Wireless Communications and Mobile Computing

Backscatter Communications for Battery-Free IoT Networks

Lead Guest Editor: Yinghui Ye
Guest Editors: Shu Fu, Zheng Chu, and Liqin Shi

Backscatter Communications for Battery-Free IoT Networks

Wireless Communications and Mobile Computing

Backscatter Communications for Battery-Free IoT Networks

Lead Guest Editor: Yinghui Ye
Guest Editors: Shu Fu, Zheng Chu, and Liqin Shi

Chief Editor

Associate Editors

Academic Editors

Contents

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9898456, 1 page
https://doi.org/10.1155/2023/9898456

Retraction

Retracted: A Lightweight Face Verification Based on Adaptive Cascade Network and Triplet Loss Function

Wireless Communications and Mobile Computing

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Lin, C. Ye, W. Liu et al., "A Lightweight Face Verification Based on Adaptive Cascade Network and Triplet Loss Function," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3017149, 10 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9830151, 1 page
https://doi.org/10.1155/2023/9830151

Retraction

Retracted: Research on Multimodal Image Fusion Target Detection Algorithm Based on Generative Adversarial Network

Wireless Communications and Mobile Computing

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Wu, X. Wu, Y. Zhu et al., "Research on Multimodal Image Fusion Target Detection Algorithm Based on Generative Adversarial Network," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1740909, 10 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9806161, 1 page
https://doi.org/10.1155/2023/9806161

Retraction

Retracted: Deep Learning and Collaborative Filtering-Based Methods for Students' Performance Prediction and Course Recommendation

Wireless Communications and Mobile Computing

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Liu, C. Yin, Y. Li, H. Sun, and H. Zhou, "Deep Learning and Collaborative Filtering-Based Methods for Students' Performance Prediction and Course Recommendation," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2157343, 13 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6718689, 13 pages
https://doi.org/10.1155/2022/6718689

Research Article

An Effective and Robust Method for Unauthorized Reader Detection Based on Tag's Energy

Ziwen Cao,[1,2] **Jinxing Xie**,[1,2] **Siye Wang**,[1,2] **Yanfang Zhang**,[1] **Yue Cui**,[1,2] **Shang Jiang**,[1,2] **and Biao Jin**[3]

[1]*Institute of Information Engineering, Chinese Academy of Sciences, China*
[2]*School of Cyber Security, University of Chinese Academy of Sciences, China*
[3]*National Security Science and Technology Evaluation Centre, Beijing, China*

Correspondence should be addressed to Siye Wang; wangsiye@iie.ac.cn

Received 23 January 2022; Revised 15 July 2022; Accepted 27 August 2022; Published 7 October 2022

Academic Editor: Yinghui Ye

With the rapid development of the Internet of Things, ultra-high frequency (UHF) passive radio frequency identification (RFID) technology plays a vital role in various fields. UHF RFID faces unauthorized access attacks due to its long identification distance. Unauthorized readers can hide within a certain distance and use standard commands to read or modify tags. However, existing methods require additional equipment or are susceptible to environmental influences. In this paper, we make a novel attempt to counterattack unauthorized access. We propose a new method for Unauthorized Reader Detection based on Tag's Energy, called URDTE, to detect unauthorized readers by observing the energy of the tag. The competitive advantage of URDTE is that it is fully compatible with the RFID standard EPCglobal Gen 2, which makes it more applicable and scalable in practice. Besides, it takes the electrical energy stored in a tag's resistor-capacitor (RC) circuit as the detection principle, which is robust to environmental changes such as tag position, communication distance, and transmit power. We implement URDTE using commercial off-the-shelf (COTS) RFID devices without requiring firmware or hardware modifications. Extensive experiments show that URDTE can detect unauthorized readers with an accuracy of up to 99%.

1. Introduction

Radio frequency identification (RFID) is a noncontact automatic identification technology widely used in commercial automation, industrial automation, transportation, and many other fields, such as intelligent traffic control systems, access control systems, and warehouse management [1]. This technology uses the backscattering characteristics of radio frequency signals to achieve automatic identification [2] and mainly relies on two devices: tags (which can emit radio signals encoding identifying information) and readers (which detect the signals emitted by tags). RFID technology realizes item identification, inventory, and positioning by sticking tags on different items and placing the reader in the appropriate position. RFID tags are divided into active tags and passive tags. Active tags are expensive and only used in a few scenarios. Passive tags are widely used in all walks of life because they do not need a built-in power supply and have the advantages of low price, small volume, and long service life.

Due to the simple internal structure of the passive electronic tag chip and the weak computing power, the security protection ability of the RFID tag is poor. Many intrusion attacks against RFID systems are conducted against RFID tags. The most common attack is the unauthorized reading attack. The root cause of this problem is that there is no stipulation for tag-reader authentication in the existing RFID communication specification (i.e., EPCglobal Class 1 Generation 2 protocol [3, 4], referred to as EPCglobal Gen 2 in the following). In other words, RFID tags do not have the ability to authenticate the reader's identity, causing the tag to respond to any reader that accesses itself, which will lead to the following two risks. On the one hand, RFID tags send data in clear text. When a malicious reader compatible with

this protocol get close to the tag, the plaintext information of the tag can be obtained, resulting in the risk of privacy leakage. On the other hand, attackers use malicious readers to impersonate legal RFID readers and create tags with the same tag code, resulting in the risk of data tampering. As an example, the commercial spy may use the high-power reader to read the commercial rival warehouse's RFID tags to understand the rival situation and its business decision [5].

Meanwhile, security experts put forward protection measures to prevent unauthorized reading attacks, including data encryption, security protocol design, and RFID air interface intrusion detection. The data encryption method encrypts the communication data between the RFID tags and the readers so that the malicious RFID readers cannot parse the data. However, this requires a higher power excitation signal to drive the tag circuit, which dramatically shortens the reading distance of the tag and limits the usage scenarios [6]. Regarding RFID security protocols, scholars have proposed a series of improved security protocols for the security loopholes of existing protocols [7, 8], but these protocols are not universal. In the research of RFID air interface intrusion detection, Razm and Alavi [9] proposed a watchdog reader method to find the abnormal data when the system is working. However, the cost of this method is high due to the need to add additional RFID readers. Ding et al. [10] utilized USRP devices to monitor electromagnetic signals and proposed a fingerprint matching method to detect unauthorized readers. However, this method is not compatible with COTS RFID devices and lacks practicality. In addition, the electromagnetic fingerprint will change continuously with the changes in the environment, resulting in a decrease in detection accuracy.

To solve the unauthorized reading problem, we propose an unauthorized reader detection method based on the tag's power. The core idea is to observe the internal energy of the tag. Passive electronic tags use the electromagnetic signal emitted by the reader antenna to couple with its own antenna to generate electricity. Electrical energy is stored in the tag's internal capacitors to power the tag's internal volatile storage [11]. The energy stored in its internal capacitor lasts for a short period after the tag is accessed. Therefore, when it is detected that the internal circuit of the tag is abnormally charged, it means that the tag has been read without authorization. We designed URDTE (Unauthorized Reader Detection based on Tag's Energy) to detect unauthorized readers based on EPCglobal Gen 2, combined with the scenario of unauthorized reader detection. The specific approach is first to collect the tag's persistence time. Then, we construct models to calculate the persistence time and estimate the best model parameters. Finally, we carry on the real-time unauthorized reader detection. The final experimental results show that URDTE can detect unauthorized reading with a high accuracy rate. The main contributions are as follows:

(1) We explore a new method for malicious reader detection based on the tag's energy, called URDTE. The competitive advantage of URDTE is that it is fully compatible with the RFID standard, which makes it more applicable and scalable in practical applications. Besides, it is based on the power of the tag and is robust to various environmental conditions

(2) We propose a new metric called persistence time to detect malicious readers indirectly. Furthermore, we measure the persistence time by flipping and observing the flag in the tag's volatile memory

(3) We implemented a URDTE prototyping system based on the EPCglobal Gen 2 standard. Extensive experiments show that our method has high detection accuracy on average of 99%

The main structure of this paper is organized as follows: Section 2 describes the related work. Section 3 presents some background knowledge necessary to understand the methodology of this paper, including the EPCglobal Gen 2 protocol, persistence time, and the reading behavior model for unauthorized readers. Section 4 presents the basic principles and the overview of URDTE. Section 5 describes the design of URDTE in detail. Section 6 evaluates the effectiveness of URDTE. Finally, Section 7 concludes the paper.

2. Related Work

Through extensive research, we found that the unauthorized reading problem is mainly solved from two perspectives: defense and detection.

There are two main categories of defensive methods. The first category is to increase the access control protocols to the tag. The reader can only read the tag if it is authenticated by the tag. Burmester and Munilla, Qian et al., and Fan et al. [12–14] propose a lightweight RFID authentication protocol providing powerful authentication capabilities for authentication between tags and readers. Ma and Saxena [15] propose an authentication method based on scene context, where tags only allow access if they have sensed a specific scenario to defend against unauthorized reading. However, whether it is the authentication between the tag and the reader or the authentication with the environmental context, all the methods need to modify the RFID communication protocol or tag construction, which is very difficult for the already large-scale commercial used passive RFID system. Another defense category is to interfere with or intercept the unauthorized reading at the physical layer. Juels et al. [16] proposed a "blocker tag" concept. This "blocker tag" can be a kind of advanced tag or special radio frequency equipment. Such equipment will simulate the behavior of the real tag to interfere with the tag responding to the reader. Although this method can prevent unauthorized readers from accessing the real tag, the "blocker tag" also prevents legal readers from reading the tag and even turn into a DoS attack on the RFID system.

The approach from the detection perspective is to warn when unauthorized reading occurs. There are two main categories of this method. The first type of detection is based on physical layer signal characteristics. Ding et al. and Zhang

et al. [10, 17] utilize USRP devices to monitor electromagnetic signals in the physical space continuously and uses physical layer signal characteristics to detect the presence of unauthorized readings. The second category is the use of application layer data for detection. Sun et al. [18] use the changes in throughput to determine whether there is an unauthorized reading. However, the detection based on physical layer signal characteristics requires additional USRP devices, which leads to system costs increasing. In addition, the use of application layer data for detection is susceptible to interference from various factors, including the environment, which results in poor stability.

In contrast, our method follows the EPCglobal Gen 2 standard and does not require modification of the RFID communication protocol or additional dedicated equipment. As a result, URDTE can run directly on commercial RFID devices, is not easily affected by the environment, and has high detection accuracy and robustness.

3. Preliminaries

To help better understand URDTE, we present some preliminaries to the URDTE approach in this section, mainly including the EPCglobal Gen 2 protocol, the persistence time of RFID tags, and the modeling of unauthorized reading behavior.

3.1. EPCglobal Gen 2 Protocol. The EPCglobal Gen 2 (Gen 2) protocol is a worldwide UHF RFID standard that defines the physical interactions and logical operating procedures between the readers and tags [4]. We highlight the relevant functions involved by URDTE below based on Gen 2.

3.1.1. Session and Inventoried Tag. The EPCglobal Gen 2 standard stipulates that the reader can communicate with the tag through four sessions, respectively *S*0, *S*1, *S*2, and *S* 3. Under each session, there will be *A* or *B* two kinds of inventory flag. The inventory flag is actually a one-bit indicator of the tag's volatile memory. Volatile memory requires power to maintain stored information. Once the power falls below a certain threshold, the stored data is quickly lost and reverts to the default state *A* on each power-up. Tags can use only one session in each round of inventory, and the states under each session do not interfere with each other. Each session needs different power levels to maintain its state, so the persistence time of each inventoried flag is different. Table 1 shows the persistence time of different sessions.

The table shows that in sessions *S*2 and *S*3, the inventory flag will be maintained when the power is applied and maintained for a persistence time greater than 2 seconds after charging has stopped.

3.1.2. Select. Select is a command that is executed first in each round of inventory. This command allows the reader to select the tags to be inventoried. Aside from tag selection, the Select command can also assert or deassert a tags selected (SL) flag, or set a tags inventoried flag to either *A* or *B*. These flags determine whether a tag may respond to the reader or not. There are three core parameters of the Select command, which are as follows:

Table 1: Persistence time under different sessions.

Session	Required persistence time
S0	Tag energized: indefinite
	Tag not energized: none
S1	Tag energized: 500 ms-5 sec
	Tag not energized: 500 ms-5 sec
S2	Tag energized: indefinite
	Tag not energized: >2 sec
S3	Tag energized: indefinite
	Tag not energized: >2 sec

(1) Mask. Mask is used to match with the target tag and is often set to the EPC of the target operation tag.

(2) Target. Target determines whether the Select command operates on SL or four session flags. 0, 1, 2, 3, and 4 represent the operation objects of Session0, Session1, Session2, Session3, and SL, respectively.

(3) Action. The Select command selects the tags according to the rules. It performs different actions on the tags that match and do not match the rules. The specific actions are determined by the parameter Action, which is shown in Table 2.

3.1.3. Query. The Query command is used to initiate a round of inventory. This command is used on the set of tags selected in the previous Select step, or can be used individually. After the reader sends out the Query command, the tag that receives the command will send its information to the reader. After the reader queried the tag using the Query command, it will automatically flip the flag in its current session (from *A* to *B* or *B* to *A*). The Query command mainly consists of three core parameters:

(1) Session. Session determines the session to be used for this round of inventory.

(2) Target. Target determines which status of tags will participate in this round of inventory, where 0 indicates the tags with the session flag being *A* and 1 indicates *B*.

(3) Sel. This parameter is represented by two binary digits and determines which SL state the tag can reply to. 00_2 and 01_2 indicate all matching tags in the previous Select command; 10_2 indicates tags with deasserted SL flag (~ SL); and 11_2 indicates tags with asserted SL flag (SL).

3.2. Persistence Time. When the voltage of the internal energy storage capacitor of the RFID tag reaches above the operating voltage V_0 of the chip circuit, it can supply power to the tag [11]. When the storage capacitor starts to supply power, its supply voltage drops. When it drops below the chip's operating voltage V_0, the storage capacitor loses its power supply capability, the chip will not continue to work, and the data saved in its internal volatile storage area will be lost and reverted to its default value. The period from the decay of a fully charged capacitor to the voltage operating

Table 2: Eight actions of Select.

Action	Tag matching	Tag not-matching
000	Assert SL or inventoried $\longrightarrow A$	Deassert SL or inventoried $\longrightarrow B$
001	Assert SL or inventoried $\longrightarrow A$	Do nothing
010	Do nothing	Deassert SL or inventoried $\longrightarrow B$
011	Negate SL or ($A \longrightarrow B$, $B \longrightarrow A$)	Do nothing
100	Deassert SL or inventoried $\longrightarrow B$	Assert SL or inventoried $\longrightarrow A$
101	Deassert SL or inventoried $\longrightarrow B$	Do nothing
110	Do nothing	Assert SL or inventoried $\longrightarrow A$
111	Do nothing	Negate SL or ($A \longrightarrow B$, $B \longrightarrow A$)

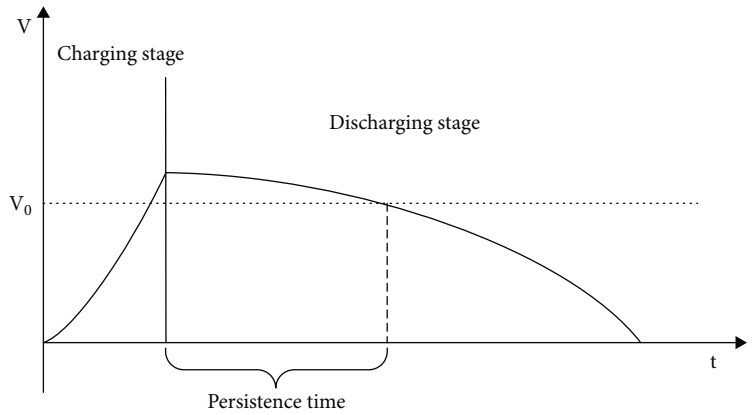


Figure 1: Persistence time.

threshold V_0 is defined as the "persistence time" [19]. Figure 1 shows the persistence time of tag charging and discharging. During this time, the tag's internal capacitive power supply continuously charges the internal volatile memory, maintaining its internal data. According to the EPCglobal Gen 2 standard, the time of the process from charging to the full charge of the tag is no more than 2 ms [3, 4]. The inventory flag under the four sessions and the SL flag are kept by the tag's internal volatile storage area. They will be lost due to power outages and reverting to the default state (state A or state $\sim$SL) upon the power supply. Chen et al. [19] propose various methods to measure the tag charging duration using this property. To capture the tag's persistence time efficiently and accurately, we present an optimized method for stepwise capturing persistence time.

3.3. Unauthorized Readers Behavior. This subsection analyzes the reading behavior of unauthorized readers. The core parameters of the Query command mainly have three parameters: Session, Target, and Sel, which control the session used in this round, the state of the tag, and the tag SL flag, respectively. For example, the command Query : {Session = 2, Target = A, Sel = $\sim$SL} is to query the tags with state A under Session 2 and Selected Flag $\sim$SL. Since the Sel parameter can be ignored, we perform a Cartesian product combination of Session and Target to obtain the full read pattern of the reader, as shown in Table 3.

Table 3: Unauthorized readers' behaviors.

Num	Session	Target
1	0	A
2	0	B
3	1	A
4	1	B
5	2	A
6	2	B
7	3	A
8	3	B

4. URDTE Overview

4.1. Basic Idea. The core idea of detecting unauthorized readings is to detect whether there is an unauthorized reader to access (charge) the tag by taking advantage of the tag's feature of sustained power after charging. The core detection process of URDTE is divided into three parts: first, we start the reader and launch the radio frequency signal to charge the tag. Second, we close the reader to stop the signal transmission. Third, we check the tag's power status after waiting for t_{gap} seconds. t_{gap} is the time it takes for the tag to run out of power. If the tag still has the remaining power after the t_{gap}, it means that unauthorized readers have carried out the charging process to the tag. It is not easy to detect the voltage and current information of the internal capacitance

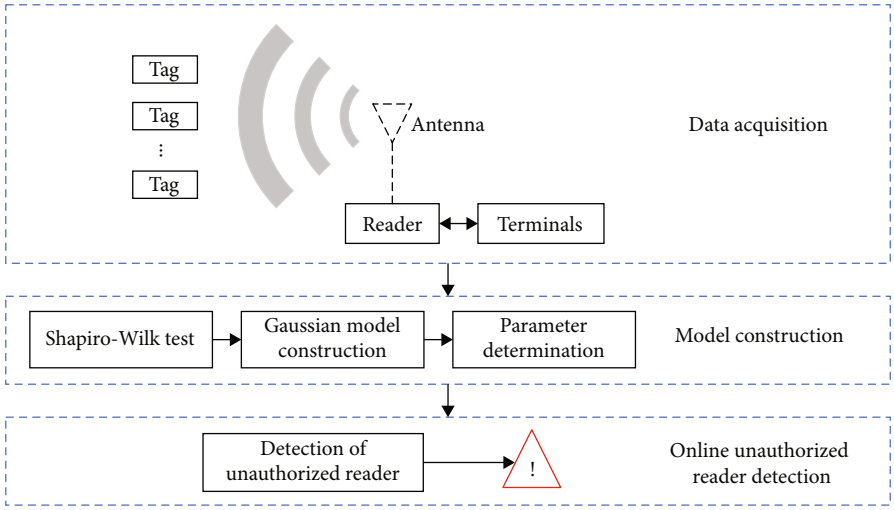


Figure 2: Framework of URDTE.

of the tag. We use the concept of persistence time to measure the ability of the tag to store power internally and use the inventory flag defined in the EPCglobal Gen 2 standard to detect whether the tag is in a state of power. The tag's electrical energy can be obtained by flipping this flag and continuously checking its status.

