9, normal AODV exhibits lower overhead because we do not consider an attacker node. While BH_AODV receives the packets from neighboring nodes and changes its contents, then, the sender nodes again send their copies. Hence, the BH_AODV results in higher overhead. D_BH_AODV detects and prevents the black hole nodes from sending copies. So, D_BH_AODV exhibits lower overhead than BH_AODV as shown in Figures 8 and 9.

In this research, the nodes are placed randomly on a specified simulation area and the environment is simulated

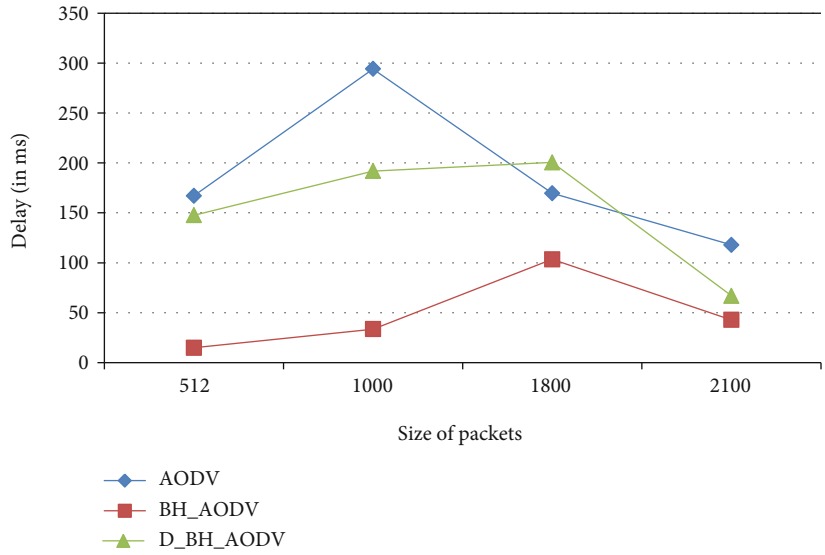


Figure 7: Delay with varying packets.

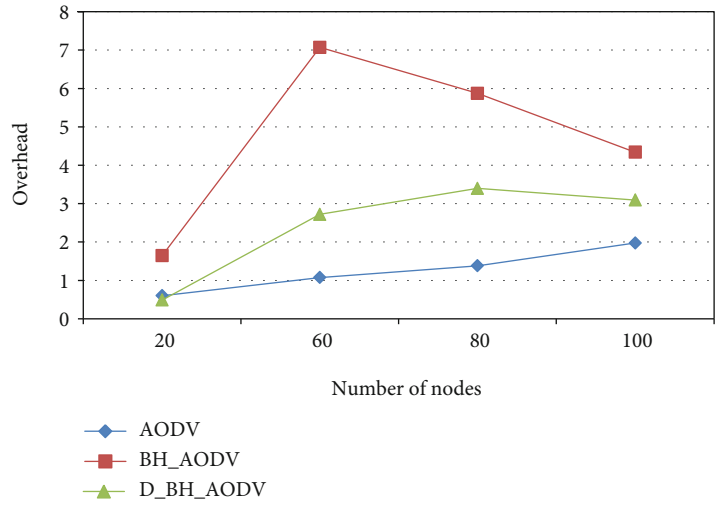


Figure 8: Overhead with varying nodes.

through RWP (random waypoint) mobility model. The simulation results depict the state of normal AODV, BH_AODV, and D_BH_AODV for the same network environment. Here, the graphical analyses depict that for varying the number of nodes and packet sizes, the black holes degrade the performance over normal AODV and the D_BH_AODV outperforms BH_AODV in terms of PDR, delay, and overhead ratio. However, in some cases, we cannot find the expected simulation results due to nodes' misbehaving [52]. Misbehaving nodes also known as selfish nodes have full access to the medium that tries to get favored from other nodes but ignoring to forward other node packets can severely reduce the whole network's performance. In this case, the prevention of different attackers is the proper solution to network improvement.

6.3. Black Hole Attack Detection in AODV Using Digital Signature and Analysis. In this section, we will see the comparison between the BH_AODV and the D_BH_AODV using xgraph. Here, an effective encryption technique (digital signature) is used to detect black hole attacks. As the previous discussion of generating trace files in Figure 3, the TCL scripts of the BH_AODV and D_BH_AODV are evaluated and xgraphs are plotted to sketch the QoS parameters with respect to the simulation time in second(s). Figures 10, 11, and 12, respectively, show the improved performance of D_BH_AODV over BH_AODV attack in the AODV protocol using xgraph in terms of PDR, average latency or delay, and overhead ratio with varying the simulations times.

For varying the simulation times in seconds as shown in Figure 10, we observe that D_BH_AODV exhibits higher packet delivery ratio compared to the black hole AODV because D_BH_AODV detects the attackers' nodes and does not forward the packets to the black hole attackers' nodes. Hence, it ensures higher delivery (Figure 10) by lowering the packet replications, i.e., lower overhead ratio (Figure 12).

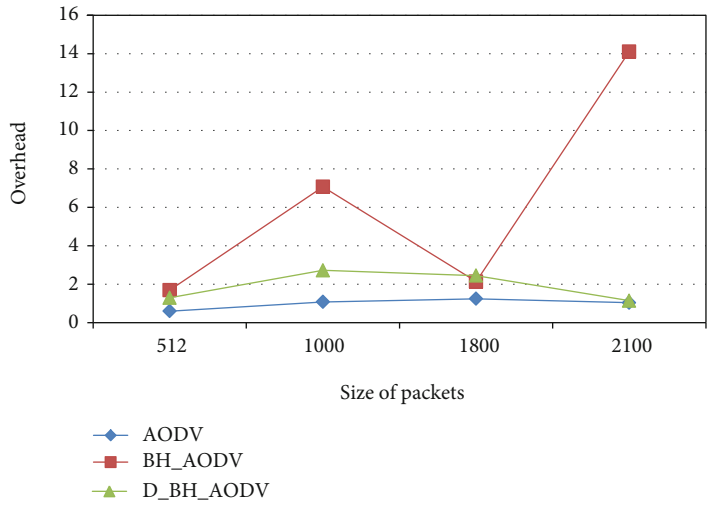


Figure 9: Overhead with varying packets.

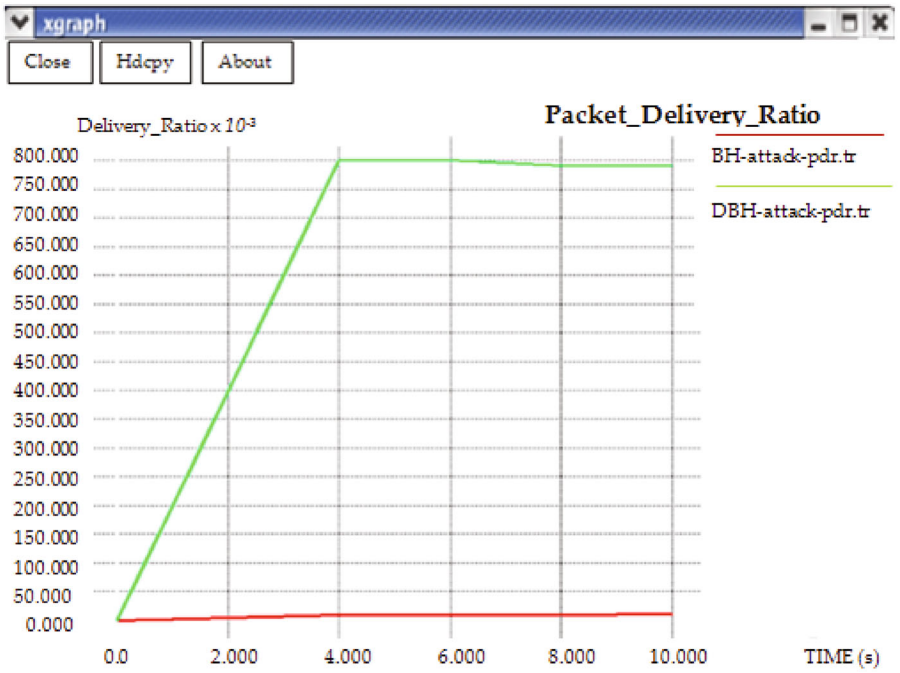


Figure 10: PDR with varying time for black hole and detected black hole attack.

For varying the simulation times in seconds as shown in Figure 11, we observe that the average delay of the BH_AODV is very high compared to the average delay of D_BH_AODV. Therefore, D_BH_AODV significantly improves wireless networks' performance by ensuring higher delivery (Figure 10) and lower delay (Figure 11) compared to the BH_AODV.

As the above discussion of Figure 10, using D_BH_AODV reduces the packet replication and ensures lower overhead as shown in Figure 12 indicated by the green line, while BH_AODV exhibits very high overhead compared to D_BH_AODV because of its uncontrolled attacking behavior.

Figure 10 to Figure 12 illustrate that PDR, delay, and overhead are desirable for D_BH_AODV over BH_AODV because D_BH_AODV ensures a higher delivery ratio, lower delay, and lower overhead ratio compared to BH_AODV. Therefore, we can say that the D_BH_AODV improves the performance over the black hole affected by AODV protocol (BH_AODV) in terms of QoS parameters under consideration.

In a concise discussion, it is clear that the D_BH_AODV routing can detect the black hole nodes and prevent those nodes from participating in further communication. Hence, D_BH_AODV uses IDS and digital signature methods to ensure higher delivery, lower delay, and lower transmission

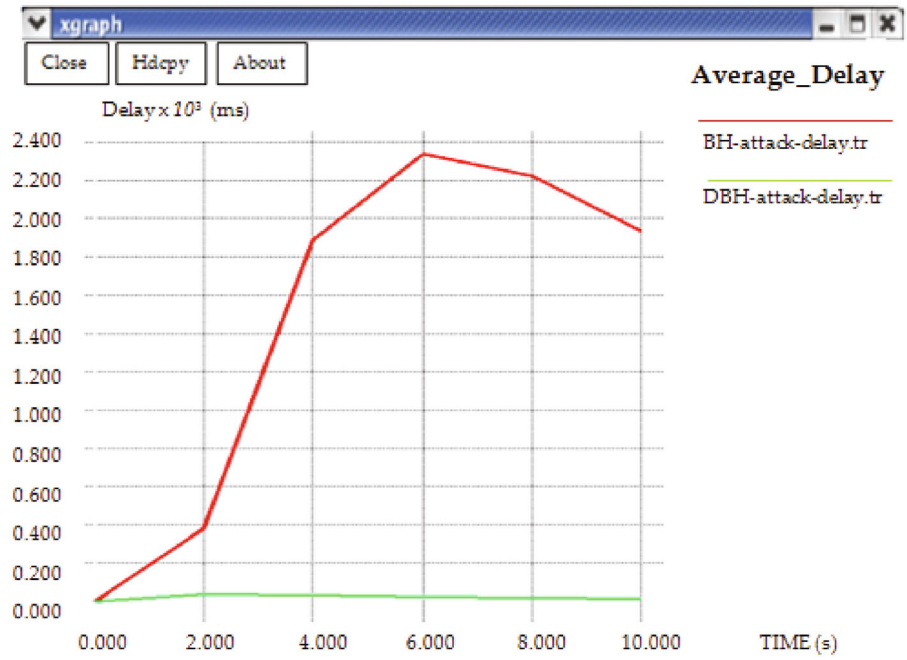


Figure 11: Delay with varying time for black hole and detected black hole attack.

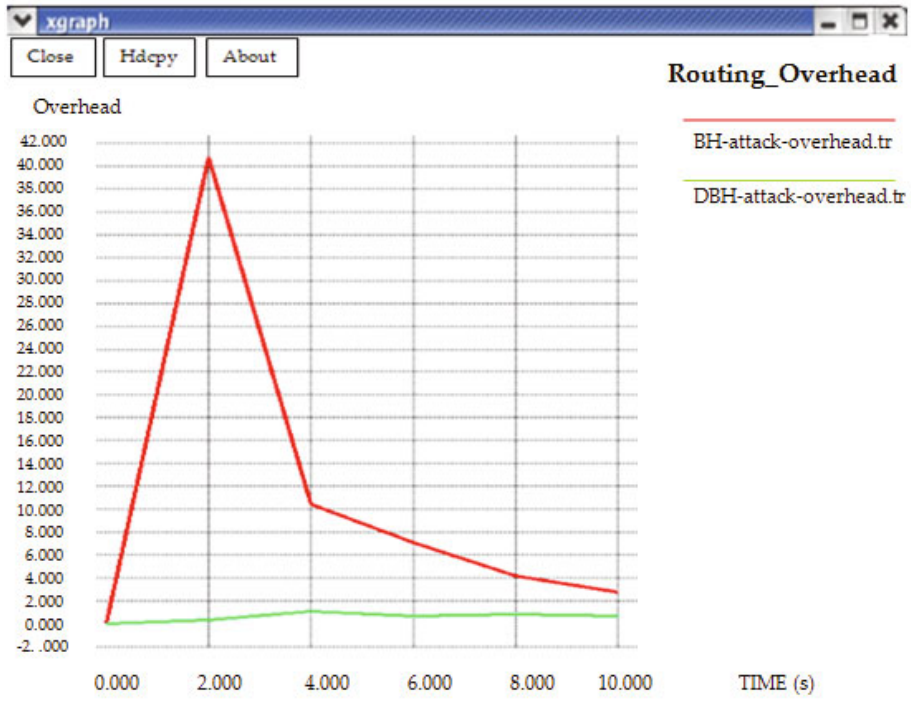


Figure 12: Overhead with varying time for black hole and detected black hole attack.

cost (i.e., lower overhead ratio) because D_BH_AODV does not forward/send message copies or packets to an attacker's node. The above investigations (from Figure 4 to Figure 12) exhibit that D_BH_AODV outperforms BH_AODV by ensuring good delivery, lower delay, and lower overhead in case of digital signature and IDS wherein D_BH_AODV shows higher delay than BH_AODV but lower delay than AODV in case of using IDS (as shown in Figures 6 and 7).

7. Conclusion

MANETs are vulnerable to different attacks that badly affect the wireless networks while establishing a secure routing. In this study, we implement a black hole attack within the AODV protocol by modifying AODV and the trust value of nodes. We detect attackers through a trust mechanism using IDS that requires a time stamp and the encryption technique

using a digital signature. In IDS, we make a graphical comparison among AODV, BH_AODV, and D_BH_AODV. Moreover, in the encryption technique, we make a comparison between BH_AODV and D_BH_AODV black hole AODV. In both cases, the analysis is done in terms of PDR, average delay, and overhead ratio for varying the number of nodes, packets' size, and simulation times. Our investigated results exhibit that under the consideration of AODV routing, the BH_AODV degrades the performance of AODV by lowering the delivery ratio and maximizing the overhead ratio for varying the number of nodes, the size of packets, and the simulation times. It also verifies in both cases, i.e., IDS and digital signature, whereas the D_BH_AODV shows higher delivery and lower overhead compared to the BH_AODV. Although the D_BH_AODV exhibits a higher delay compared to the BH_AODV in case of using IDS, our proposed and implemented D_BH_AODV shows a lower average delay than the original AODV routing for the above variation. In the case of using a digital signature, we observe that the D_BH_AODV routing exhibits a lower delay compared to the BH_AODV. Therefore, the BH_AODV sharply degrades the performance, and the D_BH_AODV improves the networks' overall performance.

8. Future Work

In this network scenario, the variation in the number of nodes is from 20 to 100, whereas the variation of packets is from 512 to 2100 bytes along with the 6 m/s of mobility speed. However, the possibility of having a higher number of nodes and packets with higher mobility can likely happen in real-world scenarios; therefore, this work can be extended to explore the scalability of the network. A secure AODV protocol would be established to prevent various attacks such as wormhole and jellyfish within wireless networks through encryption techniques to guarantee a good trade-off among PDR, average delay, overhead ratio, and energy consumption. Also, this research can also lead to other security services and domestic appliances. It can be used to prevent multiple black hole attacks. AODV protocols can also help in various IoT applications by designing different AODV extensions based on numerous criteria, e.g., quality, reliability, energy, security, and routing strategies [53]. For example, an optimized AODV (OAODV) can be designed to ensure low energy consumption of IoT sensors [54]. Also, an energy-aware secure AODV routing can be implemented by using better route maintenance approaches for large networks [55]. Furthermore, advanced AODV approaches such as collaborative black hole attack-AODV routing protocol (CBHA-AODV) [56] can be implemented for real-time IoT-based civil construction application.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

The authors would like to acknowledge the support of the Network Communication Technology (NCT) Research Groups, FTSM, Universiti Kebangsaan Malaysia. This paper is supported under the Dana Impak Perdana UKM (DIP-2018-040) and Fundamental Research Grant Scheme (FRGS/1/2018/TK04/UKM/02/17). The authors also would like to thank the ICT Division of the Bangladesh Government for awarding a research fellowship to Md Ibrahim Talukdar for this research.

References

[1] I. Mohd Zaki and H. Rosilah, "The implementation of Internet of Things using test bed in the UKMnet environment," *Asia-Pacific Journal of Information Technology and Multimedia*, vol. 8, no. 2, pp. 1–17, 2019.

[2] Z. Ismail and R. Hassan, "A performance study of various mobility speed on AODV routing protocol in homogeneous and heterogeneous MANET," in *in the 17th Asia Pacific Conference on Communications*, IEEE, 2011.

[3] T. Salam and M. S. Hossen, "Performance analysis on homogeneous LEACH and EAMMH protocols in wireless sensor network," *Wireless Personal Communications*, vol. 113, no. 1, pp. 189–222, 2020.

[4] M. S. Hossen, "DTN routing protocols on two distinct geographical regions in an opportunistic network: an analysis," *Wireless Personal Communications*, vol. 108, no. 2, pp. 839–851, 2019.

[5] M. Singh, C. Kumar, and P. Nath, "Challenges and protocols for P2P applications in multi-hop wireless networks," in *in 2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 310–316, IEEE, 2018.

[6] T. Qiu, N. Chen, K. Li, D. Qiao, and Z. Fu, "Heterogeneous ad hoc networks: architectures, advances and challenges," *Ad Hoc Networks*, vol. 55, pp. 143–152, 2017.

[7] C. S. R. Murthy, *Ad Hoc Wireless Networks: Architestures and Protocols*, Pearson Education India, 2004.

[8] S. M. Adam and R. Hassan, "Delay aware reactive routing protocols for QoS in MANETs: a review," *Journal of applied research and technology*, vol. 11, no. 6, pp. 844–850, 2013.

[9] S. Malathy, V. Porkodi, A. Sampathkumar et al., "An optimal network coding based backpressure routing approach for massive IoT network," *Wireless Networks*, vol. 26, no. 5, pp. 3657–3674, 2020.

[10] H. M. Haglan, S. A. Mostafa, N. Z. M. Safar et al., "Analyzing the impact of the number of nodes on the performance of the routing protocols in MANET environment," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 1, pp. 434–440, 2020.

[11] S. Yan and Y. Chung, "Improved ad hoc on-demand distance vector routing (AODV) protocol based on blockchain node detection in ad hoc networks," *International Journal of*

Internet, Broadcasting and Communication, vol. 12, no. 3, pp. 46–55, 2020.

[12] R. K. Mohapatra, S. Samantaray, A. Sahoo et al., "Performance analysis of reactive routing protocols in MANET under CBR traffic using NS2," in *in 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pp. 352–356, IEEE, 2018.

[13] A. K. Biswas and M. Dasgupta, "AODV-DSR hybrid reactive routing protocol and its generalization for mobile ad-hoc networks," in *in 2019 3rd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, pp. 1–5, IEEE, 2019.

[14] V. Sharma, B. Alam, and M. Doja, "An improvement in DSR routing protocol of MANETs using ANFIS," in *in Applications of Artificial Intelligence Techniques in Engineering*, pp. 569–576, Springer, 2019.

[15] K. L. Arega, G. Raga, and R. Bareto, "Survey on performance analysis of AODV, DSR and DSDV in MANET," *Computer Engineering and Intelligent Systems*, vol. 11, no. 3, pp. 23–32, 2020.

[16] F. T. AL-Dhief, N. Sabri, M. S. Salim, S. Fouad, and S. A. Aljunid, "MANET routing protocols evaluation: AODV, DSR and DSDV perspective," in *in MATEC Web of Conferences*, vol. 150, p. 06024, EDP Sciences, 2018.

[17] A. Kulkarni, R. Bukate, and S. Nanaware, "Study of various attacks and routing protocols in MANETS," in *in 2018 International Conference on Information, Communication, Engineering and Technology (ICICET)*, pp. 1–3, IEEE, 2018.

[18] A. M. Fahad and R. C. Muniyandi, "Harmony search algorithm to prevent malicious nodes in mobile ad hoc networks (MANETs)," *Information Technology Journal*, vol. 15, no. 3, pp. 84–90, 2016.

[19] Z. Pooranian, A. Barati, and A. Movaghar, "Queen-bee algorithm for energy efficient clusters in wireless sensor networks," *World Academy of Science, Engineering and Technology*, vol. 73, pp. 1080–1083, 2011.

[20] S. H. H. Nazhad, M. Shojafar, S. Shamshirband, and M. Conti, "An efficient routing protocol for the QoS support of large-scale MANETs," *International Journal of Communication Systems*, vol. 31, no. 1, article e3384, 2018.

[21] M. S. Pathan, J. He, N. Zhu, Z. Ali, M. Qasim, and A. Azmat, "An efficient scheme for detection and prevention of black hole attacks in AODV-based MANETs," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 1, pp. 243–251, 2019.

[22] M. B. M. Kamel, I. Alameri, and A. N. Onaizah, "STAODV: a secure and trust based approach to mitigate blackhole attack on AODV based MANET," in *in 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 1278–1282, IEEE, 2017.

[23] G. K. Wadhwani, S. K. Khatri, and S. K. Mutto, "Trust framework for attack resilience in MANET using AODV," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 23, no. 1, pp. 209–220, 2020.

[24] S. El Jay and A. Hasbi, "Security in mobile ad hoc networks (MANETs) and WSNs (wireless sensor networks)," *International Journal of Computer Science and Network Security (IJCSNS)*, vol. 16, no. 9, p. 118, 2016.

[25] M. Y. Thanoun and A. M. Aleesa, "Routing, significant and applications of mobile ad-hoc wireless sensor networks," *Journal of Computational and Theoretical Nanoscience*, vol. 17, no. 2, pp. 850–854, 2020.

[26] V. Tilwari, M. D. N. Hindia, K. Dimyati, F. Qamar, and M. S. A. Talip, "Contention window and residual battery aware multipath routing schemes in mobile ad-hoc networks," *International Journal of Technology*, vol. 10, no. 7, pp. 1376–1384, 2019.

[27] V. L. Narayana and C. Bharathi, "Identity based cryptography for mobile ad hoc networks," *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 5, p. 1173, 2017.

[28] M. Rath and B. K. Pattanayak, "Security protocol with IDS framework using mobile agent in robotic MANET," *International Journal of Information Security and Privacy*, vol. 13, no. 1, pp. 46–58, 2019.

[29] A. Chaudhary, V. N. Tiwari, and A. Kumar, "Design an anomaly-based intrusion detection system using soft computing for mobile ad hoc networks," *International Journal of Soft Computing and Networking*, vol. 1, no. 1, pp. 17–34, 2016.

[30] E. V. Balan, M. K. Priyan, C. Gokulnath, and G. U. Devi, "Fuzzy based intrusion detection systems in MANET," *Procedia Computer Science*, vol. 50, pp. 109–114, 2015.

[31] J. Manoranjini, A. Chandrasekar, and S. Jothi, "Improved QoS and avoidance of black hole attacks in MANET using trust detection framework," *Automatika*, vol. 60, no. 3, pp. 274–284, 2019.

[32] Y. M. Khamayseh, S. A. Aljawarneh, and A. E. Asaad, "Ensuring survivability against black hole attacks in MANETS for preserving energy efficiency," *Sustainable Computing: Informatics and Systems*, vol. 18, pp. 90–100, 2018.

[33] G. Usha, M. R. Babu, and S. S. Kumar, "Dynamic anomaly detection using cross layer security in MANET," *Computers & Electrical Engineering*, vol. 59, pp. 231–241, 2017.

[34] S. Gurung and S. Chauhan, "A dynamic threshold based approach for mitigating black-hole attack in MANET," *Wireless Networks*, vol. 24, no. 8, pp. 2957–2971, 2018.

[35] M. Thebiga and R. SujiPramila, "A new mathematical and correlation coefficient based approach to recognize and to obstruct the black hole attacks in MANETs using DSR routing," *Wireless Personal Communications*, vol. 114, no. 2, pp. 975–993, 2020.

[36] S. Kumar, M. Goyal, D. Goyal, and R. C. Poonia, "Routing protocols and security issues in MANET," in *in 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions)(ICTUS)*, pp. 818–824, IEEE, 2017.

[37] S. Gurung and S. Chauhan, "Performance analysis of black-hole attack mitigation protocols under gray-hole attacks in MANET," *Wireless Networks*, vol. 25, no. 3, pp. 975–988, 2019.

[38] U. Singh, M. Samvatsar, A. Sharma, and A. K. Jain, "Detection and avoidance of unified attacks on MANET using trusted secure AODV routing protocol," in *in 2016 Symposium on Colossal Data Analysis and Networking (CDAN)*, pp. 1–6, IEEE, 2016.

[39] R. K. Singh and P. Nand, "Literature review of routing attacks in MANET," in *in 2016 International Conference on Computing, Communication and Automation (ICCCA)*, pp. 525–530, IEEE, 2016.

[40] S. Shrestha, R. Baidya, B. Giri, and A. Thapa, "Securing black-hole attacks in MANETs using modified sequence number in AODV routing protocol," in *in 2020 8th International Electrical Engineering Congress (iEECON)*, pp. 1–4, IEEE, 2020.

[41] S. Hossain, M. S. Hussain, R. R. Ema, S. Dutta, S. Sarkar, and T. Islam, "Detecting black hole attack by selecting appropriate

routes for authentic message passing using SHA-3 and Diffie-Hellman algorithm in AODV and AOMDV routing protocols in MANET," in *in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–7, IEEE, 2019.

[42] D. A. F. B. H. INTRUSION, "Effect of clustering in designing a fuzzy based hybrid intrusion detection system for mobile ad hoc networks," *Journal of Computer Science*, vol. 9, no. 4, pp. 521–525, 2013.

[43] S. R. Deshmukh, P. Chatur, and N. B. Bhople, "AODV-based secure routing against blackhole attack in MANET," in *in 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).*, pp. 1960–1964, IEEE, 2016.

[44] V. Savkare and N. Kazi, "AODV and DSR routing protocol performance comparison in MANET using network simulator (NS2)," *Int. Res. J. Eng. Technol*, vol. 6, no. 9, pp. 7–10, 2019.

[45] F. Abdel-Fattah, K. A. Farhan, F. H. Al-Tarawneh, and F. AlTamimi, "Security challenges and attacks in dynamic mobile ad hoc networks MANETs," in *in 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, pp. 28–33, IEEE, 2019.

[46] R. Skaggs-Schellenberg, N. Wang, and D. Wright, "Performance evaluation and analysis of proactive and reactive MANET protocols at varied speeds," in *in 2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 981–985, IEEE, 2020.

[47] A. Pramanik, B. Choudhury, T. S. Choudhury, W. Arif, and J. Mehedi, "Behavioral study of random waypoint mobility model based energy aware MANET," in *in 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 624–629, IEEE, 2016.

[48] R. Thiagarajan and M. Moorthi, "Efficient routing protocols for mobile ad hoc network," in *in 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, pp. 427–431, IEEE, 2017.

[49] H. Moudni, M. Er-rouidi, H. Mouncif, and B. E. Hadadi, "Black hole attack detection using fuzzy based intrusion detection systems in MANET," *Procedia Computer Science*, vol. 151, pp. 1176–1181, 2019.

[50] Z. Ahmad and A. Bansiya, "Survey on security by using intrusion detection system in MANET," *A RKDF University Journal of Science and Engineering*, vol. 2, no. 1, pp. 21–25, 2019.

[51] S. Sivanesh and V. S. Dhulipala, "Accurate and cognitive intrusion detection system (ACIDS): a novel black hole detection mechanism in mobile ad hoc networks," *Mobile Networks and Applications*, 2020.

[52] V. Nancy, "A security for MANET interruption recognition & preclusion approaches for network layer attacks," *International Journal of Applied Engineering Research*, vol. 13, no. 12, pp. 10702–10706, 2018.

[53] T. K. Saini and S. C. Sharma, "Recent advancements, review analysis, and extensions of the AODV with the illustration of the applied concept," *Ad Hoc Networks*, vol. 103, p. 102148, 2020.

[54] A. Zrelli, H. Khlaifi, and T. Ezzedine, "Performance evaluation of AODV and OAODV for several WSN/IoT applications," in *in 2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1–6, IEEE, 2019.

[55] N. Kamboj and M. Rai, "A new secure ad-hoc on demand distance vector routing protocol to ensure less power consumption in mobile ad-hoc network," *Journal of Computational and Theoretical Nanoscience*, vol. 17, no. 6, pp. 2483–2487, 2020.

[56] T. A. S. Srinivas and S. M. Manivannan, "Preventing collaborative black hole attack in IoT construction using a CBHA–AODV routing protocol," *International Journal of Grid and High Performance Computing*, vol. 12, no. 2, pp. 25–46, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6698246, 19 pages
https://doi.org/10.1155/2021/6698246

Research Article

From Digital Divide to Information Availability: A Wi-Fi-Based Novel Solution for Information Dissemination

Muhammad Faran Majeed[1], **Irshad Ahmed Abbasi**[2], **Sikandar Ali**[3,4], **Elfatih Elmubarak Mustafa**[2], **Ibrar Hussain**[1], **Khalid Saeed**[1], **Muhammad Faisal Abrar**[5], **Mah E. No**[6] and **Muhammad Kashif Khattak**[7]

[1]*Department of Computer Science, Shaheed Benazir Bhutto University, Sheringal, Dir (U), Khyber Pakhtunkhwa, Pakistan*
[2]*Department of Computer Science, Faculty of Science and Arts at Belgarn, University of Bisha, P.O. Box 60, Sabt Al-Alaya 61985, Saudi Arabia*
[3]*Department of Computer Science and Technology, China University of Petroleum-Beijing, Beijing 102249, China*
[4]*Beijing Key Lab of Petroleum Data Mining, China University of Petroleum-Beijing, Beijing 102249, China*
[5]*Department of Computer Science, University of Engineering & Technology, Mardan, Khyber Pakhtunkhwa, Pakistan*
[6]*Department of Computer Science, CECOS University of IT & Emerging Sciences, Peshawar, Khyber Pakhtunkhwa, Pakistan*
[7]*Southern Punjab Redeemers College, Taunsa Shareef, Punjab, Pakistan*

Correspondence should be addressed to Sikandar Ali; sikandar@cup.edu.cn

Received 31 December 2020; Revised 27 January 2021; Accepted 10 February 2021; Published 28 February 2021

Academic Editor: Mohammad Hossein Anisi

Digital divide means unequal access to the people for information and communication technology (ICT) facilities. The developed countries are comparatively less digitally divided as compared to developing countries. This study focuses on District Chitral considering its geographical conditions and high mountainous topography which plays a significant role in its isolation. Aside from the digital divide, the situation in Chitral is even more severe in terms of the absence of basic ICT infrastructure and electricity in the schools. To address this issue, especially in female secondary and higher secondary schools, we designed a project to bridge the digital divide via Wireless Local Area Network on Raspberry Pi3 for balancing the ICT facilities in the targeted area. The Wi-Fi-Based Content Distributors (Wi-Fi-BCDs) were provided to bridge the digital divide in rural area schools of Chitral. The Wi-Fi-BCD is a solar-based system that is used to deliver quality educational contents directly to classroom, library, or other learning environments without electricity connection and Internet wire as these facilities are available by default in it. The close-ended questionnaire was adopted to collect data from the students, teachers, and headmistresses of girl secondary and higher secondary schools in Chitral. The procedure of validity, reliability, regression, correlation, and exploratory factor analysis was used to analyze the obtained data. The technology acceptance model (TAM) was modified and adopted to examine the effects of Wi-Fi-BCD for bridging the digital divide. The relationship of the modified TAM model was examined through regression and correlation to verify the model fitness according to the data obtained. The result analysis of this study shows that the relationship of the modified TAM model with its variables is positively significant, while the analysis of path relationship between model variables and outcomes from the questionnaire shows that it motivates learners to use Wi-Fi-BCD.

1. Introduction

In this section, we discuss the background of the study, ICT, its role, ICT education for girls in Pakistan, digital divide, and its effects.

1.1. Background of the Study. The limited access to the Internet is the main reason for the huge digital divide in the rural areas of Pakistan. This digital divide is quite visible in the case of female students. The major objective of this study, therefore, is to bridge this gap and provide accessibility through

Wi-Fi-BCD along with tablet computers to enable the female students to get benefits from the technology offered.

Our study focused on District Chitral. Chitral is surrounded by Dir (upper) in its South and connected via Lawari Pass (10,500 ft) and Gilgit in the North via Shandoor Pass (12,201 ft). Both routes to Chitral are closed due to snowfall during winter making it almost impossible to reach the city. The geographical conditions and high mountainous topography of the district play a significant role in its isolation. Aside from the digital divide, the situation in Chitral is even more severe in terms of the absence of basic infrastructure (computer labs) and electricity in schools. Due to the preliminary requirements, there exists a list of barriers that resists the inclusion of ICT in developing countries like Pakistan [1]. Due to the lack of basic infrastructure, the schools of Chitral are not attractive for the teachers as well [2]. The majority of the teachers lack basic ICT skills and thus contribute to the digital divide. Moreover, the inclusion of ICT in the curriculum has never been given importance due to digital divide becoming severe in District Chitral.

According to Pakistan Education Statistics [3], 22.64 million children are out of school. Among these children, 49% are girls and 40% are boys. Pakistan is having the second-highest ratio of girls not enrolled in schools in the entire world [4]. This makes the scenario worse. This situation has given rise to the female education crisis in recent times as the tribal communities are reluctant to female education, and another factor is the militancy that has now been overcome. However, parents in Khyber Pukhtunkhwa (KP) province are reluctant to send their daughters to schools, and the same is the case in the Federal Administrative Tribal Area (FATA) and Balochistan province. The main reasons for keeping the girls out of the schools are cultural traditions, militancy in past recent decades, unsuitable schooling environments, and financial constraints of the parents.

According to the Annual Statistical Report Government Schools in Khyber Pakhtunkhwa for the year, there are only 11,697 schools for girls compared to 21,893 for boys in the province of KP [5]. This shows the indifference of the government policies to consider female education equally important in the country. The objective of this research is to integrate ICT and to address the ICT gap related to girl education in the District Chitral. Moreover, the Wi-Fi-BCDs were given to the female schools of Chitral for evaluation purposes.

1.2. Information and Communication Technology (ICT). Information and communication technologies (ICTs) have greatly affected human life nowadays. ICT (telephones, television, computers, and Internet) gives opening and easy access towards global shared knowledge around the world in almost all fields of life. Unfortunately, ICT is not accessible all over the world [6].

ICT technologies have an important role in making human life enhanced and modernized in this modern era. It is essential to avail all ICT applications to cope with the modern challenges in the current society. Similarly, ICT brings treasured and worthwhile information in the fields of technology, social, political, and economic philosophy in the world of academic pursuit. The sectors of education, academics, and research are also being affected positively due to substantial changes in the field of ICT [7].

1.3. Role of ICT in Education. The growing development in ICT remarkably revolutionized the world in the 21st century to enhance and modernize it in this modern era. ICT is the demand of the day to make progress in the world of academics, education, and research. Hence, it is indispensable to use ICT in academics to explore and teach the learners updated materials and skills with the latest techniques. Educationalists and policymakers try to restructure their whole academic setting to fill the technical gap in the process of learning and teaching. This restructuring process requires actual implementation of technology in the existing educational setting to provide students with subject-related knowledge and to enhance profession maturity in learners or scholars [8].

The initiation of ICT in academics ensures communication with each other via email, mailing lists, and meeting rooms. ICT provides faster and convenient access to extensive and current information; ICT can also be used to perform complex mathematical and statistical calculations. In addition, ICT is the best tool to publish research findings and conclusions [7].

According to [9], using ICT for teaching in learning improves students' performance, increases motivation, enhances confidence, and improves positive attitude [10], suggesting that ICT being well incorporated in teaching and learning produces a positive impact on learning outcomes and improves motivation and creates enthusiasm in learners.

According to [11], the outline of ICT on teaching in learning supports better assignment and deep learning. Using ICT makes the learner attitude positive and boosts confidence. ICT is a kind of soft power to enhance economic skills that can develop technical strategies of the learners to get the opportunities of employment in the future. The development of ICT culture in schools is necessary to develop the strategic skills of the students.

1.4. ICT for Girls (In Digital Pakistan Policy 2017). The federal cabinet finally approved the long-awaited IT Policy of Pakistan and has issued it in the name of "Digital Pakistan Policy 2017." ICT in education can be used to empower young girls. The following are the main findings of ICT Education for girls in Digital Pakistan Policy 2017.

(1) Encourage the use of ICT among girls for their empowerment and to overcome the digital divide

(2) The initiation of ICT programs in girls' schools that can educate them in software skills. This education would be fruitful to find white-collar jobs in the future. The establishment of women empowerment centers and IT laboratories in schools that can build their (young girls) IT skills (computing, coding, and communication) in collaboration with other private sectors

(3) The provision of special incentives to boost up digital services and strengthen IT culture in the region

(4) Strengthening international collaboration in ICT accessibility so that women and girls may have active partaking in digital society

1.5. Digital Divide: A Concept. The society of human beings from the very beginning is divided into so many aspects such as gender, caste, color, race, and ethnicity. The rapid advancement of ICT in the 20th century divides the people of the world once again and is named the digital divide. Digital divide means an unequal opportunity for people to access knowledge shared on the Internet. Digital divide exists also among the countries and within the countries. The developed countries are comparatively less digitally divided as compared to developing countries [12].

The Organization for Economic Cooperation and Development defines the term digital divide as "the gap among individuals, businesses, households, and geographic areas at different socioeconomic levels with regard to both their opportunities to access ICT and their use of the Internet for a wide variety of activities." The digital division shows differences within and among the countries. The digital divisions in homes depend upon income, education, and some other factors such as the size and type of house, age, gender, ethnicity, and linguistic backgrounds [12].

Basic ICT skills and infrastructure are necessary for people to use ICT facilities. The availability of ICT infrastructure positively contributes to overcoming the digital divide. The problem with rural communities in third world countries is the lack of access to information. The lack of infrastructure basically contributes to the lack of access to information in rural areas. The Wi-Fi-BCD and tablet computer will overcome the lack of ICT facility. Various studies have highlighted the role of individual use in measuring the digital divide [13].

These studies have revealed a positive relation between the individual use of ICT facilities (access and effective utilization of high and low digital divide in a society).

1.6. Barriers in Accessing ICT Services. Van Dijk and Hacker [14] find four types of barriers to access including mental access, material access, skill access, and usage access. Mental access is due to lack of initial digital experience, interest, computer worries, and disinterest in new technology. Material access is due to the nonavailability of computer and network connection, such as hardware, software, applications, and networks. The major portion of the research in digital divide is the study of physical access to personal computers and the Internet [15]. Skill access rises because of the lack of digital skills due to inadequate user-friendliness and inadequate education or social support. One needs access skills to use and command digital media technology; some people call these as digital skills [15]. Usage access arises due to the lack of significant usage opportunities. Some researchers claim that this gap is the diversity of application usage between different people, education levels, age, and gender [14].

1.7. Bridging the Digital Divide in Chitral. Considering the mentioned condition of the area and being part of the jurisdiction of the university, our institute recently established a specific project named Bridging the Digital Divide via Wireless Local Area Network on Raspberry Pi3 (Wi-Fi-BCD) worth Rs. 2.5 million in order to balance the inequality of ICT facilities at the Government Girls Secondary and Higher Secondary Schools of the District Chitral.

As a first step of the project which is the development and installation of Wi-Fi-BCD in their schools, thousands of high school students and their teachers will be benefited from the project. They will also know about the advantages of being connected through the project. The teachers will prepare their lectures with the help of the updated material shared by the international education experts. Moreover, the practitioner's confidence, the positive impact on teaching and learning, and the quality of education provided at these schools will also be enhanced which is one of the major objectives of this research project.

1.8. Organization of the Paper. The rest of the paper is organized as follows. Section 1 has introduced the concepts related to the study thoroughly. In Section 2, we describe the development of Wi-Fi-BCD. Section 3 throws light on the conceptual model, while Section 4 describes the research methodology adopted for this specific study. Section 5 presents the results. Finally, Section 6 concludes the paper.

2. Wi-Fi-Based Content Distributor (Wi-Fi-BCD)

The Wi-Fi-BCD is a digital content distributor, based on the Raspberry Pi. The beauty of Wi-Fi-BCD is that the contents are accessible without having an Internet connection. Furthermore, the Wi-Fi-BCD is a solar-based innovation; the contents are accessible even without electricity and Internet connection. Even if the area has no Internet connection and electricity, the Wi-Fi-BCD is the solution to distribute the contents as an offline library. With the small size and solar-based properties of the Wi-Fi-BCD, one can deploy anywhere in the school and college to distribute the preloaded digital content. The contents can easily be accessed using a tablet, mobile, computer, and laptop.

2.1. Wi-Fi-BCD Development. The Wi-Fi-BCD is a portable server, which is plug-and-play as well as offline. It stores educational contents and other educational material and makes the contents available offline via a wireless connection. It acts as a library of digital contents, which can be available by just pressing a button. The digital contents can be accessed via a smartphone or PC by typing the IP address in order to connect to the Raspberry Pi server.

When you are connected to the server, you can easily access the world's best free educational websites offline including Wikipedia and Khan Academy. You can play and download the videos and audios, preview, and download books.

2.2. Specifications of the Wi-Fi-BCD. The specifications of the Wi-Fi-BCD are the following:

(i) Raspberry Pi3, 64-bit 1.2 GHz quad-core SBC

(ii) 64 GB SD card

(iii) Wi-Fi Dongle

(iv) 5-ampere charge controller

(v) 12 V DC battery inside/20-watt 12 V solar panel

The Wi-Fi-BCD is preloaded with thousands of open-source video content including offline KA Lite; Khan Academy content for math, physics, biology, chemistry, and several other subjects; RACHEL with world-famous books; Wikipedia for Schools; digital world map; health care; advanced school curricula; and many textbooks.

2.3. Raspberry Pi. The Raspberry Pi is developed by the Raspberry Pi Foundation in the UK. It is a single-board computer made for the purpose of enhancing basic computer science teaching in schools, colleges, and universities. Several generations of R-Pi are released. All versions contain a Broadcom system on a chip (SoC), CPU, and GPU. The speed of the CPU ranges from 700 MHz to 1.2 GHz for the R-Pi3, and the memory capacity is from 256 MB to 1 GB. SD cards are used to store OS and other programs. It also has the capabilities of HDMI, composite video output, 3.5 mm phone jack for audio, GPIO pins, Ethernet port, Wi-Fi 802.11n, and Bluetooth. Figures 1 and 2 show the R-Pi without and with a protective case.

2.4. RACHEL Server on Wi-Fi-BCD. The RACHEL (Remote Area Community Hotspot for Education and Learning) server is a portable offline server that can store educational contents, websites, and other educational material and make the content available offline over any local wireless network. The RACHEL server can make deploying a digital content library very easy, an initiative to make available rich educational contents to places where no Internet connection is available (worldpossible, 2018) (Connecting offline learners to the world's knowledge: URL: https://worldpossible.org).

2.5. System Development Phases. The Wi-Fi-BCD is built on the Raspberry Pi3 (Model B), on a 64 GB micro SD card with preloaded RACHEL content. Wi-Fi-BCD does not need an electricity connection or battery backup. It is recommended for an area with unreliable power. Table 1 shows the development phases of Wi-Fi-BCD, and Figure 3 shows the pictorial view of the network.

3. Conceptual Research Model

Technology acceptance is "the assessment of usage or rejection of ICT service" [16]. Diverse models with explicit elements set have been developed by information technology (IT) and information system (IS) scientists. The technology acceptance model (TAM), Theory of Reasoned Action (TRA), unified theory of acceptance and use of technology (UTAUT), and Diffusion of Innovation Model (DOI) are

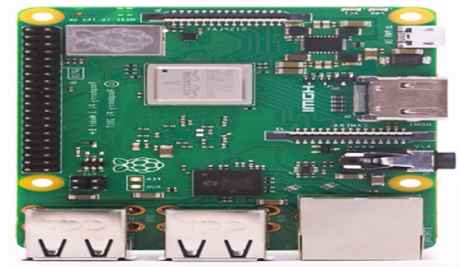


Figure 1: Raspberry Pi board.

the famous IT and IS acceptance models. These models are discussed as follows.

3.1. Theory of Reasoned Action (TRA). Theory of Reasoned Action (TRA) proposed by [17] is the most appropriate model regarding the study of user behavior; it is used in different research domains to predict user behavior. The TRA predicts the factors which affect human behavior to accept a particular system.

3.2. Unified Theory of Acceptance and Use of Technology (UTAUT). Unified Theory of Acceptance and Use of Technology (UTAUT) is generally used in the field of technology acceptance models. The unified theory of acceptance and use of technology explains user intentions to use a particular system and the usage behavior. The UTAUT model explains the influence of individual differences in the use of technology, thus adding control variables such as gender, age, and experience. The predictors of UTAUT are behavioral intention or usage, effort expectancy, performance expectancy, social influence, and facilitating conditions [18].

3.3. Technology Acceptance Model (TAM). The technology acceptance model (TAM) is the most cited and famous model to examine the user acceptance and use of IT and IS. The TAM model was developed by [19] from the TRA. The TAM model further explains why users will accept technology and what factors influence the acceptance of technology on users. This model has two cognitive factors perceived usefulness (PU) and perceived ease of use (PEOU). According to [19], PU means "the extent to which the person believes that using a particular system would enhance the performance of the individual." [19] defines PEOU as an individual's perception regarding the ease of use of a particular system. Simply, PU means the usefulness of a system, and PEOU means the easy use of the system.

3.4. Reasons to Choose TAM Model. The TAM model is used in studies based on IT systems to evaluate their effectiveness for the users. The TAM was proposed by [19] and is widely used to explore users' acceptance of IT and IS. The model is explicitly designed to know the acceptance of Computer-Based Technologies (CBT) in a workplace. This study, therefore, is constructed to include

(1) perceived usefulness (PU)

(2) perceived ease of use (PEOU)

Figure 2: Raspberry Pi casing.

(3) attitude towards using (ATU)

(4) behavioral intention (BI) to use

The main objective of this study is to develop and evaluate a model of the influence of system characteristics related to the acceptance of computer-based ISs by users. The TAM model is developed with two major objectives: to improve our perception of user acceptance processes and to provide the theoretical basis for testing this framework of user acceptance which will enable the system developers to test their proposed systems before launching them for implementation. This prototype helps further improvement of the system.

In the TAM model as shown in Figure 4, the external variables are the design features of the IT or IS. The PEOU and PU are the cognitive responses of users, the ATU is an effective or accepted response, and the BI is the actual use of the system. The PEOU, PU, and ATU are the user motivation.

The TAM model consisting of six major elements which include external variable (EV), PEOU, PU, BI, ATU, and actual use (AU) and a set of relations among them is hardly used by scholars without modifications [20, 21]. The researchers often modify the model to fulfill the requirements of particular subject and research questions. Also, the researchers need to modify the model by choosing and concentrating on a portion of the initially proposed components and relations, extending the model by including new components, and recommending new relations among the components [22, 23].

No doubt, the TAM model is famous and is a cited research model regarding the acceptance of technology [24]. The TAM model is widely tested in different scenarios and content and verified to be reliable and valid by explaining the acceptance of IS and IT [25, 19, 18]. In the light of the above discussion, a modified research model based on TAM is designed, which is discussed as follows.

3.5. Modified TAM Model for Our Research. Luhamya et al. [26] have discussed different models and theories about ICT in education. They discussed the TRA model developed by Ajzen and Fishbein (1980). In fact, the TRA model has Actual Behavior (AB) as its main variable. Actual Behavior is an individual's response which can be observed in each situation with respect to a given target.

The related literature is showing six theories regarding ICT, such as the TAM, the TRA, the Theory of Planned Behavior (TPB), the Technology-Organization-Environment (TOE) framework, the Technological Pedagogical Content Knowledge (TPACK) framework, and the UTAUT [26]. All the stated above theories have focused on student's motivation, practitioner's confidence, and improvement in teaching and learning in one or another way. Our modified TAM model is shown in Figure 5.

In fact, the Wi-Fi-BCD is a kind of ICT in education that helps to improve students' outcome. The development, distribution, and installation of Wi-Fi-BCD in Girls Secondary and Higher Secondary Schools of Chitral will motivate students for higher outcomes, will develop confidence in using ICT, and will make effective and convenient the teaching and learning process.

3.5.1. Demographic and Topographic Factors. Access to ICT is also dependent on demographic or topographic factors; also, the geographical location plays an important role in accessing ICT. According to [14], the major measure of the research of the digital divide is dedicated to the observation of physical access to personal computers and the Internet based on demographic factors that include gender, age, income, and education. According to [27], the weightier cultural, social, and mental reasons behind the digital divide have not been addressed so far, and findings show that income is the most significant factor behind physical access followed by age and education, respectively.

3.5.2. Student Motivation (SM). Motivation has been one of the key concepts of pedagogical psychology for a long time. It is a process which involves objectives as well as physical and mental activities. On a general level, it encompasses various events that cause shifts and lead to appropriate measures. Järvelä [28] claimed that it was not prior to 1980 that motivation was transferred from laboratories to the natural learning environment and that until then studies on motivation had very little impact on the teaching process. Nowadays motivation is crucially related to the learning process and learning achievements. In fact, it is of great interest to the students in using Wi-Fi-BCDs for high achievements.

3.5.3. Positive Impact of Wi-Fi-BCD on Teaching and Learning (PI). Primarily, the content server was used to facilitate student learning. Teachers have increased their personal and professional uses of computers [29]. Alongside, these increases in teachers' professional uses are increased in the reported instructional uses of computers in the classroom.

3.5.4. Practitioner's Confidence in Using Live Technology (PC). As with other professionals, it is expected that teachers use technology in ways that extend and increase their effectiveness. Here, practitioners' confidence means how boldly the teachers and students can use computers and tablets to browse data and to write, save, or delete data files.

Table 1: Wi-Fi-BCD development phases.

Phase	Deliverable
Phase 1	Need a Raspberry Pi
Phase 2	64 GB SD card (to host the operating system and RACHEL Pi)
Phase 3	An open-source RACHEL Pi image
Phase 4	Download RACHEL Pi from FTP: ftp.worldpossible.org on FTP Client Filezilla
Phase 5	The downloaded RACHEL Pi image will be written on SD card
Phase 6	Insert the SD card into the Raspberry Pi and power it on
Phase 7	Look for a network named RPI, connect it
Phase 8	Enter 10.10.10.10 into browser address bar, the RACHEL content will be displayed

Figure 3: Pictorial view of network architecture.

3.5.5. Wi-Fi-BCD. This is thoroughly discussed in Section 2.

3.5.6. Perceived Ease of Use (PEOU). Davis et. al [19] define PEOU as an individual's perception regarding the ease of use of a particular system, i.e., whether it is easy to understand or difficult to learn and use. It is a concept which refers to users' perception of how an IS is difficult to use. Moreover, it is the evaluation of the degree of using the technology with minimum effort.

3.5.7. Perceived Usefulness (PU). According to [19], PU means "the extent to which the person believes that using a particular system would enhance the performance of the individual." In this study, we must find the use of Wi-Fi-BCD beneficial, particularly in terms of bridging the digital divide more effectively. Furthermore, according to [30], the concept of perceived usefulness is the perception of a user of a given technology to achieve his work goals.

3.5.8. Attitudes towards Using (ATU). Davis et al. [19] define attitude as the positive or negative feelings of a user of a system. The ATU, in this study, is to assess the attitudes towards using the Wi-Fi-BCD for bridging the digital divide. Positive attitude certainly helps to overcome the digital divide and improve the contents of the Wi-Fi-BCD. The more the use of the Wi-Fi-BCD, the more the improvement.

3.5.9. Behavioral Intention (BI). The BI is the degree of the strength of one intent to do a specified task (behavior). The BI is recognized as an exceptionally main construct in the TAM study as it is proposed to be a key measure in user acceptance study [31]. It was devised from the TRA by [17]. Earlier research has shown that BI has a direct relationship to the actual use of a given technology. The features of this design are important as they have the capability to influence the perception of the users; they directly touch the PU and PEOU. Without affecting the attitude or behavior, these features indirectly influence the perception of the users regarding its usefulness and comfort.

3.6. Relationship between Model Variables. The modified TAM model used in this study is aimed at examining the effect of Wi-Fi-BCD for bridging the digital divide in Girls School of Chitral. The required relationship between SM, PI, PC, PEOU, PU, ATU, and BI is examined to find the effects of Wi-Fi-BCD for bridging the digital divide. All the relationships between the above variables are tested; if found positively significant, the effects of the Wi-Fi-BCD will be positive on bridging the digital divide.

3.6.1. Relationship of Wi-Fi-BCD with SM, PI, and PC. In the modified TAM model for the study, the Wi-Fi-BCD (WCD) is a dependent variable for student motivation for higher outcomes (SM), practitioners' confidence (PC) in using live-technology, and positive impact of Wi-Fi-BCD on teaching and learning (PI). A significant relationship of the Wi-Fi-BCD with SM, PI, and PC means that the Wi-Fi-BCD effect is positive to bridge the so-called digital divide.

3.6.2. Relationship of Wi-Fi-BCD with PEOU and PU. The Wi-Fi-BCD is a dependent variable for the PEOU and PU. A positive association of Wi-Fi-BCD with PEOU and PU means that the Wi-Fi-BCD effect is positive to bridge the digital divide. A positive relationship of PEOU and PU with Wi-Fi-BCD means that teachers and students found the Wi-Fi-BCD easy to use and found it useful.

3.6.3. Relationship between PEOU and PU. Davis et al. [19] accomplished several experimental studies to confirm the TAM model, based on the independent variable PEOU and PU with the dependent variable ATU. In the results of the experimental studies, Davis et al. found the correlation of PU positively significant with PEOU and ATU. Davis et al. also found the correlation of PEOU with ATU positively significant. But the correlation of PU with ATU was more

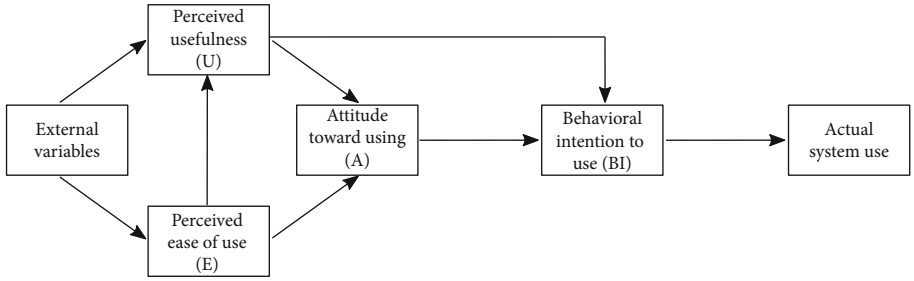


Figure 4: TAM model proposed by Davis (1986).

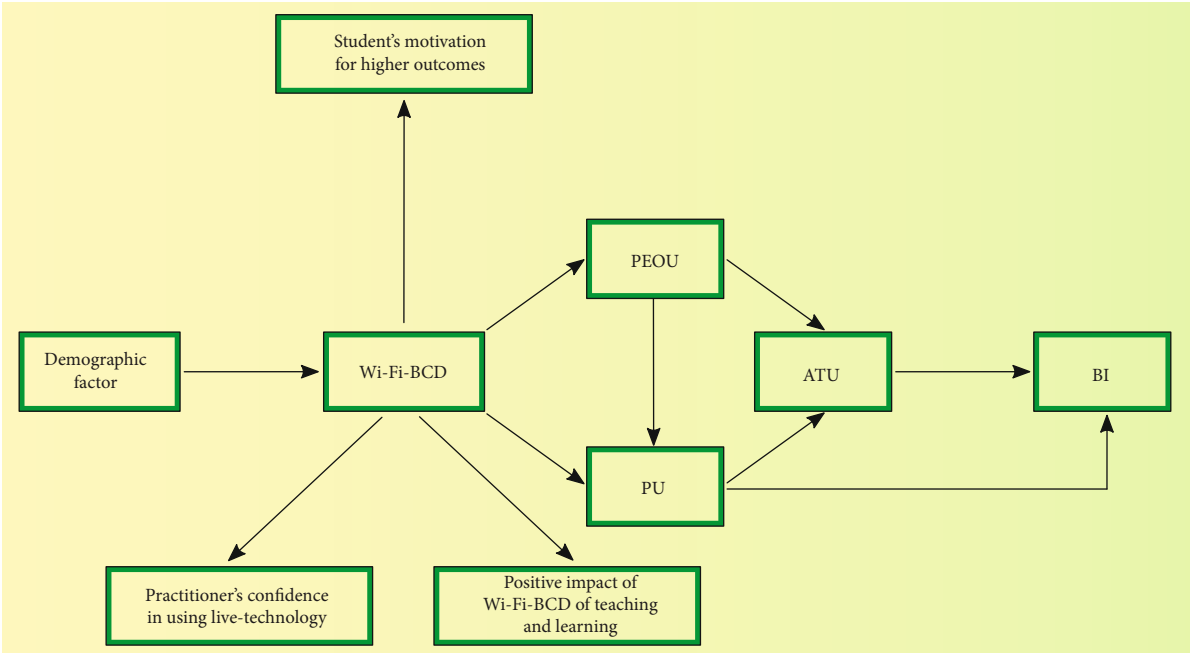


Figure 5: Modified TAM model for our research.

strongly positively significant. According to [19], the PEOU is also a dependent variable for the PU; the high PEOU of a system increases the PU of that system [30, 19]. Based on [19] assumption, the perceived ease of use of the Wi-Fi-BCD will result in PU of the Wi-Fi-BCD to bridge the digital divide.

3.6.4. Relationship between PU and ATU. According to [32], the PU of IT and IS is the most important factor for ATU. According to [33], PU has a significantly strong association with ATU than PEOU. According to [19], the PU has a strong effect on ATU and is verified in several previous and current studies; the ATU is the most important factor to determine the actual usage of a system. For the Wi-Fi-BCD, the higher the effect of PU on ATU, the higher will be the usage of the Wi-Fi-BCD.

3.6.5. Relationship between PEOU and ATU. Vallerand et al. [17] proposed that the ATU of the system is determined from the ease of use of that system. PEOU is an independent variable for ATU and a dependent variable for PU. According to the assumption of [19], PEOU and PU will predict attitude towards using the Wi-Fi-BCD.

3.6.6. Relationship between ATU and BI. According to Leonard et al. [34], the ATU has a direct effect on behavioral intention to use. BI is a dependent variable for ATU and PU; from the discussion above, the behavioral intention to use is mainly dependent on ATU and PU. The attitude towards using and PU will predict the BI of the Wi-Fi-BCD (actual usage of Wi-Fi-BCD) for bridging the digital divide.

3.7. Aims of the Adopted TAM Model. The study is to find the effects of Wi-Fi-BCD to bridge the digital divide. The relationship between the variables Wi-Fi-BCD, SM, PI, PC, PEOU, PU, ATU, and BI is analyzed for the purpose, to know the effects of Wi-Fi-BCD for bridging the digital divide. As per the modified TAM model, the following path relationship is tested to know the effects of Wi-Fi-BCD for model fitness on the data collected.

(1) The relationship of Wi-Fi-BCD with SM, PI, and PC

(2) The relationship of Wi-Fi-BCD with PEOU and PU

(3) The relationship of PEOU and PU

(4) The relationship of PEOU and PU with ATU

(5) The relationship of ATU and BI

4. Research Methodology

This section describes the methodology, which has been followed to accomplish the objectives and answer the research questions of this study. Saunders and Lewis [35] define the research methodology as the collection of data, interpretation of data, and conclusion of information with a distinct objective. Methodology plans the techniques for collection of data and answering research questions [36]. This section describes the strategies for the research design and the research method used in this study.

4.1. Research Design. The research design is the process of conducting the research. It shows the plan of how the research is carried out. The objective of a research design is to propose the plan to make the experimental analysis that is used to answer the research question and provides sound results. A quantitative research approach is used in this study, which was initially developed in natural science to study science phenomena [37]. Today, quantitative research is well accepted in the education and social sciences. This method makes use of numerical analysis. The problem is determined from existing literature. The data collected from the target population through the survey tool is analyzed by statistical techniques such as descriptive statistics (percent distribution), Pearson correlation, standard multiple regression, and exploratory factor analysis.

4.2. Reasons for Choosing the Quantitative Research Method. The quantitative research method is selected on the ground that it depends on measurements of various scales to produce numbers that can be analyzed, using statistical methods, i.e., descriptive, and inferential statistics. The quantitative method is objective, unbiased, and value-free [37].

4.3. Research Strategy. The research strategy used for this research is a case study. According to [38], a case study is a research strategy not a specific method. According to [39], a case study explores properties, attitudes, effects, and actions of individual or group characteristics with the help of a questionnaire, observation, interviews, and document analysis. A case study is used for in-depth understanding of a certain phenomenon or problem.

4.4. Research Site. This research is carried out at Government Girls Secondary and Higher Secondary Schools of District Chitral, Khyber Pukhtunkhwa, Pakistan. The data is collected from the schools where the Wi-Fi-BCDs are fully functionally equipped. The Deputy Commissioner Chitral and District Education Officer (female) gave written permission to the researcher to collect the data from these schools. The data was collected from the headmistress, teachers, and students.

Table 2: Population and sample.

Respondent	Population (N)	Sample size (n)
Headmistress	25	20
Teacher	232	89
Student	3317	432

4.5. Population and Sample. According to [37], a population is a big group of individuals that obey explicit criteria, for which the scholar planned to simplify the results of the study. The number of schools and teachers is taken from the Annual Statistical Report Government School [4]. There are a total of 232 female teachers and 3317 students enrolled in 25 Girls Secondary and Higher Secondary Schools of Chitral. The data is shown in Table 2.

According to [40], at least 10% of the population has a reasonable representation for a large population. In the case of a small population, then 20% representation may be required. To determine the sample size, 40% is used for teachers. The number of students is large; for student sample size, [41] suggests that a 15% sample size is acceptable. A stratified random sampling technique is used because stratified random sampling is the best sampling technique as it produces a minimum sample error [42].

4.6. Data Collection Instrument. For collection of data, a close-ended questionnaire was developed. It comprises two parts: part A: demographic data (for teacher only) and part B: Wi-Fi-BCD from items 1 to 4 (4 items), SM from items 5 to 8 (4 items), practitioner confidence in using live technology (PCLT) from items 9 to 11 (3 items), and positive impact of content server on teaching (PICST) from items 12 to 14 (3 items), similarly, PEOU from 15 to 18 (4 items), PU from 19 to 22 (4 items), the ATU from 23 to 24 (2 items), and the BI from 25 to 26 (2 items). These (26) items were perceived through a Likert scale ranging from 1 to 5 options ("1: strongly disagree, 2: disagree, 3: undecided, 4: agree, and 5: strongly agree").

4.7. Reliability. Reliability is the measure of the degree to which a tool produces repeatable results subsequently for several trials [43]. To measure the reliability of the data gathered for research, the investigator uses a test-retest procedure in which the tool is checked over and again for related subjects. For reliability, this study examined the value of the following tests. For internal reliability, the value of Cronbach alpha ≥ 0.7044.

4.8. Validity. Validity is concerned that the findings and results are likely the same as expected. Validity is defined as a method measuring accurately what we suppose to measure. Validity is the accuracy and meaningfulness of research results [35]. Validity refers to the ability of an instrument that should be measured for a structure. According to [44], four tests that face effectiveness, convergence, build, and distinguish effectiveness are important to ensure the effectiveness of any structure.

4.9. Normality of the Data. The degree to which the distribution of sample data corresponds to a normal distribution [44], while normal distribution refers to such a distribution in which the horizontal axis shows the distribution of possible values of a variable. Apart from this, normality was also checked through KMO value and Bartlett test of sphericity.

4.10. Data Collection Procedure. For data collection, a quantitative approach will be used aimed at the generalization of research findings. In this research study, the constructs named as (1) demography, (2) Wi-Fi-BCD, (3) student motivation for higher outcomes, (4) practitioner confidence in using live technology, (5) positive impact of the content server on teaching and learning, (6) PEOU, (7) PU, (8) ATU, and (9) BI are conceptualized. A closed-ended questionnaire is used having statements with choices from 1 to 5 ("1: strongly disagree, 2: disagree, 3: undecided, 4: agree, and 5: strongly agree"). The data is used to collect data from the students, teachers, and headmistress of Girls Secondary and Higher Secondary Schools of District Chitral. The questionnaires were distributed and collected by the researcher to assure a best response rate.

4.11. Data Analysis. The data collected from both teachers and students is analyzed using the Statistical Package for the Social Sciences (SPSS) software. Data analysis is the procedure for providing the structure, order, and interpretation to collected data [45]. The data analysis will be carried out with descriptive and inferential statistics. Descriptive statistics is used to represent the data in the form of frequencies, percentages, mean, and standard deviations. The descriptive statistic represents the data in tabular and graphical method, shrinking a huge amount of information into smaller and sample representation for easy interpretation [37]. Descriptive statistics is used to analyze the demographic factor of teachers (age and experience). The inferential statistics include the regression, correlation, and the exploratory factor analysis. For research question 1, correlation, regression, and exploratory factor analysis are used to answer. For research question 2, descriptive statistics (frequencies, percentage, mean, and standard deviations) and regression analysis are used to answer. Similarly, for research question 3, the Pearson correlation is used to answer.

4.12. Research Questions. The following are the research questions.

(1) The relationship between variables of the modified TAM model (Wi-Fi-BCD, SM, PI, PC, PEOU, PU, ATU, and BI are the variables of the modified TAM model) is tested for the model fitness on the collected data.

To find the model fitness on the data collected, path analysis through Pearson correlation, regression analysis for model fit, and exploratory factor analysis (EFA) are used.

(2) Does the Wi-Fi-BCD improve students' motivation for high-quality lesson outcome?

To find the effect of Wi-Fi-BCD on student motivation, descriptive statistics (frequencies, percentages, mean, and standard deviation) and regression analysis are used.

(3) Is there any significant correlation between Wi-Fi-BCD, SM, PI, PEOU, PU, PEOU, ATU, and BI (SM = student motivation, PI = positive impact, PC = practitioner confidence, PEOU = perceived ease of use, PU = perceived usefulness, ATU = attitude towards using, and BI = behavioral intention) contributing to bridge the digital divide?

To find a significant relationship among Wi-Fi-BCD, SM, PC, PI, PU, PEOU, ATU, and BI, Pearson correlation analysis is used.

5. Results and Discussion

This section presents analysis, presentation, and interpretation of results and findings. The data is analyzed about the variables such as demographics, student motivation for higher outcomes, practitioner confidence in using live technology, positive impact of Wi-Fi-BCD on teaching and learning, Wi-Fi-BCD, PEOU, PU, ATU, and BI. Summaries are presented from the information, obtained from the questionnaires collected from students, teachers, and headmistress. Descriptive statistics, regression, and correlation are used to evaluate the relationship between these factors and to check the model appropriateness for the study. Results are obtained from the data on the following statistical methods.

5.1. Reliability. Reliability is the measure of the degree to which a tool produces repeatable results subsequently for several trials [43]. To measure the reliability of the data gathered for research, the investigator used a test-retest procedure in which the tool is checked over and over again for related subjects. Data collected initially was a pilot test of the tools to confirm that they produce the required data or results. The tools were ensured to be reliable after they produced a reliability coefficient greater than 0.7 for its entire constructs. According to [44], Cronbach alpha greater than 0.7 is observed as sufficient and having good reliability. Cronbach alpha is used in the study to measure reliability, which is commonly used to measure internal consistency. The main objective of reliability is to make sure of the internal consistency of the instrument and its construct [46]. The measured values of different indices for reliability are given in Table 3, which shows that all the constructs are Cronbach alpha greater than 0.7.

5.2. Validity. Validity means the ability of an instrument that should be measured for a structure. Validity refers to the degree of correctness, significance, and meaningfulness of inference based on research results [43]. Validation of the data is completed via content validity. According to [44], four tests that face effectiveness, convergence, build,

Table 3: Reliability of the construct.

Construct name	Symbol	Cronbach alpha
Wi-Fi-Based Content Distributor	Wi-Fi-BCD	0.786
Student motivation	SM	0.804
Practitioner confidence on live technology	PC	0.749
Positive impact of Wi-Fi-BCD	PI	0.884
Perceived ease of use	PEOU	0.806
Perceived usefulness	PU	0.801
Attitude towards using	ATU	0.807
Behavioral intention	BI	0.783

and distinguish effectiveness are important to ensure the effectiveness of any structure.

5.2.1. Content Validity. According to [47], content validity measuring of instruments is essential. Content validity helps to ensure construct validity, which provides confidence to the researchers about the tool. Content validity shows the degree that the tool covers the content that it is developed to measure. According to Lynn [48], researchers calculated two types of CVIs: the content validity of individual items (I-CVI) and the content validity of the overall scale (S-CVI). The content validity index (CVI) is the percentage of respondents that rated the item as 4 or 5 (in the rating scale from 1 to 5 where 5 represents the highest agreement) [48]. The formula for calculating the I-CVI and S-CVI is given below:

$$\text{CVI} = \frac{\text{Number of judge rate options 4 or 5}}{\text{Total number of judges}} * 100. \quad (1)$$

According to [49], the acceptable CVI values should be 1 for a few expert three or four judges, 0.80 for 5 experts, and 0.78 for more than 5 or larger expert groups. The acceptable CVI for this study is determined to be 0.78. The I-CVI value for the study ranges from 0.08 to 0.09, and the S-CVI value for the study ranges from 0.798 to 0.92.

5.2.2. Discriminant Validity. According to [50], if discriminant validity is not performed and factor analysis is misinterpreted, then results and measurement scales used in research may not function properly, and the assumptions made about relationships between variables under research may be incorrect. For example, if a relationship is confirmed between variables and in real it is no relationship, or the strength or weakness of the relationship is miscalculated (Type II error). The Pearson correlation matrix is used to determine the discriminant validity, which measures the linear dependence between two variables [36], suggesting a score between +1 and -1. Table 4 shows that all values are in the range of +1 and -1; hence, the study meets the validity requirement [36].