In practical implementation, selecting the waiting interval t_{gap} has a decisive influence on the detection performance. The best choice for t_{gap} is to make it consistent with the actual persistence time of the tag. In theory, the persistence time of a tag is only related to the capacitive circuitry inside the tag and is a stable value. However, when we collect the persistence time of the tags, the data collected each time will be slightly different. The model of the reader and external factors during the acquisition process will slightly affect the accuracy of the persistence time. To eliminate the effects caused by random factors during the acquisition process, we collected persistence time several times and constructed a Gaussian model to describe persistence time as accurately as possible to select the optimal waiting interval t_{gap}.

4.2. URDTE Overview. The unauthorized reader detection method based on the tag's energy consists of three main phases. Figure 2 shows the general framework of URDTE.

The first phase is the data acquisition. This part collects the persistence time of the tag for subsequent modeling. It mainly includes the tag attached to the object, the reader reading the tag, the antenna to transmit the RF signal, and the terminal to handle the collected data.

The second phase is the model construction. In this phase, to cope with the variability of the actual environment, a distribution test of the collected persistence time data is required before the Gaussian model is constructed. By counting the data collected in the previous phase and using the Shapiro-Wilk test, we test whether the collected persistence time data conformed to the Gaussian distribution. If they do not match, we return to the first stage for data collection again. If they match, we accept this data and use it to construct a Gaussian model. After constructing the Gaussian model, we calculate the tolerance factor Δt to estimate the minimum discharge time. The false negative rate will increase when increasing Δt. Whereas Δt decreases, the false positive rate will increase. Therefore, the determination of Δt is critical. We will provide a detailed theoretical analysis of the selection of Δt.

The last phase is the online unauthorized reader detection stage, which is the core part of URDTE. In this phase, the reader carries on the charge to the tag. After waiting for the predicted time when the tag power should be exhausted, it detects whether the tag power is exhausted. If it detects the tag still has the remaining power, it indicates that the tag is read by the unauthorized reader and has completed the charging process. URDTE keeps repeating the above process to judge whether there is an unauthorized reader.

5. URDTE Design

5.1. Data Acquisition. The first module of the URDTE is the data acquisition module, whose primary function is to collect the persistence time of the tag. The primary basis for the persistence time collection is the change of the tag's inventory flag. The tag's inventory flag is stored in an internal volatile memory area, which causes the tag to lose inventory flag data when the battery is depleted and automatically reset to its initial state when recharging. As we introduced in Section 3, there are four session modes within the tag, each with its independent flag, and the flag exists in two states, *A* or *B*. The default state is *A*. Therefore, we can first set the tag inventory flag to *B*. After waiting for t_{gap} seconds, the Query command is used to inventory the tag in state *B*, and if a tag response is received, it means that the tag is still in state *B*, i.e., there is still power to maintain its inventory flag. Keep increasing t_{gap} until there is no tag response, where t_{gap} is the persistence time of the tag.

The selection of the time interval t_{gap} is crucial. We use a stepwise approach to collect it. The specific approach is to

Input: Whether the Query command get tags: True or False
Output: The persistence time: pt
s = rand (2 or 3)
step =1//the step value in current iteration
pt =0//the persistence time
t_{gap} =1//the waiting time in this round
While step > 0.01 **do**
  Reader: Select(session = s, action = 100_2)
  Stop the reader
  Sleep t_{gap}
  If reader: Query(session = s, target = B, Sel = 00_2)
  Get tags **then**
    $t_{gap} = t_{gap}$ + step
  Else
    pt = t_{gap} − step
    step = step ∗ 0.1
  End if
End while
Return pt

Algorithm 1:Persistence time acquisition algorithm.

use a large granularity time interval to collect rough persistence time in the initial stage, lock the range, and then gradually reduce the time granularity. The specific algorithm pseudocode is shown in Algorithm 1.

5.2. Model Construction. This section introduces the construction of the detection model with three core parameters: the Gaussian model of the persistence time, the maximum persistence time $PT_{\max}$, and Δt. First, we perform Gaussian model using the persistence time. Then, the Shapiro-Wilk test is used to check whether the data conforms to the Gaussian distribution [20]. Finally, we describe the determination of parameters Δt.

5.2.1. Shapiro-Wilk Test. The Shapiro-Wilk test tests the null hypothesis that a sample $\{\mathrm{pt}_1, \cdots, \mathrm{pt}_n\}$ comes from a normally distributed population. We use n to denote the amount of persistent time data. First, we arrange the collected persistence time in order.

$$\mathrm{pt}_1 \le \mathrm{pt}_2 \le \cdots \le \mathrm{pt}_i \le \cdots \le \mathrm{pt}_n. \tag{1}$$

Then, we calculate the value of the statistic W according to Formula (2).

$$W = \frac{\left(\sum_{i=1}^{(n/2)} a_i (\mathrm{pt}_{n+1-i} - \mathrm{pt}_i)\right)^2}{\sum_{i=1}^{n} (\mathrm{pt}_i - \bar{\mathrm{pt}})^2}. \tag{2}$$

The coefficients a_i are given by

$$(a_1, a_2, \cdots, a_n) = \frac{m^T V^{-1}}{C}, \tag{3}$$

where C is a vector norm,

$$C = \left|V^{-1} m\right| = m^T V^{-1} m^{(1/2)}, \tag{4}$$

and the vector m,

$$m = (m_1, \cdots, m_n)^T, \tag{5}$$

where m is constructed from the anticipated values of the order statistics of independently distributed random variables selected from the standard normal distribution, and V is the covariance matrix for those statistics.

Finally, we obtain the critical value W_α at the significance level $\alpha = 0.1$ and compare the magnitude of the calculated W with the critical value W_α. If $W \ge W_\alpha$, then the original data conform to the normal distribution. Otherwise, it is necessary to return to the acquisition phase to recollect the persistence time.

5.2.2. Gaussian Model Construction. When the persistence time $\mathrm{PT} = \{\mathrm{pt}_1, \cdots, \mathrm{pt}_n\}$ accords with a Gaussian distribution, it can be expressed as follows:

$$\mathrm{PT} \sim N(\mu, \sigma^2). \tag{6}$$

The mean of the persistence time data can be expressed as

$$\mu = \frac{1}{n} \sum_{i=1}^{n} \mathrm{PT}_i. \tag{7}$$

In addition, we also need to record the maximum data $PT_{\max}$ in the persistence time data sample, which determines the waiting time t_{gap} together with the tolerance factor Δt.

5.2.3. Parameter Determination. In order to better describe the setting of Δt, Figure 3 shows the model of the tag state changing with time. The tag state is set to B and the SL state is set to positive at t_1 and reaches time t_3 after T_{gap} seconds have elapsed. The algorithm detects whether there is a tag

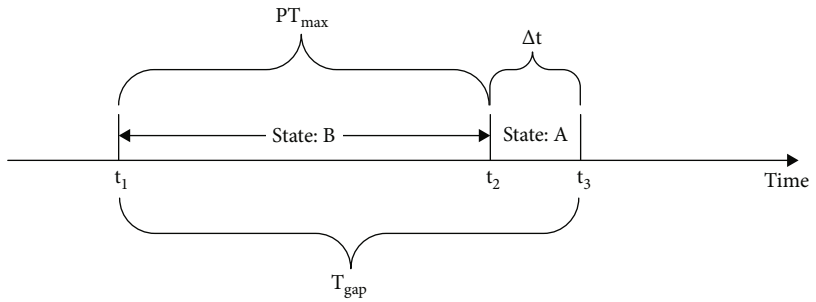


Figure 3: Model of tag state.

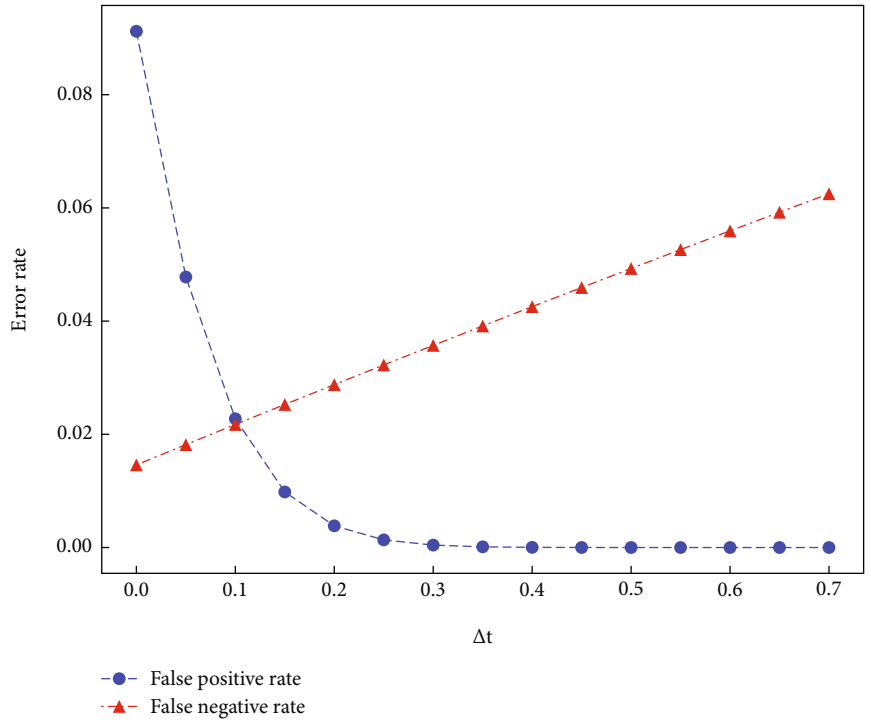


Figure 4: Tolerance factor Δt setting.

with state B at this time. After detecting that there is no tag with state B, it detects whether there is a tag with state SL as positive. If the tag power is not consumed at t_3, it will cause false positives. The setting of the tolerance factor Δt ensures that the tag power has been exhausted.

The probability of false positive alarm prob_{FP} caused by the tag's power not being consumed at t_3 is

$$\text{prob}_{\text{FP}} = P(\text{PT} > \text{PT}_{\max} + \Delta t). \tag{8}$$

However, if Δt is too large, it will lead to a longer t_{gap} and affect the detection accuracy. For example, the unauthorized reader reads the tag immediately after t_1. Due to the long t_{gap}, although the tag is read by an unauthorized reader and completes charging during this process, it may still run out of power before the detection at t_3, resulting in a missed alarm. That is, if an unauthorized reader reads within the time range of $(t_1, t_3 - \text{pt})$, it will run out of power before the detection at t_3. The probability of a false negative prob_{FN} in this case is

$$\text{prob}_{\text{FN}} = \frac{(\text{PT}_{\max} + \Delta t) - \mu}{\text{PT}_{\max} + \Delta t}. \tag{9}$$

To strike a balance between the two false alarm rates, we take the intersection of the false negative rate curve and the false positive rate curve as the value of Δt, as shown in Figure 4.

Therefore, we use the following equations to solve the Δt.

$$\begin{cases} \text{prob}_{\text{FP}} = P(\text{PT} > \text{PT}_{\max} + \Delta t), \\ \text{prob}_{\text{FN}} = \dfrac{(\text{PT}_{\max} + \Delta t) - \mu}{\text{PT}_{\max} + \Delta t}, \\ \text{prob}_{\text{FP}} = \text{prob}_{\text{FN}}. \end{cases} \tag{10}$$

Input: The maximum value of the persistence time: PT_{max}
Tolerance factor: Δt
Output: The result of unauthorized reader detection: res
s = rand (2 or 3)
$T_{gap} = PT_{max} + \Delta t$
While True **do**
  Reader:Select(session = s, action = 001_2)
  Reader:Query(session = s, Target = A, Sel = 00_2)
  Stop the reader
  Sleep t_{gap}
  If reader:Query(session = s, target = B, Sel = 00_2)
  Get tags **then**
    res = True
  Else if reader:Query(session = s, target = A, Sel = 11_2) get tags **then**
    res = True
  Else
    res = False
  End if
  If res = True **then**
    Return res and alarm
  Else
    Return res
  End if
End while

Algorithm 2: Unauthorized reader detection algorithm.

5.3. Unauthorized Reader Detection. The basic idea is that the tags are inventoried (charged) at regular intervals, and after a specific interval, the tags are checked to see if they have been inventoried (charged) during this interval. We found that the state under *S*2 and *S*3 sessions in the tag, as well as SL, rely on the power saved by the capacitor inside the tag to maintain its state and will revert to the default state after the tag runs out of power. We detect whether an unauthorized reader accesses the tag through the state change of these flags. The steps are as follows:

(1) Start the reader, adopt the Select command to change the tag status under Session 2 to state *B*, and state SL is set to positive, finish charging, and stop the reader

(2) After waiting for T_{gap} seconds, use the Query command to query whether there is a tag with Session 2 state *B*. If a tag response exists, it indicates the presence of the unauthorized reader. Otherwise, continue to the next step

(3) Use the Query command to query whether there is a tag whose SL flag is true to indicate the presence of an unauthorized reader

When the unauthorized reader adopts *S*0, *S*1, or *S*3 session to carry on the tag inventory (unauthorized reader mode as 1-4, 7-8 in Table 3), it will charge the tag and directly increase the tag's flag persistence time after successfully communicating with the tag. Therefore, it can be detected by step 2.

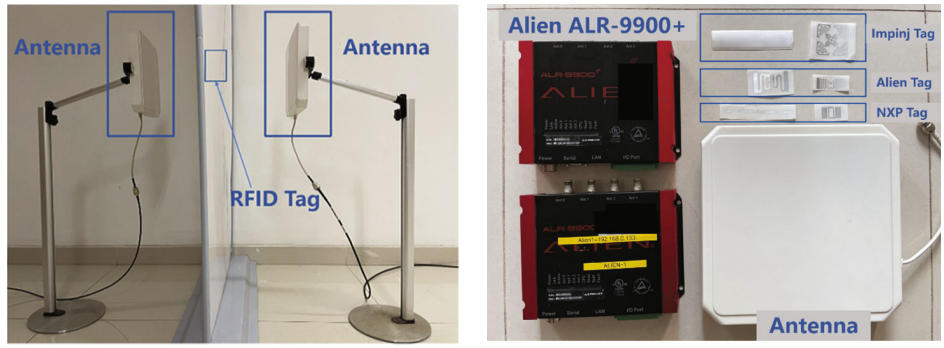
When an unauthorized reader uses an *S*2 session for unauthorized reading (unauthorized reader mode is 5-6 in Table 3), since the unauthorized reader uses the same session as the legal reader, this affects our flag state and needs to be discussed in detail. On the one hand, when the unauthorized reader adopts the *S*2 session and reads the tag of state *A* (unauthorized reader mode corresponds to 5 in Table 3), the tag will not respond because the tag state has been modified to state *B*. However, since other Query commands charge the tag, this read increases the persistence time of the tag. On the other hand, when the unauthorized reader uses the *S*2 session and reads the tag in state *B* (the unauthorized reader mode corresponds to 6 in Table 3), the unauthorized reader can successfully read the tag because the tag state has already been modified to state *B*, and the tag state is automatically flipped to state *A* at this time. The second step of URDTE cannot effectively detect this situation because it will query the tag of state *B*, and if there is no state *B* tag, it is considered that there is no unauthorized reader. This situation needs to go to the third step to detect.

The third step detects the SL flag. Although the unauthorized reader mode 6 uses the same session with legal readers, the read operation does not affect the SL flag so that the unauthorized reader mode 6 can be detected accurately.

In addition, we can randomly choose to use *S*2 or *S*3 to enhance the randomness of the detection algorithm and reduce the probability of collision with the read session used by unauthorized readers. The pseudocode for the detection process is shown in Algorithm 2.

6. Device Deployment and Experimental Result

In this section, we implement a prototype of URDTE in a commodity RFID system. Besides, we evaluate the performance of



(a) (b)

Figure 5: Experimental scene. (a) System prototype. (b) Equipment and tags.

URDTE through extensive experiments in terms of robustness to environmental changes and detection accuracy.

6.1. System Deployment. The system prototype is shown in Figure 5(a), where the equipment and tags used are shown in Figure 5(b). The reader used in the system is Alien ALR-9900+, and the tags of three manufacturers are selected: Impinj, Alien, and NXP. We follow LLRP [21] (Low-Level Reader Protocol for EPCglobal Gen 2 standard) for reader development and do not need to make any changes to the reader hardware or firmware. The back-end computer is equipped with an Intel Core i5-8250U 1.6 GHz CPU and 8 GB RAM.

6.2. Detection Performance. In the experiment, we tested the influence of tag chip, angle, power, and distance of unauthorized reader on URDTE detection effectiveness. The detection of unauthorized readers is essentially a binary classification problem. Therefore, we evaluate URDTE using classical metrics from machine learning. Our prototype system feeds back the detection results at a fixed frequency with positive or negative to indicate the presence of unauthorized readings. There are four cases as follows:

(1) If there is an unauthorized reader intrusion, and URDTE detects this intrusion, it is a true positive (TP)

(2) If there is an unauthorized reader intrusion, but it is not detected by URDTE, it is a false negative (FN)

(3) If there is no unauthorized reader intrusion, and URDTE feedback results in a normal state, it is a true negative (TN)

(4) If there is no unauthorized reader intrusion, but URDTE feedback results in unauthorized reader intrusion, that is a false positive (FP)

False positive (FP) and false negative (FN) in these misclassifications are our focus, where false positive (FP) can cause false alarms, while false negative (FN) can cause missed alarms for intrusions, resulting in a security risk. In addition, the accuracy rate is also the focus of our attention. Therefore, we will use these three metrics to evaluate the effectiveness of the prototype system, which are defined as follows:

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN} + \mathrm{FP}}, \tag{11}$$

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{TN} + \mathrm{FP}}, \tag{12}$$

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TN} + \mathrm{FP} + \mathrm{FP} + \mathrm{FN}}. \tag{13}$$

6.2.1. Effect of Tag's Chip on Accuracy. UHF RFID reader and tag follow the specifications of the EPCglobal Gen 2 protocol for communication. Although the EPCglobal Gen 2 protocol specifies that the tag should have four communication sessions, the requirements for session 2 and session 3 are vague and only require a persistence time greater than 2 seconds. Through experiments, we found that different brand tags' persistence times are different, even if the chips of different models of the same brand are different. Therefore, we tested different tag chips for their accuracy.