5.3. Normality. A normality test is used to check that the data set is having a normal distribution. The skewness and kurtosis are checked to find the distribution of the data set. The normality test can also be performed with the peak (kurtosis) and the skewness of a distribution through a graphical representation [51]. The results are obtained with descriptive statistics of the factors to indicate that all the factors are skewed (either positive or negative), and the same has for kurtosis (either positive or negative). A test of normality is important as a normal distribution is usually essential for most inferential analyses [51].

Table 5 shows the estimations of skewness and kurtosis (for normal data, the z value should be in the range of -1.96 to +1.96). In view of Table 5, the skewness and kurtosis are in normal span for the variable in this study (Wi-Fi-BCD, SM, PC, PU, PEOU, ATU, and Wi-Fi-BCD) are regularly appropriated. KMO value = 0.846 ($p \leq 0.01$). The value of KMO indicates that the sample for the study is adequate [52]. Next, the ordinarily utilized Kaiser-Meyer-Olkin (KMO) measure of examining ampleness and Bartlett's test of sphericity was raced to see whether the specimen estimate is fitting for factor investigation and the quality of the relationship among the factors is critical [53].

5.4. Demographic Information of the Respondents. For in-depth understanding of the questionnaire responses and the data examination, it is worth investigating the respondent's demographic information. The demographic information of respondent (teachers) age and teaching experience is shown in Tables 6 and 7.

Table 6 shows the demographic data of the 89 female teachers. Their ages range from 25 to 45 years. Out of these respondents, 5% were of the age of 25, followed by 28% teachers whose ages were in the range of 26-30 years. Similarly, the teachers aging between 31 and 35 years were 47%. Furthermore, teachers who were in the age range of 36-40 were 15%. Lastly, teachers having age of more than 40 were 5%. This can be seen in Figure 6. Table 7 shows that only 1% of the teachers have an experience less than 1 year; 23.6% of the teachers have an experience from 1 to 5 years. Similarly, 31.5% of the teachers have an experience in the range of 6-10 years, where 31.5% of the teachers have experience from 11 to 15 years. Furthermore, 10.1% of the teachers were in the range of an experience from 16 to 20 years. Lastly, 2.2% of the teachers have an experience of 21 years and above. Figure 7 represents the experience of the teachers.

5.5. Research Questions. To extract the information from the data collected and to find the effect of Wi-Fi-BCD for bridging the digital divide in Government Girls Secondary and Higher Secondary Schools of Chitral, the result of the research questions of this study will be analyzed and presented.

5.5.1. Research Question #1. Does the relationship between variables (Wi-Fi-BCD, SM, PC, PI, PEOU, PU, ATU, and BI) of the model fit and significant on the data collected?

(1) Path Analysis for the Model Fitness. As per the modified TAM model, the following path relationship is tested to know the effects of Wi-Fi-BCD for model fitness on the data collected.

Table 4: Pearson correlation matrix.

	Wi-Fi-BCD	SM	PC	PI	PEOU	PU	ATU	BI
Wi-Fi-BCD	1							
SM	0.794*	1						
PC	0.821*	0.430*	1					
PI	0.807*	0.967*	0.826*	1				
PEOU	0.590*	0.470*	0.309*	0.333*	1			
PU	0.794*	1.000*	0.430*	0.967*	0.470*	1		
ATU	0.487*	0.424*	0.300*	0.4004	0.541*	0.424*	1	
BI	0.644*	0.772*	0.453*	0.726*	0.458*	0.772*	0.443*	1

Table 5: Normality of the data.

Variable	Wi-Fi-BCD	SM	PC	PI	PEOU	PU	ATU	BI
Mean	3.64	3.73	3.77	3.77	3.87	3.73	3.64	3.97
Std. deviation	0.926	1.09	0.927	1.13	1.10	1.09	1.04	1.07
Variance	0.859	1.19	0.861	1.29	1.22	1.19	1.09	1.14
Skewness	-1.093	-1.28	-1.006	-1.30	-1.50	-1.28	-1.14	-1.14
Kurtosis	0.689	0.385	0.545	0.372	0.899	0.385	0.705	1.24

Table 6: According to age.

Gender	Freq	% age
Female	89	100
25 years	5	5
26-30 years	25	28
31-35 years	42	47
36-40 years	13	15
41-45 years	4	5

Table 7: According to experience.

Gender	Freq	% age
Less than year	01	1.1
1-5 years	21	23.6
6-10 years	28	31.5
11-15 years	28	1.5
16-20 years	09	10.1
21 years above	02	2.2

(1) Relationship of Wi-Fi-BCD with SM, PI, and PC

(2) Relationship of Wi-Fi-BCD with PEOU and PU

(3) Relationship of PEOU and PU

(4) Relationship of PEOU and PU with ATU

(5) Relationship of ATU and BI

The above path relationships are tested to know the effects of Wi-Fi-BCD for model fitness on collected data. To find out the relationship, Pearson's correlation test is used, to examine the (positive or negative, strong, or weak) association between the variables. Table 8 displays the Pearson correlation matrix.

(2) Relationship of Wi-Fi-BCD with SM, PI, and PC. Table 8 presents that Wi-Fi-BCD has a highly significant positive correlation between SM, PI, and PC, Wi-Fi-BCD with SM ($r = 0.794$, $p \leq 0.01$), Wi-Fi-BCD with PI ($r = 0.807$, $p \leq 0.01$), and Wi-Fi-BCD with PC ($r = 0.821$, $p \leq 0.01$).

(3) Relationship of Wi-Fi-BCD with PEOU and PU. Table 8 presents that Wi-Fi-BCD has a highly significant positive correlation between PEOU and PU, Wi-Fi-BCD with PEOU ($r = 0.590$, $p \leq 0.01$), and Wi-Fi-BCD with PU ($r = 0.794$, $p \leq 0.01$).

(4) Relationship of PEOU and PU. Table 8 shows PEOU and PU relationship with $r = 0.470$, $p \leq 0.01$; the PEOU and PU have significant positive relationships.

(5) Relationship of PEOU and PU with ATU. Table 8 shows PEOU and ATU relationship with $r = 0.541$, $p \leq 0.01$; the relationship of PU and ATU is $r = 0.424$, $p \leq 0.01$. The relationship of PEOU and PU with ATU is significantly positive.

(6) Relationship of ATU and BI. Table 8 shows ATU and BI relationship with $r = 0.443$, $p \leq 0.01$; the relationship of ATU with BI is significantly positive. All of the above relationships are highly positively significant with each other, which shows that the stated path relationships mentioned in the previous section are supported by the data collected. The model fitness for the data collected is also tested with regression analysis for R^2 and F value.

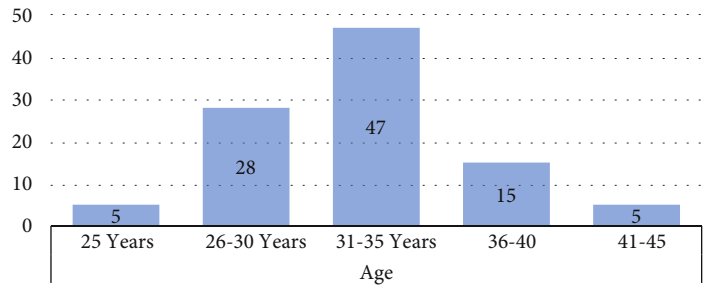


Figure 6: Classification according to age.

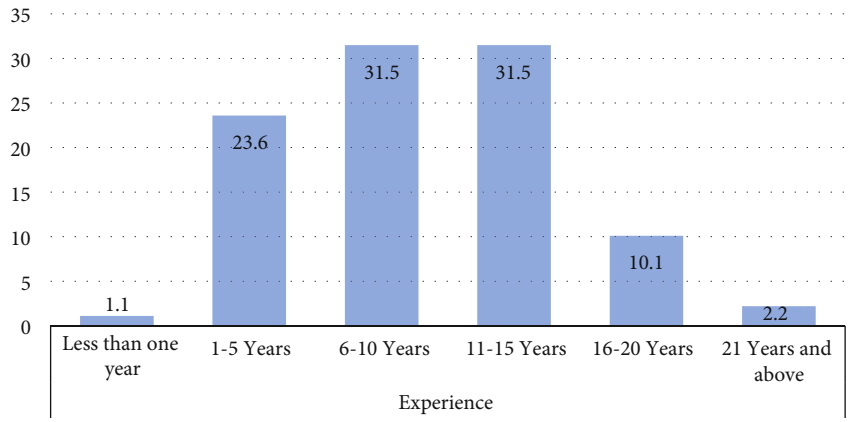


Figure 7: Classification according to experience.

Table 8: Pearson correlation matrix.

	Wi-Fi-BCD	SM	PC	PI	PEOU	PU	ATU	BI
Wi-Fi-BCD	1							
SM	0.794*	1						
PC	0.821*	0.430*	1					
PI	0.807*	0.967*	0.826*	1				
PEOU	0.590*	0.470*	0.309*	0.333*	1			
PU	0.794*	1.000*	0.430*	0.967*	0.470*	1		
ATU	0.487*	0.424*	0.300*	0.400*	0.541*	0.424*	1	
BI	0.644*	0.772*	0.453*	0.726*	0.458*	0.772*	0.443*	1

5.5.2. Regression Analysis for Model Fitness. Model fitness can be checked through model summary obtained through regression analysis. Similarly, p value in ANOVA can also confirm model fitness. The variable Wi-Fi-BCD is entered as a dependent variable, whereas PEOU, PU, ATU, and BI are entered as independent variables (predictors). From Table 9, it is found that Wi-Fi-BCD is entered as a dependent variable, whereas PU, PEOU, ATU, and BI are entered as independent variables. Results of the model summary show that the value of R^2 (0.777) shows that the proposed model is 77% found to fit with the collected data. Similarly, Table 10 shows the p value (0.01) in ANOVA results, which also confirms model fitness on the collected data.

5.5.3. Exploratory Factor Analysis (EFA). Sample size is important for EFA analysis; according to Field (2005), a sample size of 300 is fit for EFA analysis, while other researchers proposed 10 samples for each item [54], proposed for each item of 5 cases. The sample size for our study is 521, which is fit to perform EFA. The KMO value must be greater than 0.5; if the KMO value is less than 0.5, it means that the sample is not fit to perform the EFA.

Exploratory factor analysis is used to find meaningful patterns in large data sets; EFA has the advantage as it determines which of the items hang together in the form of sets. The EFA method is used to highlight many relationships among different items in a more simple and parsimonious way.

Table 9: Model summary.

Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.882a	0.777	0.767	0.44773

Table 10: Table of ANOVA.

Model	Sum of squares	Df	Mean square	F	Sig.	
1	Regression	58.772	4	14.693	73.295	0.000b
	Residual	16.839	84	0.200		
	Total	75.611	88			

5.5.4. Scree Plot. The scree plot in Figure 8 shows the variation of the curve. The scree plot shows the variance explained by each item in the EFA analysis. The y-axis is the eigenvalue, and the x-axis is the factors.

5.5.5. KMO and Bartlett's Test. The Kaiser-Meyer-Olkin measure of sampling adequacy value should be greater than 0.70. The significant value in the last row should be significant (less than 0.05), indicating that the correlation matrix is significantly different from an identity matrix, in which correlations between variables are all zero. The KMO and Bartlett test of sphericity are used to explain that correlations between items were appropriately high for EFA. Table 11 shows KMO and Bartlett test.

The purpose of EFA analysis is to determine empirically whether the participants in our survey responded similarly to questions based on PEOU, PU, ATU, and BI. For this purpose, we run factor analysis using the principal axis factoring method and specify the number of factors to be four (because our conceptualization of the underlying phenomena is based on a scale of four variables, i.e., PEOU, PU, ATU, and BI. Table 12 shows the EFA correlation matrix.

A positive correlation between two variables means that both are in the same direction, while the negative correlation means that if the value of one variable increases, the other value decreases.

5.5.6. Communalities. The basic purpose of communalities is to represent the relation between all the questions (i.e., the squared multiple correlations between the item and all other items). Table 13 shows the communalities of EFA; communalities are said to be R^2 for each of the variables included in the EFA analysis. Table 13 shows the percentage of variance of each item that is explained by the elements.

5.5.7. Rotated Factor Matrix. This is one of the important matrices in EFA analysis. Actually, this is where we see a number of factors underlying a construct. First, we observe the table and count physically how many clusters there are. So, if there are three clusters, it means that there are three variables. If there are four clusters, it means that there are four variables. Ideally, we should first count the clusters, then analyze each question within the cluster, and based on thematic analysis, we infer information about the variable. In simple words, we read all the questions within each cluster again and again to see what common area they are focusing on. In our study, there are four clusters which mean that the items are loading exactly the way they were expecting. Table 14 shows that all factor loading is greater than 0.6 except the PEOU1 and PEOU4, according to [55], for a good fitness sample size not matter, but the factor loading shall be greater than 0.6. The KMO value is 0.860 ($p \leq 0.01$), and the Cronbach alpha of all the constructs is greater than 0.07. According to Table 14, the model is a reliable one to fit the collected data. Table 15 shows the descriptive statistics of EFA Analysis.

5.6. Research Question #2. Does the Wi-Fi-BCD effect improve students' motivation for high-quality lesson outcomes?

5.6.1. Descriptive Analysis of the Question. To organize, summarize, and reduce, a large set of information descriptive statistics is used, which transforms a set of observations into numbers [37]. Graphical representation of information and calculation of range, minimum, maximum, mean, mode, median, standard deviation, and variance are based on descriptive statistics. The graphical representation technique is used to reduce a large scale of information into smaller representation for easy interpretation [37].

Based on the exploratory nature of the research question, descriptive statistics is used to analyze the data. While presenting the descriptive reports, only high percentages have reported against each statement. Six statements with five response options are asked from the respondent with the objective to know if the Wi-Fi-BCD improves students' motivation.

Table 16 presents the concise results in which the respondent is either strongly disagree, disagree, undecided, agree, or strongly agree to the research question that the Wi-Fi-BCD improves student motivation. The respondents strongly agree, and agree responses against each question are as follows.

(i) Learning becomes fun (75%)

(ii) Enhances students' interest (71%)

(iii) Motivates to learn new things (73%)

(iv) Most of the students take interest in using (72%)

Table 17 shows that the mean value of all the questions is greater than 3, which reflects that the response of the majority respondents is either strongly agree or agree that the Wi-Fi-Based Content Distributor improves student motivation.

5.6.2. Inferential Analysis of the Question. According to [50], regression analysis is an appropriate method for TAM models to measure the relationship between model variables. According to [20], a linear regression model is frequently used in TAM model analysis. Linear regression is based to

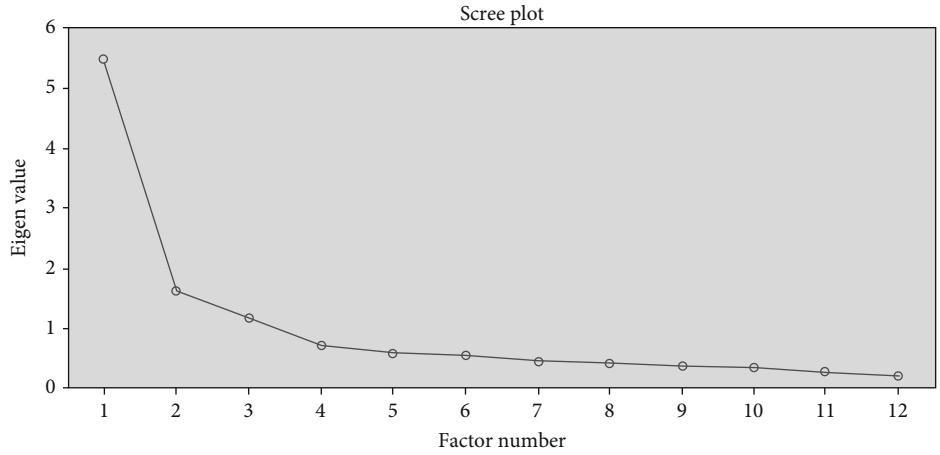


Figure 8: Scree plot.

Table 11: KMO and Bartlett test.

Kaiser-Meyer-Olkin measure of sampling adequacy		0.860
Bartlett's test of sphericity	Approx. chi square	3103.525
	Df	66
	Sig.	0.000

find the relationship between variables and provides in-depth and complex estimation of the relationship between variables. In view of the above discussion, regression analysis supports comparison of results and used in similar research. Therefore, in this study, it is also used.

Table 17 presents the regression analysis of the dependent variable Wi-Fi-BCD and independent variables (SM). The result shows that independent variables (SM) affect Wi-Fi-BCD positively. The table further explains that the predictor SM (beta = 0.794) best predicts the outcome Wi-Fi-BCD and has a high effect on the dependent variable (Wi-Fi-BCD).

6. Discussion

From the above descriptive analysis, Wi-Fi-BCDs affect students' motivation. Based on the observations of the respondents, the Wi-Fi-BCD enhances individual experience, develops digital skills to access ICT services, and makes learning fun. Moreover, the Wi-Fi-BCD enhances student interest and motivates them to learn new things. The majority of the students take interest in interacting with Wi-Fi-BCD via a tablet computer.

From the inferential statistical analysis, the regression result shows that independent variables (SM) affect Wi-Fi-BCD positively (SM, beta = 0.794) b, which shows a high effect on the dependent variable (Wi-Fi-BCD).

The majority of the government globally invests a huge amount in ICT to improve teaching and learning in schools [8, 56] suggesting that teachers can make their lessons motivating and enriching, with the help of easy access to digital materials. The Wi-Fi-BCD provides easy access to digital materials for teachers and students. The main state in the United Nations (US) formed a one-to-one laptop environment for 5,000 teachers and over 42,000 middle school students.

A survey was conducted to gather information about teachers' use of the laptops and to know teachers' perceptions of their students' use of laptops. About 80% of the teachers stated that students are more actively engaged and involved in learning new things for high-quality work when using laptops. Students' motivation and class participation improved by using ICT, and the same were confirmed by both the principal and teachers [57].

The teachers and students also affirm that Wi-Fi-BCD increases student motivation. According to DfES (2003) (Department for Education and Skills, United Kingdom), research projects find that ICT plays a major role in students' motivation and encourages them to learn new things, inside and outside the classroom. The path relationship of Wi-Fi-BCD with SM is tested and found significant positive. The findings of the Wi-Fi-BCD are in line with findings of previous research; thus, it is concluded that the Wi-Fi-BCD improves student motivations.

6.1. Research Question #3. Is there any significant correlation relationship between Wi-Fi-BCD, SM, PC, PI, PU, PEOU, ATU, and BI contributing to bridging the digital divide?

According to [58], correlation coefficients measure the linear relationship between two or more variables. It also measures the strength and direction of the relationship between variables. It shows the range of association or dependence between variables. The correlation coefficient value (r) ranges from +1 to -1; zero means no relationship between variables. The negative value means the opposite relationship between variables if one increases and the other decreases. A positive correlation means that both the variables move in the same direction. If one of the variables increases, the other also increases.

Table 12: EFA correlation matrix.

	PEOU1	PEOU2	PEOU3	PEOU4	PU1	PU2	PU3	PU4	ATU1	ATU2	BI1	BI2
PEOU1	1											
PEOU2	0.506	1										
PEOU3	511	0.609	1									
PEOU4	0.550	0.436	0.439	11								
PU1	0.380	0.278	0.317	0.394	1							
PU2	0.203	0.351	0.327	0.228	0.518	1						
PU3	0.273	0.259	0.362	0.248	0.497	0.565	1					
PU4	0.328	0.265	0.234	0.356	0.559	0.423	0.453	1				
ATU1	0.322	0.533	0.445	0.281	0.275	0.387	0.377	0.279	1			
ATU2	0.337	0.455	0.484	0.283	0.260	0.371	0.349	0.207	0.678	1		
BI1	0.334	0.339	0.328	0.295	0.484	0.500	0.674	0.525	0.408	0.435	1	
BI2	0.367	0.312	0.294	0.375	0.520	0.443	0.481	0.759	0.339	0.290	0.643	1

Table 13: EFA communalities.

Variable	Initial	Extraction
PEOU1	0.454	0.582
PEOU2	0.521	0.577
PEOU3	0.498	0.551
PEOU4	0.391	0.499
PU1	0.475	0.498
PU2	0.459	0.469
PU3	0.559	0.558
PU4	0.621	0.652
ATU1	0.544	0.607
ATU2	0.53	0.615
BI1	0.618	0.638
BI2	0.668	0.676

Table 14: Factor loading matrix.

Item	Factor				Cronbach alpha
	PEOU	PU	ATU	BI	
PEOU1	0.599				0.806
PEOU2	0.627				
PEOU3	0.625				
PEOU4	0.560				
PU1		0.648			0.801
PU2		0.626			
PU3		0.668			
PU4		0.662			
ATU1			0.634		0.807
ATU2			0.612		
BI1				0.735	0.783
BI2				0.721	

Table 15: Descriptive statistics.

	Exploratory factor analysis	
Variable	Mean	Std. deviation
PEOU1	4.16	1.058
PEOU2	3.92	1.007
PEOU3	3.9	1.066
PEOU4	4.07	0.962
PU1	3.93	1.047
PU2	3.80	1.051
PU3	3.81	1.102
PU4	3.95	1.101
ATU1	3.74	0.969
ATU2	3.63	1.034
BI1	4.18	0.959
BI2	4.22	0.960

To find out the correlation significance between Wi-Fi-BCD, SM, PC, PI, PEOU, PU, ATU, and BI, Pearson's correlation coefficient is used.

According to [54, 59], a correlation range of $r = \pm 0.1$-± 0.29 is considered small effect, a range of correlation $r = \pm 0.30$-± 0.49 is considered moderate, while $r = \pm 0.50$-± 1.0 is considered high.

6.2. Discussion. According to [19, 30], PEOU and PU have been presented to be the key factors of information technology acceptance and usage. According to [60] in their studies using the TAM model, it emphasized a positive relationship between PEOU and PU [30, 60], suggesting that PU has effect on BI. According to [46] finding, PU directly positively affects teacher's behavior and intention to use technology.

According to [19, 30], the most commonly acknowledged factors of the TAM model are PEOU, PU, and ATU; the PEOU and PU are independent variables to predict the ATU.

Table 16: Wi-Fi-BCD improves student motivation.

Statement	Response	Freq	% age	Mean	Std. div
Learning becomes fun	Strongly disagree	23	4.4	3.73	1.25
	Disagree	32	6.1		
	Undecided	75	14.4		
	Agree	225	43.3		
	Strongly agree	166	31.9		
Enhances student interest	Strongly disagree	22	4.2	3.87	1.29
	Disagree	48	9.2		
	Undecided	79	15.2		
	Agree	238	45.7		
	Strongly agree	134	25.7		
Motivates to learn new things	Strongly disagree	23	4.4	3.72	1.23
	Disagree	58	11.1		
	Undecided	62	11.9		
	Agree	229	44		
	Strongly agree	149	28.6		
Most of the students take interest in using	Strongly disagree	27	5.2	3.64	1.37
	Disagree	30	5.8		
	Undecided	84	16.1		
	Agree	187	35.9		
	Strongly agree	187	35.9		

Similarly, the PU is an independent variable to predict the BI [61, 30], suggesting that BI is the main element that decides whether the users will accept and utilize the technology.

Based on the above discussion, Table 18 shows that the correlation between Wi-Fi-BCD and SM is 0.794 with $p \leq 0.01$, between Wi-Fi-BCD and practitioner confidence (PC) is 0.821 ($p \leq 0.01$), between Wi-Fi-BCD and positive impact on teaching and learning (PI) is 0.807 ($p \leq 0.01$), between Wi-Fi-BCD and PEOU is 0.590 ($p \leq 0.01$), between Wi-Fi-BCD and PU is 0.794 ($p \leq 0.01$), between Wi-Fi-BCD and ATU is 0.487 ($p \leq 0.01$), and between Wi-Fi-BCD and BI is 0.644 ($p \leq 0.01$).

The correlation between the major constructs of the TAM model, PEOU, PU, ATU, and BI also shows a significant positive relationship.

Table 18 shows that the correlation between PEOU and PU is 0.470 ($p \leq 0.01$), between PEOU and ATU is 0.541 ($p \leq 0.01$), between PU and ATU is 0.424 ($p \leq 0.01$), between PU and BI is 0.772 ($p \leq 0.01$), and between ATU and BI is 0.443 ($p \leq 0.01$).

6.3. Path Analysis. As per the modified TAM model for the study, the path relationships set earlier are tested to examine the (positive or negative, strong, or weak) association between the variables. Table 19 displays the results of the path analysis.

Table 19 displays that all paths are strongly positively significant, the Wi-Fi-BCD positively affects its dependent variables, and PEOU, PU, and ATU also affect the dependent variables. It means that the effect of Wi-Fi-BCD is significantly positive for bridging the digital divide. As per the above findings, the result is in line with previous studies [17–19, 60, 61].

Table 17: Regression analysis of Wi-Fi-BCD and SM.

Model	Unstandardized coefficients	Standardized coefficients		T	Sig
	B	Std. error	Beta		
SM	1.125	0.216	0.794	5.221	0.000
	0.673	0.055		12.165	0.000

The easiness and acceptance of Wi-Fi-BCD positively affect the PEOU and PU. Similarly, the PEOU positively affects the PU. The PEOU and PU as independent variables positively predict the dependent variable ATU. The PU is an independent variable for the BI, positively affecting the BI. Similarly, the independent variable ATU for BI also predicts the dependent variable BI.

The ATU is the use of the Wi-Fi-BCD. The research is in line with the previous research of [62]. Based on the above discussion, it is concluded that the PEOU and PU of the Wi-Fi-BCD are positively correlated with the Wi-Fi-BCD. It shows that the students and teachers found the Wi-Fi-BCD easy and useful. Similarly, the correlation between PEOU and PU positively correlates with the ATU; it shows the attitude towards using the Wi-Fi-BCD by teachers and students. The ease of use of the system affects the users to make positive ATU of that system [30].

Table 18: Pearson correlation matrix.

	Wi-Fi-BCD	SM	PC	PI	PEOU	PU	ATU	BI
Wi-Fi-BCD	1							
SM	0.794**	1						
PC	0.821**	0.430**	1					
PI	0.807**	0.967**	0.826**	1				
PEOU	0.590**	0.470**	0.309**	0.333**	1			
PU	0.794**	1.000**	0.430**	0.967**	0.470**	1		
ATU	0.487**	0.424**	0.300**	0.400**	0.541**	0.424**	1	
BI	0.644**	0.772**	0.453**	0.726**	0.458**	0.772**	0.443**	1

Table 19: Path analysis of modified TAM model.

Dependent variable	Independent variable	r	P	Test
Wi-Fi-BCD	SM	0.794	0.000	Positively significant
Wi-Fi-BCD	PC	0.821	0.000	Positively significant
Wi-Fi-BCD	PI	0.807	0.000	Positively significant
Wi-Fi-BCD	PEOU	0.590	0.000	Positively significant
Wi-Fi-BCD	PU	0.794	0.000	Positively significant
PEOU	PU	0.740	0.000	Positively significant
PEOU	ATU	0.541	0.000	Positively significant
PU	ATU	0.424	0.000	Positively significant
ATU	BI	0.443	0.000	Positively significant

7. Conclusion and Future Work

This research investigated the effects of Wi-Fi-BCD to bridge the digital divide in Government Girls Secondary and Higher Secondary Schools of Chitral. The project of Wi-Fi-BCD provides facilities for female students in Chitral to use ICT and to get hands-on experience in the field. The Wi-Fi-BCD is developed for the Government Girls Secondary and Higher Secondary School of Chitral to provide ICT facilities successfully. The modified TAM model is used to find the effects of Wi-Fi-BCD for bridging the digital divide. The path relationship is found significantly positive. The effects of Wi-Fi-BCD on student's motivation, practitioner confidence, teaching, and learning are meaningfully positive. The effects of Wi-Fi-BCD on its PEOU and PU are found pointedly positive. The PEOU and PU positively affect the ATU; this indicates that the Wi-Fi-BCD is easy to use and it is useful. The ATU positively affects the BI; this indicates the actual usage of the Wi-Fi-BCD. The results of this research show that Wi-Fi-BCD has quite positive effects on bridging the digital divide. The teachers agree that the use of Wi-Fi-BCD is applicable in practical situations, and it may help the female learners in the field of ICT. Future research may include developing local language contents in Wi-Fi-BCD to remove the language barriers from the subject of content distributor.

Data Availability

The data collected during the data collection phase will be provided upon request to the authors.

Conflicts of Interest

The authors declare no potential conflict of interests.

Acknowledgments

The work presented in this paper was supported by the China University of Petroleum-Beijing, Fundamental Research Funds for the Central Universities under Grant number 2462020YJRC001. The authors also acknowledge the support provided by Prof. Dr. Khan Bahadar Marwat (SI) and Prof. Dr. Lutfullah Kakakhel throughout the execution of this project.

References

[1] A. Zafar, "Woes of primary education in Khyber Pakhtunkhwa," 2018, https://chitraltoday.net/2018/02/15/woes-of-primary-education-in-khyber-pakhtunkhwa/.

[2] K. Waqas, "ICT in education, reached? Nowhere!," 2018, https://archive.pakistantoday.com.pk/2017/06/18/ict-in-education-reached-nowhere/.

[3] Report, "Annual Statistical Report Government Schools 2014-15," 2016, http://library.aepam.edu.pk.

[4] ASER, *Annual Status of Education Report: ASER Pakistan 2015 National (Urban)*, Federal Ministry of Education, Pakistan, 2015.

[5] Statistic Pakistan, *Annual Statistical Report Government Schools*, Department of Elementary & Secondary Education, Khyber Pakhtunkhwa, Pakistan, 2015.

[6] CIDA, *CIDA Strategy on Knowledge for Development through ICT*, Canadian International Development Agency, 2006.

[7] M. O. Yusuf, "Information and communication technology and education: analysing the Nigerian national policy for information technology," *International Education Journal*, vol. 6, no. 3, pp. 316–321, 2005.

[8] C. Buabeng-Andoh, "Factors influencing teachersâ adoption and integration of information and communication technology into teaching: a review of the literature," *International*

Journal of Education and Development using ICT, vol. 8, no. 1, 2012.

[9] M. Andrade-Aréchiga, G. López, and G. López-Morteo, "Assessing effectiveness of learning units under the teaching unit model in an undergraduate mathematics course," *Computers & Education*, vol. 59, no. 2, pp. 594–606, 2012.

[10] R. J. Khobo, *The effect of using computers for the teaching and learning of Mathematics to grade 10 learners at secondary school, [M.S. thesis]*, University of South Africa, Pretoria, South Africa, 2015.

[11] M. Marshman and P. Grootenboer, "Scissors, papers rock: old-world technologies for future-proofing pedagogy. Re-engaging students in mathematics classrooms," in *Transformative Approaches to New Technologies and Student Diversity in Futures Oriented Classrooms*, pp. 139–158, Springer, 2012.

[12] OECD, "Organization for economic co-operation and development," 2001, http://www.oecd.org/dataoecd.

[13] H. Ono and M. Zavodny, "Digital inequality: a five country comparison using microdata," *Social Science Research*, vol. 36, no. 3, pp. 1135–1155, 2007.

[14] J. Van Dijk and K. Hacker, "The digital divide as a complex and dynamic phenomenon," *The Information Society*, vol. 19, no. 4, pp. 315–326, 2003.

[15] J. A. Van Dijk, "Digital divide: impact of access," *The International Encyclopedia of Media Effects*, vol. 1, pp. 1–11, 2017.

[16] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, "Past, present, and future of decision support technology," *Decision Support Systems*, vol. 33, no. 2, pp. 111–126, 2002.

[17] R. J. Vallerand, P. Deshaies, J.-P. Cuerrier, L. G. Pelletier, and C. Mongeau, "Ajzen and Fishbein's theory of reasoned action as applied to moral behavior: a confirmatory analysis," *Journal of Personality and Social Psychology*, vol. 62, no. 1, pp. 98–109, 1992.

[18] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: toward a unified view," *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, 2003.

[19] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models," *Management Science*, vol. 35, no. 8, pp. 982–1003, 1989.

[20] P. Legris, J. Ingham, and P. Collerette, "Why do people use information technology? A critical review of the technology acceptance model," *Information & Management*, vol. 40, no. 3, pp. 191–204, 2003.

[21] D. W. Turner III, "Qualitative interview design: a practical guide for novice investigators," *The Qualitative Report*, vol. 15, no. 3, pp. 754–760, 2010.

[22] J.-H. Wu and S.-C. Wang, "What drives mobile commerce?: an empirical evaluation of the revised technology acceptance model," *Information & Management*, vol. 42, no. 5, pp. 719–729, 2005.

[23] S. R. Hill and I. Troshani, "Adoption of personalisation mobile services: evidence from young Australians," *BLED 2009 Proceedings*, vol. 35, 2009.

[24] Y. Lee and K. A. Kozar, "The technology acceptance model: past, present, and future," *Larsen*, pp. 752–780, 2003.

[25] K. Mathieson, "Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior," *Information Systems Research*, vol. 2, no. 3, pp. 173–191, 1991.

[26] A. N. Luhamya, F. E. K. Bakkabulindi, and P. B. Muyinda, "Using the theory of planned behaviour to explain the integration of ICT in teaching and learning by educators in public teacher training colleges in Tanzania," *International Journal of Computing & ICT Research*, vol. 11, no. 2, 2017.

[27] J. A. G. M. van Dijk, O. Peters, and W. Ebbers, "Explaining the acceptance and use of government Internet services: a multivariate analysis of 2006 survey data in the Netherlands," *Government Information Quarterly*, vol. 25, no. 3, pp. 379–399, 2008.

[28] S. Järvelä, "Shifting research on motivation and cognition to an integrated approach on learning and motivation in context," *Motivation in learning contexts: Theoretical advances and methodological implications*, pp. 3–14, 2001.

[29] P. A. Ertmer and A. T. Ottenbreit-Leftwich, "Teacher technology change," *Journal of research on Technology in Education*, vol. 42, no. 3, pp. 255–284, 2010.

[30] V. Venkatesh and M. G. Morris, "Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior," *MIS Quarterly*, vol. 24, no. 1, pp. 115–139, 2000.

[31] S. Taylor and P. A. Todd, "Understanding information technology usage: a test of competing models," *Information Systems Research*, vol. 6, no. 2, pp. 144–176, 1995.

[32] N. Jahangir and N. Begum, "The role of perceived usefulness, perceived ease of use, security and privacy, and customer attitude to engender customer adaptation in the context of electronic banking," *African Journal of Business Management*, vol. 2, no. 2, pp. 032–040, 2008.

[33] H. Amin, "Internet banking adoption among young intellectuals," *The Journal of Internet Banking and Commerce*, vol. 12, no. 3, pp. 1–13, 1970.

[34] M. Leonard, S. Graham, and D. Bonacum, "The human factor: the critical importance of effective teamwork and communication in providing safe care," *BMJ Quality & Safety*, vol. 13, Supplement 1, pp. 85–90, 2004.

[35] M. N. K. Saunders and P. Lewis, *Doing Research in Business & Management: An Essential Guide to Planning your Project*, Pearson, 2012.

[36] N. Malhotra and D. Birks, *Marketing Research: An Applied Approach: 3rd European Edition*, Pearson Education, 2007.

[37] J. H. McMillan and S. Schumacher, *Research in Education: Evidence-Based Inquiry, MyEducationLab Series*, Pearson, 2010.

[38] R. K. Yin, "The case study crisis: some answers," *Administrative Science Quarterly*, vol. 26, no. 1, pp. 58–65, 1981.

[39] H. C. J. Godfray, J. R. Beddington, I. R. Crute et al., "Food security: the challenge of feeding 9 billion people," *Science*, vol. 327, no. 5967, pp. 812–818, 2010.

[40] L. R. Gay, G. E. Mills, and P. Airasian, *Educational Research: Competencies for Analysis and Application*, Pearson, 4th edition, 1992.

[41] N. J. Kathuri and D. A. Pals, *Introduction to Educational Research*, Scientific Research Publishing, Njoro Kenya, 1993.

[42] Y. P. Chua, *Research Method*, McGraw-Hill Companies, Malaysia, 2011.

[43] O. M. Mugenda and A. G. Mugenda, *Qualitative and quantitative approaches*, Research Methods Africa Center for Technology Studies (Acts) Press, Nairobi Kenya, 2003.

[44] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham, *Análise multivariada de dados*, Bookman Editora, 2009.

[45] C. Marshall and G. B. Rossman, "The "what" of the study: building the conceptual framework," *Designing Qualitative Research*, vol. 3, no. 3, pp. 21–54, 1999.

[46] T. Teo, *Technology Acceptance Research in Education*, Springer, 2011.

[47] F. Yaghmaie, "Content validity and its estimation," *Journal of Medical Education*, vol. 3, no. 1, 2003.

[48] M. R. Lynn, "Determination and quantification of content validity," *Nursing Research*, vol. 35, no. 6, 1986.

[49] D. F. Polit, C. T. Beck, and S. V. Owen, "Is the CVI an acceptable indicator of content validity? Appraisal and recommendations," *Research in Nursing & Health*, vol. 30, no. 4, pp. 459–467, 2007.

[50] A. M. Farrell and J. M. Rudd, "Factor analysis and discriminant validity: a brief review of some practical issues," in *Australia-New Zealand Marketing Academy Conference*, Australia-New Zealand Marketing Academy, 2009.

[51] S. M. Fox-Wasylyshyn and M. M. El-Masri, "Handling missing data in self-report measures," *Research in Nursing & Health*, vol. 28, no. 6, pp. 488–495, 2005.

[52] Y. Akbulut, S. Şendağ, G. Birinci, K. Kılıçer, M. C. Şahin, and H. F. Odabaşı, "Exploring the types and reasons of Internet-triggered academic dishonesty among Turkish undergraduate students: development of Internet-triggered academic dishonesty scale (ITADS)," *Computers & Education*, vol. 51, no. 1, pp. 463–473, 2008.

[53] B. Norman, *Analyzing Quantitative Data: From Description to Explanation*, Sage, 2003.

[54] J. Pallant, *SPSS Survival Manual: A Step by Step Guide to Using SPSS for Windows (Version 12)*, Allen & Unwin, New South-Wales, Australia, 2005.

[55] N. Opitz, T. F. Langkau, N. H. Schmidt, and L. M. Kolbe, "Technology acceptance of cloud computing: empirical evidence from German IT departments," in *2012 45th Hawaii International Conference on System Sciences*, Maui, HI, USA, 2012.

[56] A. Balanskat, R. Blamire, and S. Kefala, *The ICT Impact Report: A Review of Studies of ICT Impact on Schools in Europe*, European Communities, 2006.

[57] D. L. Silvernail and D. M. M. Lane, *The Impact of MaineâĂŹs One-to-One Laptop Program on Middle School Teachers and Students*, Maine Education Policy Research Institute (MEPRI), University of Southern Maine, 2004.

[58] B. G. Tabachnick and L. S. Fidell, "Principal components and factor analysis," *Using Multivariate Statistics*, vol. 4, pp. 582–633, 2001.

[59] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Routledge, 2013.

[60] H. Van der Heijden, "Factors influencing the usage of websites: the case of a generic portal in The Netherlands," *Information & Management*, vol. 40, no. 6, pp. 541–549, 2003.

[61] K. McKinnon and A. Igonor, "Explaining eLearning perceptions using the technology acceptance model and the theory of planned behavior," in *Proceedings of E-Learn 2008-World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, Las Vegas, Nevada, USA, 2008.

[62] M. Kim, M. Park, and D. Jeong, "The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services," *Telecommunications Policy*, vol. 28, no. 2, pp. 145–159, 2004.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6627771, 16 pages
https://doi.org/10.1155/2021/6627771

Research Article

Proposing a Density-Based Clustering Approach (DBCA) to Aggregate Data Collected from the Environment in Arid Area for Desertification

Zhihao Peng,[1] **Raziyeh Daraei**,[2] **Seyed Mojtaba Ahmadpanahi**,[2] **Amir Seyed Danesh**,[3] **Safieh Siadat**,[4] **Poria Pirozmand**,[1] **and Rozita Jamili Oskouei**[2]

[1]*School of Computer and Software, Dalian Neusoft University of Information, Dalian, China*
[2]*Department of Computer Science and Information Technology, Mahdishahr Branch, Islamic Azad University, Mahdishahr, Iran*
[3]*Faculty of Technology and Engineering, East of Guilan, University of Guilan, Rudsar-Vajargah, Iran*
[4]*Department of Computer Engineering and Information Technology, Payame Noor University (PNU), Tehran, P.O. Box 19395-4697, Iran*

Correspondence should be addressed to Safieh Siadat; siadat@pnu.ac.ir

Received 30 November 2020; Revised 31 December 2020; Accepted 12 January 2021; Published 10 February 2021

Academic Editor: Suleman Khan

Nowadays, the expansion of desert areas has become one of the main problems in arid areas due to various reasons such as rising temperatures and vegetation fires. Establishment of wireless sensor networks in these areas can accelerate the process of environmental monitoring and integrate temperature and humidity information sending to base stations in order to make basic decisions on desertification. The main problem in this regard is the energy limitation of sensor nodes in wireless sensor networks, which is one of the main challenges in using these nodes due to the lack of a fixed power supply. Because the node consumes the most energy during data transmission, the node that transmits the most data or sends the packets over long distances runs out of energy faster than the others and the network work process is disrupted. Therefore, in this study, a density-based clustering approach is proposed to integrate data collected from the environment in arid areas for desertification. In the proposed method at each step, the node that has the most residual energy and is highly centralized will be selected to transfer information. The results of experiments for evaluating the performance of the proposed method show that the proposed method balances the energy consumption of the nodes and optimizes the lifespan of the nodes in the wireless sensor network installed in the arid area.

1. Introduction

In arid areas, due to the high temperature of the area and direct sunlight, the existing plant tissue is at risk of fire. When plant tissue is destroyed in desert areas and that area becomes a desert, the habitat in these areas will be severely damaged. On the other hand, human intervention and environmental manipulation in desert areas is also a potential threat to the plant tissue of the region. Human beings may cause fires in these areas due to negligence, in which case they will increase the area of deserts and waterless and grassy areas. Therefore, controlling such areas using wireless sensor networks to detect temperature changes at different times of the day can be a useful solution to prevent fires.

Wireless sensor networks (WSNs), as part of the Internet of Things (IoT) that can collect data from different areas, are widely used today in various aspects of human life [1]. The wireless sensor network consists of a set of sensor nodes that can sense environmental conditions such as temperature, pressure, humidity, and pollution. The sensor node has three main functions: it senses the data, processes the collected data, and transmits it to other nodes or a base station (BS)

[2]. One of the most common uses of WSNs is to monitor changes in the field of interest. The changes may belong to any physical variable such as heat, sound, and light. The sensor node detects these changes and processes them to become a consistent format of the data. Now, this data must reach the hole effectively [3]. Recent advances in wireless network technology with optimal energy consumption have created the technical conditions for the construction of multifunctional small sensor devices that can be used to sense and observe physical phenomena. The main purpose of environmental monitoring is not only to collect data from different places but also to provide the necessary information for scientists, planners, and policymakers to be able to optimally decide on management and improvement of the environment and provide useful information to end users [4].

The motivation of this paper is as follows:

(i) Wireless sensor networks suffer from limited resources, especially shortages of fixed power supplies. Therefore, sensor nodes can consume their limited energy resources when performing a process or transmitting information in a wireless environment. Therefore, it is essential to use different methods to save energy when performing processes and communications

(ii) Further, the lifespan of sensor nodes is highly dependent on battery life; therefore, clustering of wireless sensor nodes in order to transfer the sensed information to the nearest hole and the base station can be very effective in reducing energy consumption, which is a reason for the need for this research

The contributions of this paper are as follows:

(i) Proposing a novel clustering method for wireless sensor nodes in order to transfer the sensed information faster to the hole and the base station

(ii) The use of a proposed method for reducing energy consumption and increasing the lifespan of sensors in WSN, which are the main purposes of this research

(iii) We have used density-based clustering in environmental observer sensor networks in the direction of desertification in which the selection of the cluster head node in a comparative way to create a balance between distance and residual energy between nodes

This research is organized in five sections. In the second section, we explain some of the researches which are done by other researchers related to this research. In the third section, we explain the proposed method in this research. In the fourth section, we will explain in detail the simulation performed to evaluate the performance of the proposed method, and finally, the fifth section will include the conclusion.

2. Related Works

Environmental monitoring applications by wireless sensor networks can be broadly classified into internal and external monitoring. Consequences of internal oversight typically include oversight of buildings and commercial offices. These applications include measuring temperature, light, humidity, and air quality. Other important indoor applications may include habitat monitoring, road traffic monitoring, earthquake detection, volcanic eruption detection, flood detection, and weather forecasting. Sensor nodes are also widely used in arid and desert areas and can play a very important role in monitoring humidity and temperature in these areas [5].

Clustering in wireless sensor networks (WSNs) is an important step in communicating between sensor nodes. Many clustering techniques are presented with different features. Their main purpose is to facilitate conscious communication of energy between a large number of established nodes. One of the most important factors affecting the clustering process is the distribution of nodes. In many cases, the distribution of nodes is random. This type of distribution creates a network with different regions and with different densities. A different number of nodes in each subarea of the network means different communication loads and therefore different energy consumption.

Abdellatief et al. [6] proposed a distributed density-based clustering method called "spatial density-based clustering for WSNs" to reduce power consumption and ultimately increase grid life. This method was performed with the help of a simple initial spatial analysis for the distribution of nodes before the clustering process and divides the network into different regions according to their density level. Clusters are formed in each region separately, according to the amount of density measured.

Sinha et al. [7] developed a new forest fire prevention technique that detects the High Active Area (HA) (near the center of the fire) in the forest and records all sensed data through wireless communication that will be transferred to the base station as soon as possible so that the fire department can take the necessary measures as soon as possible to prevent the spread of fire. For this purpose, sensors have been deployed in the forest area to sense the various data needed to detect forest fires and divide that area into different clusters. A law-based semiregulatory classification model is proposed in this study to identify whether the cluster area is high active, medium active (MA), or low active (LA) in the forest. This model is trained in such a way that only one parameter of the data sensed by the sensor nodes is transferred to the primer (hole) of that area due to energy limitation, the primer can state (HA, MA, and LA) to predict. The region with 96% accuracy transmits all its sensor nodes in the HA cluster through the cap to the base station, and the greedy transfer technique is continuously used. This proposed method efficiently and quickly transfers the measured data from the HA area to the base station in the forest and saves the energy of all sensor nodes in the forest.

Grover et al. [8], in order to monitor the forest, divided its environment into square clusters, each of which has a sensor system. Node localization is done using satellite communications to reduce cover holes and ensure maximum range with the least delay. These nodes link the data to a monitoring station with a specific location, and an alarm is generated according to the sensed thresholds based on the new logic algorithm.

Moussa et al. [9] proposed a method for comparing fault-tolerant routing protocols to develop a method for forest fire detection that measures network longevity and network response time to an event, taking into account characteristics. Considers various issues in this regard. To test this method, the researchers simulated multilevel fault-tolerant routing protocols, disconnected heterogeneous routing protocol (HDMRP), and advanced ant-based QoS-based routing protocol in heterogeneous wireless sensor networks (EAQHSeN). Castalia WSNs were performed and tested in this research. The simulation results showed that HDMRP and EAQHSeN use twice as much and almost three times longer network response time than the multilevel protocol, respectively.

Ravikumar [10] proposed a method involving the exceptional component of remote control sensors for detecting forest fires. The sensor data is aggregated using the Arduino board and transmitted to a remote base station. In addition, an alert is generated and sent using the GSM module.

Kadir et al. [11] considered the development of the use of wireless sensor networks (WSNs) to identify the source of forest fires in Indonesia. WSN technology, which is used for the ground-based sensor system to collect environmental data, records any changes in reporting time to the data center for analysis. Data training is performed to identify fire points in the data center to determine the focus of the fire. The sensors will be installed in several places where the fire has already occurred and the next possible fire location is predicted. Mathematical analysis was used in this study to model the number of sensors required to be located in the forest area.

Aksamovic et al. [12] proposed a system for early detection of forest fires using a WSN simulator based on the proposed sensor model and the developed WSN model. The WSN emulator covers important design issues (including coverage of the monitored area in relation to the initial sensor deployment, the number of sensors required for targeted deployment, and the change of coverage as a function of time).

Darabkh et al. [13] proposed a density-based and energy-aware clustering and routing protocol (EA-DB-CRP) for data collection in wireless sensor networks. Its purpose is to distribute the load between the sensor nodes, which in turn balances the energy consumption in the network and thus prolongs the life of the network.

Yuan et al. [14] proposed a compact sensor-based cyclic routing data collection (CS-CARDG) to improve network life. The key technology adopted by the CS-CARDG scheme is that the data is collected by the cluster head node and the network first forms a cluster and each node in the cluster sends the data packet to the head of the cluster, and finally, each cluster forms M-dimensional data according to the requirements of intensive measurement technology to ensure data retrieval.

Abdullah et al. [15] proposed a common data collection algorithm (CDCA). The simulation results of the algorithm presented in this research showed that using this algorithm, the latency in a small number of mobile elements is significantly reduced. In addition, the performance of the CDCA algorithm is compared with the area division algorithm (ASA). As a result, CDCA showed superior performance in terms of network latency, load balance, and required number of mobile elements.

Edla et al. [16] proposed a clustering algorithm based on particle swarm optimization and a new fit function taking into account the average distance of clusters, gate load, and number of overhead gates in the network suggested.

Mahdi et al. [17] first briefly reviewed the various clustering methods, and then, they reviewed the node clustering methods that have been proposed by several researchers for tracking targets in wireless sensor networks and are based on data aggregation. They explained and briefly stated the advantages and disadvantages of each one of these methods. Finally, these researchers tabulated all the methods studied in this study in terms of their performance in tracking targets in wireless sensor networks and compared their advantages and disadvantages.

Khan et al. [18] focused on saving energy and extending the life of nodes in wireless sensor networks and proposed the Energy-Efficient Multistage Routing Protocol (EE-MRP) for energy saving in wireless sensor networks. This protocol has 3 parts including (a) a new multistage routing algorithm for data transmission in the network, (b) a new algorithm for efficient selection of cluster head, and (c) a new scheme for cluster formation. The simulation results of the proposed protocol in this research, using MATLAB software, showed that, compared to other researches, this protocol increases the performance of wireless sensor networks by reducing energy consumption, streamlining the selection of cluster head, and increasing the lifespan and operational capacity of the network.

Mahdi et al. [19] focused on tracking identified targets in wireless sensor networks and proposed a fully distributed algorithm named as Endocrine Inspired Sensor Activation Mechanism for Multi-Target Tracking (ESAM) in which node placement is self-organizing and energy efficient. This operation of this algorithm is limit to the number of activated nodes around the targets that are being tracked, so that they can reduce energy consumption and increase network life. The proposed algorithm reflects the features of the real-time activation system of sensors based on the general information flow in the endocrine system in the human body. The amount of hormones can directly affect the regulation of sleep and wakefulness of the nodes. In this algorithm, special rules were defined that can enable hormone levels to regulate the sleep-wake cycle of the nodes in a way that reduces the wake-up time of the nodes and ultimately reduce the energy consumption of these nodes. The simulation results of the proposed algorithm showed that this algorithm can control the performance of nodes without having to receive commands from the central controller, and in addition, compared to other algorithms for tracking an attacking object in wireless sensor networks, the results show that this algorithm is more efficient and stable.

Mahdi et al. [20] proposed a routing strategy aimed at maximizing overlap paths for efficient data aggregation and linking cost issues in clustered WSNs, known as the weighted data aggregation protocol (WDARS). The proposed method

was evaluated simultaneously to evaluate the energy consumption, network life, production capacity, and package delivery ratio, and its performance was compared with the InFRA and DRINA protocols (which are cluster-based routing protocols that aim only to maximize routes overlap for efficient data aggregation). The results of analysis and simulation showed that WDARS has better performance and more reliable than other methods and can further increase the lifespan of the network.

Ali et al. [21] categorized all research conducted by other researchers based on important parameters such as their objectives, applications, communication technology, types of data sets used, discovery, and types of data. In addition, several case studies were examined to demonstrate the role of sensor clouds in providing high computational capabilities. In addition, they outlined some of the challenges of collecting data in sensor clouds.

Wang et al. [22] proposed the advanced Power Efficient Gathering in Sensor Information Systems (EPEGASIS) algorithm to reduce the hot spot problem. This algorithm has two stages; the operations performed in the first stage are determining the optimal communication distance to reduce energy consumption, setting the threshold to protect the dying nodes, and using mobile sink technology to balance energy consumption between nodes, and the operation performed in the second step is setting the communication range of nodes to the sink node. The results obtained from the simulation of the proposed algorithm of these researchers showed that the proposed algorithm has a better performance compared to other methods and algorithms in terms of increasing life, reducing power consumption, and reducing latency in the network.

Wang et al. [23] proposed a new algorithm to optimize the performance of the PSO algorithm. In the proposed algorithm, the nodes are first randomly placed in fixed geographical areas. Then, a network is created between these nodes. The whole network is then divided into several subnetworks, and the amount of coverage and energy consumption for each subnetwork is calculated. The results obtained from the simulation of the proposed algorithm show that this algorithm is able to create a balance between the amount of network coverage and energy consumption in that network.

Ge et al. [24] on the similarities and differences between the big data technologies used in different areas of the Internet of Things (such as smart cities, industry, smart homes, smart transportation, and healthcare) discussed and proposed a platform for mutual understanding of these differences and similarities and mentioned that some of these technologies can be used in several IoT domains. Finally, these researchers proposed a conceptual framework to guide other researchers in selecting Big Data technologies in various areas of the Internet of Things.

Wang et al. [25] studied an empower Hamilton loop-based data collection algorithm with a mobile agent for WSNs, then proposed the combination of the PEGASIS algorithm and the Hamilton loop algorithm. Through a combination of single-hop and multihop mechanisms and a moving agent added a mobile agent node (MA), this authors proposed a design of an optimal empower Hamilton loop using a local optimization algorithm. The simulation results of the proposed algorithm showed that this algorithm can increase the network lifespan and reduce the propagation delay in the network and balance the amount of resource consumption compared to the application of each algorithm alone (PEGASIS and Hamilton loop algorithm).

Table 1 summarizes the research-related articles reviewed in this study.

3. The Proposed Method

This study presents a density-based clustering approach for collecting critical data in wireless sensor networks. In wireless sensor networks, resource constraint is one of the main challenges facing the network. In most cases, the installation of nodes in the network is done randomly. These nodes are spread in different areas by helicopter, and access to these nodes is not easily possible. For this reason, the limited energy in wireless sensor nodes has become a major issue for the longevity of wireless sensor networks. Wireless sensor networks are evaluated according to the balance of energy consumption and network life. Therefore, researchers are trying to optimize energy consumption in wireless sensor networks and a lot of work has been done in this field. Wireless sensor nodes use the following three modes of power when networked:

(i) Receive data and information from the environment

(ii) Data processing

(iii) Sending information

Optimization of energy consumption in wireless sensor networks depends on providing optimal methods to reduce energy consumption in the above situations. At the stage of receiving data and information from the environment, each of the sensor nodes embedded in the network environment has a constant energy consumption, and reducing energy consumption at this stage requires the design of special-purpose sensor nodes, which may incur design costs. There are more nodes than normal sensor nodes, and in this case, using wireless sensor networks may not be cost-effective. Optimization of energy consumption in the data processing stage depends on the policy of sending data to the hole node. Wireless sensor nodes send the collected data to the node based on periodic and event-based policies. In the periodic data transmission policy, wireless sensor nodes periodically send the collected data to the hole node at regular intervals. In this policy, the amount of energy consumption is higher, because the collected data, in addition to the energy required to process the data, also consume the energy required to send the data in each intermittent period of time, if it may not have changed at all in the environmental parameters and the energy consumed to send the data is actually wasted. In this regard, an event-based policy has been proposed to reduce energy consumption. In this policy, when something special happens in the environment under the supervision of the wireless sensor network, the wireless sensor nodes send information. Wireless sensor nodes, like periodic policies,

Table 1: Summarization of related works reviewed in this research.

Related	Author name	Year	Brief explanation
Proposed algorithm based on PSO	Edla et al. [16]	2019	Proposed a clustering algorithm based on particle swarm optimization and a new fit function taking into account the average distance of clusters, gate load, and number of overhead gates in the network suggested.
	Wang et al. [23]	2018	Proposed a new algorithm to optimize the performance of the PSO algorithm. In the proposed algorithm, the nodes are first randomly placed in fixed geographical areas. Then, a network is created between these nodes. The whole network is then divided into several subnetworks, and the amount of coverage and energy consumption for each subnetwork is calculated.
Survey	Mahdi et al. [17]	2016	First briefly reviewed the various clustering methods and then they reviewed the node clustering methods that have been proposed by several researchers for tracking targets in wireless sensor networks and are based on data aggregation. They explained and briefly stated the advantages and disadvantages of each one of these methods.
	Ali et al. [21]	2018	Categorized all research conducted by other researchers based on important parameters such as their objectives, applications, communication technology, types of data sets used, discovery, and types of data. In addition, several case studies were examined to demonstrate the role of sensor clouds in providing high computational capabilities. In addition, they outlined some of the challenges of collecting data in sensor clouds.
	Ge et al. [24]	2018	Studied on the similarities and differences between the big data technologies used in different areas of the Internet of Things discussed and proposed a platform for mutual understanding of these differences and similarities and mentioned that some of these technologies can be used in several IoT domains. Finally, these researchers proposed a conceptual framework to guide other researchers in selecting Big Data technologies in various areas of the Internet of Things.
Proposed algorithm based on PEGASIS algorithm	Wang et al. [22]	2018	Proposed the advanced Power Efficient Gathering in Sensor Information Systems (EPEGASIS) algorithm to reduce the hot spot problem.
	Wang et al. [25]	2019	Proposed the combination of the PEGASIS algorithm and the Hamilton loop algorithm. Through a combination of single-hop and multihop mechanisms and a moving agent added a mobile agent node (MA), these authors proposed a design of an optimal empower Hamilton loop using a local optimization algorithm.
Reduce energy consumption	Mahdi et al. [19]	2016	Proposed a fully distributed algorithm named as Endocrine Inspired Sensor Activation Mechanism for Multi-Target Tracking (ESAM) in which node placement is self-organizing and energy efficient. The main purpose of the ESAM is to reduce energy consumption and increase network lifespan.
	Mahdi et al. [20]	2016	Proposed a routing strategy aimed at maximizing overlap paths for efficient data aggregation and linking cost issues in clustered WSNs, known as the weighted data aggregation protocol (WDARS). The main purpose of WDARS is to reduce energy consumption, increase network lifespan, increase production capacity, and package delivery ratio.

collect data from the environment at regular intervals and process the collected data. When the processing result is true under a certain condition, the wireless sensor nodes send the data. In this case, only when a special event occurs in the network and the network-controlled environment, the necessary energy is consumed to send data.

In the proposed method, a wireless sensor network has been used for desertification. In desert and arid regions, due to the high temperature of the region and direct sunlight, the existing plant tissue is at risk of fire. When plant tissue is destroyed in arid areas that area becomes a desert and a lot of damage is done to the habitat in these areas. On the

other hand, human intervention and environmental manipulation in desert areas is also a potential threat to the plant tissue of the region. By setting fires in these areas, humans may cause fires in these areas due to negligence, in which case it will lead to an increase in the area of deserts and waterless and grassy areas. Therefore, controlling such areas using wireless sensor networks to send temperature data at different times of the day can be a useful solution. Wireless sensor networks aggregate area temperature information and send it to the base station. In order to desertify the proposed method, sensor nodes equipped with a temperature sensor have been used. In desert areas, when the temperature exceeds a certain level, it may damage the vegetation of the area and cause the loss of vegetation and the expansion of desert areas. Therefore, in the proposed method to collect information from desert areas, the temperature of the area is collected at regular intervals. In the proposed method, in order to optimize energy consumption in the process of information processing and data aggregation and sending to the hole, the event-based data aggregation policy is used. In this method, the ambient temperature is compared with the temperature threshold specified in the central station and if the temperature obtained is higher than the average, it is recognized as critical data and this critical data is sent to the hole node. Critical data in wireless sensor networks are aggregated in the proposed method when plant tissue in these areas is compromised. The danger in the proposed method is felt when the temperature of the area rises to such an extent that it exceeds the tolerable threshold for plant tissue in the area and there is a possibility of loss of plant tissue. In such circumstances, it is necessary to send a warning to the Central Desert Monitoring Station to take the necessary measures to prevent damage to plant tissue. Therefore, in the proposed method, considering that the temperature of desert areas increases only in limited hours of the day, critical data detection and transfer of data and information is done only when necessary, and energy consumption will be its lowest value. Therefore, the proposed method considers energy consumption optimization in the data processing stage according to the information processing model and critical data transmission.

In the data transmission stage, the energy consumption of the nodes is determined based on the distance between the source node and the destination node. The greater the distance between the nodes, the more force is required to send the nodes and the energy consumption will naturally increase. Therefore, methods based on sending data packets and information directly to the hole can have the highest energy consumption in routing wireless sensor networks. For this purpose, in order to optimize energy consumption at this stage, multistep methods are used instead of single-step sending. In multistep transmission methods, the node with the most residual energy is selected as the next node to transfer data at each stage. One of the most common methods introduced to optimize energy consumption is clustering. In clustering methods, the nodes that have the most residual energy are selected as cluster nodes and the task of aggregating data from the wireless sensor nodes of the cluster member and sending it to the hole are in charge. In these methods, the energy in the wireless sensor nodes is balanced when the data is transferred to the hole and the lifespan of the wireless sensor network is increased. In the proposed method, the density-based clustering method is used. Due to the sensitivity of clustering methods to the initial central points, clustering methods based on density have been developed to create clusters without the need for a cluster center and with the desired shape. These methods usually identify clusters as dense areas of nodes in the network. The density of nodes in the network reduces the intercluster distances as much as possible and optimizes the energy consumption in the wireless sensor network. One of the most popular density-based clustering algorithms is the DBSCAN algorithm, which expands clusters according to density-based connection analysis. In wireless sensor networks, the distances of the nodes are different due to the fact that the nodes are randomly distributed in the network. Thus, the clustering method places the density of nodes that are close to each other in a cluster. In the following, we present the basic concepts of density-based clustering.

Density-based application spatial clustering (DBSCAN) is one of the most popular density-based clustering algorithms. This algorithm covers areas of the network with sufficient density in the clusters and forms irregularly shaped clusters in the network space with nodes far from the rest of the nodes as noise. This algorithm defines each cluster as a set of maximum connected points with high density [26]. The main idea of density-based clustering is based on neighborhood definitions and connected nodes in the neighborhood.

3.1. Proposed Routing in This Research. In the proposed method, in order to aggregate data from desert areas, a wireless sensor network is used, in which we use the density-based clustering method to send data for routing. In the proposed method, the proposed sensor network is first simulated based on the random distribution of nodes in the desert region. According to the definitions of the density-based clustering algorithm in wireless sensor network, the proposed method clusters wireless nodes in the network. In the proposed method for density-based clustering, several random nodes are selected. The neighborhood radius and the threshold are initially randomly selected, but by changing these parameters based on trial and error, we will reach the optimal value for the neighborhood radius and the threshold number of wireless sensor nodes in the neighborhood. Then, according to the neighborhood radius and the threshold of the minimum nodes in the neighborhood of ε-node, the neighboring nodes of each of the random nodes are searched and the nodes in the neighborhood radius of each are identified. If the number of these nodes exceeds the minimum threshold of the number of nodes in the neighborhood, the initial random node is selected as the main member of the cluster and the same process is repeated for all neighboring nodes of the main node. Thus, clusters of sensor nodes are formed in the network. After density-based clustering in the network to transmit sensed data, the wireless sensor nodes after processing the data, if data is detected as critical data, sends the hole or node that has the most value based on the selection function in the cluster. The amount of

competence in the proposed method is determined based on a combination of three parameters (including the distance of the nodes from each other, the direct distance of the node to the hole, and the remaining energy of the nodes). The node that has the shortest distance to its neighbors in the neighborhood radius and its direct distance to the hole is less than the rest and also has more residual energy will have the highest amount of merit. The node with the most competence in the cluster is selected as the cluster head node in each step so that the energy consumption of the nodes in the cluster is balanced. If the cluster head nodes collide with another cluster head node on their way to the hole, they send the data to that node to take advantage of the multistep transfer. Otherwise, if the head node is in the closest cluster to the hole, it sends the data directly to the hole.

In wireless sensor networks, after determining the density-based clusters, some nodes may not be in the neighborhood of any other node and also do not have the appropriate number of nodes in their neighborhood, so they do not belong to a cluster and remain as individual nodes in the network. Such nodes send their data to the hole in a single step if they have a critical data to send to the hole, if these nodes are close to the hole; otherwise, they send their data to the nearest node in the nearest cluster to route between the cluster member node and the cluster head node. By sending data to the head node of a cluster, the routing process from the head node to the hole node is done in a single step or multiple steps.

3.2. Select the Cluster Head Node. In the proposed method, after clustering the wireless sensor nodes based on the DBSCAN algorithm, in the routing and data transfer stage, nodes that contain critical data send cluster head data into nodes. In order to select the cluster head node in the proposed method, three factors (including distance within the cluster, distance to the hole, and residual energy) are examined. The node with the most energy remaining in the cluster, the shortest distance to the other nodes in the cluster, and the shortest distance to the hole is considered the best option for selection as the cluster head. The main difference between the proposed method and the DBSCAN clustering algorithm is that at each step, the cluster head node is evaluated and the node that is more qualified is selected as the cluster head node. Therefore, the cluster head nodes are constantly updated and changed, in which three factors (distance within the cluster, distance to the hole, and residual energy) are examined at each stage. Given the proposed wireless sensor network, it is assumed that the nodes are fixed. Therefore, the main factor determining the cluster head node will be the amount of energy remaining, and if the nodes are considered mobile, the values of all three factors are constantly changing and changes in the selection of cluster head nodes at each stage can be very significant. In the following, we will model the factors of the cluster head node selection function.

3.2.1. Intracluster Distance D_{nc}. In the proposed method, cluster member nodes must send information to the cluster head node. When the cluster member nodes in each cluster surround the cluster head, it means that the distance between the cluster member nodes and the cluster head node is half, and the distance between the data packets and data transmission is the shortest. In fact, when the cluster head node is almost in the middle of the other nodes in the cluster, it will be approximately the same distance from all nodes or will have the minimum distance from all nodes. Thus, the transmission of data in the shortest distance requires the least amount of energy. The intracluster distance model is expressed in Equation (1) [27]:

$$D_{\mathrm{nc}} = \min \left(\sum_{m=1}^{M} \left(\sum_{n=1}^{N} d_{\mathrm{ncluster}} \right) \right). \tag{1}$$

In Equation (1), M represents the number of clusters, N represents the number of members of each cluster, and d_{ncluster} represents the Euclidean distance of cluster head nodes to cluster member nodes.

3.2.2. Distance of Cluster Head Node to D_{cs} Hole Node. In clustering-based protocols, the cluster head node combines data and information on the received data and information and sends them to the hole node. In fact, in a cluster head node, a data processing step is performed to eliminate incomplete and duplicate information. The remaining information is then sent to the hole node as useful information. In the initial protocols, nodes are sent from the cluster head to the hole in a single step, and information is sent directly from each cluster head node to the hole. Such a strategy wastes too much energy on a cluster head. Therefore, using multistep data transfer is a solution to solve this problem, which is used in the proposed method. As shown in Figure 1, cluster head nodes send data and information to the destination through other cluster head nodes to transfer data.

As shown in Figure 1, cluster head nodes use a multistep approach to transmit data and information. In this case, the shorter the distance between the cluster head node and the hole node is, the shorter the path, and as a result, energy consumption is reduced. The distance model between the cluster head node and the cavity node is expressed in

$$D_{\mathrm{cs}} = \min \sum_{m=1}^{M} d_{\mathrm{csink}}, \tag{2}$$

where d_{csink} shows the Euclidean distance from the cluster head node to the hole node.

3.2.3. Energy Consumption. The energy consumption model in WSN is divided into two general parts, which we will review and model in the following.

(1) Total Energy Consumption of the Energy1 Network. The total energy consumption of the network in the clustering stage is as follows.