We, respectively, place each kind of tag 0.5 meters directly in front of the legal reader, and the unauthorized reader antenna is placed at 0.5 meters at the back of the tag, and carry out 500 times unauthorized reading for each kind of tag, respectively.

The results are shown in Figure 6. The detection accuracy of all tags was high (>97%), and the false negative rate was low (<3%), indicating that URDTE has a high detection rate and is not affected by the tag chip type.

6.2.2. Effect of Tag Angle on the Accuracy. We place the tag at 0.5 meters in front of the legal reader and change the angle between the tag and the reader. We set the angle between the tag and the legal reader antenna to 0, 45, 90, 135, 180, 225, and 270, respectively. The unauthorized reader performs illegal readings in front of the tag 500 times for each angle separately.

The experimental results are shown in Figure 7. At the 90 and 270, the RF signal emitted from the antenna cannot successfully activate the tag due to the specificity of the angle, resulting in the system not being able to read the tag, i.e., the charging process of the tag cannot be completed.

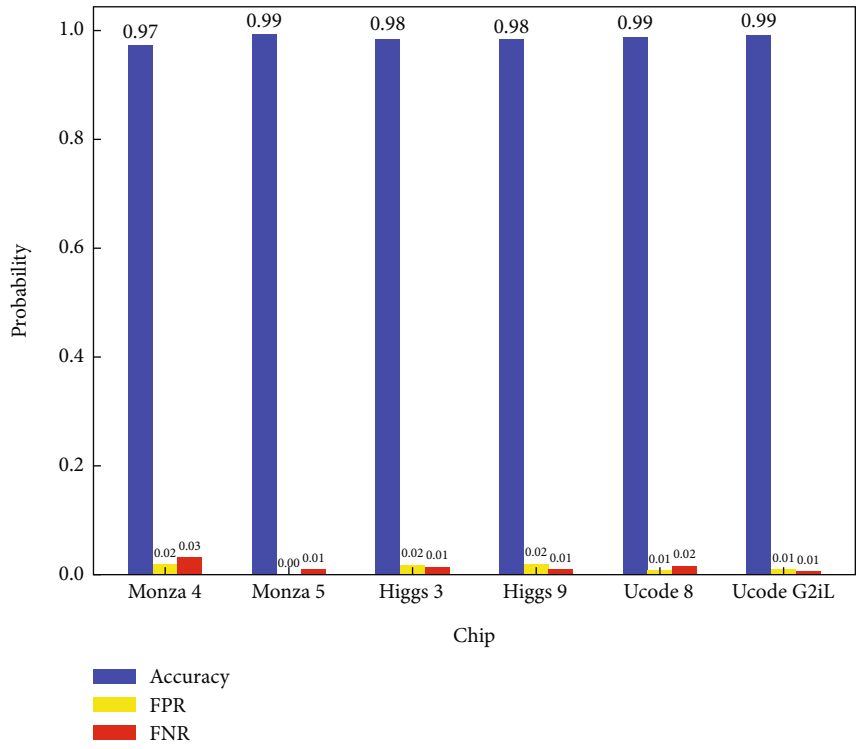


Figure 6: Different chips.

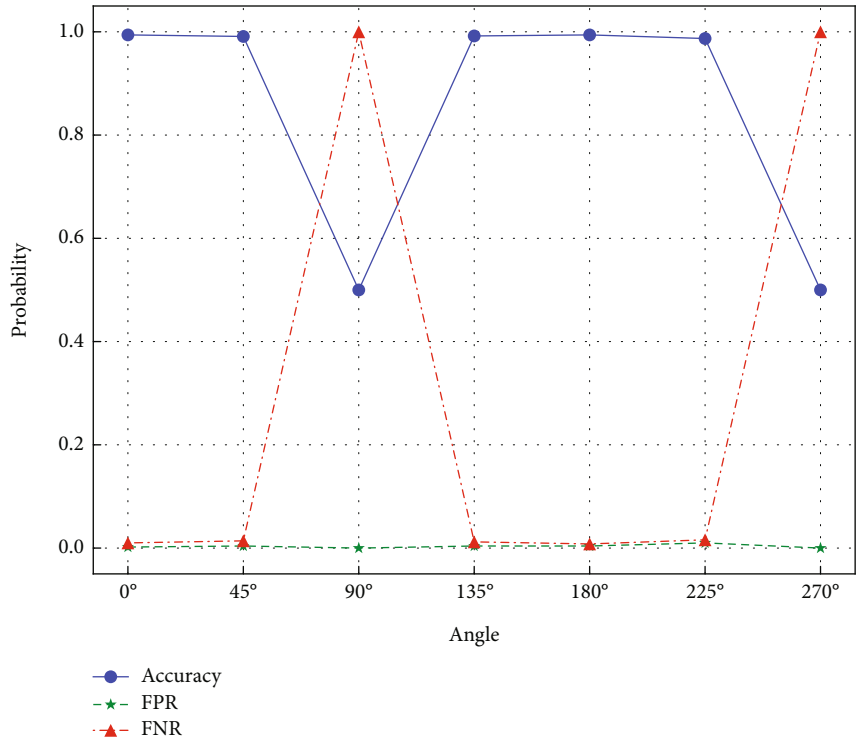


Figure 7: Different angles.

For the other angles, it can be found that the error rate and accuracy do not change much with the change of angle, and both have good detection results. Therefore, when we deploy and apply the system, we should pay attention to the legal antenna should try to form a parallel angle with the tag to get the best detection effect, and if the tag signal is found to disappear during the detection process, the angle between the tag and the antenna should be adjusted appropriately.

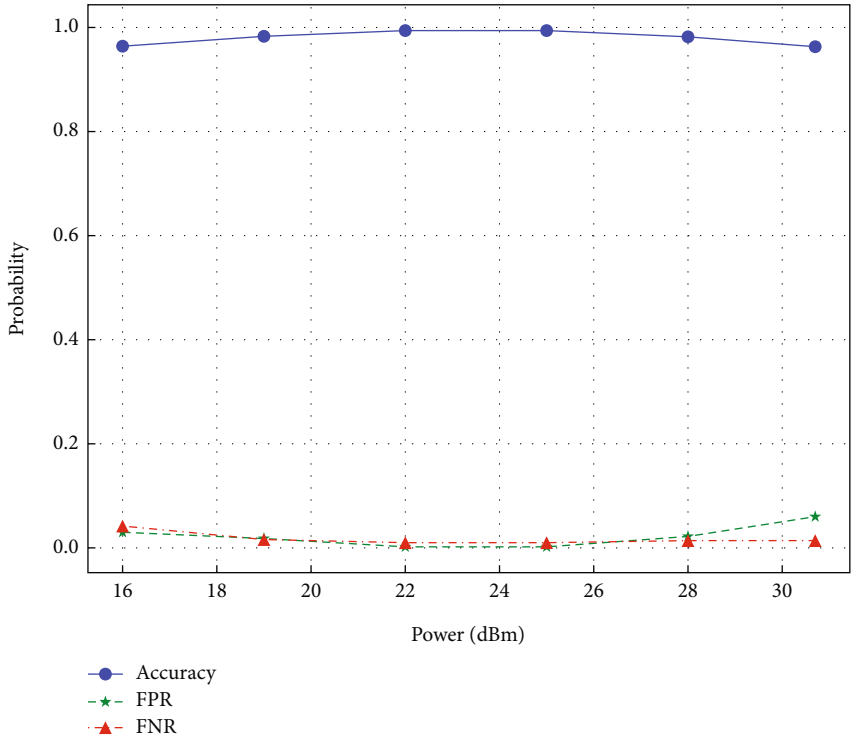


Figure 8: Different power.

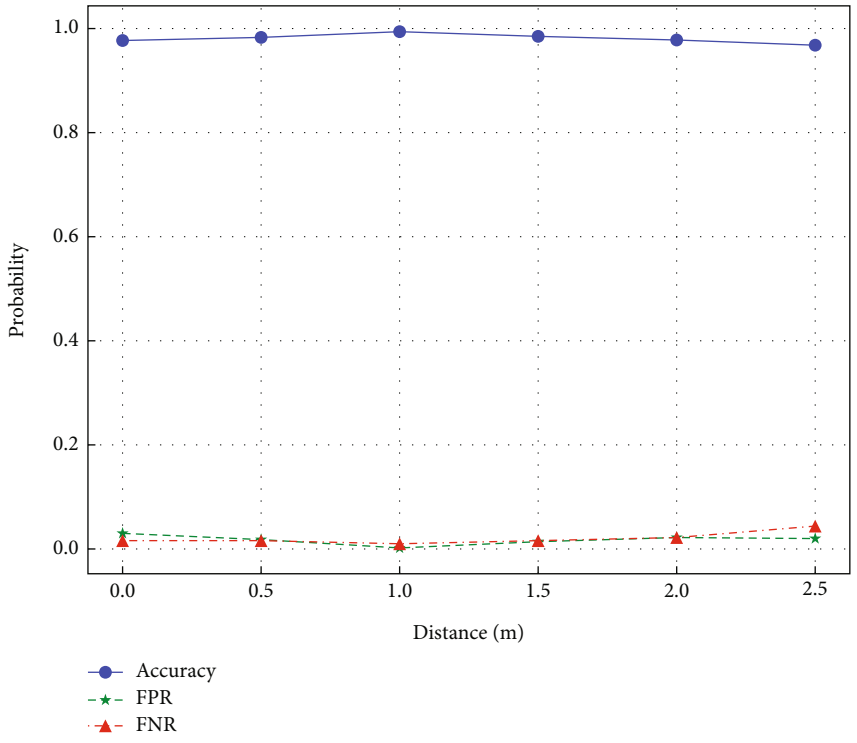


Figure 9: Different distances.

6.2.3. Effect of Unauthorized Reader Power on Accuracy. In this experiment, we investigate the effect of unauthorized reader transmit power on detection accuracy. We place the tag between legal and unauthorized readers, and the tag is 0.5 meters away from both legal and unauthorized readers. Unauthorized readers perform unauthorized readings from16 dBm to 30 dBm. The experimental results are shown in Figure 8. We found that the detection accuracy will be

high at the appropriate power. However, the detection accuracy will be suppressed to a certain extent under small or large power. The results show that the URDTE can detect the intrusion with high accuracy regardless of the power of the unauthorized reader.

6.2.4. Effect of Unauthorized Reader Distance on Accuracy. In this experiment, we move the unauthorized reader and change the distance between the unauthorized reader and the tag to investigate the effect of distance on the detection accuracy. We first place the unauthorized reader at a distance of 2.5 meters from the tag and constantly reduce the transmit power of the unauthorized reader. We found that when the power is less than 18 dBm, the unauthorized reader will no longer read the tag. We deliberately set the transmit power of the unauthorized reader to 18 dBm. That is, at 2.5 meters, 18 dBm is the minimum power to be able to read the tag. We keep the distance between legal readers and tags at 0.5 meters, then move unauthorized readers to distances of 0, 0.5, 1, 1.5, 2, and 2.5 meters from the tags for testing and perform 500 reads at each distance.

Figure 9 shows that even if the unauthorized reader performs unauthorized reading at low power at a remote location, our system can still accurately detect this intrusion and still has a high accuracy rate. Regardless of the distance and the power of the unauthorized reader, as long as the unauthorized reader can read the tag, it will complete the charging of the tag, which makes our detection method extraordinarily stable and robust. However, if the distance is too large or too small, it will slightly impact the detection accuracy. The reason is that the charging and discharging times of the tags are slightly different at different distances.

In summary, the experimental results show that the URDTE algorithm has extremely high detection accuracy and robustness for unauthorized readers with different chips, powers, and distances on the premise that legal readers can read tags. Moreover, the algorithm does not rely on special detection equipment. The detection method is low cost and has high practical application value for discovering unauthorized readers and protecting the security of air interface data of the RFID system.

7. Conclusion

This paper proposes a method for detecting unauthorized reading based on the tag's power. The core idea of this method is to determine whether an unauthorized reader has accessed a tag by detecting the tag power. We implemented this method on a commercial reader following the EPCglobal Gen 2 standard. Extensive experiments have shown that URDTE has high accuracy and strong robustness in detecting unauthorized reading, and this method effectively enhances RFID system security, which is essential for preventing air interface intrusion.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Strategic Priority Research Program of Chinese Academy of Sciences, Grant No. XDC02040300.

References

[1] K. Domdouzis, B. Kumar, and C. Anumba, "Radio-frequency identification (RFID) applications: a brief introduction," *Advanced Engineering Informatics*, vol. 21, no. 4, pp. 350–355, 2007.

[2] Y. Ye, L. Shi, X. Chu, and G. Lu, "On the outage performance of ambient backscatter communications," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7265–7278, 2020.

[3] E. O. R. Implied, T. This, D. Is et al., *Specification for RFID Air Interface EPC Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 Mhz 960 Mhz*, 2008.

[4] EPCglobal, *Epc Radio-Frequency Identity Protocols Generation-2 UHF RFID Standard, Specification for RFID Air Interface Protocol for Communications at 860 Mhz960 Mhz*, 2018.

[5] T. Karygiannis, B. Eydt, G. Barber, L. Bunn, and T. Phillips, *Guidelines for Securing Radio Frequency Identification (RFID) Systems*, Special Publication 800-98, National Institute of standards and Technology, Technology Administration U.S. Department of Commerce, 2007, http://csrc.nist.gov/publications/nistpubs/800-98/SP800-98_RFID-2007.pdf .Accessed15.

[6] H. Y. Chien, "Sasi: a new ultralightweight RFID authentication protocol providing strong authentication and strong integrity," *IEEE Computer Society Press*, 2007.

[7] Q. Jia, P. Chen, X. Gao, L. Wei, and B. Zhao, "Lightweight anti-desynchronization RFID mutual authentication protocol," *Journal of Central South University(Science and Technology)*, vol. 46, no. 6, pp. 2149–2156, 2015.

[8] Z. Rong, L. Zhu, X. Chang, and Y. Yi, "An efficient and secure RFID batch authentication protocol with group tags ownership transfer," *Collaboration and Internet Computing*, 2015.

[9] A. Razm and S. E. Alavi, "An intrusion detection approach using fuzzy logic for RFID system," *Advances in Information Science and Applications*, vol. 2, 2014.

[10] H. Ding, J. Han, Y. Zhang et al., "Preventing unauthorized access on passive tags," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 1115–1123, Honolulu, HI, USA, 2018.

[11] Y. Zhang, L. T. Yang, and J. Chen, *RFID and Sensor Networks: Architectures, Protocols, Security, and Integrations*, CRC Press, 2009.

[12] M. V. D. Burmester and J. Munilla, "Lightweight RFID authentication with forward and backward security," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, pp. 1–26, 2011.

[13] Q. Qian, Y.-L. Jia, and R. Zhang, "A lightweight RFID security protocol based on elliptic curve crytography," *International Journal of Network Security*, vol. 18, no. 2, pp. 354–361, 2016.

[14] K. Fan, N. Ge, Y. Gong, H. Li, R. Su, and Y. Yang, "An ultra-lightweight RFID authentication scheme for mobile commerce," *Peer-topeer Networking and Applications*, vol. 10, no. 2, pp. 368–376, 2017.

[15] D. Ma and N. Saxena, "A context-aware approach to defend against unauthorized reading and relay attacks in RFID systems," *Security and Communication Networks*, vol. 7, 2695 pages, 2014.

[16] A. Juels, R. L. Rivest, and M. Szydlo, "The blocker tag: selective blocking of RFID tags for consumer privacy," in *Proceedings of the 10th ACM conference on Computer and communications security*, pp. 103–111, New York, NY, USA, 2003.

[17] W. Zhang, S. Zhou, J. Luo, H. Cheng, and Y. Liao, "A lightweight detection of the RFID unauthorized reading using rf scanners," in *2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing*, pp. 317–322, New York, NY, USA, 2015.

[18] D. Sun, Y. Cui, Y. Feng, J. Xie, S. Wang, and Y. Zhang, "Urtracker: unauthorized reader detection and localization using cots RFID," in *International Conference on Wireless Algorithms, Systems, and Applications*, pp. 339–350, Cham, 2021.

[19] X. Chen, J. Liu, X. Wang, H. Liu, D. Jiang, and L. Chen, "Eingerprint: robust energy-related fingerprinting for passive {RFID} tags," *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20)*, pp. 1101–1113, 2020.

[20] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965.

[21] GS1, *Low level reader protocol*, 2008, https://www.gs1.org/standards/epc-rfid/llrp/1-1-0.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8063190, 9 pages
https://doi.org/10.1155/2022/8063190

Research Article

Energy-Efficient Resource Allocation in Cognitive Wireless-Powered Hybrid Active-Passive Communications

Jianjun Luo,[1,2] **Ming Li**,[1,2] **and Xin Ning**[1,2]

[1]*School of Astronautics, Northwestern Polytechnical University, Xi'an, China*
[2]*National Key Laboratory of Aerospace Flight Dynamics, Northwestern Polytechnical University, Xi'an, China*

Correspondence should be addressed to Xin Ning; ningxin@nwpu.edu.cn

Received 30 March 2022; Revised 5 May 2022; Accepted 27 May 2022; Published 18 July 2022

Academic Editor: Liqin Shi

Integrating hybrid active-passive communications into cognitive radio can achieve a spectrum- and energy-efficiency information transmission, while the resource allocation has not been well studied particularly for the network with multiple secondary users (also termed as the Internet of Things (IoT) users). In this article, we formulate an optimization problem to maximize the energy efficiency of all the IoT nodes in a cognitive wireless-powered hybrid active-passive communication network by taking the interference from the IoT node to the primary link, the energy causality constraint, and the minimum throughput constraint per IoT node. By using the Dinkelbach method and introducing auxiliary variables, we devise an iterative algorithm to optimally solve the formulated problem. Computer simulations are provided to validate the quick convergence of the iterative algorithm and the advantages of the proposed scheme in terms of the energy efficiency.