In the first step, the cluster head node plays a message informing the others that it is the cluster head node. The cluster node table is also updated and distributed among the cluster member nodes. This table is sent to the nodes of the

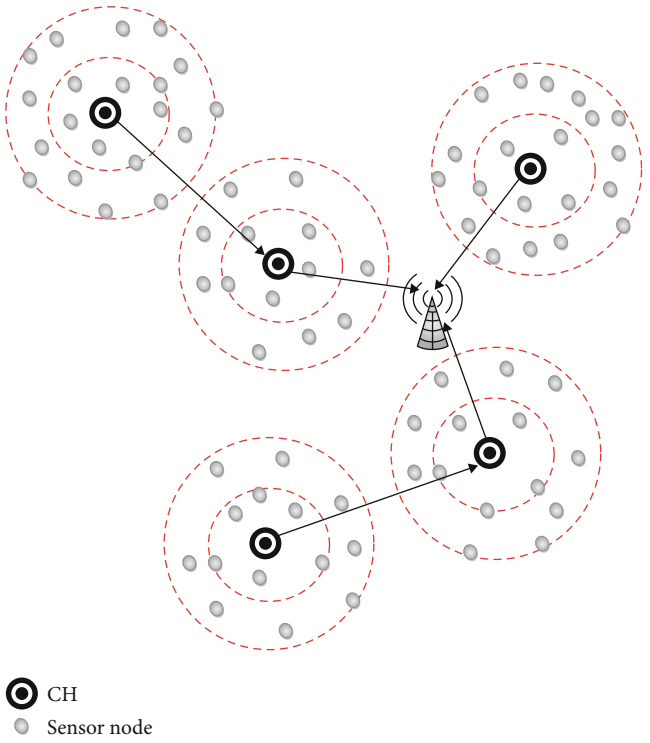


Figure 1: Multistage transfer from cluster head to hole node [19].

cluster, and the value of the transmitted data in this case is equal to t bits. The energy consumption of the cluster head node when it sends information is calculated from Equation (3) [27].

$$E_{\mathrm{cn}}(t.d_{\mathrm{cn}}) = \begin{cases} t\left(E_{\mathrm{elec}} + \tau_f d_{\mathrm{cn}}^2\right).d_{\mathrm{cn}} < d_c, \\ t\left(E_{\mathrm{elec}} + \tau_m d_{\mathrm{cn}}^4\right).d_{\mathrm{cn}} \geq d_c. \end{cases} \tag{3}$$

In Equation (3), E_{elec} represents the energy consumption by the cluster head node for the transmission of 1 bit of data, τ_f and τ_m represent the signal amplifier energy consumption when transmitting 1 bit of data at each unit distance in open space, respectively. Multiple fade models. d_{cn} shows the Euclidean distance of the current cluster members to the cluster head node. Threshold $d_0 = \sqrt{\tau_f/\tau_m}$ is used for conversion between communication channel models. The cluster member node then receives the t-bit information and the cluster table from the cluster head node and, according to the same table, sends the t-bit data to the cluster head node to verify the cluster head node. In this process, the energy consumption of cluster member nodes is calculated using Equation (4) [27]:

$$E_{\mathrm{non-cn}}(t.d_{\mathrm{cn}}) = \begin{cases} t\left(E_{\mathrm{elec}} + \tau_f d_{\mathrm{cn}}^2\right) + t \times E_{\mathrm{elec}}.d_{\mathrm{cn}} < d_c, \\ t\left(E_{\mathrm{elec}} + \tau_m d_{\mathrm{cn}}^4\right) + t \times E_{\mathrm{elec}}.d_{\mathrm{cn}} \geq d_c. \end{cases} \tag{4}$$

Finally, the process energy consumption for the cluster head node to accept the cluster member nodes to send the packet to them is calculated through Equation (5) [27]:

$$E_{\mathrm{cn}}(t.d_{\mathrm{cn}}) = tE_{\mathrm{elec}} \times \left(\frac{N}{M} - 1\right). \tag{5}$$

In summary, the total energy consumption of the Energy_1 grid in the clustering phase is summarized by Equation (6) [27].

$$\mathrm{Energy}_1 = \begin{cases} \min\left(tE_{\mathrm{elec}} \times \left(\frac{N+2}{M} - 1\right) + \tau_f d_{\mathrm{cn}}^2 \times \left(\frac{N}{M} - 1\right)\right).d_{\mathrm{cn}} < d_c, \\ \min\left(tE_{\mathrm{elec}} \times \left(\frac{N+2}{M} - 1\right) + \tau_m d_{\mathrm{cn}}^4 \times \left(\frac{N}{M} - 1\right)\right).d_{\mathrm{cn}} \geq d_c. \end{cases} \tag{6}$$

(2) Energy Consumption Balance of Energy2 Network. The network energy balance of the network has two parts, D_{no} and D_{en}. The frequency of the number of node members in each D_{no} cluster is as follows.

The smaller the cluster size, the higher the average number of nodes per cluster, meaning that the load on each cluster head is more balanced [27].

$$D_{\mathrm{no}} = \frac{\sum_{i=1}^{m}(v_i - u)^2}{m}, \tag{7}$$

where v_i is the number of node members in the i cluster and u is the average number of nodes in each cluster in the network.

Energy consumption per cluster is D_{en}. The smaller size of each cluster, then the average energy consumption in the clusters is calculated from

$$D_{\mathrm{en}} = \frac{\sum_{i=1}^{m}(E_i - u_e)^2}{m}, \tag{8}$$

where E_i is the total energy consumption in clusters i and u_e is the average energy consumption of each cluster.

In summary, the network energy balance is calculated from Equation (9) [27]:

$$\mathrm{Energy}_2 = \min\left(D_{\mathrm{no}} + D_{\mathrm{en}}\right). \tag{9}$$

Thus, in order to optimize energy consumption in WSN, it is possible to select a suitable cluster head node, based on the mentioned factors, which can improve energy consumption and wireless sensor network life. Figure 2 shows the proposed method flowchart.

4. Simulation and Evaluation

In order to implement the proposed method based on multistep routing and density-based clustering, we first simulate the environmental monitoring network in desert areas. The proposed network is configured in a monitored desert environment in a space of 100×100 that the number of nodes distributed in this environment is equal to 100 sensor nodes

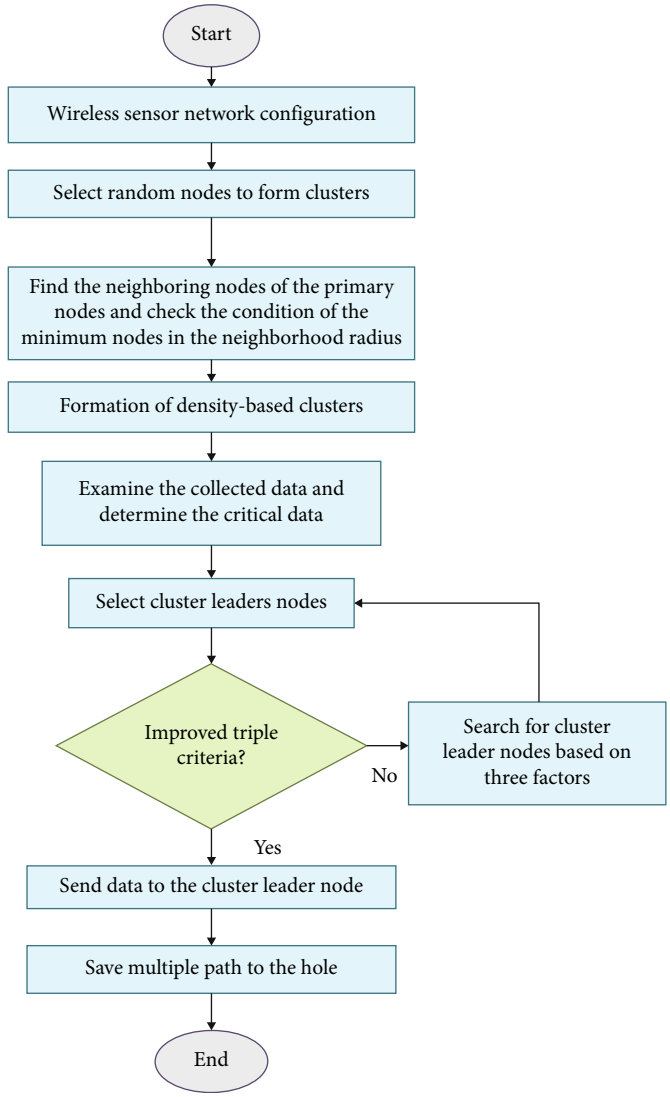


Figure 2: Flowchart of the proposed method.

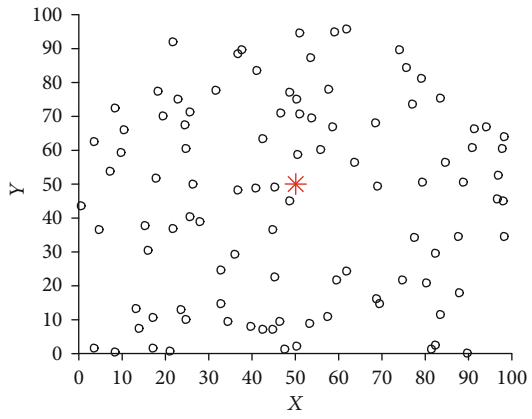


Figure 3: Configuration of the desert environment sensor network.

that these nodes are randomly in the monitored environment are scattered. Due to the scalability of the proposed method and review of improvements made by this method, the number of sensors can be changed in different scenarios. Recommended wireless sensor network settings include antenna and transmission channel settings, bandwidth, power consumption, packet length, queue length for sensor nodes, and other network configuration parameters according to the standards listed in various studies conducted by numerous researchers have been considered. In order to implement the proposed method, MATLAB software version 2015 has been used. The proposed method is theoretically done in MATLAB software with the aim of using it in desert areas in order to aggregate critical data, which can be generalized for use in other environments. Figure 3 shows the initial configuration of the proposed sensor network.

As shown in Figure 3, a hole node is embedded in the middle of the desert area with coordinates $(50 * 50)$ that can be accessed by all wireless sensor nodes and the distance between the wireless sensor nodes is from a balanced cavity. Wireless sensor nodes in the monitored area surround the hole node to collect sensed information in the desert area, including the temperature of the area, and send it to the hole.

Wireless sensor nodes are clustered around the hole to classify nodes close together, and a cluster head node is specified to communicate with the hole. Cluster head nodes usually lose their energy with each transfer of information, so the cluster is updated at each step of the data transfer, and the node provides the best conditions for communication with other cluster head nodes. They exchange information with the hole, selected as the cluster head node. In this way, the energy consumption of the nodes in each cluster is balanced and according to the clustering of all nodes in the network, the energy consumption in the whole network will be balanced.

In the proposed method, density-based clustering is used. Density-based clustering is done according to the presence of nodes around the cavity node. There are two basic conditions for selecting a node as a member of a cluster (including neighborhood radius and number of wireless sensor nodes in the vicinity of a wireless sensor node). The value of the parameters related to the neighborhood radius and the threshold of the number of sensor nodes in the neighborhood radius vary according to different applications and the area covered. Thus, in the proposed method, in order to obtain the value of these parameters, several tests have been performed to determine the neighborhood radius of each node in the desert area under monitoring and the number of neighboring sensor nodes.

Based on the test results, the best value of the neighborhood radius parameter is equal to 10 meters and the best value for the threshold number of sensor nodes in the neighborhood is equal to 3 neighboring sensor nodes. Figure 4 shows the density-based clustering for wireless sensor nodes in the desert region.

As shown in Figure 4, the wireless sensor nodes in the desert area are clustered based on density. In Figure 4, you can see the wireless sensor nodes are divided into 9 separate clusters. One of the most important advantages of the density-based clustering method is that there is no need to select the primary central node. In partial clustering methods, the basis of clustering is based on the selection of the initial central node and the distance of the spheres is calculated based on these central points. In these methods, each wireless sensor node is assigned to the nearest center of the cluster and is placed in the cluster that most likely belongs to that cluster. In split methods, incorrect selection of the primary node can lead to incorrect clustering of wireless sensor nodes. For this purpose, to eliminate the disadvantage of partition clustering methods, in the proposed method, the density-based clustering method is used, which is more efficient in dividing nodes and clustering nodes in the wireless sensor network of the region, which is under supervision. In Figure 4, the sensor nodes are clustered based on their neighborhood radius and the number of wireless sensor nodes located in the neighborhood radius, and each node is assigned to one of the clusters based on the density connection chain. Some nodes do not belong to any of the clusters. To transfer data and information from these nodes, if they are close to the hole, send data directly to the hole or send data and information to the nearest node from the nearest cluster. In the proposed method, first, the cluster head nodes

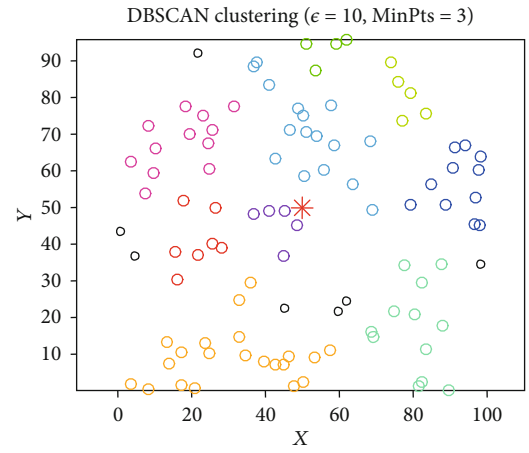


Figure 4: The density-based clustering for wireless sensor nodes in the desert region.

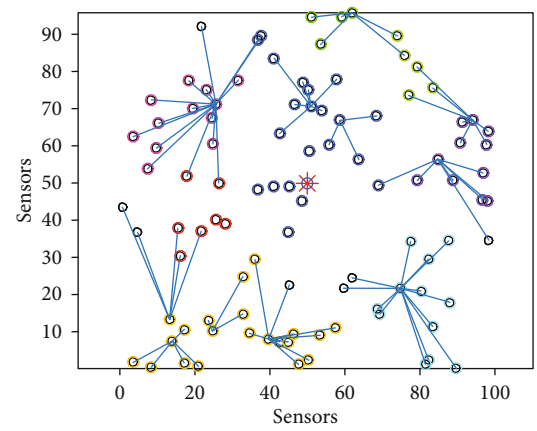


Figure 5: Sending information from cluster member nodes to cluster head nodes.

are selected according to the value of their proportionality function. Cluster head nodes are the nodes that have the most residual energy, the shortest distance to other nodes in the cluster, and the shortest distance to the hole. Given that the energy of all nodes is equal in the first step of transmitting information in the network and identifying the cluster head nodes and the energy of any node is not currently consumed, in this step to determine the cluster head nodes from two factors (including distance to cluster member nodes and distance to hole) were used. Cluster head nodes first update the cluster head node information in the routing table and send a message in the form of a routing message to all nodes in the cluster, in addition to declaring themselves heads, receive information about wireless sensor nodes. Accordingly, the wireless sensor nodes send a routing response message to the cluster head node in order to recognize and validate the cluster head node. After receiving the routing response messages, the cluster head nodes also update the routing table and determine the cluster member nodes that are responsible for receiving sensed information from the desert area covered by these nodes and send this information to the node of the hole. Figure 5 shows the sending of

Table 2: Information about cluster head nodes.

90	80	76	70	63	47	34	25	4	Index of cluster head nodes
0.49995	0.49995	0.49995	0.49995	0.49995	0.49995	0.49995	0.49995	0.49996	Residual energy
40.38	41.49	31.18	41.87	33.65	41.14	46.67	42.52	54.65	Distance to the hole

information from cluster member nodes to verify cluster head nodes.

As shown in Figure 5, cluster member nodes send their information in the form of routing response messages to cluster head nodes. Cluster head nodes are selected according to the residual energy factors, distance to cluster member nodes, and distance to hole node. Cluster member nodes also send their packets to the node when they have critical information. In some clusters, the distance from some cluster member nodes to the cluster head node may be greater than the distance from the cluster member node to the adjacent cluster head node. In this case, the cluster member nodes send their packets to the adjacent cluster head node. The aim of the proposed method is to balance energy consumption in wireless sensor nodes, and for this purpose, cluster head nodes that have a better position in terms of energy and distance within the cluster and distance to the hole are less than other nodes in the cluster, which are used to send data to the hole. Table 2 shows the selected cluster head nodes, the amount of energy remaining, and the distance between them and the hole node.

As shown in Table 2, the cluster head nodes are selected according to the residual energy and the distance to the hole node. These nodes are responsible for communicating with the hole and receive the sensed information from the wireless nodes, aggregate it, and send the processed information to the hole in a single step or multiple steps. Single-step sending, if the cluster of the critical data sending node is the closest cluster to the hole node. Otherwise, the data is sent in several steps between the cluster head nodes in order to send to the hole. Finally, the path created in the routing table is updated and declared as the current path. The next step in sending information is when one of the wireless sensor nodes installed in the desert area has critical data. In this case, the wireless sensor node needs to send information to the cluster head node. At this stage, the previous cluster head node is evaluated based on the residual energy factor. If the nodes in the network are more suitable in terms of residual energy factors, distance to other nodes of the cluster, and distance to the node of the hole compared to the current head node, the clusters are updated in order to balance energy consumption. New cluster heads are selected. Otherwise, the old head node remains in place and continues to send information to the hole. Figure 6 shows the steps for updating clusters and transmitting information over a wireless sensor network every 200 repetitive steps until the nodes are energized.

As shown in Figure 6, at each step of receiving the critical message and sending information from the wireless sensor nodes to the cluster head nodes, the cluster head nodes are updated and the node with the best merit from the factors mentioned are selected as the cluster head node.

Further, in Figure 6, we have the following:

(a) The nodes of the cluster, the cluster head node, and the hole node are repeated in the 201st step, as we see in this section

The position of the nodes inside all the clusters and the cluster heads related to each of the clusters in every 200 repetitions are shown in Figures 6(b) to 6(g).

(b) Cluster member nodes, cluster head node, and hole node are repeated in the 401st step

(c) Cluster member nodes, cluster head node, and hole node are repeated in the 601st step

(d) Cluster member nodes, cluster head node, and hole node are repeated in the 801st step

(e) Cluster member nodes, cluster head node, and hole node are repeated in the 1001st step

(f) Cluster member nodes, cluster head node, and hole node are repeated in the 1201st step

(g) Cluster member nodes, cluster head node, and hole node are repeated in the 1401st step

In all of these forms, we see that in every 200 repetitions, the amount of energy of the nodes decreases, so the cluster head is updated and replaced by another node that has more energy and less distance than the hole and other nodes. In all of these cases, the number of dead nodes is zero, that is, until 1401 repetitions, no node died, or in other words, the energy of no node was zero until 1401 repetitions.

The transmission of information in wireless sensor networks continues until most of the sensor nodes within the network lose their energy and a coverage gap in the network occurs. When a node in the network loses its energy and is not able to collect data from the monitored environment, according to the use of density-based clustering method in the proposed method, the nodes connected to the desired node, they are responsible for collecting data from the area covered by the dead node. This will continue until the cover hole phenomenon occurs in the network, but in the event of a hole in the network cover, the wireless sensor network will not be able to collect data from the desert area. Thus, the network performance is disrupted and the so-called network life ends.

Figure 7 shows the proposed wireless sensor network after the end of the network life. As shown in Figure 7, the proposed wireless sensor network is unable to collect data from the monitored environment after the death of a large number of sensor nodes embedded in the network. In this

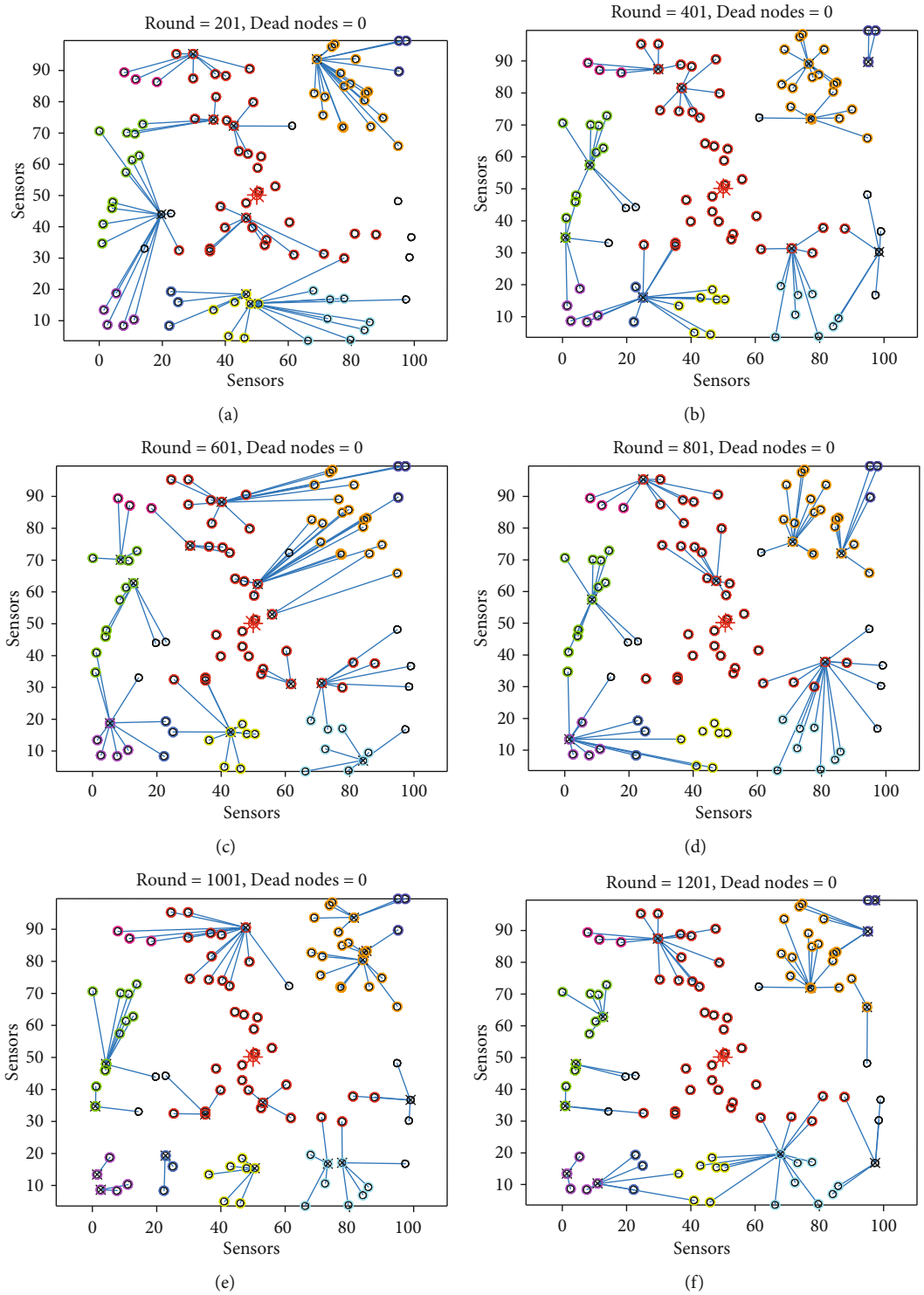
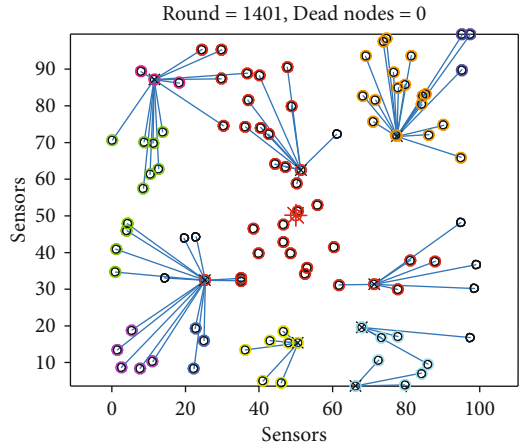


Figure 6: Continued.



(g)

Figure 6: Steps to update clusters and cluster head nodes.

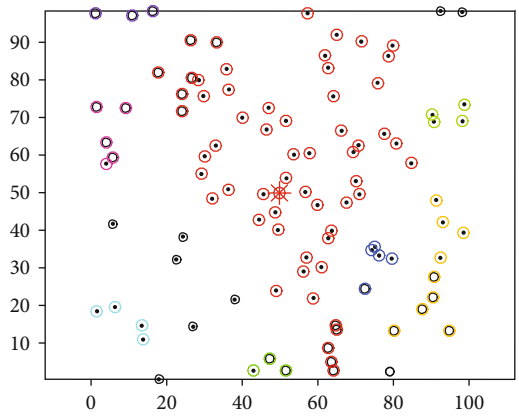


Figure 7: Wireless sensor network in desert areas after the end of network life.

case, the life of the network ends and the transfer of information to the network temperature monitoring stations is not possible. According to Figure 7, the dead nodes can be seen as small black dots, which have lost the ability to transmit information. Thus, with the death of sensor nodes in the network, the life of the network ends.

4.1. Evaluation of the Proposed Method. After implementing the proposed method to evaluate the amount of energy efficiency and other network parameters, we evaluate the proposed method. The evaluation of the proposed method includes the study of the behavior of nodes in the network and the amount of energy consumption by them and the effect of the density-based clustering method on the performance of nodes in the network. Due to the increasing use of wireless sensor networks for environmental monitoring, many parameters have been introduced in wireless sensor networks that need to be improved and optimized. In the proposed method, according to the research objectives, which include reducing the energy consumption of nodes, increasing the lifespan, and increasing the throughput and data delivery rate in the wireless sensor network, the performance of the proposed wireless sensor networks in terms of these three, the invoice is checked.

According to the clustering methods in wireless sensor networks, which is based on the transmission of information at close distances, the transmission of information at short distances in the wireless sensor network, in addition to creating a balance in energy consumption in network and increase network life, can reduce end-to-end latency between wireless sensor nodes and the number of packets lost in the network. Reducing the number of lost packets in the network increases the throughput rate and data delivery. Therefore, first, the amount of energy consumed in the wireless sensor nodes in the network is evaluated in the proposed method. The amount of energy consumed for each wireless sensor node in the proposed wireless sensor network is calculated and shown in the graph. In this diagram, the energy consumption increases linearly, which indicates the balanced energy consumption of the wireless sensor nodes in the proposed method. Due to the fact that the energy of all nodes runs out of time in close proximity to each other and there is no node in the diagram whose energy is exhausted earlier than the rest and disrupts the performance of the network, therefore, the average energy consumption in the network increases in a balanced way. Therefore, the average energy consumption of nodes in the proposed wireless network increases in a balanced way due to the use of the density-based clustering method. The density-based clustering method divides the wireless sensor nodes based on the density in the wireless sensor networks. Dense clusters of nodes in different areas of the wireless sensor network select nearby nodes and transmit data at the closest distance. Thus, energy consumption in wireless sensor networks is balanced and the average energy consumption in the network per step of data transmission in the network is 2.548 joules.

The second goal of the proposed method is to increase the lifetime of the network. Network longevity is one of the basic criteria in the network that most routing methods try to improve this criterion. Network lifetime is considered as the time when wireless sensor nodes are available and collect desert temperature data and send it to the hole. In a wireless

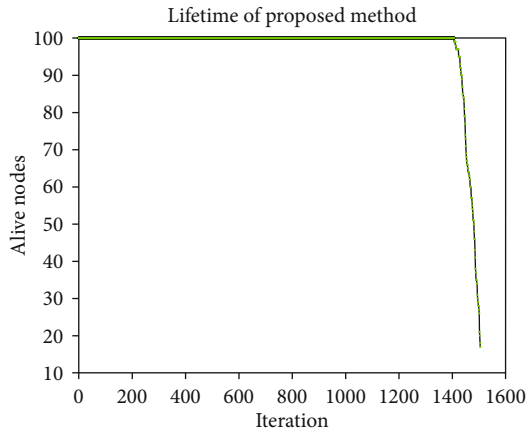


Figure 8: Graph of network lifetime in the proposed method.

Figure 9: Delivery rate in the proposed method.

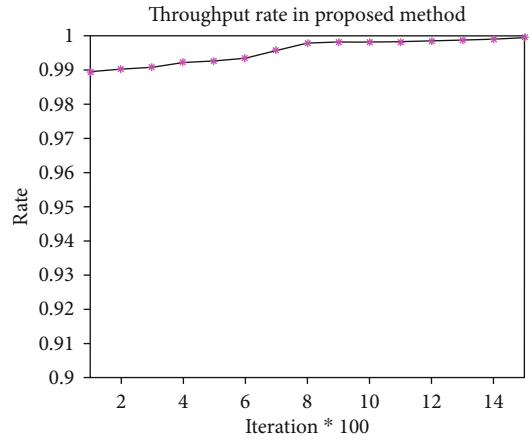


Figure 10: Permeability rate in the proposed method.

sensor network, some nodes may lose energy, but due to the use of density-based clustering, other connected nodes are responsible for collecting data, but in some cases, the number of dead nodes in the network may be so great that the network coverage is disrupted. In this case, the operation of wireless sensor networks is disrupted and the life of the network ends. Figure 8 shows the network lifetime diagram in the proposed method.

As shown in Figure 8, the lifetime of the network is completed after 1525 cycles of data and information transfer and critical data aggregation. The energy of the wireless sensor nodes ends after the first death of the nodes, respectively, and this indicates the balanced energy consumption in the proposed network. All wireless sensor nodes consume the same energy, and after the first node completes its energy in 1429, the other nodes die almost identically, indicating optimized energy consumption and longevity.

Other criteria evaluated in the proposed method are data delivery rate and network throughput. The rate of data delivery in the network is the ratio of packets delivered to the hole from total packets sent by wireless sensor nodes.

Also, the pass rate is the ratio of packages delivered per unit of time. Therefore, in order to calculate these criteria in order to optimize the proposed method based on data delivery rate and throughput rate, the rate of lost packets in the network must be first evaluated. Loss in the network increases with a very slight slope, the total number of lost packets during the life of the network is equal to 99 packets, and the loss rate of data packets is equal to 0.0198%. The proposed method, due to the use of the density-based clustering method, transmits information at the closest distance, and this reduces the rate of data packet loss and increases the rate of delivery and network throughput. Figure 9 shows a graph of data delivery rates in the proposed method. Also shown in Figure 10 is the network throughput diagram in the proposed method.

As shown in Figures 9 and 10, the proposed method reduces the data loss rate and increases the data delivery rate and throughput rate in the network by considering the data transfer at the closest distances between nodes. In the proposed method, the average data delivery rate is 93% and the average throughput is 99.55%.

4.2. Comparison of the Proposed Method with Previous Methods. After evaluating the proposed method, in order to validate and evaluate the improvement resulting from the use of the density-based clustering method, we will compare the proposed method with other previous methods. Due to the importance of energy consumption in routing methods in wireless sensor networks, various criteria for evaluating and optimizing these methods have been introduced in numerous studies conducted by researchers. One of the most important evaluation criteria introduced is the method of average energy consumption and the number of live nodes in information transmission periods in the network. Therefore, the proposed method can be compared with the methods presented in other researches based on these two criteria [6]. Figure 11 shows a comparison of the proposed method with the methods presented in other studies based on the average energy consumption.

As shown in Figure 11, the proposed method has a lower average energy consumption than the previous methods due to the use of the density-based clustering method. Figure 12 shows a comparison of the proposed method with other previous methods in terms of the number of live nodes in different periods of information transmission in the network.

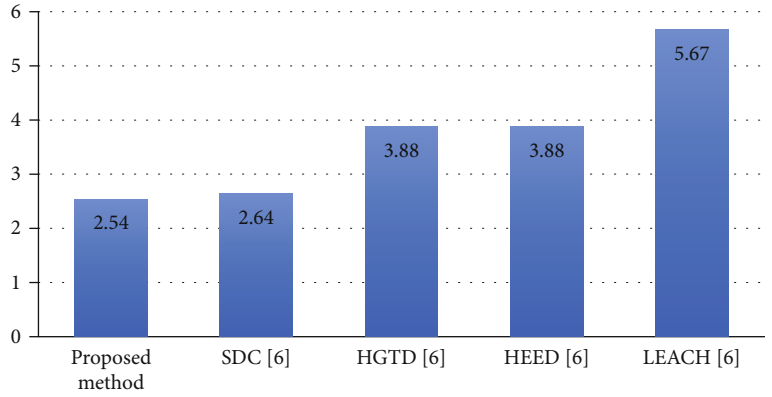


Figure 11: Comparison of the proposed method with previous methods in terms of average energy consumption.

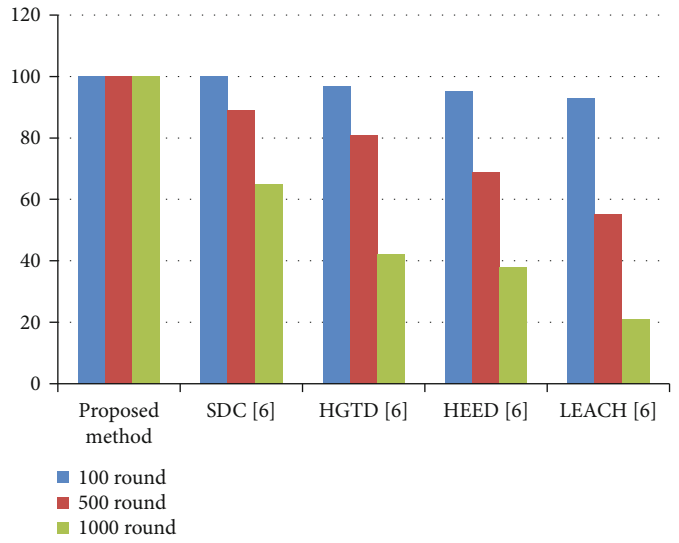


Figure 12: Number of live nodes in different cycles of information transmission in the network.

As shown in Figure 12, the proposed method has a better value than the previous methods in terms of the number of live nodes in different periods of information transmission in the network.

5. Conclusion

In this study, a density-based clustering method for collecting climate data, using wireless sensor networks that are scattered in different places in arid areas, was suggested that this method could be used for desertification-related purposes by identifying areas with critical climatic conditions. Ambient temperature information is received and analyzed at user stations. A normal threshold is defined for ambient temperature, which if the ambient temperature exceeds this threshold indicates the possibility of hazards (such as fire hazard, loss of plant tissue, and desert expansion). This study also proposes a new method to optimize energy consumption and increase the lifespan of wireless sensor networks in desert areas. The simulation results of the proposed method show that in addition to optimizing energy consumption in the network and increasing the life of the network, this method has improved the rate of packet loss and the rate of delivery and throughput in the network. In addition, the proposed method, considering the use of the density-based clustering method and the transfer of information through the closest distances, compared to other previous methods that are performed by other researchers in this field, in terms of average energy consumption and number of alive nodes on different cycles of information transmission, our proposed method has shown better results in the WSN.

Given the importance of clustering in wireless sensor networks for optimal energy routing, future suggestions for the present study include the following:

(i) Using combined metaexploration methods in order to find the optimal cluster head

(ii) Using mobile cavity nodes to reduce energy consumption

(iii) Using different density-based clustering methods for clustering wireless sensor nodes in the network

Data Availability

There is no database; we already mentioned all simulated parameters and data in this paper in the part of method simulation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Priyadarshi, B. Gupta, and A. Anurag, "Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues," *The Journal of Supercomputing*, vol. 76, no. 9, pp. 7333–7373, 2020.

[2] S. K. Singh and P. Kumar, "A comprehensive survey on trajectory schemes for data collection using mobile elements in WSNs," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 1, pp. 291–312, 2020.

[3] S. Abdollahzadeh and N. J. Navimipour, "Deployment strategies in the wireless sensor network: a comprehensive review," *Computer Communications*, vol. 91-92, pp. 1–16, 2016.

[4] T. Alhmiedat, "A survey on environmental monitoring systems using wireless sensor networks," *Journal of Networks*, vol. 10, no. 11, 2016.

[5] P. Sharma, "Wireless sensor networks for environmental monitoring," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 1, no. 9, pp. 36–38, 2014.

[6] W. Abdellatief, O. Youness, H. Abdelkader, and M. Hadhoud, "Balanced density-based clustering technique based on distributed spatial analysis in wireless sensor network," *International Journal of Wireless Information Networks*, vol. 26, no. 2, pp. 96–112, 2019.

[7] D. Sinha, R. Kumari, and S. Tripathi, "Semisupervised classification based clustering approach in WSN for forest fire detection," *Wireless Personal Communications*, vol. 109, no. 4, pp. 2561–2605, 2019.

[8] K. Grover, "WSN-based system for forest fire detection and mitigation," in *Emerging Technologies for Agriculture and Environment*, pp. 249–260, Springer, 2020.

[9] N. Moussa, A. el Belrhiti el Alaoui, and C. Chaudet, "A novel approach of WSN routing protocols comparison for forest fire detection," *Wireless Networks*, vol. 26, no. 3, pp. 1857–1867, 2018.

[10] V. Ravikumar, "Forest fire detection using wireless sensor network," *International Journal for Research in Science Engineering & Technology*, vol. 5, no. 4, pp. 1–5, 2018.

[11] E. A. Kadir, H. Irie, and S. L. Rosa, "Modeling of wireless sensor networks for detection land and forest fire hotspot," in *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, Auckland, New Zealand, January 2019.

[12] A. Aksamovic, M. Hebibovic, and D. Boskovic, "Forest fire early detection system design utilising the WSN simulator," in *2017 XXVI International Conference on Information, Communication and Automation Technologies (ICAT)*, Sarajevo, Bosnia-Herzegovina, October 2017.

[13] K. A. Darabkh, S. M. Odetallah, Z. al-qudah, A.'. F. Khalifeh, and M. M. Shurman, "Energy-aware and density-based clustering and relaying protocol (EA-DB-CRP) for gathering data in wireless sensor networks," *Applied Soft Computing*, vol. 80, pp. 154–166, 2019.

[14] Y. Yuan, W. Liu, T. Wang, Q. Deng, A. Liu, and H. Song, "Compressive sensing-based clustering joint annular routing data gathering scheme for wireless sensor networks," *IEEE Access*, vol. 7, pp. 114639–114658, 2019.

[15] A. Abdullah, A. Khaled, and A. Maamoun, "Data collection algorithm for wireless sensor networks using collaborative mobile elements," *International Journal of Electrical & Computer Engineering*, vol. 9, no. 3, article 2131, 2019.

[16] D. R. Edla, M. C. Kongara, and R. Cheruku, "SCE-PSO based clustering approach for load balancing of gateways in wireless sensor networks," *Wireless Networks*, vol. 25, no. 3, pp. 1067–1081, 2019.

[17] O. A. Mahdi, A. W. Abdul Wahab, M. Y. Idna Idris et al., "A comparison study on node clustering techniques used in target tracking WSNs for efficient data aggregation," *Wireless Communications and Mobile Computing*, vol. 16, no. 16, p. 2676, 2016.

[18] M. K. Khan, M. Shiraz, K. Zrar Ghafoor, S. Khan, A. Safaa Sadiq, and G. Ahmed, "EE-MRP: energy-efficient multistage routing protocol for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6839671, 13 pages, 2018.

[19] A. Mahdi, O. Wahab, A. W. Abdul et al., "ESAM: endocrine inspired sensor activation mechanism for multi-target tracking in WSNs," *Proceedings of the SPIE*, vol. 9902, article 99020B7, 2016.

[20] O. Adil Mahdi, A. W. Abdul Wahab, M. Y. I. Idris, A. Abu Znaid, Y. R. B. al-Mayouf, and S. Khan, "WDARS: a weighted data aggregation routing strategy with minimum link cost in event-driven WSNs," *Journal of Sensors*, vol. 2016, Article ID 3428730, 12 pages, 2016.

[21] I. Ali, A. Gani, I. Ahmedy, I. Yaqoob, S. Khan, and M. H. Anisi, "Data collection in smart communities using sensor cloud: recent advances, taxonomy, and future research directions," *IEEE Communications Magazine*, vol. 56, no. 7, pp. 192–197, 2018.

[22] J. Wang, Y. Gao, X. Yin, F. Li, and H.-J. Kim, "An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 9472075, 9 pages, 2018.

[23] 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," *2018 Tech Science Press, CMC*, vol. 56, no. 3, pp. 433–446, 2018.

[24] M. Ge, H. Bangui, and B. Buhnova, "Big data for internet of things: a survey," *Future Generation Computer Systems*, vol. 87, pp. 601–614, 2018.

[25] J. Wang, X. Gu, W. Liu, A. K. Sangaiah, and H.-J. Kim, "An empower Hamilton loop based data collection algorithm with mobile agent for WSNs," *Human Centric Computing and Information Sciences*, vol. 9, no. 1, 2019.

[26] D. Wu, S. Geng, X. Cai, G. Zhang, and F. Xue, "A many-objective optimization WSN energy balance model," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 2, 2020.

[27] I. S. Akila and R. Venkatesan, "A cognitive multi-hop clustering approach for wireless sensor networks," *Wireless Personal Communications*, vol. 90, no. 2, pp. 729–747, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6628226, 11 pages
https://doi.org/10.1155/2021/6628226

Research Article

RBM: Region-Based Mobile Routing Protocol for Wireless Sensor Networks

Muhammad Fahad Mukhtar,[1,2] Muhammad Shiraz,[2] Qaisar Shaheen,[3] Kamran Ahsan,[2] Rizwan Akhtar,[4] and Wang Changda[1]

[1]*School of Computer Science and Communication Engineering, Jiangsu University, China*
[2]*Department of Computer Science, Federal Urdu University of Arts, Science & Technology, Islamabad, Pakistan*
[3]*Department of Computer Science, Superior College, Lahore, Pakistan*
[4]*Department of IT and Computer Science, Pak-Austria Fachhochschule Institute of Applied Sciences and Technology, Pakistan*

Correspondence should be addressed to Qaisar Shaheen; qaisar.shaheen2002@gmail.com
and Wang Changda; changda@ujs.edu.cn

Received 21 October 2020; Revised 4 January 2021; Accepted 18 January 2021; Published 4 February 2021

Academic Editor: Ihsan Ali

Wireless sensor networks (WSNs) are employed for different applications for the reason of small-sized and low-cost sensor nodes. However, several challenges that include a low powered battery of the sensor nodes restrict their functionality. Therefore, saving energy in the routing process to extend network life is a serious concern while deploying applications on WSN. To this end, the key technology is clustering, which helps maximize scalability and network lifecycle. Base station (BS) collects data, aggregates it, and extracts the required information. To obtain the maximum outcome, the lifetime of the network is maximized by the use of different techniques and protocols. Data transmissions consume most of the network energy, and the transmissions over normal ranges require less energy as compared to transmissions over long ranges. Moreover, the nodes closer to the BS deplete their energy faster as compared to distant nodes because of traffic overload. The proposed protocol is aimed at reducing energy consumption and increasing the network lifetime. For this purpose, the network is divided into two regions: region 1 closer to the BS communicating directly, whereas region 2 farther away from the BS having routing nodes to communicate with the BS. Routing nodes do not take part in sensing function but will only move in region 2 collecting data and forwarding it to BS. MATLAB is used as the simulation tool for evaluation, and the results are compared with the existing optimized region-based efficient routing (AORED) and low-energy adaptive clustering hierarchical protocol (LEACH) techniques. The comparison showed that energy conservation and lifetime increased by 15%, and throughput is increased by more than 5% approximately.

1. Introduction

During the past few years, wireless sensor networks (WSNs) have earned a great attraction of researchers. WSN comprises of a large number of small-sized and low-cost sensor nodes connected to base station (BS); they are used to gather information from their environment and forward this information to the BS. They have a wide range of applications in different fields like medical, engineering, agriculture, environmental monitoring, surveillance, and battlefields. There are many types of sensor nodes to monitor different physical quantities like thermistors are used for heat sensing, photodiodes for light sensing, GPS sensors to pinpoint the location, etc. The WSNs have a large number of applications but there are some limitations, which restrict their functioning. As the nodes are small in size and low in cost, they are not equipped with a large amount of power supplies. As they are mostly used in open and vast fields, their power supplies cannot be replaced or they cannot be connected with continuous power supplies. This limited power supply also limits the abilities of sensor nodes, like limited life, limited processing power, limited storage capacity, and lower communication ranges. It is

required to make the WSN function in such a way that it consumes minimum energy resources. This is done by implementing different routing protocols and routing strategies. The communication in WSNs can either be direct (single-hop communication) or there can be an intermediary node to forward the collected data to BS (multihop communication) [1]. Nodes have a specific communication range within which they can communicate with normal power consumptions, but to communicate over these ranges, extra power is consumed. In the case of large networks clusters are formed and a cluster head (CH) is elected in each cluster, this CH collects and forwards the data of its cluster. Some challenges are faced by WSNs that restrict their working. Some of the main challenges faced by WSNs include limited power supplies, limited processing power, and less memory size. Earlier in the field of WSNs, the nodes were kept static but their topology and protocols kept on changing to gain the maximum output from the network. But during the past few years, mobility of nodes is being introduced in the WSNs to enhance the output of the WSN. Now, many researches are being done on the mobility of nodes in the network.

One solution to enhance the WSN lifetime is to move the sink to the areas in which nodes have more energy. Using a mobile sink is useful for saving energy only when we have an efficient routing strategy towards it. The reason to make the sink mobile was to transfer the load of data processing and power consumption during transmissions, from the sensor nodes to the sink node, so that the network lifetime can be enhanced. As mobile sink moves actively in the network, it moves closer to sensor nodes to reduce the transmission distance, and, hence, there are only a few intermediary nodes left to forward the data. Hence, the consumption of the energy is more equally distributed in the WSN with mobile sink as compared to the conventional WSN with static sink. The main contributions of this paper are

(1) This paper focused on decreasing the communication load on cluster heads that are at a farther distance from the base station by using mobile routing nodes

(2) Enhancing the stability and lifetime of the network by conserving the energy of sensor nodes

The proposed protocol, region-based mobile routing protocol also focuses on the mobility of nodes in the WSN. In the proposed protocol, mobile routing nodes are introduced which will move in the network and will be responsible for communication between CHs and BS but they will not take part in the sensing functions of the network. The protocol operates by dividing the network into two regions, one closer to the sink without mobility and the other farther from the sink with mobile routing nodes. The aim of this study is to enhance the network stability by increasing its lifetime and energy conservation. MATLAB was used as a simulation tool to evaluate the functioning of the proposed protocol. Some other existing techniques were used to compare the results of our proposed technique with existing techniques, i.e., optimized region-based efficient routing (AORED) [2] and low-energy adaptive clustering hierarchical protocol (LEACH) [3]. The comparison showed that the goals set for the proposed technique were achieved as our RBM performed better than LEACH and AORED. The comparison showed that energy conservation and lifetime increased by 15%, and throughput is increased by more than 5% approximately. The results showed that the proposed technique was successful in achieving its goals.

The rest of this article is structured as follows. A brief summary of the literature review is presented in Section 2. Section 3 explains the details of the proposed region cluster-based mobile routing protocol strategy for sensor energy consumption in wireless sensor networks. In Section 4, the performance of evaluation of the proposed region cluster-based mobile routing protocol strategy for sensors energy consumption in wireless sensor networks is given. The paper is concluded in Section 5 with the direction for future work.

2. Literature Review

Wireless sensor networks have an extensive variety of applications and a very complex structure. Thus, many problems are faced by the researchers while dealing with WSNs. Different researches were done to resolve these issues and hence different solutions and routing protocols were introduced. Low-energy adaptive clustering hierarchical protocol (LEACH) is the base of different other protocols. To maintain balance in power consumption of all nodes in LEACH, clusters are formed and a cluster head (CH) is elected for every cluster which is responsible for communication between nodes of its clusters and BS. Afterward, many researchers introduced some changes in LEACH to overcome deficiencies in it [3].

Stable election protocol (SEP) [4] was introduced with some changes in LEACH, in which CHs were chosen on the basis of remaining energy of the nodes. The remaining energy of the node over the average energy of the network was estimated to elect CHs in distributed energy-efficient clustering (DEEC); it was also introduced as the successor of LEACH [5]. In [6], LEACH was improved by introducing the master head and shortest path algorithm, in which data will be forwarded to BS using MIMO. In [7], H-LEACH was introduced in which the nodes that are unable to communicate are made to die; also, a record of alive nodes is maintained.

Multiple mobile data collectors were introduced in [8] to overcome the issue of blockage in the surrounding of the sink because of extraordinary traffic load. Mobile data collectors consisted of mobile sinks (destination for collected data) and mobile relays (agents to carry collected data to destination). In [9], the authors proposed an effective grid deployment method for mobile sensor nodes deployment, in which the plot is separated into many separate grids and environmental factors like predeployed nodes, boundaries, and obstacles are used to estimate the weight of every grid and mobile node focus on the grid having minimum values. In [10], the authors first proposed an approach on election-based mobile collection and framed it into an optimization problem called bounded relay hop mobile data gathering

(BRH-MDG), in which a subgroup of sensor nodes will be designated as polling points and will save locally collected data and upload it to the mobile collector on its arrival.

In [11], a hybrid (mixed) routing protocol was proposed for WSNs containing mobile sinks. The suggested routing protocol is a mixture of proactive and reactive routing protocols for low-power networks with multiple mobile sinks. DAG (directed acyclic graph) is maintained by the nodes within a specific zone nearer to the sink. But the nodes outside this zone do not use DAG, instead they use on-demand sink discovery to discover the sink at the nearest possible distance. But in the situation of high sink mobility path to sink will change repeatedly, so it is better to create small zones. But if the mobility is slow, then, in this case, bigger zones can be created for quick data transfer to the sink node. Maximum Amount Shortest Path (MASP) [12] resolves the issue of a path constrained mobile sink having constant speed because the communication time to gather data from randomly deployed nodes is limited. This issue arises the issues like the amount of collected data and power conservation. The network sensor field is partitioned into two parts, the direct communication area for nearby sensors and the multihop communication area for sensors at a greater distance. In [13], the authors defined distributed heuristics and scalable models for the synchronized movement of multiple sinks in WSN. It was shown in this paper that better performance was observed with controlled and coordinated mobile sinks as compared to static sink or uncontrolled mobility of the sink node. In [14], the authors have proposed a mobility-based clustering (MBC) protocol for WSN containing mobile nodes, a sensor in this protocol elects itself as CH on the basis of its mobility and remaining energy. During the cluster formation process by taking into consideration its connection time with cluster head, its residual energy, and its distance from CH, a noncluster head node will maintain the link stability with CH during the process of cluster formation. By restricting the moving distance of the sink node, the data loss during the movement of the sink node from one location to another can be minimized [15]. In hotspot problem, sensor nodes near the sink have to transmit the data of faraway nodes, and as a consequence, they drain their energy fast, creating a gap in the network reducing the network lifetime. In [16], the authors have addressed this issue and purposed Mobile Sink based Routing Protocol (MSRP), in which a mobile sink will move through the network and collect data from CHs within its locality. Only the CHs in the locality of the mobile sink will transmit their data to the sink, and the remaining will wait for the mobile sink to be in their locality. In [17] is given a cooperative localization algorithm that studies the presence of hurdles in networks with mobile elements. To improve the location performance, the mobile sink will move actively and cooperate with static nodes. In [18], the authors proposed a scheme to increase the lifetime of the delay-tolerant WSN by using a mobile sink. In delay tolerant WSN, a node needs not to transmit the data immediately after it becomes available but it temporarily saves the data and sends it when the mobile sink is at a suitable location. In [19], the authors limit the mobile sink to a specific number of locations and study the joint sink mobility and routing issues. The authors first developed the primal-dual algorithm for a single sink and then generalized it for multiple sinks. Packet delivery ratio along with energy conservation is the issues caused by the mobile nodes in the network. To overcome these issues in [20], the authors have used a cross-layer design between MAC and network layers and proposed a cluster-based routing protocol for mobile sensor nodes (CBR-Mobile).

Purely location information is used in geographic routing protocol instead of global topology information, hence, is considered as simple, efficient, and scalable routing protocol. Frequent location updates are sent to receive data frequently but these frequent updates cause more energy consumption and collisions in wireless transmission. In [21], the authors proposed a new geographic routing protocol named elastic routing to tackle this issue. During the movement of the sink, the location update is sent to the source along a backward geographic routing path. Source will get the location of mobile sink and forwards data, but after the sink has moved to a new location, the last hope forwarding node detects its new location and renews this information in a received packet to a new location. Transmission of this modified packet is overheard by the second last hop forwarding node and alters location information in afterward received packets and forwards them to a new location. As a result, the new location is forwarded to the source. The authors in [2] proposed a new technique using the idea of limiting the topology of communicating nodes using connected dominating set (CDS) of nodes. The authors proposed an optimized region-based efficient routing (AORED) protocol for WSNs to overcome the energy issue. Using CDS, communicating nodes build a virtual backbone in the network by minimizing the number of communicating nodes. In this technique, the authors divided the complete network field into two parts to reduce the transmission energy, direct and indirect communication parts. In [22], the unique characteristics of social relationship with MSN (mobile social network) give rise to different protocol design issues are explained. In [23], a new computing paradigm for the optimization of parameters in adaptive beamforming using fractional processing is given. In [24], modeling and optimization of microwave filter by ADS-based KBNN is presented. The authors in [25] presented the computation and analysis of energy-efficient multirelay and multihop communication scheme in wireless sensor networks.

Vehicular ad hoc networks (VANETs) were studied in [26–28], where two important factors are defined, communication lifetime and distance; these factors were used to find the closest node and time estimation for being in the communication range of forwarder. An efficient method for cluster head selection was adopted, and a reliable and efficient routing protocol was proposed to address the routing issues in [29]. WDARS was proposed in [30], in which a new technique of weighted data aggregation routing was used by studying the prevailing issues, and this technique used a hop-tree to get maximum data aggregation. Enhanced Power-Efficient Gathering in Sensor Information System (EPEGASIS) was introduced in [31] with some changes in PEGASIS to overcome the hot spot issue by following the four steps, calculating optimal communication distance,

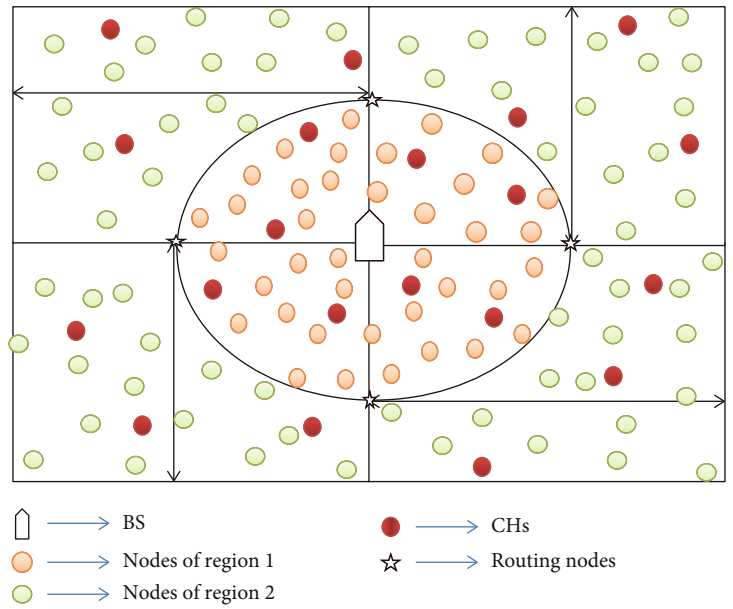


Figure 1: Proposed network architecture and mobility of routing node overview.

using mobile nodes, setting threshold for dying nodes, and communication range can be set by the node itself. Trajectory scheduling algorithm based on coverage rate for mobile sinks (TSCR-M) was introduced in [32], where authors used particle swarm optimization (PSO) along with mutation operator to search for parking positions with best coverage rate and then authors used genetic algorithm (GA) to plan the moving route for mobile sinks. In [33], an affinity propagation-based self-adaptive (APSA) clustering method was introduced, combination of advantage of machine learning algorithm, K-medoids with affinity propagation (AP) was used to get sensible clustering performance, AP calculates the number of cluster heads and enhanced K-medoids form the network topology by iteration.

3. Proposed RBM: Region-Based Mobile Routing Protocol

In this study, sensor nodes were deployed in some prespecified area, i.e., $100 \times 100\,m^2$, and the base station (BS) was installed at the center of the field. Then, the whole network field was divided into two parts; region 1 and region 2. Region 1 consisted of an area of 30 m of radius from the base station, and the rest of the area was named region 2. The sensor nodes at a distance of 30 m or less from BS were included in the region 1, and the nodes that were farther from this distance were part of the region 2. Clusters were formed, and CHs were elected among nodes based on their residual energy. As the sensor nodes can communicate with normal energy consumptions over a range of approximately 30 meters, that is why such region formation was done. Now as the CHs of the region 1 communicated directly with BS with normal energy consumptions, while for CHs of region 2, four routing nodes were placed at the boundary of regions 1 and 2 along x-axis and y-axis as depicted in Figure 1 to conserve the energy of nodes of region 2 as they were out of normal communication range. The routing nodes installed at the boundary of two regions comprised of same abilities as the other nodes but they did not take part in sensing or other normal functions of the network but they were employed for the specific purpose of gathering data from the CHs of the region 2 and transmit it to the BS. As these routing nodes were not taking part in any other function and only have to transmit data over normal ranges, hence, did not run out of energy fast and also helped the nodes of region 2 to preserve their energy. Also, these nodes had mobility over a specific range of distance where they were free to move while communicating with the BS. The proposed technique consisted of rounds, and every round comprised of two phases named setup and steady phase.

While moving in region 2, the routing nodes will update their position to CHs of region 2; these CHs will then forward the data to routing nodes when they will be in their vicinity at a closer distance. The installation positions of the four routing nodes are prespecified. As depicted in Figure 1, the entire area of the network is $100 \times 100\,m^2$ with BS at the center position (50, 50). The installation positions of mobile routing nodes are (25, 50), (50, 25), (75, 50), and (50, 75), and their movement directions are also mentioned in Figure 1.

Cluster heads in this phase are randomly selected. These CHs are chosen on basis of remaining energy of the node and threshold $T(n)$. Every node during cluster head selection process determines a random number from 0 to 1. Comparison of this number is then made with threshold $T(n)$, and the node with a number smaller than the threshold will become CH for the current round. The formula to calculate the threshold is

$$T(n) = \frac{1}{1 - (P * r \bmod (1/P))} \quad \text{if } n \in G, \tag{1}$$

$$T(n) = 0 \quad \text{if } n \notin G, \tag{2}$$

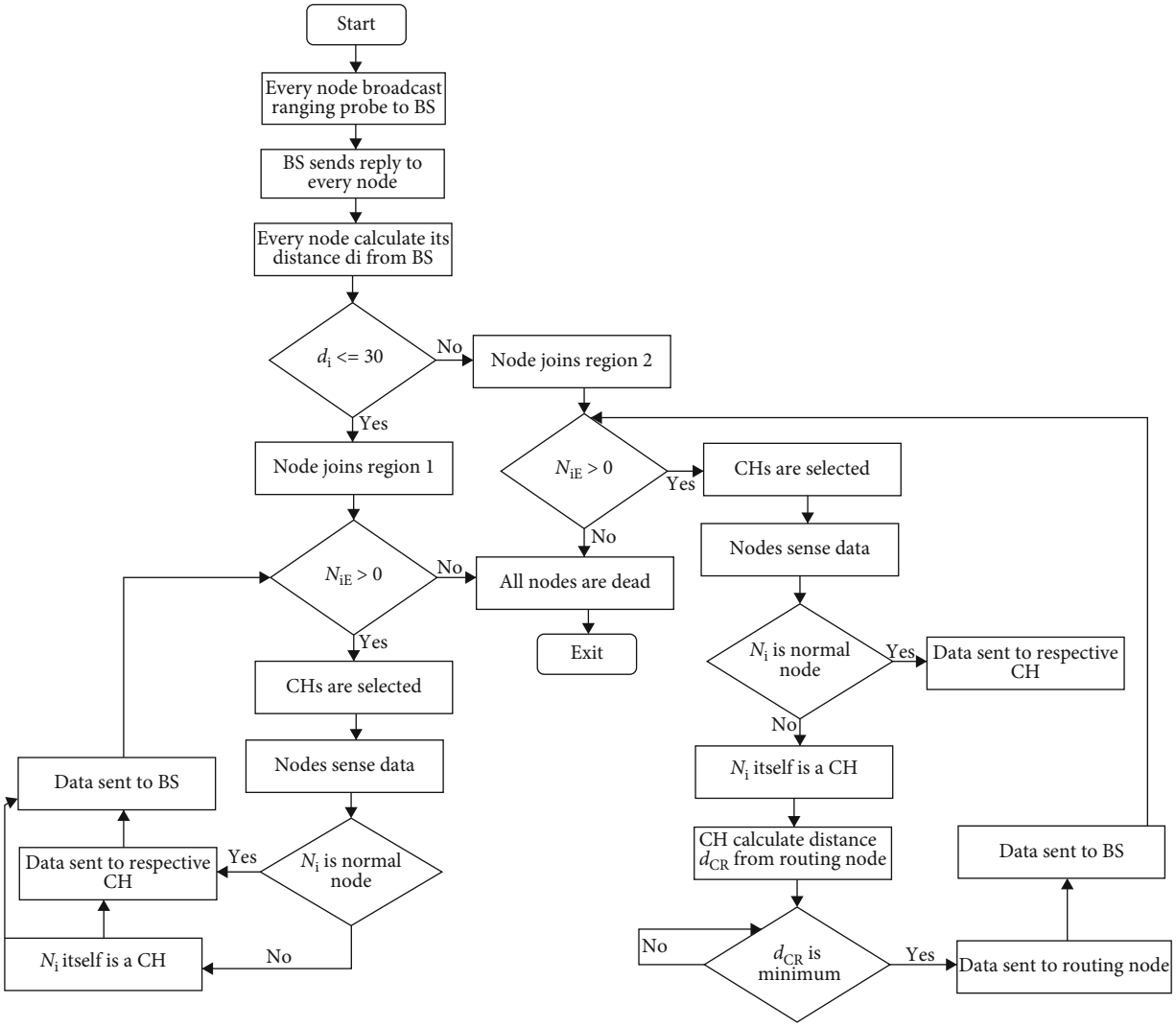


Figure 2: Flowchart of the proposed RBM.

where P is the probability to become a CH, r is the current round number, set of sensor nodes that have not been elected as CH in the last $1/p$ rounds is called G. The purpose of this algorithm is to ensure that in $1/p$ rounds, every node becomes CH only once. These two equations show the criteria for a node to become CH. If a node has not been a CH for the last $1/P$ rounds, then, there is a probability that it can become a CH; otherwise, it will not be considered as candidate for CH.

In the steady phase, the sensor nodes start communicating with their respective cluster heads. Elected CHs will broadcast their status, and on the basis of the strength of these broadcast signals, the sensor nodes will select their CH. After the cluster formation, the CHs will fix a TDMA schedule for members of its cluster and will broadcast this schedule. In region 1, the elected CHs aggregate sensed data and forward the data to BS directly. CHs of region 2 forward the received sensed aggregated data to the mobile routing nodes when they arrive in their locality. Mobile routing nodes send the received data to BS. The complete working of the proposed protocol is depicted in Figure 2.

Consider a two-dimensional area deployed with n number of nodes. According to the present study, these nodes will be categorized into three types, normal nodes, routing nodes, and cluster heads, where each type will perform a specific function. Function like sensing in the network, collection of data after sensing, and transmission of this data to CHs are performed by the normal nodes. Then comes the CHs which will perform the function of transmitting the received data, the CHs of region 1 will forward the data directly to BS while CHs of region 2 will communicate with routing nodes to forward them the received data. Finally, the routing nodes will forward this data of CHs to the BS.

Some mathematical formulas regarding the network's maximum throughput are presented in the literature. According to literature maximum, throughput can be obtained using the equation:

$$d_{\text{Total}} = \sum_{i=1}^{\text{NN}} u_t^i + u_r^i + \int_0^{2\pi} \int_0^r p\left(\pi r^2\right) r dr d\theta, \tag{3}$$

where u_t^i is number of bits transmitted by node i, and u_r^i is the number of bits received by node i. According to equation (3), if the output of the network is to be increased, then the lifetime of the network must be prolonged. To achieve a long network lifetime, the nodes should live for maximum duration. This equation shows that the total number of bits received and transmitted will be greater if the nodes live for a longer period which will only be possible if the energy consumption of the nodes is reduced.

$$\sum_{i=1}^{n} u_t^i + u_r^i \leq E_{\text{Total}} \quad \forall i \in i, \tag{4}$$

where u_t^i the number of bits is transmitted by node i, and u_r^i is the number of bits received by node i. This equation shows that the total energy consumed by the network for receiving and transmission of u bits will always be less than the total energy of the networks. This equation shows that the energy consumed to transmit u bits by the node i has an upper bound by the total energy provided to the network.

$$\sum_{i=1}^{\text{CH}} \sum_{i=1}^{\text{NN}} f_{ci} \leq F \quad \forall c, i \in n, \tag{5}$$

where f_{ci} is data flow between CHs and nodes, and F is the total flow of the network. According to equation (5), the flow of data between CHs and normal nodes has an upper bound by the total flow of the network. This equation shows that the total data flow between nodes to CHs and from CHs to normal nodes will always be less than the total data flow within the network.

$$f_{ci} \leq C_{ci} \quad \forall c, i \in n. \tag{6}$$

The relationship between flow and total capacity of some specific link is explained in equation (6). Where f_{ci} is the data flow between CH and node, and C_{ci} is the capacity of their link. This equation shows that the data flow through a link will always be less than the capacity of that link.

In the network, there are three different types of nodes; each type has different specific functions. Therefore, the energy consumption of each type of node will be different from the others depending upon their roles. The following equation used to calculate the energy consumed by the normal nodes:

$$E_{\text{NN}} = e_s + e_t(d_{\text{NC}}). \tag{7}$$

In equation (7), E_{NN} gives the energy consumed by the normal nodes, while sensing and transmitting their own data, e_s is energy spent during sensing the data, e_t shows the energy consumption in transmitting the data, and d_{NC} represents the distance from normal node to its CH. This equation shows that the total energy spent by some specific node N is the sum of energies spent in sensing the data and transmitting this data to CH over a distance d_{NC}. Equation (7) is a specific equation for some specific single node. We can generalize it for all normal nodes of the network using the following equation:

Table 1: Simulation parameters.

Parameter	Values
Number of nodes	104
Network size	100×100 m
Packet size	1 byte
Initial energy	500 mJ
Data aggregation energy cost	50 pj/bit
Transmission energy	50 nJ/bit
Receiving energy	50 nJ/bit
Simulation rounds	6000

$$E_{\text{NN}} = \sum_{\text{NN}}^{n} e_s + e_t(d_{\text{NC}}). \tag{8}$$

Equation (8) is a generalized form of equation (7). It shows that the total energy consumption of all normal nodes while sensing the data and transmitting this data to their respective CHs over some distance d_{NC}.

Assumption is made for CHs that each cluster head shares equal load from normal nodes of its cluster. First, we will take into account the energy consumed by the CHs of region 1, and their energy consumption is depicted in the equation given below:

$$E_{\text{CH},R1} = e_s + e_t(d_{\text{CB}}) + \frac{\text{pn}_{R1}}{m_{R1}} [e_r + e_t(d_{\text{CB}})], \tag{9}$$

where e_s is the energy spent by CH while sensing the data $e_t(d_{\text{CB}})$ is the energy spent while transmitting its own sensed data and also the data from other nodes to the BS, e_r is the energy spent in receiving the data, d_{CB} shows the average path distance from CH to the BS.

For the energy consumption of CHs of region 2, the following equation will be used:

$$E_{\text{CH},R2} = e_s + e_t(d_{\text{CR}}) + \frac{\text{pn}_{R2}}{m_{R2}} [e_r + e_t(d_{\text{CR}})], \tag{10}$$

where e_s is the energy spent by CH while sensing the data $e_t(d_{\text{CR}})$ is the energy spent while transmitting its own sensed data and also the data from other nodes to the routing node, e_r is the energy spent in receiving the data, and d_{CR} shows the average path distance from CH to the nearest routing node.

These equations give the estimation of the energy consumed by the CH. The energy consumption of CH will be greater than the other normal nodes. CHs of region 2 will sense the data and forward their own data along with the data of other connected nodes to the routing nodes. Energy consumption of CHs of region 2 is the sum of energies spent in sensing and forwarding its own data and energy spent in receiving and forwarding the data from other nodes of the same cluster.

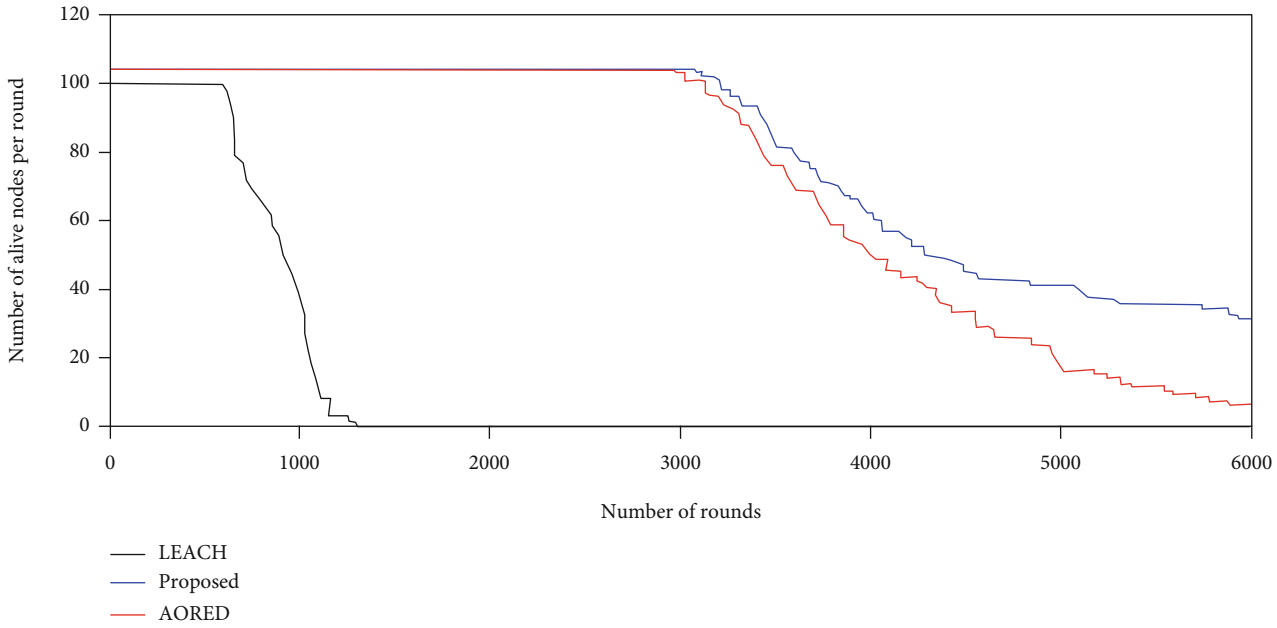


Figure 3: Number of alive nodes per round.

It is assumed about routing nodes that equal load is shared from cluster heads of region 2. The following equation will show the energy consumed by the routing nodes:

$$E_{\mathrm{RN}} = e_t(d_{\mathrm{RB}}) + \frac{\sum_{k=2}^{K} \mathrm{pn}_k}{m_{R1}} [e_r + e_t(d_{\mathrm{RB}})], \tag{11}$$

where d_{RB} is the average distance of routing node from BS, e_t is the energy consumed while transmitting the data over distance d_{RB}, and e_r is the energy consumed in receiving the data. This equation gives the total energy consumption of routing nodes which includes all the energies spent in receiving the data from CHs of region 2 and transmitting this data to the BS.

4. Evaluation of the Proposed Region-Based Mobile Routing Protocol

The performance evaluation of the proposed RBM Protocol was done by comparing its results with the results of the AORED and LEACH. Some performance evaluation criteria were considered to determine the performance of the proposed model. These performance evaluation criteria are described briefly. *Stability Period:* this is the period till the first node in the WSN is dead. So we will observe the time period when the first node of the WSN becomes exhausted and remains no more a part of our network. *Lifetime:* it is the time period from the start of network function till the expiry of all nodes in the network is termed as lifetime of the network. We will observe for the network lifetime of our network. *Number of Elected Cluster Heads per Round:* the number of elected cluster heads per round that are capable of gathering and aggregating the data will also be observed to evaluate the performance of the proposed scheme. *Number of Dead Nodes:* to evaluate the network performance, the number of nodes that are dead after some particular round is also observed. *Number of Packets to Base Station:* data sensed by nodes in the network is forwarded to CHs, where this data is aggregated and sent to BS either directly or indirectly. After the network lifetime has expired, the total number of packets received by BS is observed to evaluate the performance of the network. Simulation parameters are given in Table 1.