1. Introduction

In the past decade, it has been witnessed that the Internet of Things (IoT) technology has wide applications in our daily lives particularly in the smart factory. To realize smart applications, a large number of tiny IoT nodes should be deployed to collect data from the environment and then send the collected data to the information fusion, resulting in a huge need for spectrum resource [1–3]. It is reported by the European Union that just eHealthCare IoT connectivity requires at least 5.2 GHz bandwidth if dedicated spectrum is allocated to each tiny IoT node [4]. However, most of the spectrum resources have been allocated, leading to the shortage of spectrum resources.

To relieve the conflict between the increasing demand for spectrum and the limited spectrum resources, cognitive radio (CR) has been proposed as an efficient solution for this problem by letting IoT nodes share the same spectrum resource as the primary user [5, 6]. In CR, the tiny IoT node is allowed to access the spectrum allocated to the primary user in the opportunistic or spectrum sharing manner, while ensuring the Quality of Servers (QoS) of the primary user. On the other hand, due to the cost and form factor constraints, the tiny IoT nodes are powered by the battery with a limited capacity that can be quickly drained by information transmissions, thus limiting the battery life of these tiny IoT nodes. Recall that the primary signal can function as the energy and information sources simultaneously. Wireless-powered transfer is introduced into CR, yielding a cognitive wireless-powered communication [7].

In previous studies on cognitive wireless-powered communication (see [7–11] and reference therein), it was considered that the IoT node firstly harvests energy from the primary signal and subsequently uses the harvested energy to transmit signal by accessing the spectrum of the primary user via active radios (AR). In AR, the IoT node needs to generate the carrier signal and modulate its information on the carrier signal. Such an approach requires power-consuming components, e.g., oscillator [12–14]. Accordingly, AR achieves a high transmission rate but at the cost of a high power consumption. Since the energy consumed by the IoT node is constrained by its harvested energy, the IoT node should allocate a large proportion of time period to harvest energy and leave a limited time for AR, which

may lead to a low throughput [15–18]. Recently, passive communication has received much attention due to its low power consumption. The key idea of passive communications is allowing IoT node encoding information on the incident signal and reflecting the encoded signal to the receiver, thus removing the need of power-consuming components and realizing a low-power communication [12–14]. Due to this, the passive communication has been introduced into cognitive radio for addressing the above challenge [19]. However, the rate of the passive communication enabled IoT node is still low. Recall that both AR and passive communication have different tradeoffs between the communication rate and power consumption [15–18], which can be exploited to achieve efficient data transmissions for IoT nodes in cognitive wireless-powered communications. The above combination is referred as the cognitive wireless-powered hybrid active-passive communication in this paper.

In this conference paper [20], the authors considered that the cognitive wireless-powered hybrid active-passive communication operates in the overlay mode and maximized the IoT node's throughput by optimizing the tradeoff between passive communication and AR, subject to the constraint where the harvested energy of the IoT node is not less than that consumed by itself. Subsequently, this conference paper was extended into a journal paper [21], where the same problem was studied in both the overlay and underlay modes. In [22], the authors considered the cognitive wireless-powered hybrid active-passive communication with multiple IoT nodes, and the main contribution was to maximize the sum throughput of all the IoT nodes by jointly optimizing the energy harvesting time, the passive communication time, and the AR time for each IoT node. The authors in [23] proposed another wireless-powered cognitive hybrid active-passive communication network, where the power beacon is deployed for increasing the harvested energy of the IoT node and optimized the time for energy harvesting, passive communication, and AR of the IoT node. The above works [20–23] focused on the throughput maximization and did not optimize the backscatter coefficient. Such a gap was filled by [24]. Since the energy efficiency is of significance for wireless communications, the authors proposed to maximize the energy efficiency of the IoT node in an overlay-based cognitive wireless-powered hybrid active-passive communication, subject to the maximum tolerated interference to the primary link and the imperfect spectrum sensing constraints. The authors of [25] studied the multi-IoT nodes in cognitive wireless-powered hybrid active-passive communication and maximized the energy efficiency of all the IoT nodes, while considering the energy causality constraint and the minimum throughput constraint per IoT node. However, this work largely ignored the interference from the IoT node to the primary link; thus, the designed resource allocation may not work in practical cognitive wireless-powered hybrid active-passive communications and this should be fixed.

In this article, we consider a cognitive wireless-powered hybrid active-passive communication with multiple IoT nodes and propose to maximize the energy efficiency, while considering the maximum tolerated interference to the primary link, the energy causality constraint, and the minimum throughput constraint per IoT node. The formulated problem is optimally solved by our designed Dinkelbach-based iterative algorithm. Finally, the simulation results are provided to support our work.

2. System Model

As shown in Figure 1, we consider a cognitive wireless-powered hybrid active-passive communication network, which consists of a legacy transmitter (LT), a legacy receiver (LR), K IoT nodes, and an information gateway. All the devices are equipped with a single antenna. In order to harvest energy from the signals transmitted by the LT and encode and backscatter legacy signals for information transmission, it is assumed that both the radio frequency (RF) energy harvesting circuit and the backscatter circuit are equipped at each IoT node. Besides, the active transmission circuit is also equipped at each IoT node so that each IoT node can choose to transmit its own information via hybrid active-passive communications. Suppose that the perfect channel state information is known by the information gateway before the whole information transmission by the information exchange among the LT, the LR, IoT nodes, and the information gateway. Therefore, the information gateway can design the optimal resource allocation scheme based on all obtained channel state information and then transmit the designed scheme to IoT nodes so that each IoT node can operate by following the designed scheme. To obtain the performance bound, we assume perfect channel state information (CSI) and the details on how to obtain CSI can be referred to [26].

In the following part, we will clarify how to realize the legacy transmission and IoT nodes' transmissions in our considered network. Specifically, for the legacy transmission, the whole transmission block, denoted by T, can be divided into two periods according to whether the LT transmits the legacy signal or not. The two periods are the busy period and the idle period. Let β $(0 \le \beta \le 1)$ denote the channel busy ratio. At the busy period with the duration of βT, the LT transmits the legacy signal to the LR; i.e., the channel is in the busy period. Accordingly, the LR can receive the legacy signal and obtain the legacy information by decoding the received signal. At the same time, each IoT node can harvest energy from the legacy signal and backscatter the received signal to the information gateway. At the idle period with the duration of $(1-\beta)T$, the LT stops information transmission; i.e., the channel is in the idle period, while each IoT node can use its harvested energy to transmit its information to the information gateway.

Accordingly, for the IoT nodes' transmissions, the whole time block can also be divided into two phases, which are the backscatter communication phase and the active transmission phase. The backscatter communication phase is included in the busy period. In this phase, each IoT node take turns to perform backscatter communications so as to avoid the cochannel interference among different IoT nodes. Therefore, the backscatter communication phase can be further divided into K subphases. Let $\tau_k T$ with $\sum_{k=1}^{K} \tau_k T \le \beta T$

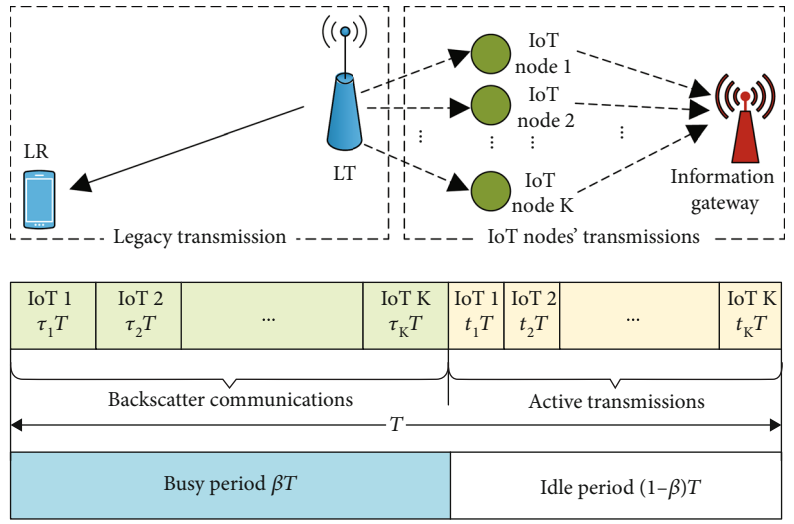


Figure 1: System model.

denote the duration of the kth subphase, where the kth IoT performs backscatter communication and the others keep harvesting energy in order to harvest energy as much as possible. The active transmission phase is included in the idle period. Likewise, in order to avoid the interference from other IoT nodes, the whole active transmission phase is also divided into K subphases. Let $t_k T$ with $\sum_{k=1}^{K} t_k T \le (1-\beta)T$ be the duration of the kth subphase in this phase, in which the kth IoT uses its harvested energy to transmit information and the others keep idle.

Let P_s denote the transmit power of the LT and $s(n)$ be the nth symbol to be transmitted by the LT with normalized power. Then, the transmitted signal at the LT is given by $x(n) = \sqrt{P_s}s(n)$. Denote $c(n)$ as the nth transmitted symbol at the IoT node with $\mathbb{E}[|c(n)|^2] = 1$ and $\alpha_k \in (0,1)$ as the normalized reflection coefficient at the kth IoT node, where a part of the received signal with ratio α_k is backscattered to the information gateway and the rest is flowed to the RF energy harvesting module. Then, the received signal at the LR in the kth subphase of the backscatter communication phase can be expressed as

$$y_{k,R} = f_0 x(n) + \sqrt{\varepsilon\alpha_k} f_k h_k c(n) x(n) + u_R(n), \tag{1}$$

where f_0 is the channel coefficient of the LT-LR link, $\varepsilon \in (0,1)$ is the backscatter efficiency, h_k is the channel coefficient between the LT and the kth IoT node ($k \in \{1,2,\cdots,K\}$), f_k denotes the channel coefficient between the kth IoT node and the LR, and $u_R(n)$ is the additive white Gaussian noise (AWGN) at the LR. Correspondingly, the signal-to-interference-plus-noise ratio (SINR) for decoding $s(n)$ at the LR is given by

$$\gamma_{k,R} = \frac{P_s|f_0|^2}{\varepsilon\alpha_k|f_k|^2|h_k|^2 P_s + \sigma^2 W}, \tag{2}$$

where σ^2 denotes the power spectral density and W is the system bandwidth.

For the kth IoT node in the kth subphase of the backscatter communication phase, the received signal can be represented as

$$y_k(n) = g_0 x(n) + \sqrt{\varepsilon\alpha_k} g_k h_k c(n) x(n) + u_G(n), \tag{3}$$

where g_0 denotes the channel coefficient of the LT-the information gateway link, g_k is the channel coefficient of the kth IoT node-the information gateway link, and $u_G(n)$ is the AWGN at the information gateway.

Obviously, the backscatter communication suffers the interference from the LT's transmission, which leads to a poor performance, since the backscattered signal is much weaker than the legacy signal due to the double-fading effect in the backscattered signal. To address this issue and improve the performance of the backscatter communication, the successive interference cancellation (SIC) is employed to decode $c(n)$ at the information gateway. Specifically, the information gateway will decode $s(n)$ first and subtract it from the received signal before decoding $c(n)$. Thus, the SINR for decoding $s(n)$ is given by

$$\gamma_{1k} = \frac{P_s|g_0|^2}{\varepsilon\alpha_k|g_k|^2|h_k|^2 P_s + W\sigma^2}. \tag{4}$$

When $s(n)$ is decoded successfully, i.e., $\gamma_{1k} \ge \gamma_{\min}$, where $\gamma_{\min}$ is the minimum required signal-to-noise ratio (SNR) to decode $s(n)$, the SNR for decoding $c(n)$ is given by

$$\gamma_{2k} = \frac{\varepsilon\alpha_k|g_k|^2|h_k|^2 P_s}{W\sigma^2}. \tag{5}$$

1: Initialize the maximum iterations $I_{\max}$ and the maximum error tolerance ε;
2: Set the maximum energy efficiency $q = 0$ and iteration index $l = 0$;
3: **repeat**
4: Solve $\mathbf{P}_4$ with a given q and obtain the optimal solution (τ^+, t^+, P^+);
5: **if** $\sum_{k=1}^{K} C_k^{(2)+} - q(\sum_{k=1}^{K} P_{c,k}\tau_k^+ + \sum_{k=1}^{K}(y_k^+ + p_{c,k}t_k^+)) < \varepsilon$ **then**
6: Flag = 1;
7: Set $\tau^* = \tau^+$, $t^* = t^+$, $P^* = y^+/t^+$, $q^* = \sum_{k=1}^{K} C_k^{(2)+}/\sum_{k=1}^{K} P_{c,k}\tau_k^+ + \sum_{k=1}^{K}(y_k^+ + p_{c,k}t_k^+)$ and return;
8: **else**
9: Set $q = \sum_{k=1}^{K} C_k^{(2)+}/\sum_{k=1}^{K} P_{c,k}\tau_k^+ + \sum_{k=1}^{K}(y_k^+ + p_{c,k}t_k^+)$, $l = l + 1$;
10: Flag = 0;
11: **end if**
12: **until** Flag = 1 or $l = I_{\max}$

Algorithm 1: Dinkelbach-based iterative algorithm for $\mathbf{P}_2$.

According to (5), the achievable throughput of the kth IoT node via the backscatter communication can be computed as

$$\begin{aligned} C_k^b &= W\tau_k T \log_2(1+\gamma_{2k}) \\ &= W\tau_k T \log_2\left(1 + \frac{\varepsilon\alpha_k |g_k|^2 |h_k|^2 P_s}{W\sigma^2}\right). \end{aligned} \tag{6}$$

Please note that $c(n)s(n)$ may not follow the Gaussian distribution. However, for analytical tractability, we assume that $c(n)s(n)$ follows the Gaussian distribution such that the throughput of the backscatter communication can be approximated by using Shannon capacity [14–17].

For energy harvesting, a more practical nonlinear energy harvesting model [26] is considered here to be more practical. Please note that our proposed Algorithm 1 can be used for any nonlinear energy harvesting model. Then, the harvested energy at the kth IoT node in this subphase is given by

$$E_k^b = \frac{E_{\max}\left(1 - \exp\left(-a(1-\alpha_k)P_s|h_k|^2\right)\right)}{1 + \exp\left(-a(1-\alpha_k)P_s|h_k|^2 + ab\right)}\tau_k T, \tag{7}$$

where $E_{\max}$ denotes the maximum harvestable power when the circuit is saturated and a and b represent the fixed parameters determined by the resistance, capacitance, and diode turn-on voltage. Let $P_{c,k}$ be the circuit power consumption of the kth IoT node when backscattering. Then, the constraint $E_k^b \geq \tau_k T P_{c,k}$ should be satisfied so that the harvested energy is enough for the circuit operation and the kth IoT node can backscatter signals to the information gateway. We note that the IoT node can also harvest energy from the signal transmitted by other IoT nodes, but it is too much smaller compared with that of LT. Thus, in this work, we assume that each IoT node only harvests energy from the signals from the PT.

Note that the harvested energy of the kth IoT node for the other subphases is used to support its active transmission in the active transmission phase. Thus, the total harvested energy for the active transmission can be calculated as

$$E_k^a = \frac{E_{\max}\left(1 - \exp\left(-aP_s|h_k|^2\right)\right)}{1 + \exp\left(-aP_s|h_k|^2 + ab\right)}(\beta - \tau_k)T. \tag{8}$$

For the kth IoT node in the kth subphase of the active transmission phase, its achievable throughput is given by

$$C_k^a = W t_k T \log_2\left(1 + \frac{P_k|g_k|^2}{W\sigma^2}\right), \tag{9}$$

where P_k is the transmit power of the kth IoT node during the active transmission phase.

3. Energy-Efficient Resource Allocation

In this section, with the practical nonlinear energy harvesting model considered, we aim to maximize the energy efficiency of all the IoT nodes in the investigated network by jointly optimizing the backscattering time $[\tau_1,\cdots,\tau_K]$ and reflection coefficients $[\alpha_1,\cdots,\alpha_K]$ of all IoT nodes in the backscatter communication phase and the transmit power $[P_1,\cdots,P_K]$ and time $[t_1,\cdots,t_K]$ of all IoT nodes in the active transmission phase, subject to the energy causality constraint, the minimum SNR requirements, etc.

3.1. Problem Formulation. The goal of this work is to maximize the energy efficiency of all the IoT nodes, which is defined as the ratio of the total achievable throughput of all the IoT nodes, denoted by C_{sum}, to all the IoT nodes' energy consumption, namely, E_{sum}. In the following part, we aim to determine the expressions of C_{sum} and E_{sum}. Based on (6) and (9), we can determine the expression of C_{sum} as

$$\begin{aligned} C_{\text{sum}} &= \sum_{k=1}^{K}\left(C_k^b + C_k^a\right) \\ &= \sum_{k=1}^{K}\left(W\tau_k T\log_2\left(1 + \frac{\varepsilon\alpha_k|g_k|^2|h_k|^2P_s}{W\sigma^2}\right)\right. \\ &\quad \left. + W t_k T\log_2\left(1 + \frac{P_k|g_k|^2}{W\sigma^2}\right)\right). \end{aligned} \tag{10}$$

As for the total energy consumption of all the IoT nodes, E_{sum} consists of the energy consumed in the backscatter communication phase and the energy consumption in the active transmission phase. Let $p_{c,k}$ denote the constant circuit power consumption at the kth IoT node in the active transmission phase. Then, E_{sum} can be computed as

$$E_{\text{sum}} = \sum_{k=1}^{K} P_{c,k}\tau_k T + \sum_{k=1}^{K} \left(P_k + p_{c,k}\right) t_k T. \tag{11}$$

Therefore, the energy efficiency maximization problem can be formulated as

$$\begin{aligned}
\mathbf{P}_1 : \quad & \max_{(\boldsymbol{\tau},\mathbf{t},\boldsymbol{\alpha},\mathbf{P})} \frac{C_{\text{sum}}}{E_{\text{sum}}} \\
\text{s.t. :} \quad & \text{C1} : \sum_{k=1}^{K} \tau_k T \le \beta T,\ \sum_{k=1}^{K} t_k T \le (1-\beta)T \\
& \text{C2} : \gamma_{1k} \ge \gamma_{\min}, \quad k \in \{1,\cdots,K\} \\
& \text{C3} : \gamma_{k,R} \ge \gamma_{\min}, \quad k \in \{1,\cdots,K\} \\
& \text{C4} : E_k^b \ge \tau_k T P_{c,k}, \quad k \in \{1,\cdots,K\} \\
& \text{C5} : \left(P_k + p_{c,k}\right) t_k T \le E_k^a, \quad k \in \{1,\cdots,K\} \\
& \text{C6} : \sum_{k=1}^{K} \left(C_k^b + C_k^a\right) \ge C_{\min} \\
& \text{C7} : 0 \le \alpha_k \le 1, \quad k \in \{1,\cdots,K\} \\
& \text{C8} : \tau_k \ge 0, t_k \ge 0, P_k > 0, \quad k \in \{1,\cdots,K\},
\end{aligned} \tag{12}$$

where $\boldsymbol{\tau} = [\tau_1,\cdots,\tau_K]$, $\mathbf{t} = [t_1,\cdots,t_K]$, $\boldsymbol{\alpha} = [\alpha_1,\cdots,\alpha_K]$, $\mathbf{P} = [P_1,\cdots, P_K]$, and $C_{\min}$ is the total minimum required throughput for all IoT nodes.

In $\mathbf{P}_1$, constraint C2 is the necessary condition for effective backscatter transmission to ensure that the SIC can be performed successfully at the information gateway. Constraint C3 is to ensure that the LR can decode $s(n)$ successfully under the IoT nodes' interferences. Constraints C4 and C5 are the energy causality constraints, which ensure that the energy consumption of each IoT node in the backscatter communication and active transmission phases cannot be larger than its harvested energy. Constraint C6 ensures the total minimum throughput requirement for all IoT nodes.

It is obvious that problem $\mathbf{P}_1$ is a nonconvex fractional optimization problem and is very challenging to solve since the coupling relationships among different optimization variables, i.e., P_k and t_k, τ_k, and α_k, exist in both the objective function and the constraints, leading to a nonconvex objective function and several nonconvex constraints, e.g., C4, C5, and C6.

3.2. Solution to $\mathbf{P}_1$. In order to address $\mathbf{P}_1$, Proposition 1 is provided to obtain the optimal reflection coefficients as follows.

Proposition 1. *For any given system parameters and optimization variables, the optimal reflection coefficient for the kth IoT node is given by* $\alpha_k^* = \alpha_{\max}^k, k \in \{1,\cdots,K\}$, *where* $\alpha_{\max}^k$ *is given by* $\alpha_{\max}^k = \min\left(\left(P_s|g_0|^2 - \gamma_{\min} W\sigma^2\right)/\left(\varepsilon|g_k|^2 |h_k|^2 P_s \gamma_{\min}\right), \left(P_s|f_0|^2 - \gamma_{\min}\sigma^2 W\right)/\left(\varepsilon|f_k|^2|h_k|^2 P_s\gamma_{\min}\right), 1 - \left(1/aP_s|h_k|^2\right) \ln\left(\left(E_{\max} + P_{c,k}e^{ab}\right)/\left(E_{\max} - P_{c,k}\right)\right)\right)$.

Proof. When $\boldsymbol{\tau}$, $\mathbf{t}$, and $\mathbf{P}$ are given, it is obvious that the objective function of $\mathbf{P}_1$ increases with the increasing of α_k. On the other hand, by combining constraints C2, C3, and C4, the upper bound for α_k, denoted by $\alpha_{\max}^k$, is obtained. Thus, in order to achieve the maximum energy efficiency of all the IoT nodes, we have $\alpha_k^* = \alpha_{\max}^k, k \in \{1,\cdots,K\}$.

The proof is completed. □

Substituting $\alpha_k^* = \alpha_{\max}^k, k \in \{1,\cdots,K\}$ in to $\mathbf{P}_1$, the optimization problem $\mathbf{P}_1$ can be revised as

$$\begin{aligned}
\mathbf{P}_2 : \quad & \max_{(\boldsymbol{\tau},\mathbf{t},\mathbf{P})} \frac{\sum_{k=1}^{K} C_k^{(1)}}{\sum_{k=1}^{K} P_{c,k}\tau_k T + \sum_{k=1}^{K}\left(P_k + p_{c,k}\right) t_k T} \\
\text{s.t. :} \quad & \text{C1 ; C8 ;} \\
& \text{C5} - 1 : \left(P_k + p_{c,k}\right) t_k \le B_k(\beta - \tau_k), \quad k \in \{1,\cdots,K\} \\
& \text{C6} - 1 : \sum_{k=1}^{K} C_k^{(1)} \ge C_{\min},
\end{aligned} \tag{13}$$

where $C_k^{(1)} = W\tau_k T \log_2(1 + A_k\alpha_{\max}^k) + Wt_k T\log_2(1 + (P_k|g_k|^2/W\sigma^2))$, $A_k = \varepsilon|g_k|^2|h_k|^2 P_s/W\sigma^2$, and $B_k = (E_{\max}(1 - \exp(-aP_s|h_k|^2)))/[1 + \exp(-aP_s|h_k|^2 + ab)]$.

In order to tackle the nonconvex fractional objective function in $\mathbf{P}_2$, the Dinkelbach method is used to obtain the optimal solutions. In particular, let q^* and $*$ denote the maximum energy efficiency and the optimal solutions for the optimization variables of $\mathbf{P}_2$. Based on the generalized fractional programming theory [27], the maximum energy efficiency q^* is obtained if and only if the following equation holds:

$$\begin{aligned}
& \max_{(\boldsymbol{\tau},\mathbf{t},\mathbf{P})} \sum_{k=1}^{K} C_k^{(1)} - q^*\left(\sum_{k=1}^{K} P_{c,k}\tau_k T + \sum_{k=1}^{K}\left(P_k + p_{c,k}\right) t_k T\right) \\
& = \sum_{k=1}^{K} C_k^{(1)*} - q^*\left(\sum_{k=1}^{K} P_{c,k}\tau_k^* T + \sum_{k=1}^{K}\left(P_k^* + p_{c,k}\right) t_k^* T\right) \\
& = 0,
\end{aligned} \tag{14}$$

where $C_k^{(1)*} = W\tau_k^* T \log_2(1 + A_k\alpha_{\max}^k) + Wt_k^* T\log_2(1 + (P_k^*|g_k|^2/W\sigma^2))$.

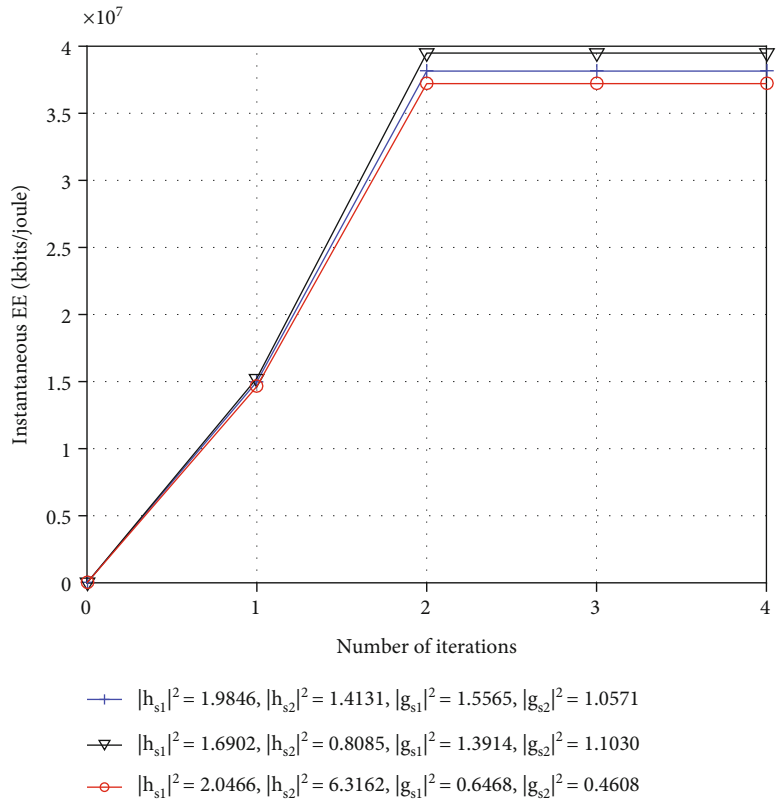


FIGURE 2: The convergence of the proposed algorithm.

Accordingly, problem $\mathbf{P}_2$ can be transformed by solving the following problem $\mathbf{P}_3$ with a given parameter q, given by

$$\begin{aligned} \mathbf{P}_3: \quad & \max_{(\boldsymbol{\tau},\mathbf{t},\mathbf{P})} \sum_{k=1}^{K} C_k^{(1)} - q\left(\sum_{k=1}^{K} P_{c,k}\tau_k T + \sum_{k=1}^{K} \left(P_k + p_{c,k}\right) t_k T\right) \\ \text{s.t.}: \quad & \text{C1, C5} - 1, \text{C6} - 1, \text{C8}, \end{aligned} \tag{15}$$

where q will be updated in each iteration.

As for $\mathbf{P}_3$, it is more tractable than $\mathbf{P}_2$, but it is still a non-convex problem due to the coupling relationship between P_k and t_k. To address this problem, we introduce a series of auxiliary variables, denoted by y_k, into $\mathbf{P}_3$.

By letting $y_k = P_k t_k, \forall k$, $\mathbf{P}_3$ can be transformed as

$$\begin{aligned} \mathbf{P}_4: \quad & \max_{(\boldsymbol{\tau},\mathbf{t},\mathbf{y})} \sum_{k=1}^{K} C_k^{(2)} - q\left(\sum_{k=1}^{K} P_{c,k}\tau_k T + \sum_{k=1}^{K} \left(y_k + p_{c,k} t_k\right) T\right) \\ \text{s.t.}: \quad & \text{C1, C8} - 1: \tau_k \geq 0, t_k \geq 0, y_k > 0, \quad k \in \{1,\cdots,K\} \\ & \text{C5} - 2: y_k + p_{c,k} t_k \leq B_k(\beta - \tau_k), \quad k \in \{1,\cdots,K\} \\ & \text{C6} - 2: \sum_{k=1}^{K} C_k^{(2)} \geq C_{\min}, \end{aligned} \tag{16}$$

where $y = [y_1,\cdots,y_K]$ and $C_k^{(2)} = W\tau_k T \log_2(1 + A_k \alpha_{\max}^k) + W t_k T \log_2(1 + (y_k |g_k|^2 / t_k W\sigma^2))$.

It is easy to prove that $\mathbf{P}_4$ is a convex problem and can be efficiently solved by many existing convex tools, i.e., the Lagrange duality method and the interior-point method. In the following part, the Lagrange duality method is used to obtain the optimal solutions to $\mathbf{P}_4$. Let P_k^+ denote the optimal transmit power of the kth IoT node during the active transmission phase, and it can be determined by Proposition 2.

Proposition 2. *In the cognitive wireless-powered hybrid active-passive communication network, the optimal transmit power P_k^+ of the kth IoT node during the active transmission phase for maximizing the energy efficiency of all the IoT nodes is given by*

$$P_k^+ = \left[\frac{T(1+\lambda)}{(qT + \mu_k)\ln 2} - \frac{1}{D_k}\right]^+, \tag{17}$$

where $D_k = |g_k|^2 / W\sigma^2$ and $\mu_k \geq 0$ and $\lambda \geq 0$ are the dual variables corresponding to C5 − 2 and C6 − 2, respectively.

Proof. See the appendix. □

Substituting P_k^+ into $\mathbf{P}_4$, we observe that $\mathbf{P}_4$ is a linear programming problem with respect to t_k and τ_k. Thus, standard linear optimization tools, i.e., the simplex method, can be employed to obtain the optimal solutions efficiently. It is

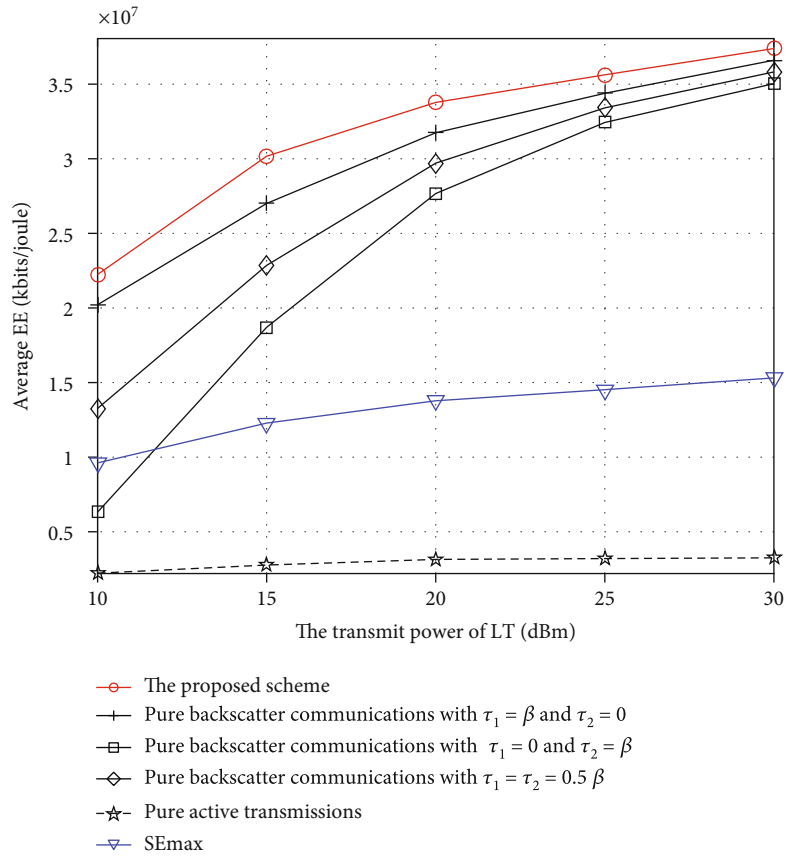


Figure 3: Average energy efficiency versus the transmit power of LT.

worth noting that $\alpha_{\max}^k$ may be less than 0. In such case, the IoT node cannot backscatter signals to the information gateway since the harvested energy is not enough for the circuit operation and $C_k^b = 0$. In order to achieve the maximum energy efficiency, we have $\tau_k^* = 0$.

3.3. Iterative Algorithm. In this subsection, a Dinkelbach-based iterative algorithm is proposed to obtain the optimal solutions to $\mathbf{P}_2$. The detailed process of the proposed algorithm is shown in Algorithm 1. Specifically, in each iteration, $\mathbf{P}_4$ with a given q should be optimally solved to obtain the optimal solution, denoted by (τ^+, t^+, P^+). Let ε denote the error tolerance. If the stop condition $\sum_{k=1}^K C_k^{(2)+} - q(\sum_{k=1}^K P_{c,k} \tau_k^+ T + \sum_{k=1}^K (y_k^+ + p_{c,k} t_k^+) T) < \varepsilon$ holds, then we have $\tau^* = \tau^+$, $t^* = t^+$, and $P^* = P^+$. Otherwise, q is updated as $q = \sum_{k=1}^K C_k^{(2)+} / (\sum_{k=1}^K P_{c,k} \tau_k^+ T + \sum_{k=1}^K (y_k^+ + p_{c,k} t_k^+) T)$. Then, repeat the above steps until the stop condition is satisfied.

4. Simulations

In this section, we verify the performance of the cognitive wireless-powered hybrid active-passive communication under the proposed scheme. Let d_1 denote the distance between the LT and the information gateway. d_{0k} and d_{k1} are denoted as the distances of the LT-the kth IoT node link and the kth IoT node-the information gateway link, respectively. In the following part, we present the basic parameter settings. We set $K = 2$, the path loss exponent $\varsigma = 3$, $P_s = 30$ dBm, $W = 10$ kHz, $\beta = 0.7$, $T = 1$ s, $\gamma_{\min} = 0$ dB, $\sigma^2 = -150$ dBm/Hz, $P_{c,1} = P_{c,2} = 10\,\mu$W, $p_{c,1} = p_{c,2} = 50\,\mu$W, $\varepsilon = 0.8$, $C_{\min} = 50$ kbps, $E_{\max} = 240\,\mu$W, $a = 5000$, and $b = 0.0002$. The distances are set as $d_{01} = d_1 = 5$ meters, $d_{02} = 8$ meters, and $d_{11} = d_{12} = 1$ meter.

Figure 2 shows the convergence of the proposed algorithm, where $|h_{sk}|^2$ and $|g_{sk}|^2$ denote the small fadings of the LT-the kth IoT node link and the kth IoT node-the information gateway link, respectively. It can be seen that with any given channel settings, the proposed algorithm can always converge to the optimal energy efficiency after only two iterations, which indicates that our proposed algorithm is computationally efficient and has a fast convergent rate.

Figure 3 shows the average energy efficiency of all the IoT nodes versus the transmit power of the LT P_s. In order to demonstrate the superiority of the proposed scheme, we compare the energy efficiency under the proposed scheme with that under three other schemes, which are the pure backscatter communications with $t_k = 0$ (denoted as pure backscatter communications), the pure active transmissions with $\tau_k = 0$ (denoted as pure active transmissions), and the throughput maximization (denoted as SEmax), respectively. As for the pure backscatter communications, we consider three ways for allocating the backscatter time which are (1)

$\tau_1 = \beta$ and $\tau_2 = 0$; (2) $\tau_2 = \beta$ and $\tau_1 = 0$; and (3) $\tau_1 = \tau_2 = 0.5\beta$. For the pure active transmissions, the transmit time and power for each IoT node are optimized to maximize the energy efficiency of all the IoT nodes under the same constraints as $\mathbf{P}_1$. As for the throughput maximization, this scheme is optimized to maximize the total achievable throughput of all the IoT nodes under the same constraints as $\mathbf{P}_1$.

From this figure, we can see that the average energy efficiency of all the IoT nodes under all the schemes will increase with the increasing of P_s. The reasons are as follows. With a larger P_s, the received legacy signal at each IoT node is stronger and the harvested energy of each IoT node increases, bringing a higher throughput achieved by all the IoT nodes. Since the total throughput grows faster than the growth of the total energy consumption, all the curves show an upward trend. By comparisons, it can be observed that the proposed scheme always achieves the best performance in terms of the energy efficiency of all the IoT nodes among these schemes. This is because the proposed scheme provides more flexibility to utilize the resource efficiently to achieve the maximum energy efficiency. More interestingly, we observe that the energy efficiency under the pure active transmissions is lowest compared with the other schemes. This is because compared to the pure backscatter communications, the pure active transmissions need more energy to achieve the same throughput.

5. Conclusions

In this work, we have investigated the energy efficiency maximization for a cognitive wireless-powered hybrid active-passive communication network, where multiple IoT nodes transmit information to the information gateway via the backscatter communications and the active transmissions. Specifically, an optimization problem was formulated to maximize the energy efficiency of all the IoT nodes by jointly optimizing the backscatter time and reflection coefficients, the transmit time, and power of all the IoT nodes, subject to the energy causality constraint, the minimum SNR requirements, etc. The formulated problem was a highly nonconvex fractional optimization problem. In order to solve it, we proposed an iterative algorithm to obtain the optimal solutions. Simulation results have verified the fast convergence of the proposed algorithm and demonstrated the superiority of our proposed scheme in terms of the energy efficiency of all the IoT nodes.