4.1. Simulation Results and Analysis. To prove the proposed system, MATLAB simulator was used as a simulation tool to study and evaluate its performance. The better working of the proposed protocol was shown by comparing its results with some existing techniques, i.e., AORED and LEACH. The working of both techniques AORED and LEACH has already been explained in the literature review section of this article.

Figure 3 gives a comparison between the number of live nodes (along y-axis) after each round and total number of rounds (along x-axis). This comparison actually gives us the stability period of the network; the time period from the start of network functioning till the expiry of the first node is called the stability period of the network. Figure shows that in the proposed protocol, the first node becomes dead approximately after 3200 rounds approximately, and in AORED, the first node becomes dead at about 2900 rounds, but in the case of LEACH, the first node is dead approximately near about 600 rounds. As the proposed RBM remained stable for more rounds as compared with other techniques, hence, RBM proved to have more stability period than other techniques. The figure depicts that the stability of the proposed protocol has been enhanced by 11% approximately when compared with AORED and more than 400% when compared with LEACH.

In Figure 4, the number of dead nodes (along y-axis) is compared with the number of rounds (along x-axis) to

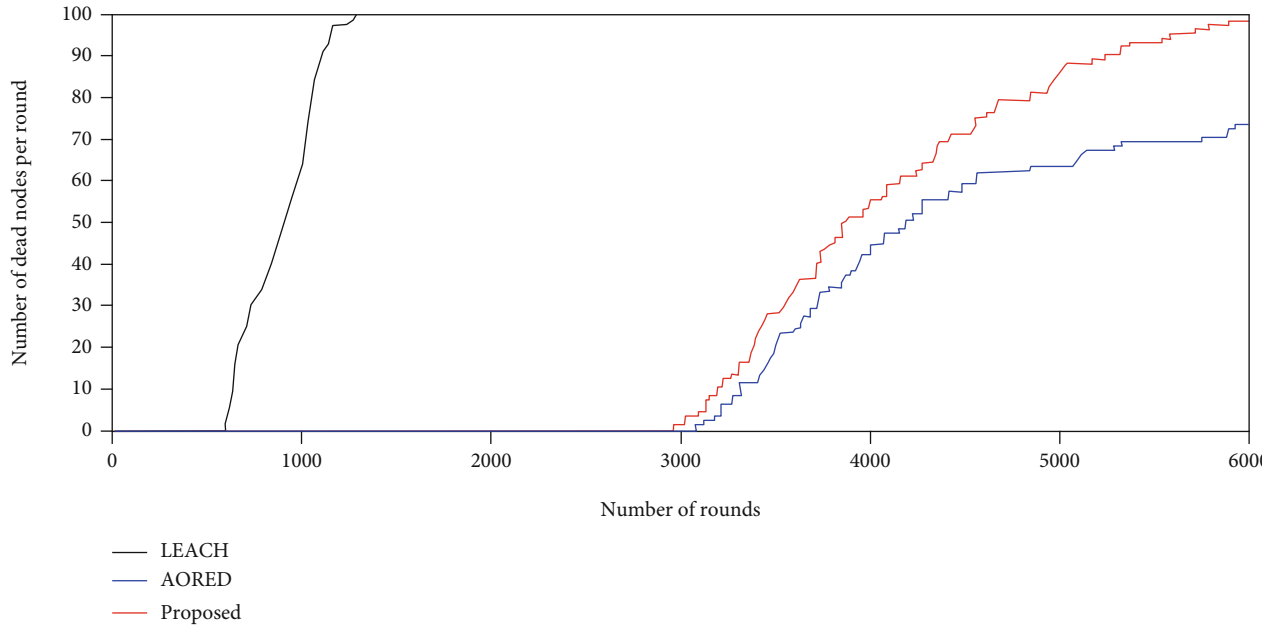


Figure 4: Number of dead nodes per round.

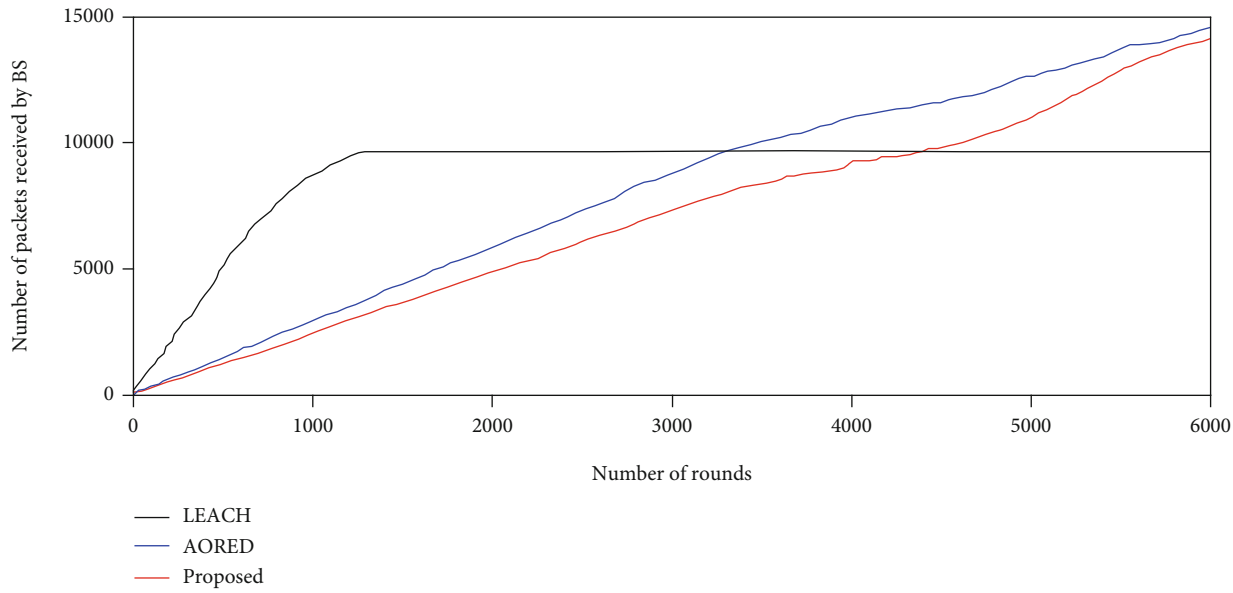


Figure 5: Number of packets sent to base station per round.

calculate the lifetime of the network. The time period from the start of network function till the expiry of all nodes in the network is termed as the lifetime of the network. The first node in case of LEACH becomes dead at near about 600 rounds, and approximately after 1300 rounds, all of the sensor nodes are dead showing that the LEACH has a lifetime of about 1300 rounds only. For AORED, it is observed from the figure that its first sensor is dead nearly at 2900 rounds, and after the completion of 6000 rounds, 99% of its sensor nodes are dead leaving the WSN of no more use. But in the case of the proposed RBM Routing Protocol, it is obvious from the figure that its first node becomes dead after about 3200 rounds, and after the completion of 6000 rounds, it can be seen that about 73% of nodes are dead leaving about 27% of the sensor nodes still in useful state and will continue their function for the WSN leaving the network still in working condition until all of their energy is drained out. These results show that in terms of lifetime, the proposed RBM showed 34% better performance than AORED and more than 4 times better performance than LEACH. Hence, achieving the desired better performance.

In Figure 5, the number of packets sent to BS (along y -axis) and number of rounds (along y-axis) are related to calculate the throughput of the WSN. It has been explained in the working of the proposed protocol that the CHs of region 2 will forward their data packets to mobile routing

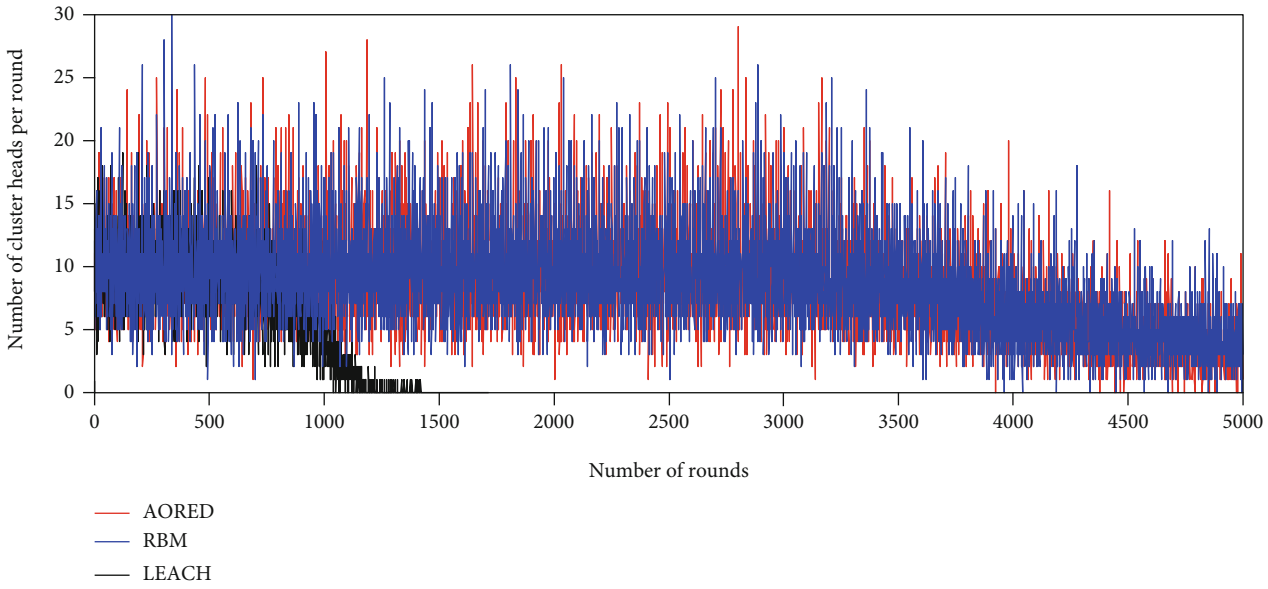


Figure 6: Number of elected cluster heads per round.

nodes. These mobile routing nodes will then send the data to BS after aggregation. In this manner, more than 50% of the data sent to BS is received and controlled through routing nodes. The packets received by BS and it is seen that in the case of AORED and proposed techniques, the packets are sent to BS in a regular manner but the packets sent by proposed are more than AORED is shown in Figure 5 graph. While in LEACH packets sent to BS suddenly reaches its peak and then stops. In 6000 rounds, the proposed RBM sent more than 14500 packets to BS while AORED sends about less than 14000 packets approximately. This shows more than 5% increase in throughput.

Figure 6 shows the number of CHs elected in each round. CHs are responsible to obtain the sensed data from sensor nodes, and after aggregating the data, they forward it to BS. New CHs are chosen after the completion of every round. This election of CHs is in a random manner to keep balance in energy consumption between the sensor nodes. If a small number of CHs are chosen, then, they have to receive, aggregate, and forward the data from more sensor nodes which consume more of their energy. Election of more CHs mean increased capability of receiving, aggregating, and forwarding of more data with less consumption of energy resources.

5. Conclusion and Future Work

In this study, a region-based mobile routing protocol has been proposed for the enhancement of network lifetime in WSN. The network field was first partitioned into two parts on the basis of distance from BS. BS was placed at the center of the network field. CHs of region 1 that were closer to BS communicated with BS directly with normal energy consumptions. Mobile routing nodes were introduced in region 2 that moved on the specified path to gather data from CHs of region 2 and sent it to BS after aggregating it, conserving the energy of nodes and consequently increasing the network lifetime. Simulations were performed using MATLAB simulator, and the results were compared to AORED and LEACH techniques. Evaluation of the proposed RBM technique on the basis of selected parameters, stability period, node lifetime, number of elected cluster heads per round, and number of packets to BS shows that RBM performed better than AORED and LEACH. The simulation results showed that the proposed technique achieved its goals of energy conservation and enhancement of network lifetime.

As future research considerations, the network field can further be divided into more than two regions and introducing some more routing nodes reducing the transmission distance, which will save energy. Also in the same model, the experimentations can be done by changing the movement path or directions of the routing nodes. This study can also be extended by adding more routing nodes in the system.

Data Availability

The data 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 National Key Research Project under grant no. 2017YFB1400703, China.

References

[1] H. Lee, M. Jang, and J.-W. Chang, "A new energy-efficient cluster-based routing protocol using a representative path in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 10, no. 7, Article ID 527928, 2014.

[2] J. Luo and J.-P. Hubaux, "Joint sink mobility and routing to maximize the lifetime of wireless sensor networks: the case of

constrained mobility," *IEEE/ACM Transactions on Networking*, vol. 18, no. 3, pp. 871–884, 2010.

[3] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless micro-sensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences*, Maui, HI, USA, USA, January 2000.

[4] G. Smaragdakis, I. Matta, and A. Bestavros, "SEP: a stable election protocol for clustered heterogeneous wireless sensor networks," in *Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, Boston, MA, USA, 2004.

[5] L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer Communications*, vol. 29, no. 12, pp. 2230–2237, 2006.

[6] A. Bharti, C. Devi, and V. Bhatia, "Enhanced energy efficient LEACH (EEE-LEACH) algorithm using MIMO for wireless sensor network," in *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–4, Madurai, India, December 2015.

[7] A. Razaque, S. Mudigulam, K. Gavini, F. Amsaad, M. Abdulgader, and G. S. Krishna, "H-LEACH: hybrid-low energy adaptive clustering hierarchy for wireless sensor networks," in *2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, pp. 1–4, Farmingdale, NY, USA, April 2016.

[8] Y. D. Yasmine-Derdour, B. K. Bouabdellah-Kechar, and M. Faycal-Khelfi, "Using mobile data collectors to enhance energy efficiency and reliability in delay tolerant wireless sensor networks," *Journal of Information Processing Systems*, vol. 12, no. 2, pp. 275–294, 2016.

[9] M. Dhivya, K. Divya, R. Keerthi, and K. P. Kumar, "Sink mobility for data collection in wireless sensor network life cycle," *International Journal of Computer Science and Information Technologies*, vol. 6, no. 2, pp. 1015–1018, 2015.

[10] M. S. Vidhya, V. Subhashini, N. V. Banu, A. Vaishnavi, M. Jayachithra, and J. Jayarajan, "Localisation algorithm based efficient controlled sink mobility for wireless sensor network," *International journal of electrical engineering and telecommunications*, vol. 1, no. 1, 2015.

[11] F. El-Moukaddem, E. Torng, and G. Xing, "Mobile relay configuration in data-intensive wireless sensor networks," *IEEE transactions on mobile computing*, vol. 12, no. 2, 2013.

[12] Y. Gu, Y. Ji, J. Li, and B. Zhao, "ESWC: efficient scheduling for the mobile sink in wireless sensor networks with delay constraint," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, 2013.

[13] L. Shi, B. Zhang, H. T. Mouftah, and J. Ma, "DRP: an efficient data-driven routing protocol for wireless sensor networks with mobile sinks," *International Journal of Communication Systems*, vol. 26, pp. 1341–1355, 2013.

[14] R. C. Luo and O. Chen, "Mobile sensor node deployment and asynchronous power management for wireless sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 5, pp. 2377–2385, 2012.

[15] M. Zhao and Y. Yang, "Bounded relay hop mobile data gathering in wireless sensor networks," *IEEE transactions on computers*, vol. 61, no. 2, pp. 265–277, 2012.

[16] V. Safdar, F. Bashir, Z. Hamid, H. Afzal, and J. Y. Pyun, *A Hybrid Routing Protocol for Wireless Sensor Networks with Mobile Sinks*, ISWPC, 2012.

[17] S. Gao, H. Zhang, and S. K. Das, "Efficient data collection in wireless sensor networks with path-constrained mobile sinks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 4, pp. 592–608, 2011.

[18] S. Basagni, A. Carosi, C. Petrioli, and C. A. Phillips, "Coordinated and controlled mobility of multiple sinks for maximizing the lifetime of wireless sensor networks," *Wireless networks*, vol. 17, no. 3, pp. 759–778, 2011.

[19] S. Deng, J. Li, and L. Shen, "Mobility-based clustering protocol for wireless sensor networks with mobile nodes," *IET Wireless Sensor Systems*, vol. 1, no. 1, pp. 39–47, 2011.

[20] W. Liang, J. Luo, and X. Xu, "Prolonging network lifetime via a controlled mobile sink in wireless sensor networks," in *2010 IEEE global telecommunications conference GLOBECOM*, pp. 1–6, Miami, FL, USA, December 2010.

[21] B. Nazir and H. Hasbullah, "Mobile sink based routing protocol (MSRP) for prolonging network lifetime in clustered wireless sensor network," in *2010 international conference on computer applications and industrial electronics (ICCAIE 2010)*, Kuala Lumpur, Malaysia, 2010.

[22] R. Akhtar, Y. Shengua, Z. Zhiyu et al., "Content distribution and protocol design issue for mobile social networks: a survey," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, 2019.

[23] M. A. Z. Raja, R. Akhtar, N. I. Chaudhary, Z. Zhiyu, Q. Khan, and A. U. Rehman, "A new computing paradigm for the optimization of parameters in adaptive beamforming using fractional processing," *The European Physical Journal Plus*, vol. 134, no. 6, p. 275, 2019.

[24] C. Yi, T. Yubo, and L. Mingjun, "Modeling and optimization of microwave filter by ADS-based KBNN," *International Journal of RF and Microwave Computer Aided Engineering*, vol. 27, no. 2, 2017.

[25] Z. Xie, Q. Shen, Y. Hu, Y. Su, and Y. Wang, "The computation and analysis of energy-efficient multirelay and multihop communication scheme in wireless sensor networks," *International Journal of Communication Systems*, vol. 31, no. 7, article e3435, 2017.

[26] A. I. Ahmed, A. Gani, S. H. Ab Hamid, S. Khan, N. Guizani, and K. Ko, "Intersection-based distance and traffic-aware routing (IDTAR) protocol for smart vehicular communication," in *IEEE, 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 489–493, Valencia, Spain, June 2017.

[27] Y. R. Al-Mayouf, N. F. Abdullah, M. Ismail, S. M. Al-Qaraawi, O. A. Mahdi, and S. Khan, "Evaluation of efficient vehicular ad hoc networks based on a maximum distance routing algorithm," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016.

[28] Y. R. Al-Mayouf, M. Ismail, N. F. Abdullah et al., "Efficient and stable routing algorithm based on user mobility and node density in urban vehicular network," *PLoS One*, vol. 11, no. 11, article e0165966, 2016.

[29] M. K. Khan, M. Shiraz, K. Z. Ghafoor, S. Khan, A. S. Sadiq, and G. Ahmed, "EE-MRP: energy-efficient multistage routing protocol for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6839671, 13 pages, 2018.

[30] O. A. Mahdi, A. W. A. Wahab, M. Y. I. Idris, A. A. Znaid, Y. R. B. Al-Mayouf, and S. Khan, "WDARS: a weighted data aggregation routing strategy with minimum link cost in event-

driven WSNs," *Journal of Sensors*, vol. 2016, Article ID 3428730, 12 pages, 2016.

[31] J. Wang, G. Yu, X. Yin, F. Li, and H.-J. Kim, "An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 9472075, 9 pages, 2018.

[32] J. Wang, Y. Gao, C. Zhou, R. Simon 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.

[33] J. Wang, Y. Gao, K. Wang, A. K. Sangaiah, and S.-J. Lim, "An affinity propagation-based self-adaptive clustering method for wireless sensor networks," *Sensors*, vol. 19, no. 11, article 2579, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 6660733, 12 pages
https://doi.org/10.1155/2020/6660733

Research Article

Smart Farming: An Enhanced Pursuit of Sustainable Remote Livestock Tracking and Geofencing Using IoT and GPRS

Qazi Mudassar Ilyas[1] **and Muneer Ahmad**[2]

[1]*Department of Information Systems, College of Computer Sciences and Information Technology, King Faisal University, Saudi Arabia*
[2]*Department of Information Systems, Faculty of Computer Science & Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia*

Correspondence should be addressed to Qazi Mudassar Ilyas; qilyas@kfu.edu.sa

Received 16 October 2020; Revised 9 November 2020; Accepted 4 December 2020; Published 19 December 2020

Academic Editor: Mohammad Hossein Anisi

The farmers of agricultural farms manage and monitor different types of livestock. The manual inspection and monitoring of livestock are tedious since the cattle do not stay at fixed locations. Fencing many cattle requires a considerable cost and involves farmers' physical intervention to keep an eye to stop them from crossing beyond the access points. Visual tracking of livestock and fencing is a time-consuming and challenging job. This research proposes a smart solution for livestock tracking and geofencing using state-of-the-art IoT technology. The study creates a geographical safe zone for cattle based on IoT and GPRS, where the cattle are assigned dedicated IoT sensors. The cattle can be easily remotely monitored and controlled without having any need for farmers to intervene for livestock management physically. The smart system collects the data regarding the location, well-being, and health of the livestock. This kind of livestock management may help prevent the spread of COVID-19, lower the farming costs, and enable remote monitoring.

1. Introduction

Food is a basic need for every individual, and the importance of agricultural industry cannot be overstated. The rapidly growing population of the world reduced farming area because of industrialization, exodus of farmers to urban areas, and climate change which are some of the factors that are challenging the agricultural industry to the next level. A stable and progressive agricultural industry is extremely important to feed the ever-increasing population of the world.

The world economy was driven mostly by agriculture until the 18th century. Around 1760, the first industrial revolution started with the invention of the steam engine. The large-scale mechanization resulting from this revolution started luring farmers to abandon their farms and move to urban areas for socioeconomic benefits. The second and third industrial revolutions in the next two centuries accelerated this migration process resulting in increased abandonment of farmlands. Shengfa and Li [1] analyzed farmland abandonment in various regions of the world since the 1950s. The study argues that the phenomenon is more pronounced in more advanced regions of the world, and this trend is expected to continue in the future too.

Today, we are at the cusp of the fourth industrial revolution which is driven by several disruptive technologies including but not limited to sophisticated machine learning algorithms, Artificial Intelligence (AI), Internet of Things (IoT), Unmanned Aerial Vehicles (UAVs), robotics, and quantum computing [2, 3]. Such powerful technologies have already changed our lives dramatically. They have found application in almost every domain of life, and agriculture is no exception. Artificial Intelligence and machine learning technologies have been applied in analyzing and managing soil [4], crops [5], livestock [6], and water resources [7]. Computer vision techniques have been employed for addressing several issues in agriculture such as plant disease detection [8], insect detection [9], farmland management

[10], and crop yield analysis [11]. References [12, 13] provide excellent reviews of the application of IoT technologies in agriculture. Precision farming is another technique in agriculture that has seen significant boost because of invention of cheap sensors and UAVs [14].

A satellite navigation device (commonly called a GPS receiver) can be used to determine its position using a satellite navigation system. As of today, there are four active satellite navigation systems that provide global coverage, namely, Global Positioning System (GPS) by the United States, Galileo by Europe, GLObal NAvigation Satellite System (GLONASS) by Russia, and BeiDou by China. A satellite navigation device may be attached to any object to track and monitor its position in real time. A geofence can be established by defining a closed polygon referring to a geographic area on earth. A location-aware device can then make use of this geofence to trigger alerts when the object enters or leaves the area defined by the geofence.

Internet of Things (IoT) is an extremely exciting set of technologies that is already shaping the future of humankind. IoT is based on the concept of uniquely identifiable interconnected devices (such as sensors, computers, and mechanical devices), collecting the data, and storing it in the cloud that is processed by intelligent algorithms to achieve common goals. IoT has several applications in almost all domains of life. References [15–17] provide excellent reviews of some such applications.

Livestock monitoring is another important aspect of farming. Traditionally, cattle were monitored manually and confined in farms by building physical fences. However, advanced technologies have made it possible to track and monitor the cattle automatically. Navigation satellites and Global Positioning System (GPS) are extensively used for tracking the position of cattle. UAVs have made real-time monitoring of cattle a cost-effective and hassle-free task. Radio-frequency identification (RFID), wireless sensor networks, and the Low Power Wide Area Network (LPWAN) are other potential technologies for establishing virtual fences to keep the farm animals in a confined area.

2. Related Work

Smart farming concept relates to location-aware devices to monitor the movement of animals and raise alerts when they violate the boundary of the geofence of the farm or pasture. Additionally, IoT sensors may be used to monitor the health and well-being of farm animals. References [18, 19] proposed a device that is based on the satellite navigation system to track the position of an object to which it is attached. The position is transmitted through an available wireless transmission medium such as a radio frequency, wireless, or cellular network. References [20–22] described a geofencing scheme based on the geographic area being divided into one or more grids. The proposed scheme exploits this grid structure to optimize computational resources required for location monitoring because a complex polygon requires more calculations to achieve the desired goal.

Figure 1 presents the number of research articles published from 2010 to 2020 on remote livestock tracking. It

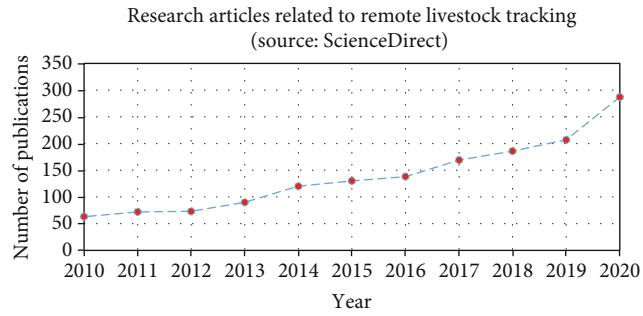


Figure 1: Research trend on remote livestock tracking.

can be seen that there is significant scope of remote monitoring of livestock employing the latest technologies. With the passage of time, the researchers have devised enhanced solutions in this problem domain. Despite a good number of cited works, still, the need to address the issues related to optimized geofencing is on the way.

References [23–25] also proposed the concept of geofencing by providing auditory feedback and light electric shocks to an animal wearing a tracking device. Through these feedback mechanisms, the device can effectively guide an animal from one to another location. It may also be used to keep an animal within a geofence defined by its owner.

References [26, 27] proposed an implant device for tracking the location of an animal in addition to monitoring its health and well-being. The device also contains enough storage to record medical information about an animal such as past surgeries, any disease, medication, and owner contact information.

Safeguarding against theft of animals is an obvious application of the ability to monitor the location of an animal in real time. To this end, researchers have proposed several systems that can be used to minimize the chances of animal theft. Reference [28] proposed such a system based on a centralized database to store livestock details, radio-frequency identification (RFID) tags, and an available communication technology such as a cellular network. The proposed system employs several heuristics to identify theft of animals. One such heuristics marks an animal as "stolen" if the animal is found in a geographic location that is considered "far" from the registered location of the animal and its geofence for gazing. Another heuristic is discovery of an animal having registered a location drastically different from that of other animals in a herd. Reference [29] proposed a similar system that uses wireless sensor networks and unmanned aerial vehicles for animal tracking and identification.

Several researchers have proposed IoT-based frameworks for geofencing as well as other aspects of smart farming. References [30–35] proposed various systems for monitoring of livestock through data recorded using sensors and network communication nodes. References [36–38], in addition to collecting the data, also proposed means for automatically analyzing the data and an interface to monitor the livestock. Reference [39] proposes the use of a long-range Low Power Wide Area Network (LPWAN) technology to collect and process several parameters related to the health of cattle as

well as their environment. The data recorded by sensors is relayed to one or more gateways through long-range end devices. This data is subsequently processed by an application server and presented to the user for visualization and analytics. A C++ simulation shows that the proposed architecture can effectively cover an area of $7\,km^2$ in a rather harsher hilly terrain. Reference [40] proposed a similar solution based on Long-range Wireless Access Network (LoRaWAN) technology. The use of LoRaWAN makes such solutions more feasible in rural areas with poor cellular or Internet coverage.

Many different satellite navigations have been considerably employed to determine the position of livestock. We can observe four active satellite navigation systems providing global coverage, namely, the Global Positioning System (GPS) by the United States, Galileo by Europe, GLObal NAvigation Satellite System (GLONASS) by Russia, and BeiDou by China. It is very contemporary to navigate through satellite nowadays. Commonly, a satellite navigation device is attached to the livestock under monitoring, and the device can promptly track and monitor the position of the livestock in a real-time scenario. In addition to this navigation, a dedicated geofence can greatly help in defining a closed shape, normally a polygon, that refers to a geographic area on earth. A location-aware device can then make use of this geofence to trigger alerts when the object enters or leaves the area defined by the geofence. A number of cited works highlight that geofencing-related solutions are more appealing for remote livestock monitoring.

Table 1 presents a brief summary and comparison of the research works discussed above. We can see many different solutions of remote livestock tracking. We can see that safeguarding the livestock appears as an obvious application of the ability to monitor the location of an animal in real time. Until now, the researchers have proposed several systems that can be used to minimize the chances of animal theft based on RFID tags, wireless sensor networks, unmanned aerial vehicles for animal tracking and identification, long-range Low Power Wide Area Network (LPWAN) technologies, GPS, IoT, and GPRS. Despite a good number of cited works, only a few researches can be identified that focus on the geofencing for livestock monitoring.

3. Methodology

Varieties of cattle in a paddock have genetically different grazing, sleeping, and playing patterns. Goats and sheep are more active, and they have different food intake and digestion systems than cows and buffalos. The current livestock management systems mostly employ IoT and GPS sensors connected to satellite and GPRS for navigation and communication, respectively. GPS and GPRS sensors consume device energy and communication bandwidth. Besides, the same set of sensors is installed for all livestock categories despite genetic diversity among animals in the same herd. In addition, the conventional tracking systems track the movements of livestock without any profound geographical boundaries that become challenging in case the animals go very far from the main access points. To address these issues, this research study proposes an enhanced management system that provides convenience to farmers to define a geographical safe zone for livestock. The farmers are notified by the system when cattle try to go beyond the defined boundary of the zone. Besides, the navigation and communication are automatically controlled according to the genetic diversity of different animals.

Figure 2 presents an overview of the conceptual framework of the proposed system. The red ellipse represents a drawn geographical safe zone for the livestock. Ultrasonic sensors installed at the elliptical boundary of the safe zone identify the movements of the cattle. The ultrasonic sound waves propagate and discover the presence of livestock, and its distance is calculated. If the distance of the cattle crosses the defined safe-distance threshold, the communication navigator is activated. The animals in the herd are equipped with navigation sensors that sense the locations of animals by navigating through the satellite. The system calculates the distance of each animal from the safe zone geographical boundary and alarms the farmer when the distance of the animal gets close to a threshold value. The proposed system glimpses the exact location of animals in case the animals are out of the safe zone for a specified period. The motion sensor suspends the navigation and communication when the animal is recorded in a static state to optimize the energy and communication bandwidth for significant utilization.

Figure 3 presents the implementation scenario of the proposed framework described in Figure 2. We can see different steps involved in tracking a whole herd or particular cattle in the herd based on sensing the location coordinates and having communication with the system through the communication channel. An elaborative presentation is presented in the following algorithm:

(1) A herd **H** may contain a variety of livestock with any number depending on the nature of business and local facilities available to farmers for keeping and managing animals. For the moment, we suppose that a typical herd **H** can contain a maximum of **N** animals defined as $\mathbf{H} = \{H_1, H_2, H_3, H_i\}$, $i \leq N$

(2) Since the currently proposed system tracks the locations of animals contained in herd **H**, we need a navigation sensor to connect individual cattle with the communication satellite to receive the location coordinates of each animal. Besides, since the system needs to analyze the latest distance of animals from the defined safe zone, we need to pass these location coordinates to the system through the communication channel. For this, we need to equip H_i animals with N_i and C_i navigation and communication sensors, where $i < N$

(3) Since primarily the system is supposed to keep the animals in a defined safe zone, in order to maximize their security, grazing, and leisure period, we define a safe zone **S** with j geographical coordination points such that $\mathbf{S} = \{S_1, S_2, S_3, \cdots, S_j\}$

Table 1: An overview of different solutions proposed for livestock tracking and management.

Study	Weaknesses identified	Proposed solutions	State-of-the-art technology adoption				Research outcomes
			GPS	GPRS	IoT	Others	
[41]	The conventional livestock tracking and management for a large herd is challenging for farmers in remote areas.	Spatial and temporal interaction of traditionally herded livestock and wildlife using GPS and GSM technologies in Northern Kenya	✓	✓	✓	Radio frequency	This study demonstrated the feasibility of tracking cattle using radio collars. It shows the complexity of spatial use for cattle and wildlife.
[42]	The tracking system available on commercial basis lacks the data storage capacity required for frequent collection of livestock data.	The study designed the Clark GPS Animal Tracking System to satisfy the needs of stakeholders attached to livestock management.	✓	✓	✓	Extended data storage	An evolving demand of ecological research requires adoption of the latest technology for tracking and managing the livestock.
[43]	The cost of GPS technology is a big barrier for efficient livestock tracking and management.	The study reviewed the GPS-based technologies being used for cattle management and suggested the improvements.	✓			Review of literature involving GPS technology	Ecologists have been employing best efforts towards livestock management using the latest technologies.
[44]	The animal behavior and grazing patterns are a very important measure for livestock health, tracking, and management.	The GPS system is used for tracking cows in six summer grazing areas having different environmental conditions and livestock managements.	✓		✓		The GPS positions of animals help to collect information relating to their grazing, resting, and playing patterns.
[37]	The traditional livestock management methods are tedious that involve human intervention and other resources.	The integrated system is comprised of tags, beacons, and base station nodes. Tag nodes communicate with other nodes to transmit location information of livestock.	✓	✓	✓	Beacon, tags	The android-based application outperformed the existing conventional tracking and cattle management systems.
[45]	The welfare breading of individual cattle is an important aspect especially in case of a large herd. Traditional methods of welfare breading are insufficient for managing a large number of cattle.	A remote monitoring system based on computer vision and wireless technologies was developed for remote monitoring of pigs in addition to measuring other parameters, i.e., humidity, temperature, and harmful gases.			✓	Computer vision	The study improved the traditional CAMShift algorithm for an enhanced tracking of pigs using computer vision and WSN for remote sensing technologies.
[46]	The livestock are under high threat of landslides, earthquakes, and other natural disasters. Remote sensing of such natural disasters is viable to save the precious cattle.	The study integrated Landsat-8 and phased array type L-band synthetic aperture radar-2 (PALSAR-2) datasets and adopted the analytical hierarchy process (AHP) method in mapping landslides in the Kelantan river basin, Peninsular Malaysia.		✓		Landsat-8 and PALSAR-2	The study demonstrated that employment of Landsat-8 and PALSAR-2 tools for remote sensing data along with GIS techniques were promising tools to map landslide assessment for tropical environments.
[47]	The conventional livestock monitoring parameters are insignificant for sustainable ranching of sheep.	The study integrated the global information system with remote sensing to analyze environmental variables to monitor and track sheep grazing.	✓	✓		Remote sensing tool	The GIS-based time analysis tool helped in collecting point data from GPS collars installed on sheep that enabled the remote sensing significantly.
[36]	The precision livestock farming in agriculture and food industries requires sustained	BOSCA and CyberBar were developed for real-time product visibility to ensure		✓	✓	Cloud computing	The tracking of agrifood products with remote environmental monitoring

Table 1: Continued.