Appendix

The Lagrangian function of $\mathbf{P}_4$ is given by

$$\begin{aligned}\mathscr{L} = &\sum_{k=1}^{K} C_k^{(2)} - q\left(\sum_{k=1}^{K} P_{c,k}\tau_k T + \sum_{k=1}^{K}\left(y_k + p_{c,k}t_k\right)T\right)\\ &+ \sum_{k=1}^{K}\mu_k\left[B_k(\beta - \tau_k) - y_k - p_{c,k}t_k\right] + \lambda\left(\sum_{k=1}^{K} C_k^{(2)} - C_{\min}\right)\\ &+ \upsilon\left(\beta - \sum_{k=1}^{K}\tau_k\right) + \rho\left(1 - \beta - \sum_{k=1}^{K} t_k\right),\end{aligned} \tag{A.1}$$

where μ_k, λ, υ, and ρ are nonnegative Lagrangian multipliers. Then, the first-order derivative of the Lagrangian with respect to y_k can be given by

$$\frac{\partial \mathscr{L}}{\partial y_k} = \frac{(1+\lambda)TD_k t_k}{(t_k + D_k y_k)\ln 2} - qT - \mu_k, \tag{A.2}$$

where $D_k = |g_k|^2/W\sigma^2$. By letting $\partial\mathscr{L}/\partial y_k = 0$, we have

$$P_k^+ = \frac{y_k^o}{t_k^o} = \left[\frac{T(1+\lambda)}{(qT+\mu_k)\ln 2} - \frac{1}{D_k}\right]^+. \tag{A.3}$$

Therefore, Proposition 2 is obtained.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: a survey on enabling technologies, protocols, and applications," *IEEE Communications & Surveys Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[2] Q. Li, "Sum-throughput maximization in backscatter communication-based cognitive networks," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7768588, 11 pages, 2022.

[3] Y. Xu, H. Sun, and Y. Ye, "Distributed resource allocation for SWIPT-based cognitive ad-hoc networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 4, pp. 1320–1332, 2021.

[4] European Commision, "Identification and quantification of key socio-economic data to support strategic planning for the introduction of 5G in Europe," https://data.europa.eu/doi/10.2759/037871.

[5] Z. Qin, X. Zhou, L. Zhang, Y. Gao, Y.-C. Liang, and G. Y. Li, "20 years of evolution from cognitive to intelligent communications," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 6–20, 2020.

[6] Y. Ye, Y. Li, G. Lu, and F. Zhou, "Improved energy detection with Laplacian noise in cognitive radio," *IEEE Systems Journal*, vol. 13, no. 1, pp. 18–29, 2019.

[7] S. Lee, R. Zhang, and K. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4788–4799, 2013.

[8] A. Shome, A. K. Dutta, and S. Chakrabarti, "BER performance analysis of energy harvesting underlay cooperative cognitive radio network with randomly located primary users and secondary relays," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4740–4752, 2021.

[9] A. F. Tayel, S. I. Rabia, A. H. A. El-Malek, and A. M. Abdelrazek, "An optimal policy for hybrid channel access in cognitive radio networks with energy harvesting," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11253–11265, 2020.

[10] L. Ni, X. Da, H. Hu, M. Zhang, and K. Cumanan, "Outage constrained robust secrecy energy efficiency maximization for eh cognitive radio networks," *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 363–366, 2020.

[11] Z. Ding, R. Schober, and H. V. Poor, "No-pain no-gain: DRL assisted optimization in energy-constrained CR-NOMA networks," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5917–5932, 2021.

[12] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: a contemporary survey," *IEEE Communications & Surveys Tutorials*, vol. 20, no. 4, pp. 2889–2922, 2018.

[13] S. H. Kim and D. I. Kim, "Hybrid backscatter communication for wireless-powered heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6557–6570, 2017.

[14] Y. Ye, L. Shi, R. Hu, and G. Lu, "Energy-efficient resource allocation for wirelessly powered backscatter communications," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1418–1422, 2019.

[15] Y. Ye, L. Shi, X. Chu, and G. Lu, "Total transmission time minimization in wireless powered hybrid passive-active communications," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, pp. 1–5, Helsinki, Finland, 2021.

[16] S. Fu, P. Jiang, and C. Ding, "Wireless powered hybrid backscatter-active communications with hardware impairments," *Physical Communication*, vol. 52, pp. 1–9, 2022.

[17] Y. Ye, L. Shi, X. Chu, and G. Lu, "Throughput fairness guarantee in wireless powered backscatter communications with HTT," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 449–453, 2021.

[18] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 2876–2891, 2022.

[19] X. Gao, L. Shi, and G. Lu, "Throughput fairness in cognitive backscatter networks with residual hardware impairments and a nonlinear EH model," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 12, p. 16, 2022.

[20] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "The tradeoff analysis in RF-powered backscatter cognitive radio networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Washington, DC, USA, 2016.

[21] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Ambient backscatter: a new approach to improve network performance for RF-powered cognitive radio networks," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3659–3674, 2017.

[22] D. T. Hoang, D. Niyato, P. Wang, and D. I. Kim, "Optimal time sharing in rf-powered backscatter cognitive radio networks," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, France, 2017.

[23] B. Lyu, H. Guo, Z. Yang, and G. Gui, "Throughput maximization for hybrid backscatter assisted cognitive wireless powered radio networks," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2015–2024, 2018.

[24] Y. Zhuang, X. Li, H. Ji, H. Zhang, and V. C. M. Leung, "Optimal resource allocation for rf-powered underlay cognitive radio networks with ambient backscatter communication," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 15216–15228, 2020.

[25] L. Shi, R. Q. Hu, J. Gunther, Y. Ye, and H. Zhang, "Energy efficiency for RF-powered backscatter networks using HTT protocol," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13932–13936, 2020.

[26] Y. Ye, L. Shi, X. Chu, and G. Lu, "On the outage performance of ambient backscatter communications," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7265–7278, 2020.

[27] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1723740, 7 pages
https://doi.org/10.1155/2022/1723740

Research Article

Real-Time Monitoring of College Sports Dance Competition Scenes Using Deep Learning Algorithms

Fei Yang,[1,2] GeMuZi Wu,[2] and HongGang Shan[2]

[1]*Graduate School, Namseoul University, Cheonan, Republic of Korea*
[2]*Department of Sports Art, Hebei Institute of Physical Education, Hebei, China*

Correspondence should be addressed to HongGang Shan; 2008006@hepec.edu.cn

Received 22 November 2021; Revised 28 March 2022; Accepted 4 May 2022; Published 14 June 2022

Academic Editor: Liqin Shi

In order to improve the real-time detection effect, therefore, a research on real-time scene detection of sports dance competition based on deep learning is proposed. The collected scene image is grayed by using the weighted average method, and the best image interpolation is calculated by using the deep learning method, so as to realize the smooth processing of sawtooth and mosaic information generated by panoramic mapping. After selecting the cube model, the processed scene information is projected to the visual plane to construct the panorama of the competition scene. Finally, combined with the three-frame difference, the changes between adjacent image frames are calculated to obtain the moving target. The test results show that the motion detection accuracy of professional dancers can reach more than 75.0% and that of amateur dancer can reach more than 64.2%.

1. Introduction

With the development of modern technology, digital media technology can integrate text, images, sound, video, etc. through computers. The logical relationship between them is established through the processes of sampling quantization, editing and modification, and encoding and compression. With its incomparable advantages of convenience, accuracy, high efficiency, convenient storage, and easy modification, the computer constantly refreshes various records. In recent years, with the rapid development of multimedia technology and the continuous upgrading of software, the wide application of computer in the field of design and performance has won broad prospects. Multimedia technology has not only brought a revolution to the stage visual performance in a new form but also brought great changes to people's aesthetic ideas. With the wide application of multimedia technology, the art of dance beauty is breaking through the limitations and closure of the traditional manual era. The stage is designed as a complex of space-time art, including the temporality and hearing of literature and music, as well as the spatiality and vision of painting and architecture [1–3]. The function of dance art should not only expand the visual scope of the audience and editors but also expand people's thinking ability. The important role of computer technology in dance beauty creation and graphic design provides sports dance practitioners with new ideas. Computers can become a new "stage language." People also have enough reasons to believe that the combination of computer and sports dance will stimulate more creative passion and artistic spirit. Due to the late start of sports dance in China, there is an urgent need to improve the competitive level, teachers, teaching, and scientific research [4, 5]. Sports dance provides a positive and healthy form of sports for college students to establish friendship, cultivate personality, exchange feelings, strengthen physique, cultivate sentiment, shape body shape, and improve skills, which is deeply loved by college students. According to the results of a questionnaire in a college, a survey was conducted on the mental health of 2000 students participating in sports dance before and after training. The results showed that sports dance had a significant effect on shaping healthy personality,

strong physique, and normal interpersonal relationship. However, the current situation is that the students' increasing learning enthusiasm does not match the teaching level and venues in colleges and universities, which specifically shows that there is a general lack of teaching venues in colleges and universities all over the country, and the instruments and equipment cannot meet the needs of sports dance teaching; the quality of sports dance teachers is generally not high, and the level of teachers is uneven; the scientific research level of sports dance lags behind seriously. Especially for the construction of sports dance–related competitions, there is an obvious lag problem. For the competition environment and the participants' action, capture cannot meet the real-time requirements. At the same time, for the controversial parts, there is a lack of more powerful confirmation of video resources [6–8]. This has restrained the development of sports dance to a certain extent. On the one hand, sports dance itself has great limitations on development, which is difficult to form a good development trend through self-competition. On the other hand, due to the lack of effective professional development and management mechanism, most students have low enthusiasm to participate in the major, and the students of their major cannot find good development in the industry after graduation. In this context, it is very necessary to use modern technology to strengthen the real-time detection effect of the game scene. Reference [9] proposes a spatiotemporal sparse RPCA moving object detection algorithm. This method spatiotemporally regularizes the sparse components in the form of the graph Laplacian. Each Laplacian corresponds to a multifeature map constructed over the superpixels of the input matrix. While minimizing the RPCA objective function, the sparse component is used as the feature vector of the spatiotemporal graph Laplacian to realize the detection of scene objects. Reference [10] proposes a 3D motion detection and long-term tracking method. This method provides a new energy function optimization framework for motion pose estimation. It can independently estimate the 3D motion pose of each object. However, considering the complexity of the scene, the effect of real-time target detection is not ideal. Based on the difficulties in sports dance teaching described above, it has increasingly highlighted the necessity and urgency of digital media technology in sports dance. First of all, its advantage is that it can greatly improve the existing competition conditions [11, 12]. At present, the sports venues in most colleges and universities, especially those suitable for sports dance competitions, are very tight, with short opening hours and a large number of participants. Secondly, real-time detection of competition scenes through digital media technology is also of great value to avoid sports injuries. In some intense and confrontational movements, athletes are often vulnerable to injury, and the vast majority of injuries occur in a state beyond the control of athletes. Using digital media technology to learn movements can avoid sports injuries at this stage. If the virtual reality technology is used for practice, students can safely and boldly analyze the actions without considering these problems, which will not cause any harm to people. At the same time, they can also point out the shortcomings of students' actions, put forward suggestions, and score comprehensively, which can improve the evaluation effect of the competition. In addition, it can also break through the limitation of time and space [13–15]. Every athlete is eager to get the support of world-class coaches, and every world-class coach is also eager to promote his training concept to every corner of the world. However, due to time and space constraints, it cannot be realized. With digital media technology, all this will not be a dream. The realization of all this is inseparable from the effective detection of the game scene.

On this basis, a real-time detection method of sports dance competition scene based on deep learning is proposed. Firstly, the collected game scene image information is preprocessed. We take advantage of the information feature extraction of deep learning to improve the detection effect of game scenes. Finally, in a random test of 10 motion dancers, the recognition rate of 5 professional motion dancers was above 75%. The recognition rate of the other 5 amateur sports dancers is also above 64%. Through this research, we also hope to provide valuable help for the development of sports dance competition.

2. Image Preprocessing

When using digital media technology to collect the scene information of sports dance competition, due to the reasons of the equipment itself or the environment in the competition scene, it is very easy to cause incomplete or fuzzy information, so the effect of scene detection is not ideal. Therefore, this paper first preprocesses the collected game scene image information [16–18]. So as to ensure that the panoramic image of the game scene constructed later can meet the detection requirements. Since most real-time image acquisition devices store information in RGB color channel model in. jpg compression format, this paper first grays the image, which can effectively reduce the extraction and processing time of SIFT features, so as to meet the requirements of real-time detection of moving objects.

Image graying is the process of making the R, G, B components of image color equal. Since the value range of R, G, B is 0~255, the gray level is only 256. That is, grayscale images can only show 256 grayscales [19, 20]. The image gray processing method used in this paper is weighted average method. It has the advantage that the importance of different components can be considered. Different weights can be assigned to the three components according to their respective importance. Then, take the weighted average as the grayscale result. Make the image grayscale more in line with the needs of practical applications. According to the importance of its color or other indicators, give different weighted values to the three values of R, G, B, and make them weighted average. Since human eyes have the highest sensitivity to green and the lowest sensitivity to blue, the specific weighting formula is

$$R = G = B = 0.299R + 0.587G + 0.114B. \quad (1)$$

Because the sports dance competition scene contains moving objects, on the basis of real-time detection, it is

required to improve the speed and ensure the accuracy as much as possible. Therefore, we should try to avoid floating-point operation in practical application. Floating point arithmetic is real arithmetic. Because computers can only store integers, floating-point operations are slow and prone to errors. Combined with this demand, this paper adjusts the weighted calculation to the full integer algorithm and scales formula (1) one thousand times to realize the integer operation algorithm. At this time, the image gray processing result is

$$\text{Gray} = \frac{(299R + 587G + 114B + 500)}{1000}, \tag{2}$$

where Gray is the game scene image after gray processing. It should be noted that the accuracy of RGB three-color channel is generally 8-bit accuracy. After scaling it a thousand times, the subsequent corresponding image operations are also 32-bit integer data operations. The division following formula (2) is integer division, and the purpose of adding 500 is to realize rounding. Because the algorithm needs 32-bit operation, the time will increase. Combined with the real-time requirements of sports dance competition scene detection, this paper further processes formula (2), and the final result is

$$\text{Gray} = \frac{(30R + 59G + 11B + 50)}{100}. \tag{3}$$

In this way, the unified preprocessing of dance competition scene image information is realized.

3. Build a Real-Time Panorama of Sports Dance Competition

3.1. Scene Image Interpolation Based on Deep Learning. After the above image graying, in order to have higher accuracy in the later stage of the game scene stitching, this paper carries out cylindrical mapping projection on the sequence images. However, in the actual processing process, it is found that the transformed point coordinates are often not integer coordinate values. If these coordinate values are simply calculated as integers, great errors will be caused, resulting in geometric distortion of the projected image [21, 22]. Therefore, this paper uses image interpolation technology. The algorithm implementation of image interpolation is generally divided into forward mapping method and backward interpolation method. The forward mapping algorithm is to map from the source image to the target image; the backward mapping algorithm is the reverse mapping from the target image to the source image. In the process of pixel transformation, the forward mapping algorithm may be projected to the outside of the image area, resulting in multiple calculations, and the gray value of some pixels of the target image will be repeatedly determined, which not only wastes the calculation time but also sometimes affects the real-time performance. The backward mapping algorithm generates the output image for each pixel without interval. The pixel gray value of each target image is determined by the color values of the pixels of four source images after interpolation algorithm, and then the output image is generated [23–25]. In this paper, the nearest neighbor interpolation algorithm is used to realize this process. In the specific operation process, firstly, the gray value of the input pixel closest to the position mapped to the target image is selected as the interpolation result. After adding geometric transformation, the corresponding coordinate values of the pixels with coordinates (x, y) on the output image on the original image are (u, v). At this time, the formula of the nearest interpolation algorithm can be expressed as

$$\begin{cases} g(x, y) = f(x, y) \\ x = u + 0.5(\text{Take an integer}), \\ y = v + 0.5(\text{Take an integer}) \end{cases} \tag{4}$$

where $g(x, y)$ represents the position information of the target image, and $f(x, y)$ represents the position information closest to the target image. This algorithm is simple and fast. However, this method will cause obvious jagged edges and mosaic in the newly generated image. However, this approach results in noticeable jagged edges in the newly generated image. Therefore, this paper uses the method of deep learning to smooth the sawtooth and mosaic information. In order to obtain the more accurate value of the output image pixel of the fuzzy part, it is not enough to only use the four nearest pixels of the input image pixel as the object of depth learning. Therefore, this paper takes the influence of the fuzzy point to the surrounding 16 nearest pixels as the learning goal. Constructing insert value learning function

$$S(x) = \sin(\pi x)/\pi x, \tag{5}$$

where $S(*)$ represents the learning function. In this way, the values of each point between collected images can be accurately obtained. Use formula (5) to iteratively approximate the best interpolation function. The specific iterative method is

$$S(x) = \begin{cases} 1 - (\lambda + 3)|x|^2 + (\lambda + 2)|x|^3, |x| < 1 \\ -4\lambda + 8\lambda|x| - 5\lambda|x|^2 + \lambda|x|^3, 1 \le |x| \le 2, \\ 0, |x| \ge 2 \end{cases} \tag{6}$$

where λ represents the learning coefficient. Finally, according to the actual image processing effect, the value of λ is -1. In this way, the fast interpolation processing of competition scene information is realized with a small amount of calculation and simple algorithm. At the same time, the influence of other adjacent pixels is considered to ensure that the gray level of the collected real-time scene image maintains obvious continuity, the loss of image quality is minimized, and the phenomenon of image sawtooth and mosaic is avoided.

3.2. Panoramic Construction of Sports Dance Competition. Panorama image generally means that the angle of view of the image is greater than the normal visual angle of human eyes, that is, it is about 90 degrees in the horizontal direction

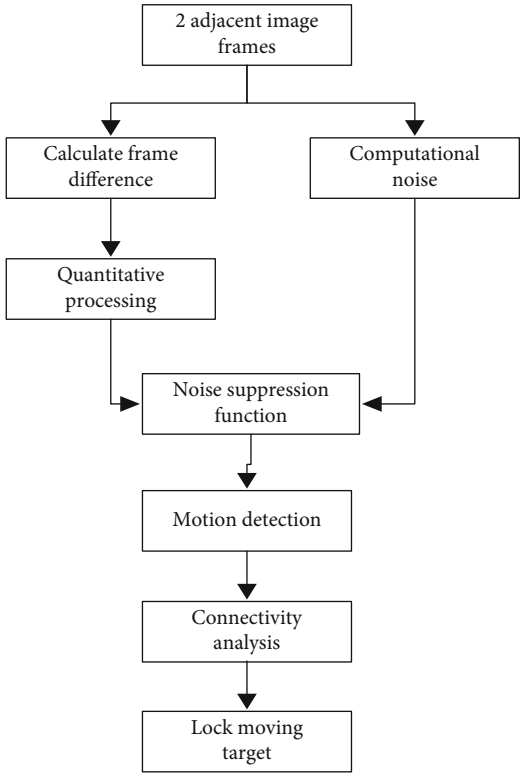


Figure 1: Scene real-time detection method.

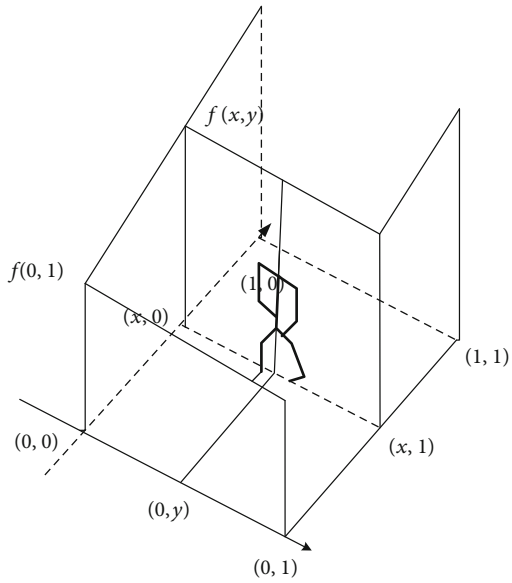


Figure 2: Portrait fitting results.

and about 70 degrees in the vertical direction. In addition to using special equipment to obtain panoramic images, panoramic images can also be obtained through image stitching and stitching technology. Generally, first, the rotating camera is used to collect the sequence video images to ensure that the sequence images are partially overlapped. Then, the mosaic algorithm is used to splice the collected sequence video images to form a panoramic mosaic image.

On the basis of the above image processing, this paper mainly includes several steps in the process of panoramic image stitching and stitching: panoramic image generation, model selection, collecting sequence images and ensuring image overlap, image stitching. For the panorama generation model, this paper selects the cube model to project the collected scene information to the visual plane. The cube model projection allows the user to view the image 180 degrees in the vertical direction and 360 degrees in the horizontal direction. The formed cube panorama is composed of six planes, which are projected onto the six surfaces of the cube and spliced. Projection transformation formula is the key formula of panoramic mosaic image cube projection transformation and inverse transformation method. Therefore, this paper establishes the projection transformation coordinate system. The cube takes the central position point o as the origin of the coordinate system, and R is the variable length of the cube, $\alpha\,(-\pi <= \alpha <= \pi)$ as azimuth and $\beta\,(-\pi/2 <= \beta <= \pi/2)$ as the pitch angle. The coordinate equation of six faces can be expressed as

Left plane

$$y = -\frac{r}{2}\left(-\frac{r}{2} \le x, z \le \frac{r}{2}\right). \tag{7}$$

Right plane

$$y = \frac{r}{2}\left(\frac{r}{2} \le x, z \le \frac{r}{2}\right). \tag{8}$$

Anterior plane

$$x = \frac{r}{2}\left(-\frac{r}{2} \le y, z \le \frac{r}{2}\right). \tag{9}$$

Posterior plane

$$x = -\frac{r}{2}\left(-\frac{r}{2} \le x, z \le \frac{r}{2}\right). \tag{10}$$

Upper plane

$$z = \frac{r}{2}\left(-\frac{r}{2} \le x, y \le \frac{r}{2}\right). \tag{11}$$

Lower plane

$$z = -\frac{r}{2}\left(-\frac{r}{2} \le x, y \le \frac{r}{2}\right). \tag{12}$$

In this way, the collected real-time scene information of the independent sports dance competition venue is transformed into a panoramic image.

4. Scene Real-Time Detection

On the above basis, the scene detection is realized by using the inter frame image difference algorithm of the collected panoramic image. The core of the algorithm is to gray the video image of the current frame, then gray the image of the previous frame or the next frame, and then make the difference

Table 1: Individual test results.

Detection result	Major					Amateur				
	1	2	3	4	5	1	2	3	4	5
Correct quantity	36	38	49	43	39	38	43	58	76	59
Number of errors	12	12	5	6	12	13	25	17	17	25
Accuracy/%	75.0	76.0	90.7	87.8	76.5	74.5	64.2	77.3	81.7	70.2

between the current frame image and the previous frame or the next frame, so as to extract the moving object. Assuming that at time point t, the current frame image is I_t and the previous frame image is I_{t-1}, the moving target can realize scene change detection by comparing the image differences of three adjacent frames. This method has strong adaptability to the dynamic environment, good robustness, small amount of computation, convenient implementation, and can quickly and effectively detect the moving target from the background. The specific algorithm flow is shown in Figure 1.

According to the method shown in Figure 1, the moving object of the difference between frame images is calculated by using the three-frame difference. At the same time, the problem of image blur and edge information loss caused by mean filtering is reduced under the action of noise suppression function, so as to achieve the purpose of accurately detecting the contour information of moving objects in the scene.

5. Scene Detection and Analysis

5.1. Experimental Data. In order to test the effect of the design detection method, the experimental data are collected by Dahua DH-IPC-HF8431E camera. The experiments were carried out on the MATLAB platform. The main control chip is the gt6 stm32f407v chip made by ST company. The processor is arm series. This paper mainly analyzes the key frame data of sports dancers when walking, that is, the movement of sports dancers when stepping, so as to realize the professional evaluation of sports dancers' walking movement. This paper takes video data of 24 professional sports dancers and 13 and nonprofessional sports dancers as experimental samples and selects effective experimental data, including 2004 key frame pictures of professional sports dancers and 958 key frame pictures of nonprofessional sports dancers. 1000 key frame data of professional sports dancers and 700 data of nonprofessional sports dancers are used as the training data set, and the other samples are used as the test data set.

5.2. Data Processing. The human joint point data are fitted with different interpolation coefficients according to the method in this paper, and the results are shown in Figure 2.

The shape of the knee is relatively simple; (0, 0) represents the coordinate origin of the knee joint. (x, y) represents the coordinate interpolation coefficient of the knee joint, and the interpolation coefficient of 1.0 can better fit, while the shape of the ankle and hip is relatively complex. For these two parts, 20 groups of data are randomly selected to calculate the average curve fitting degree of polynomial fitting, and the final interpolation coefficients are 0.6223 and 0.8592, respectively.

5.3. Experimental Results. In this paper, the 10-fold cross-validation method is used to test all samples, and the average accuracy is 71.9%. In order to compare the action recognition results of different sports dancers, this paper will test each sports dancer one by one. In the test set, all key frame data of 5 professional sports dancers and 5 nonprofessional sports dancers are randomly selected for test, and the results are shown in Table 1.

It can be seen from Table 1 that among the results of 10 sports dancers tested separately, the second professional sports dancer has the lowest recognition rate, 75.0% and 76.0%, respectively. The final grade of the professional sports dancer is also the worst among the five professional sports dancers; the second amateur sports dancer has the lowest recognition rate of 64.2%. Conversely, among the five amateur sports dancers, the sports dancer has the best performance and is closest to the professional level. In fact, the final performance of the amateur sports dancer is the best among the five amateur sports dancers. The evaluation of the walking posture of sports dancers is not determined by a certain moment, but by a series of moments. Therefore, the performance of a sports dancer should be evaluated by integrating the movement and posture of all key frames of a sports dancer's walk. In the random test of 10 sports dancers, the recognition rate of 5 professional sports dancers is more than 75%, and the recognition rate of 5 amateur sports dancers is also more than 64%. It shows that the detection method designed in this paper can effectively detect the scene.

6. Conclusion

In this paper, the academic algorithm is applied to practice, aiming at the detection of small targets in the target detection algorithm in deep learning. High detection accuracy is achieved on the experimental data set, and the feasibility of the deep learning algorithm in the field of target detection is also verified. However, in order to be applied to the actual scene, the method still needs to be further improved, especially in the optimization of data and structure; there are some problems and areas that need to be improved:

(1) The success of the deep learning algorithm is attributed to the use of large-scale well-labeled data sets. Although the algorithm of this subject has achieved good detection on the experimental data set, the amount of data in this data set is not very large, and the generalization ability of the training model on this data set remains to be investigated. The focus of this algorithm is on panoramic image

reconstruction, which makes less contribution to the face detection algorithm. In the future, we can refer to multiscale fusion to improve the face detection network, so as to improve the effect of the face detection algorithm itself. At the same time, the image reconstruction strategy of the face detection algorithm proposed in this paper needs two feature extraction and one image reconstruction, which cannot achieve real-time effect in the detection speed. In the follow-up, we need to do research on improving the detection speed

(2) Aiming at the difficulty of small target detection in complex scenes from the perspective of media tool collection and the high missed detection rate of existing algorithms, this paper proposes a detection algorithm for scene repair based on deep learning. On the basis of the data, the annotation of various target perspectives under the panoramic perspective is added. However, due to the insufficient amount of data, the effect still needs to be improved, and the expanded data set still needs to be supplemented. At the same time, the annotation of data also needs to be further standardized

(3) Aiming at the problems existing in small target detection in practical application scenarios, this topic puts forward the construction of panoramic image, applies the theory to practice, and realizes the engineering application of the algorithm. Although it can achieve good target detection effect, the algorithm may get stuck in practical application scenarios. Therefore, the detection speed of the algorithm still needs to be improved. At the same time, in the communication of real-time image acquisition data link, the fixation of some data lines and the security of the platform need to be improved

(4) In this paper, the weighted average method is used to grayscale the collected scene images. We employ a deep learning approach to compute optimal image interpolation. In future research, advanced techniques can be introduced to detect multiple moving objects more accurately

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Wang, "WITHDRAWN: Multimedia Technology Embedded Processor Optimizing Physical Education Teaching Innovation Under Internet Environment," *Microprocessors and Microsystems*, vol. 13, article 104086, 2021.

[2] Y. Li and S. J. Lu, "Research on physical education system model using multimedia technology," *Multimedia Tools and Applications*, vol. 79, no. 15-16, pp. 10461–10474, 2020.

[3] Z. H. U. Jingsi, "Application of the multimedia network technology in the informatization of art education," *International English education research: English version*, vol. 3, pp. 77–79, 2019.

[4] Y. Sun and J. Chen, "Human movement recognition in dancesport video images based on chaotic system equations," *Advances in Mathematical Physics*, vol. 2021, Article ID 5636278, 12 pages, 2021.

[5] N. Keay and A. Rankin, "Infographic. Relative energy deficiency in sport: an infographic guide," *British Journal of Sports Medicine*, vol. 53, no. 20, pp. 1307–1309, 2019.

[6] A. Dc, B. Rp, and C. Rr, "A new fast and accurate heuristic for the Automatic Scene Detection Problem," *Computers & Operations Research*, vol. 136, article 105495, 2021.

[7] T. Khan, R. Sarkar, and A. F. Mollah, "Deep learning approaches to scene text detection: a comprehensive review," *Artificial Intelligence Review*, vol. 54, pp. 3239–3298, 2021.

[8] M. Mandal, V. Dhar, A. Mishra, and S. K. Vipparthi, "3DFR: a swift 3D feature reductionist framework for scene independent change detection," *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1882–1886, 2019.

[9] S. Javed, A. Mahmood, S. Al-Maadeed, T. Bouwmans, and S. K. Jung, "Moving object detection in complex scene using spatiotemporal structured-sparse RPCA," *IEEE Transactions on Image Processing*, vol. 28, no. 2, pp. 1007–1022, 2019.

[10] S. Liu, Y. Wang, F. Dai, and J. Yu, "Simultaneous 3D motion detection, long-term tracking and model reconstruction for multi-objects," *International Journal of Humanoid Robotics*, vol. 16, no. 4, pp. 1950017–1952262, 2019.

[11] Z. Li, "Three-dimensional diffusion model in sports dance video human skeleton detection and extraction," *Advances in Mathematical Physics*, vol. 2021, Article ID 3772358, 11 pages, 2021.

[12] W. H. Nisbett, A. Kavuri, and M. Das, "On the correlation between second order texture features and human observer detection performance in digital images," *Scientific Reports*, vol. 10, no. 1, p. 13510, 2020.

[13] W. He, X. Y. Zhang, F. Yin, Z. Luo, J. M. Ogier, and C. L. Liu, "Realtime multi-scale scene text detection with scale-based region proposal network," *Pattern Recognition*, vol. 98, article 107026, 2020.

[14] W. Zhang, K. Wang, Y. Liu, Y. Lu, and F. Y. Wang, "A parallel vision approach to scene-specific pedestrian detection," *Neurocomputing*, vol. 394, pp. 114–126, 2020.

[15] X. Hua, X. Wang, T. Rui, H. Zhang, and D. Wang, "A fast self-attention cascaded network for object detection in large scene remote sensing images," *Applied Soft Computing*, vol. 94, article 106495, 2020.

[16] S. Pu, W. Zhao, W. Chen, S. Yang, D. Xie, and Y. Pan, "Unsupervised object detection with scene-adaptive concept learning," *Engineering*, vol. 22, no. 5, pp. 638–651, 2021.

[17] Z. L. Wang and G. H. Tian, "Integrating manifold ranking with boundary expansion and corners clustering for saliency detection of home scene," *Neurocomputing*, vol. 379, pp. 182–196, 2020.

[18] X. Wang, X. Feng, and Z. Xia, "Scene video text tracking based on hybrid deep text detection and layout constraint," *Neurocomputing*, vol. 363, pp. 223–235, 2019.

[19] Y. Jin, Y. Zhang, Y. Cen, Y. Li, V. Mladenovic, and V. Voronin, "Pedestrian detection with super-resolution reconstruction for low-quality image," *Pattern Recognition*, vol. 115, no. 3, article 107846, 2021.

[20] H. Li, Y. Dong, L. Xu, S. Zhang, and J. Wang, "Object detection method based on global feature augmentation and adaptive regression in IoT," *Neural Computing and Applications*, vol. 33, no. 9, pp. 4119–4131, 2021.

[21] S. K. Pal, D. Bhoumik, and D. B. Chakraborty, "Granulated deep learning and Z-numbers in motion detection and object recognition," *Neural Computing and Applications*, vol. 32, no. 21, pp. 16533–16548, 2020.

[22] L. Fan, T. Zhang, and W. Du, "Optical-flow-based framework to boost video object detection performance with object enhancement," *Expert Systems with Applications*, vol. 170, article 114544, 2021.

[23] H. Yang, Y. Lin, H. Zhang, Y. Zhang, and B. Xu, "Towards improving classification power for one-shot object detection," *Neurocomputing*, vol. 455, pp. 390–400, 2021.

[24] X. Zhong, C. Ding, X. Qu, and D. Tao, "Polysemy deciphering network for robust human–object interaction detection," *International Journal of Computer Vision*, vol. 129, no. 6, pp. 1910–1929, 2021.

[25] Y. Liu, Y. Gu, F. Yan, and Y. Zhuang, "Outdoor scene understanding based on multi-scale PBA image features and point cloud features," *Sensors*, vol. 19, no. 20, p. 4546, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3638436, 14 pages
https://doi.org/10.1155/2022/3638436

Research Article

RF-Gait: Gait-Based Person Identification with COTS RFID

Shang Jiang,[1,2] **Jianguo Jiang**,[1,2] **Siye Wang**,[1,2] **Yanfang Zhang**,[1,2] **Yue Feng**,[1,2] **Ziwen Cao**,[1,2] **and Yi Liu**[1,2]

[1]*School of Cyberspace Security, University of Chinese Academy of Sciences, China*
[2]*Institute of Information Engineering, Chinese Academy of Sciences, China*

Correspondence should be addressed to Siye Wang; wangsiye@iie.ac.cn

Received 20 January 2022; Revised 18 April 2022; Accepted 7 May 2022; Published 4 June 2022

Academic Editor: Yinghui Ye

Recently, person identification has been a prerequisite in many applications of the Internet of Things. As a new biometric identification technology, gait recognition has a wide application prospect with the advantages of long-distance recognition and difficulty to forge. However, the existing gait recognition methods have some problems, such as complex algorithm calculation, high user participation, and large equipment overhead. In this paper, we propose RF-Gait, a method that identifies a person through unobtrusive gait perception with COTS RFID. The key insight is that wireless signal fluctuation can be exploited to distinguish each person's unique gait behavior. To this end, we first collect and preprocess the gait-induced data composed of multiple RFID tags. Furthermore, multivariate variational mode decomposition is utilized to extract the intrinsic features in the spatial multichannels cooperatively. By developing a support vector machine model, we identify a person via the intrinsic walking pattern. Finally, extensive experiments show that our method can identify a person with an average accuracy of 96.3% from a group of twenty persons in a complex indoor environment.

1. Introduction

In the era of the Internet of Things, human-computer interaction has become crucial to integrating the physical world and the information world. Person identification provides a guarantee for the security of human-computer interaction. Gait, as an emerging biometric feature, refers to the human characteristic activity of moving the body through the interactive action of two feet. For human identification, gait has the following two advantages: on the one hand, from the perspective of medicine, the length of leg bones, the strength of muscles, the height of center of gravity, and the sensitivity of motor nerves determine the uniqueness and stability of gait, so it is difficult to be imitated by others in a short time; on the other hand, gait recognition is not strict on distance and protects privacy.