Study	Weaknesses identified	Proposed solutions	State-of-the-art technology adoption				Research outcomes
			GPS	GPRS	IoT	Others	
	production that is not possible by employing a traditional system.	integrity and quality. The proposed system helped stakeholders for better decision-making.					significantly assisted in timely decision-making to all stakeholders.
[48]	The monitoring of livestock for safety, security, grazing, and health aspects is always very challenging due to animal nature and habits of livestock.	An open sourced framework that could capture the health parameters of livestock was developed. A wireless location acoustic sensing system was utilized to intake the health parameters of cattle.	✓	✓		Open sourced framework	The system provides high quality support to farmers in rural areas because of low cost and portability. The system outperformed the existing systems for monitoring the livestock.
[49]	The lamb industry in Victoria, which is a big industry, requires tracking and other visualization parameters of livestock for better herd management.	A comprehensive designing of spatial-temporal location movements of livestock with respect to the environmental parameters greatly helps in designing the paddocks to improve the management and performance of cattle welfare.	✓			Remote visualization systems	The GIS-based system significantly helps the farmers to seek frequent information about nitrogen emission to meet the local and global greenhouse gas targets. This helps in improving life quality of livestock ultimately.
[50]	Wild stocking animals are often lost due to a poor and old style cattle management system that increases the farming cost considerably.	The study combined GPS and wireless mobile cell network as positioning technologies to monitor the wild animals. The system is based on mobile beacons and network base stations.	✓	✓	✓	Wireless mobile cell networks	The traditional GPS, compass, and Wi-Fi technologies are useful for common tracking and management of livestock, but these technologies are still inadequate for monitoring of wild animals.
[51]	Conventional livestock management for a large herd is challenging for farmers in remote areas.	An RFID-based system is used to track and monitor the livestock for their identity and vaccination procedures.				RFID	The analysis concentrated on the colossal measure of RFID for better query processing.

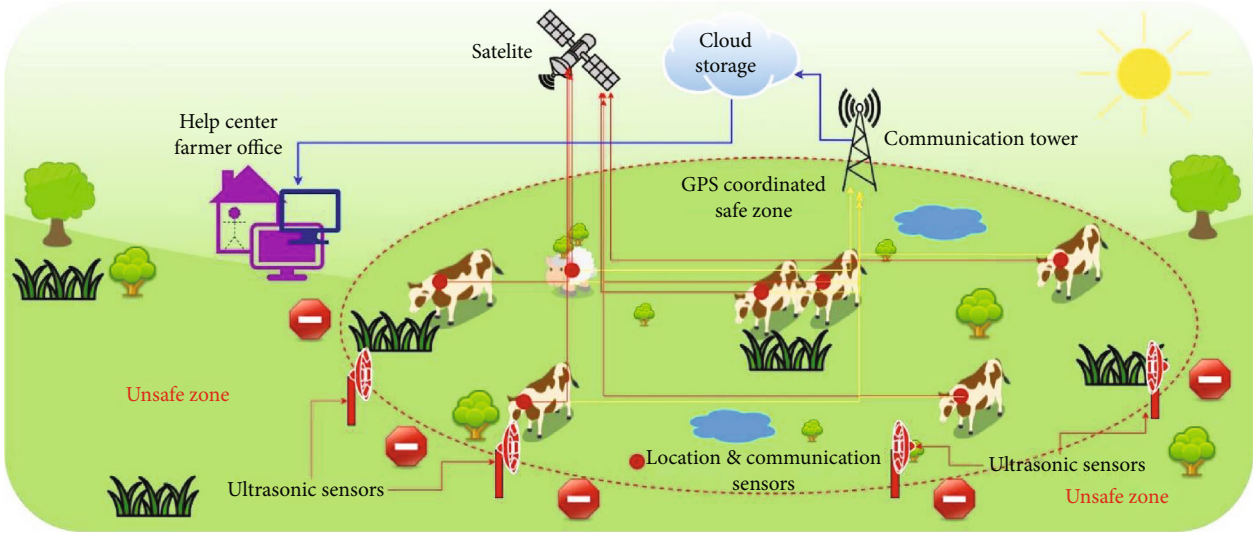


Figure 2: The conceptual framework of the proposed system.

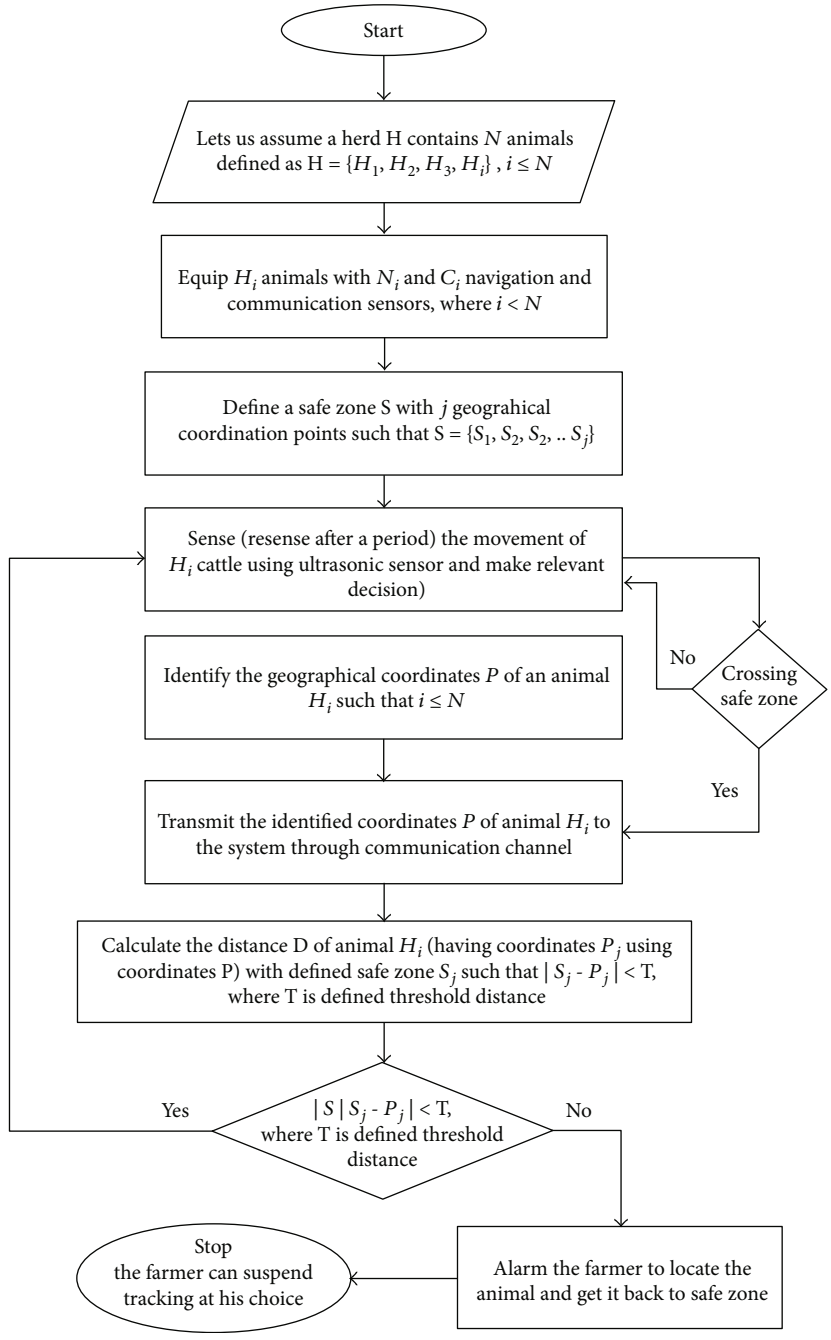


Figure 3: Implementation scenario of a conceptual framework.

(4) The ultrasonic sensors sense the cattle H_i crossing the safe zone **S** and alarms/alerts the GPS sensor to initiate detecting the location of H_i

(5) The geographical sensors installed calculate the geographic coordinates of livestock by communicating with the satellite. We identify the geographical coordinates **P** of an animal H_i such that $i \leq N$

(6) For the sake of comprehensive and timely analysis of calculated coordination points, we need to transmit the identified coordinates **P** of animal H_i to the system through the communication channel

(7) Next, we calculate the distance **D** of animal H_i (having coordinates P_j using coordinates **P)** with defined safe zone S_j such that $|S_j - P_j| <$**T**, where **T** is the defined threshold distance

(8) Now, we evaluate that the distance **D** < **T**, where **T** is defined as the threshold distance for animals to keep safe in the safe zone

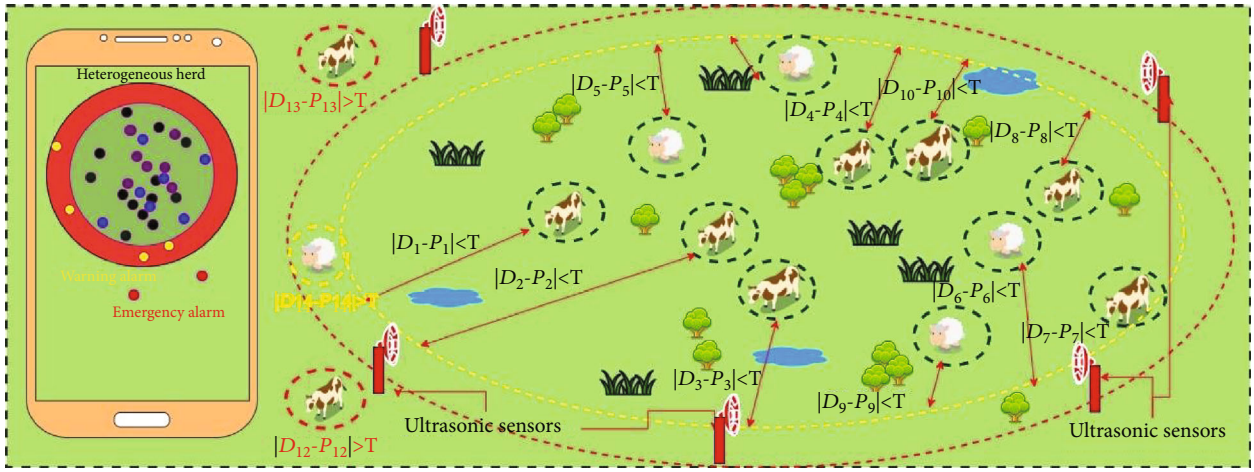


Figure 4: Monitoring of livestock by a geographical paddock in a heterogeneous herd.

Let us create the instances Δt_1, Δt_2, Δt_3, Δt_4,…, Δt_n, where $n \leq t$ (forward and backpropagation of waves from the ultrasonic sensor).

The time t can be defined by

$$f(t) = \begin{cases} t_1, & t < T, \\ t_2, & t \geq T \text{ where } T \text{ is the defined threshold} \end{cases} \quad (1)$$

(9) In this case, if the result is t_1 (calculated **D** is not less than threshold **T**), then the system alarms the farmer to locate the animal and get it back to the safe zone

(10) Otherwise, the system resenses the current location of livestock after a defined period **T**

Contrary to the conventional livestock tracking system where the farmers have to sometimes do a physical exertion for tracing the cattle that go beyond the common access points, the defined safe zone provides convenience with secure and prompt management. Besides, in the case of lazy livestock that do not change their locations more frequently, the system can significantly save power consumption and communication channel utilization. In the scenario where the farmers require a more substantial safe zone, there is very low probability for livestock to go beyond safe zones by exceeding the location threshold. The proposed system is equally suitable for the sensing devices equipped with a solar panel for power supply.

4. Experimentation

This study designed a geographical paddock to monitor the spatial, temporal behaviors of livestock. It is a contemporary phenomenon that different livestock animals have different patterns of grazing, movement, and resting events. The frequency of these events varies from animals to animals, and mostly, they are related to the genetics of individuals and the current psychological states in rare cases. The tracking phenomenon of livestock is tremendously attributed to such spatial, temporal events reflecting the allocation of appropriate sleep time units to hardware for saving energy and communication bandwidth.

Figure 4 describes the visual implementation of monitoring of livestock using a defined geographical paddock through remote sensing of their spatial, temporal activities. The yellow and red circles represent warning and alarming zones for the farmers or caretakers. The ultrasonic sensors sense the cattle H_i crossing the safe zone **S** and initiate the GPS to trace the location of H_i. The application calculates the geographical distance of each animal H_i, $i \leq N$ for an N number of animals in the herd **H**. The application also defines a distance threshold **T** that helps to match the current difference of H_i with the safe zone $\mathbf{S} = \{S_1, S_2, S_3, \cdots, S_j\}$. When H_i (having geographical coordinates P_j) approaches S_j, the absolute geographical distance $|P_j - S_j|$ is calculated and compared with threshold **T**. The application warns the farmer in case the difference of two geographical distances seeks the threshold value. Similarly, the farmer gets a warning notification when the said difference of distances exceeds the defined threshold **T**. The warning alarm helps the farmer to get an alert on the current location of H_i, and the farmer may carefully observe the movement of H_i.

Figure 5 describes the detection and tracking phases of the proposed system. The ultrasonic sensors connected to Arduino sense the presence of cattle. The ultrasonic sensor data is sensed through ThingSpeak. At first, the Arduino initializes the ultrasonic sensor to collect its data. The distance of the cattle is calculated, and the Arduino forwards the distance to the ESP8266 module using serial communication. We need to choose a threshold value for the sensors to sense the obstacle and to calculate the distance. Based on the calculated distance, the serial communication helps ESP8266 to convey the distance to ThingSpeak using a communication channel.

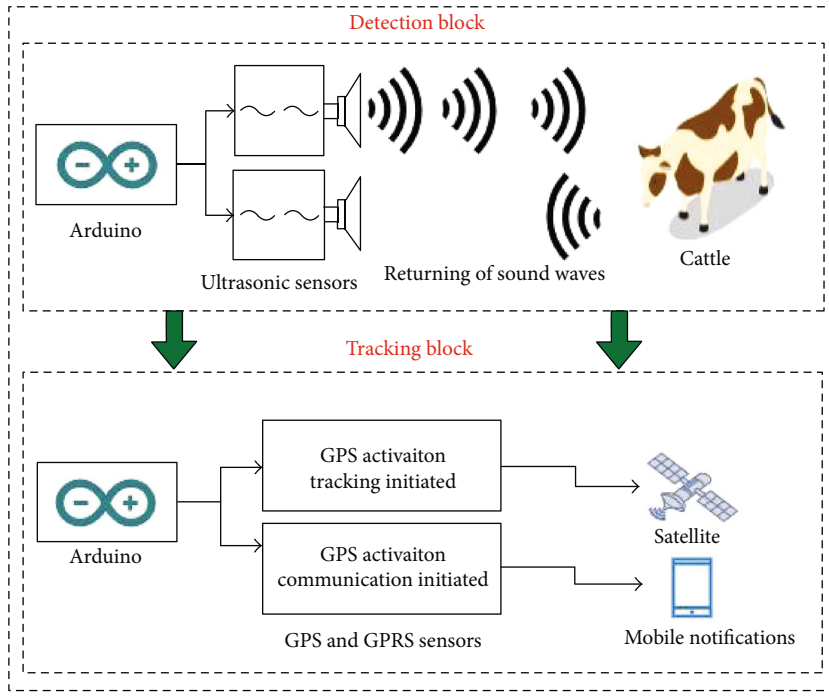


Figure 5: Detection and tracking blocks.

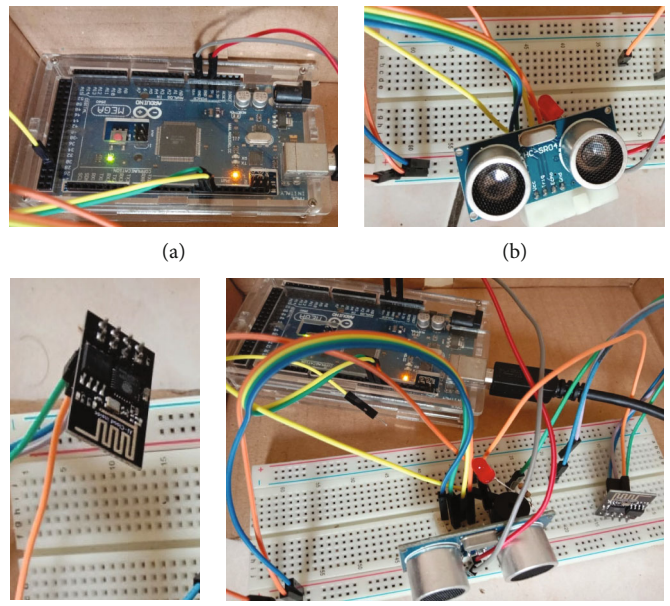
The distance of cattle from the sensors is measured using the following formula: Measured Distance = 1/2 Time × Speed; here, the level of humidity and temperature impacts the speed of the sound wave at a particular place.

Once the system detects the cattle going beyond the safe zone, the GPS module attached to Arduino is activated, and the module connects to nearby satellites and starts reading the current location coordinates of the cattle. These location coordinates are passed to the farmer's mobile through the communication channel attached to Arduino.

Figure 6 presents an overview of hardware components employed for livestock detection and tracking. Figure 6(a) depicts an Arduino Uno microcontroller board manufactured using Microchip ATmega328P. This microcontroller board has been used to connect the detection and navigation components. Figure 6(b) shows an ultrasonic HC-SR04 sensor that uses SONAR to detect the livestock and determines its distance from the safe zone. Figure 6(c) is the NRF24L01 wireless module used as a communication module in the setup. Figure 6(d) presents the hardware modules installed for livestock identification and tracking. In the experimentation, the ultrasonic module detects the presence of livestock by calculating and comparing its distance with the safe zone threshold defined in the system. The navigation module is activated to track the current location of the animal once the system identifies the cattle breaching the designated safe zone. The navigation coordinates are communicated to the network using a communication channel. The farmer can locate the current location of those cattle that are outside the safe zone. This provides automated fenceless farming to farmers by remotely identifying the current situation of all cattle in a heterogeneous herd.

Table 2 presents an experimental scenario in which a variety of livestock H_i at different instances of time access the safe zone boundary. The ultrasound sensors sense their distance comparing it with the defined threshold T. The farmers receive the mobile notifications for the cattle that cross the safe zone, and their location coordinates assist the farmers in tracking the current location of livestock. The sound waves emitted by sensors are mechanical in nature and travel using a medium. It has been noticed that certain environmental conditions, i.e., humidity and temperature, may also impact the performance of such sensors. Since these sensors are being employed for the detection of livestock pertaining to a distance threshold compared with the distance of cattle from the safe zone boundary, the environmental conditions do not impact the sensor performance degradation for the detection of cattle. The significant outcome of this detection mechanism is to detect the livestock without impacting any physical damage to them in the course of longitudinal compression of waves having a frequency of 100 kHz to 50 MHz.

The significance of the proposed livestock management system reflects in reducing the time and energy complexity of the system and integrated modules, as shown in Figure 7. A variety of livestock have different grazing patterns genetically, and thus, their physical activities are proportional to these genetic behaviors [52–54]. It has been observed that certain livestock, e.g., cows and buffaloes, are sluggish as compared to goats and sheep. Once the cattle have gone through the graze period, the physical excitement to intake more food is considerably reduced, and the livestock prefers to have rest or lay down for a long time. During this inactive or passive activity, the safe zone ensures that the cattle



(a) (b) (c) (d)

Figure 6: Simulation setup of detection and navigation equipment.

Table 2: Livestock detection and tracking activities.

Instance	Cattle	Safe zone threshold (T)	Distance (D) from the boundary	Tracker activated	Mobile notifications
T_1	H_1	5 meters	15 meters	×	×
T_2	H_2	5 meters	3 meters	✓	✓
T_3	H_1	5 meters	5 meters	✓	✓
T_4	H_3	5 meters	1 meter	✓	✓
T_5	H_4	5 meters	7 meters	×	×
T_6	H_8	5 meters	2 meters	✓	✓
T_7	H_7	5 meters	4 meters	✓	✓
T_8	H_6	5 meters	9 meters	×	×
T_9	H_6	5 meters	3 meters	✓	✓
T_{10}	H_8	5 meters	4 meters	✓	✓

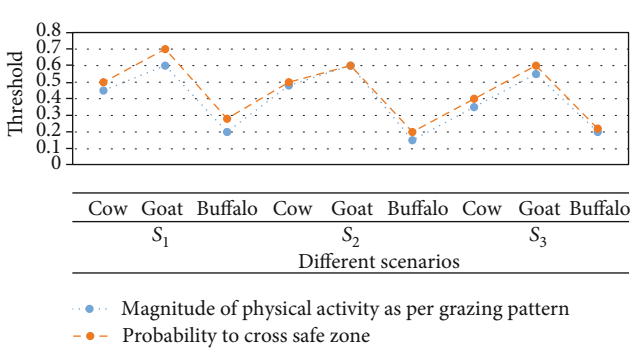


Figure 7: Physical activity and chances of crossing safe zone based on their grazing patterns.

remain inside the geographical boundary and do not step out. This way, the tracking system is not required to be initiated to record the locations of cattle since the cattle are already inside the safe zone parameters. Goats and sheep, on the other hand, are more physically active even after the intake of proper food. The probability of such livestock is higher than the sluggish or passive livestock. The tracking of active livestock might be frequently required in this case as compared to cows and buffaloes. The proposed system signifies the avoidance of unnecessary utilization of resources, i.e., time, energy, and effort. All three scenarios defined in Figure 6 describe that the probability for goat or sheep to cross the geographical threshold is much higher than cows followed by buffalos.

Figure 8 presents the average probability values for tracking different livestock under certain restrictions, i.e., considering different scenarios of the safe zone and without a safe zone. The red line in the legend describes a uniform

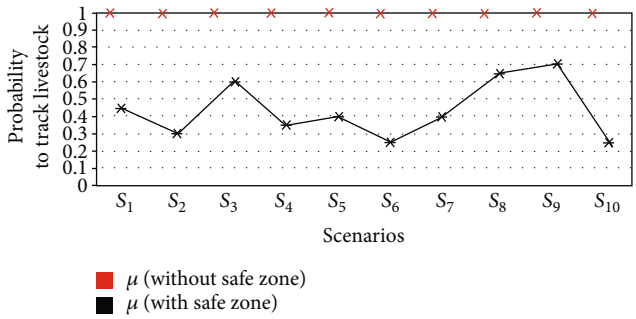


Figure 8: Probabilities of tracking different livestock in different scenarios.

probability of tracking all livestock without having consideration of a safe zone. The existing approaches, cited in the literature, track the livestock right away, syncing with the initialization of the tracking system. This phenomenon assigns the same probability values to all livestock irrespective of their genetic type and behaviors and thus results in wastage of system resources. On the contrary, the concept of adopting a geographical safe zone ensures that genetic behaviors and activities of different livestock are taken care of while tracking their movements. In all scenarios described in Figure 6, it can be observed that the average tracking probability of different livestock is quite different from each other. For instance, the average likelihood of tracking goats or sheep is higher than that of cows and buffaloes. Besides, the cows have a somehow higher average tracking probability as compared to the average likelihood of buffaloes. Hence, the proposed livestock detection and tracking system are significantly context-aware (as per genetics of different livestock).

5. Conclusion

This study proposes design of a geographical paddock to monitor spatiotemporal behaviors of livestock. In a conventional livestock tracking system, the farmers have to do physical exertion for tracing the cattle that go beyond the common access points. The proposed solution addresses these issues by providing convenience to farmers to define a geographical safe zone for livestock. The farmers are notified by the system when cattle try to go beyond the defined boundary of the zone. Besides, the navigation and communication are automatically controlled according to the genetic diversity of different animals. The system calculates the distance of each animal from the safe zone geographical boundary and alarms the farmer when the distance of the animal gets close to a threshold value. The proposed system glimpses the exact location of animals in case the animals are out of the safe zone for a specified period. The motion sensor suspends the navigation and communication when the animal is recorded in a static state to optimize the energy and communication bandwidth for significant utilization. The significance of the proposed livestock management system is reflected in reducing the time and energy complexity of the system and integrated modules.

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

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number IFT20172.

References

[1] S. Li and X. Li, "Global understanding of farmland abandonment: a review and prospects," *Journal of Geographical Sciences*, vol. 27, no. 9, pp. 1123–1150, 2017.

[2] A. Nikitas, K. Michalakopoulou, E. T. Njoya, and D. Karampatzakis, "Artificial intelligence, transport and the smart city: definitions and dimensions of a new mobility era," *Sustainability*, vol. 12, no. 7, p. 2789, 2020.

[3] H. Zahmatkesh and F. Al-Turjman, "Fog computing for sustainable smart cities in the IoT era: caching techniques and enabling technologies - an overview," *Sustainable Cities and Society*, vol. 59, p. 102139, 2020.

[4] A. A. Javadi and M. Rezania, "Applications of artificial intelligence and data mining techniques in soil modeling," *Geomechanics and Engineering*, vol. 1, no. 1, pp. 53–74, 2009.

[5] N. Kim, K. J. Ha, N. W. Park, J. Cho, S. Hong, and Y. W. Lee, "A comparison between major artificial intelligence models for crop yield prediction: case study of the Midwestern United States, 2006–2015," *ISPRS International Journal of Geo-Information*, vol. 8, no. 5, p. 240, 2019.

[6] R. S. Alonso, I. Sittón-Candanedo, Ó. García, J. Prieto, and S. Rodríguez-González, "An intelligent edge-IoT platform for monitoring livestock and crops in a dairy farming scenario," *Ad Hoc Networks*, vol. 98, p. 102047, 2020.

[7] H. Afzaal, A. A. Farooque, F. Abbas, B. Acharya, and T. Esau, "Computation of evapotranspiration with artificial intelligence for precision water resource management," *Applied Sciences*, vol. 10, no. 5, p. 1621, 2020.

[8] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, 2016.

[9] N. Larios, H. Deng, W. Zhang et al., "Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects," *Machine Vision and Applications*, vol. 19, no. 2, pp. 105–123, 2008.

[10] M. O. Adebiyi, R. O. Ogundokun, and A. A. Abokhai, "Machine learning-based predictive farmland optimization and crop monitoring system," *Scientifica*, vol. 2020, 12 pages, 2020.

[11] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: a systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.

[12] P. Asghari, A. M. Rahmani, and H. H. S. Javadi, "Internet of things applications: a systematic review," *Computer Networks*, vol. 148, pp. 241–261, 2019.

[13] J. M. Talavera, L. E. Tobón, J. A. Gómez et al., "Review of IoT applications in agro-industrial and environmental fields," *Computers and Electronics in Agriculture*, vol. 142, pp. 283–297, 2017.

[14] R. Dolci, "IoT solutions for precision farming and food manufacturing: artificial intelligence applications in digital food," in *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, Turin, Italy, 2017.

[15] L. Atzori, A. Iera, and G. Morabito, "The internet of things: a survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[16] S. Li, L. XuDa, and S. Zhao, "The internet of things: a survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.

[17] L. XuDa, W. He, and S. Li, "Internet of things in industries: a survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.

[18] A. Patrik, G. Utama, A. A. S. Gunawan et al., "GNSS-based navigation systems of autonomous drone for delivering items," *Journal of Big Data*, vol. 6, no. 1, 2019.

[19] W. Ruan, Q. Z. Sheng, L. Yao, T. Gu, M. Ruta, and L. Shangguan, "Device-free indoor localization and tracking through human-object interactions," in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Coimbra, Portugal, 2016.

[20] I. Daugela, J. Sužiedelyte Visockiene, and V. Česlovas Aksamitauskas, "Erratum to: RPAS and GIS for landfill analysis," *E3S Web of Conferences*, vol. 44, article 00203, 2018.

[21] M. Wang, X. Liu, Y. Zhang, and Z. Wang, "Camera coverage estimation based on multistage grid subdivision," *ISPRS International Journal of Geo-Information*, vol. 6, no. 4, p. 110, 2017.

[22] G. Q. Tao, X. K. Ou, Y. M. Guo et al., "Priority area identification for vegetation in Northwest Yunnan, based on protection value and protection cost," *Acta Ecologica Sinica*, vol. 36, no. 18, 2016.

[23] I. Halachmi, A. S. Tello, A. P. Fernández et al., "6.4. Discussion: PLF for automatic detection of animal health in cows," in *Precision livestock farming applications*, 2015.

[24] A. Spink, B. Cresswell, A. Kölzsch et al., "Animal behaviour analysis with GPS and 3D accelerometers," in *Precision Livestock Farming 2013 - Papers Presented at the 6th European Conference on Precision Livestock Farming, ECPLF 2013*, Leuven, Belgium, 2013.

[25] J. Wall, G. Wittemyer, B. Klinkenberg, and I. Douglas-Hamilton, "Novel opportunities for wildlife conservation and research with real-time monitoring," *Ecological Applications*, vol. 24, no. 4, pp. 593–601, 2014.

[26] M. Benjamin and S. Yik, "Precision livestock farming in swine welfare: a review for swine practitioners," *Animals*, vol. 9, no. 4, p. 133, 2019.

[27] S. Neethirajan, S. K. Tuteja, S. T. Huang, and D. Kelton, "Recent advancement in biosensors technology for animal and livestock health management," *Biosensors and Bioelectronics*, vol. 98, pp. 398–407, 2017.

[28] J. K. Siror, S. Huanye, W. Dong, and W. Jie, "Use of RFID technologies to combat cattle rustling in the East Africa," in *2009 Fifth International Joint Conference on INC, IMS and IDC*, Seoul, South Korea, 2009.

[29] P. Wamuyu, "A conceptual framework for implementing a WSN based cattle recovery system in case of cattle rustling in Kenya," *Technologies*, vol. 5, no. 3, p. 54, 2017.

[30] H. Bouazza, O. Zerzouri, M. Bouya, A. Charoub, and A. Hadjoudja, "A novel RFID system for monitoring livestock health state," in *2017 International Conference on Engineering and Technology (ICET)*, Antalya, Turkey, 2018.

[31] A. Carabús, M. Gispert, and M. Font-i-Furnols, "Imaging technologies to study the composition of live pigs: a review," *Spanish Journal of Agricultural Research*, vol. 14, no. 3, 2016.

[32] I. Kröger, E. Humer, V. Neubauer, N. Kraft, P. Ertl, and Q. Zebeli, "Validation of a noseband sensor system for monitoring ruminating activity in cows under different feeding regimens," *Livestock Science*, vol. 193, pp. 118–122, 2016.

[33] G. Mattachini, E. Riva, F. Perazzolo, E. Naldi, and G. Provolo, "Monitoring feeding behaviour of dairy cows using accelerometers," *Journal of Agricultural Engineering*, vol. 47, no. 1, p. 54, 2016.

[34] A. Peña Fernández, T. Norton, E. Tullo et al., "Real-time monitoring of broiler flock's welfare status using camera-based technology," *Biosystems Engineering*, vol. 173, pp. 103–114, 2018.

[35] B. Xu, W. Wang, G. Falzon et al., "Livestock classification and counting in quadcopter aerial images using mask R-CNN," *International Journal of Remote Sensing*, vol. 41, no. 21, pp. 8121–8142, 2020.

[36] U. McCarthy, L. Brennan, S. Ward, and G. Corkery, "Enhanced efficiencies in the poultry industry via real-time monitoring and cloud-enabled tracking," in *Precision Livestock Farming 2013 - Papers Presented at the 6th European Conference on Precision Livestock Farming, ECPLF 2013*, pp. 212–222, Leuven, Belgium, 2013.

[37] N. A. Molapo, R. Malekian, and L. Nair, "Real-time livestock tracking system with integration of sensors and beacon navigation," *Wireless Personal Communications*, vol. 104, no. 2, pp. 853–879, 2019.

[38] K. E. Veblen, D. A. Pyke, C. L. Aldridge, M. L. Casazza, T. J. Assal, and M. A. Farinha, "Range-wide assessment of livestock grazing across the sagebrush biome," U.S. Geological Survey Open-File Report 2011-1263, 2011.

[39] L. Germani, V. Mecarelli, G. Baruffa, L. Rugini, and F. Frescura, "An IoT architecture for continuous livestock monitoring using lora LPWAN," *Electronics*, vol. 8, no. 12, p. 1435, 2019.

[40] U. S. Abdullahi, M. Nyabam, K. Orisekeh et al., "Exploiting IoT and LoRaWAN technologies for effective livestock monitoring in Nigeria," *Arid Zone Journal of Engineering, Technology and Environment*, vol. 15, pp. 146–159, 2019.

[41] E. A. Raizman, H. B. Rasmussen, L. E. King, F. W. Ihwagi, and I. Douglas-Hamilton, "Feasibility study on the spatial and temporal movement of Samburu's cattle and wildlife in Kenya using GPS radio-tracking, remote sensing and GIS," *Preventive Veterinary Medicine*, vol. 111, no. 1-2, pp. 76–80, 2013.

[42] P. E. Clark, D. E. Johnson, M. A. Kniep et al., "An advanced, low-cost, GPS-based animal tracking system," *Rangeland Ecology & Management*, vol. 59, no. 3, pp. 334–340, 2006.

[43] D. L. Swain, M. A. Friend, G. J. Bishop-Hurley, R. N. Handcock, and T. Wark, "Tracking livestock using global positioning systems are we still lost?," *Animal Production Science*, vol. 51, no. 3, p. 167, 2011.

[44] H. Homburger, A. Lüscher, M. Scherer-Lorenzen, and M. K. Schneider, "Patterns of livestock activity on heterogeneous subalpine pastures reveal distinct responses to spatial autocorrelation, environment and management," *Movement Ecology*, vol. 3, no. 1, 2015.

[45] Y. Duan, L. Ma, and G. Liu, "Remote monitoring system of pig motion behavior and piggery environment based on internet of things," *Transactions of the Chinese Society of Agriculture Engineering*, vol. 31, pp. 216–221, 2015.

[46] M. Hashim, S. Misbari, and A. B. Pour, "Landslide mapping and assessment by integrating Landsat-8, PALSAR-2 and GIS techniques: a case study from Kelantan State, Peninsular Malaysia," *Journal of the Indian Society of Remote Sensing*, vol. 46, no. 2, pp. 233–248, 2018.

[47] R. R. Miller, *Utilizing GIS and remote sensing to determine sheep grazing patterns for best practices in land management protocols, [Ph.D. thesis]*, ProQuest Dissertations and Theses, 2012.

[48] H. Q. T. Ngo, T. P. Nguyen, and H. Nguyen, "Research on a low-cost, open-source, and remote monitoring data collector to predict livestock's habits based on location and auditory information: a case study from Vietnam," *Agriculture*, vol. 10, no. 5, p. 180, 2020.

[49] K. K. Benke, F. Sheth, K. Betteridge, C. J. Pettit, and J. P. Aurambout, "A geo-visual analytics approach to biological shepherding: modelling animal movements and impacts," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I-2, pp. 117–122, 2012.

[50] H. Lei and L. Yang, "Research and design of technology for tracking and positioning wild stocking animals," *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, vol. 23, 2014.

[51] V. Maria Anu and R. Aroul Canessane, "Livestock monitoring using RFID with R+ tree indexing," *Biomedical Research*, vol. 28, no. 6, 2017.

[52] D. W. Bailey and J. R. Brown, "Rotational grazing systems and livestock grazing behavior in shrub-dominated semi-arid and arid rangelands," *Rangeland Ecology & Management*, vol. 64, no. 1, pp. 1–9, 2011.

[53] D. C. Ganskopp and D. W. Bohnert, "Landscape nutritional patterns and cattle distribution in rangeland pastures," *Applied Animal Behaviour Science*, vol. 116, no. 2-4, pp. 110–119, 2009.

[54] D. B. Lindenmayer, W. Blanchard, M. Crane, D. Michael, and C. Sato, "Biodiversity benefits of vegetation restoration are undermined by livestock grazing," *Restoration Ecology*, vol. 26, no. 6, pp. 1157–1164, 2018.