Some kinds of methods have been developed to promote the recognition performance of gait, and among them, video-based and radio-based methods have contributed significantly. Typical video image methods extract the contour image of user motion for gait feature extraction and recognition. However, it is worth noting that the video image methods need complex algorithm, large amount of calculation, and require expensive computing equipment for feature extraction and comparison [1]. Therefore, researchers consider to use wireless signal for gait recognition. As we know, when a person walks in the sensing area, the wireless signal will be reflected and diffracted [2], and furthermore, the mapping relationship between each person's gait and the signal fluctuation is unique. WiFiU [3] is the first method that uses COTS WiFi devices to fulfill gait recognition. Channel State Information (CSI) is transformed to the spectrogram in the time-frequency domain, and the spectrogram signature is extracted for gait recognition. GaitSense [4] proposes a WiFi CSI-based novel method that can extract characterized gait patterns and is termed as gait body-coordinate velocity profile. In [5], the spatial gain provided by a RFID tag array is used to resist the different effects of the same person when the person is walking. However, indoor ambient noise may limit the recognition performance of these methods. Due to the fact that the subcarriers of WiFi CSI all use the same propagation paths, there are no

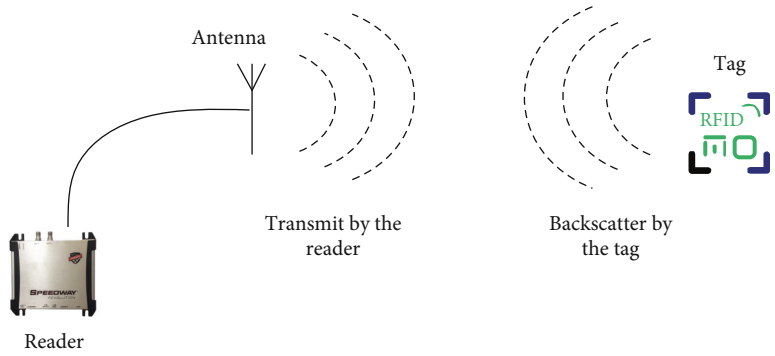


Figure 1: RFID backscatter communication process.

discernible differences among them. Additionally, RFID tags' spatial gain is not fully utilized for gait recognition. In some case, FMCW radar has been used for extract fine-grained gait features in the noisy wireless signal. However, this requires the use of customized equipment, which is costly and does not meet actual needs.

To realize accurate human gait recognition via commercial devices, we need to overcome two critical challenges. The first challenge is the RF signal composition between human gait and other indoor static objects which is highly nonlinear. In order to extract fine-grained gait signals for person identification, we adopt MVMD to decompose the received RF signal and acquire the fine-grained gait signal. The second challenge is that there are differences between each person's different walking signals. In order to eliminate the difference and accurately identify the person's identity, we construct the feature set of human gait from the time domain and frequency domain of RF signals and train the SVM machine learning model for human identification. Moreover, we propose RF-Gait, which use COTS RFID for gait recognition and the cheap RFID tags can provide rich spatial diversity for fine-grained gait feature extraction. The key insight of this paper is to adopt multivariate variational mode decomposition (MVMD) [6] to extract intrinsic features from noisy RFID measurements so as to further improve the accuracy of gait recognition. Our main contributions can be summarized as the following aspects:

(1) We propose a new fine-grained gait feature extraction method with COTS RFID. Owing to device defects, the constant change of indoor environment, and diversity in personnel walking, these noises are nonstationary and nonlinearly mixed with the gait based signal fluctuation. As a result, the proposed method decomposes the signal and aligns the decomposed signals with common frequencies across modes

(2) We implement RF-Gait and comprehensively evaluate the performance in a complex office building under various conditions. Extensive experiments show that our method can identify a person with an average accuracy of 96.7% from a group of ten persons in a real world scenario where people can freely walk and the surrounding environment will change

The reminder of this paper is organized as follows. We first explain some background knowledge and present the system overview in Section 2. The details of our method design are described in Section 3. The implementation and evaluation of the proposed method are presented in Section 4. We review the related works in Section 5. Finally, Section 6 concludes this paper and provides some suggestions for future.

2. Preliminaries and System Overview

2.1. Working Principle of RFID. UHF RFID usually consists of three parts, including reader, antenna, and tag. Figure 1 shows the schematic diagram of RFID backscatter communication. The reader transmits radio frequency continuous wave signal through the antenna. After the tag is activated by the signal sent by the reader, it will return its own information to the reader through the backscatter link. General commercial readers, such as Impinj Speedwayr420, can not only obtain the ID of the tag but also give the indicators of the reflected signal of the tag: signal strength (RSSI), signal phase, and signal Doppler frequency shift. Since the Doppler frequency shift provided by the reader is very noisy [7], researchers usually use two indicators, RSSI and phase for RFID sensing.

In free space, the signal needs to go through two transmissions between the reader and the tag: the forward link from the reader to the tag and the backscatter link from the tag to the reader. When the distance between the reader antenna and the tag is d, the tag RSSI and phase received by the reader can be expressed as [8, 9]

$$R(dB) = 10 \log \left[\frac{C \cdot \lambda^4}{d^4}\right], \tag{1}$$

$$\theta = \left(\frac{2\pi}{\lambda} 2d + \theta_R + \theta_T\right) \bmod 2\pi, \tag{2}$$

where $C = C_G \cdot C_M \cdot C_P / 4\pi d^4 \cdot \xi^2$ and is determined by environment and equipment deployment conditions. λ is the signal wavelength, and θ_R and θ_T represents the phase offset brought by the reader and the tag itself, respectively.

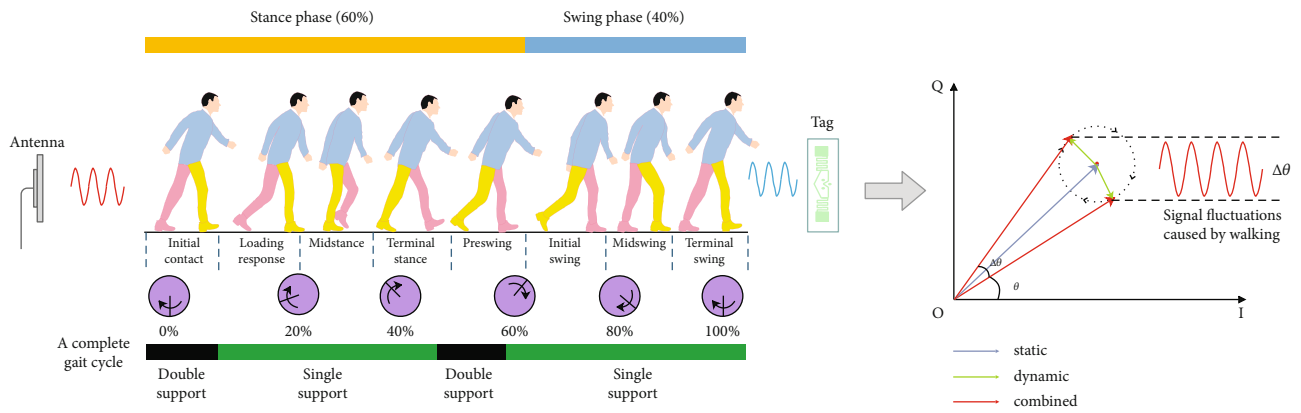
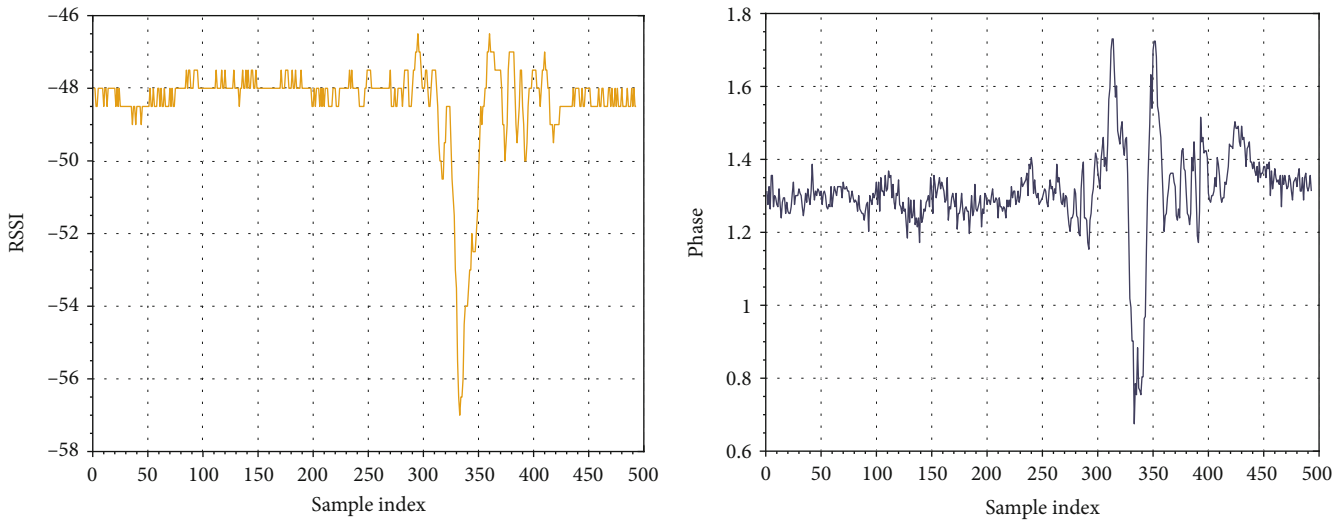


Figure 2: Principle of sensing gait using RFID signal.



(a) RSSI profiles when a person crosses the reader-tag area

(b) Phase profiles when a person crosses the reader-tag area

Figure 3: Phase and RSSI variation when a person crosses the reader-tag area.

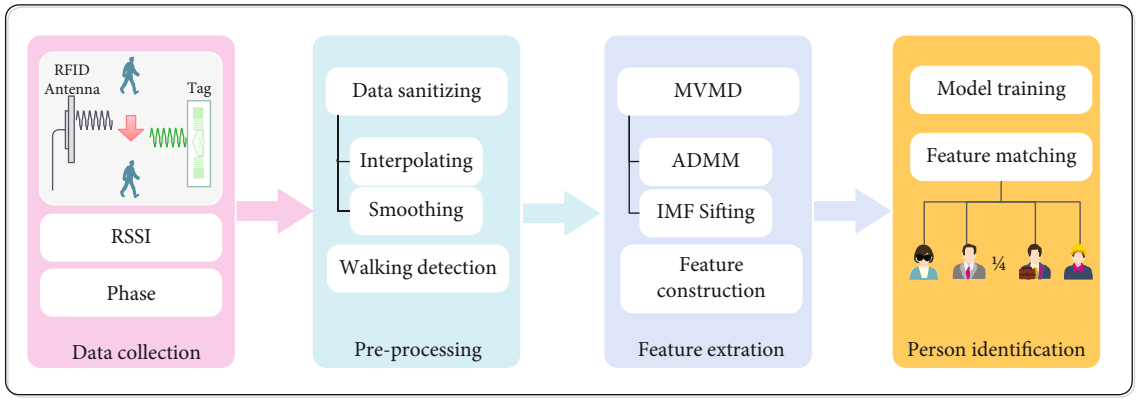
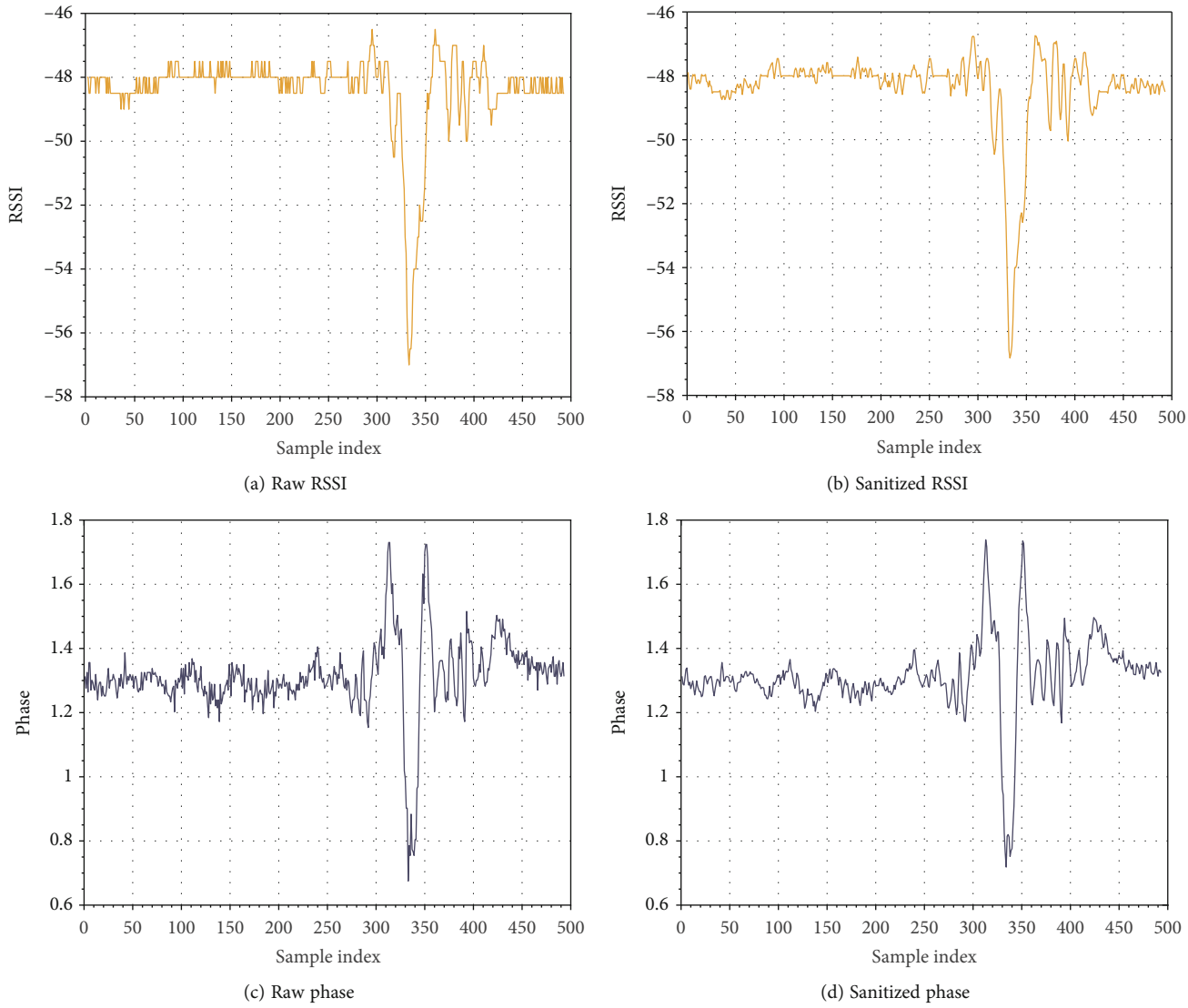


Figure 4: The walking detection and segmentation.

2.2. Feasible Study. The leading attraction of the human gait as a biological feature is contactless, and more importantly, gait is a unique feature of individuals. Murray et al. discovered that the standing, swinging, and gait-related body movements of the same participant are surprisingly similar, and walking has a strong periodicity [10]. What is more, the above mentioned gait characteristics of different participants present obvious individual differences. Specifically, as shown



(a) Raw RSSI

(b) Sanitized RSSI

(c) Raw phase

(d) Sanitized phase

Figure 5: Data sanitizing for raw RSSI and phase.

in Figure 2, when the reader antenna and a tag array (In the experiment part, we discussed the setting of array size and finally set it to 4 rows and 2 columns.) are placed on both sides of the personnel walking route, the human body can be modeled as a conducting cylinder [11]. Meanwhile, the wavelength of UHF RFID is about 32 *cm*, which is roughly close to the width of the human body. Therefore, the person is able to reflect and diffract the RF signal when he passes between the reader antenna and the tags, and there is a one-to-one mapping relationship between the impact of each person's gait on the RF signal. As a result, on the far right of Figure 2, the red sine curve represents the influence of personnel walking on RF signal which is the component we are trying to extract, the green line represents the influence of indoor static environment on RF signal, and the actually received RF signal which is red line can be ideally expressed as [12]

$$S_i(t) = A_i(t)e^{j\theta_i(t)}, A_i(t) = \sqrt{10^{R_i(t)/10^{-3}}}, \quad (3)$$

where $i \in I$, $R_i(t)$, and $\theta_i(t)$ represent the measured RSSI and phase of the *ith* tag, corresponding to the polar angle and polar radius in the I-Q plane, respectively.

Moreover when a person walks in the RFID sensing area, the received signal can be decomposed into

$$S = S_s + S_d = A_s e^{-j\theta_s} + \sum_{k=1}^{M} A_k e^{-j2\pi(d_k/\lambda)}, \quad (4)$$

where the received signal is divided into static and dynamic components in Equation (4). S_s represents the superposition of all static propagation paths, and S_d represents the superposition of all dynamic propagation paths. M is the number of dynamic propagation paths, and d_k is the length of the kth dynamic propagation path. As shown in Figure 3, we find that the periodic dynamic component of the received signal, which including both the phase and RSSI measurements, can be utilized to exploit the unique gait pattern of different persons.

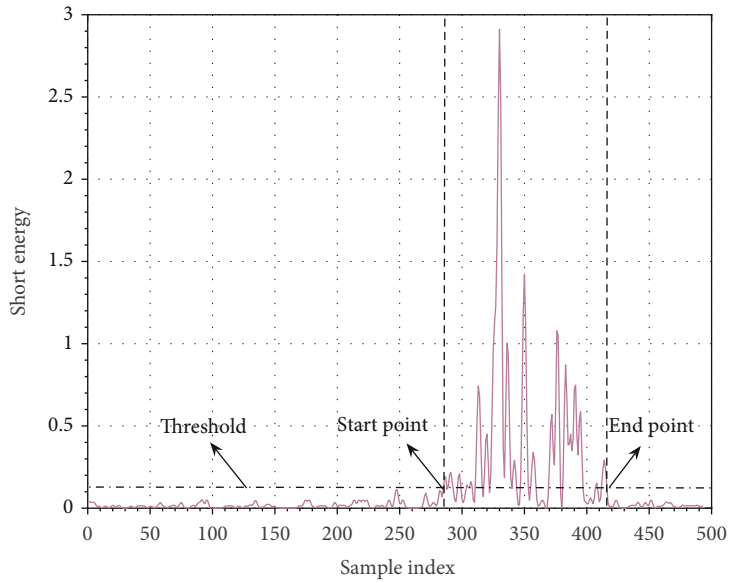


Figure 6: The walking detection and segmentation.

Input: $\{\sum_{i=1}^{I}\widehat{S}_i(t)\}$:$\{\sum_{i=1}^{I}\widehat{R}_i\}$, $\{\sum_{i=1}^{I}\widehat{\theta}\}$, β
Output: $\sum_{k=1}^{K}C_k(t)$,ϵ.
1: Initialize: $c_{k,i}^1$,ω_k^1,λ_i^1,α^1,n.
2: Repeat
3: $n \longleftarrow n+1$
4: Fix other variables and update$c_{k,i}$ as (10);
5: Fix other variables and update ω_k as (11);
6: Until$(\sum_k\sum_i(\|u_{k,i}^{n+1}-u_{k,i}^n\|_2^2/\|u_{k,i}^n\|_2^2)<\beta)$

Algorithm 1: Gait embedded multichannel RFID signal decomposition using MVMD.

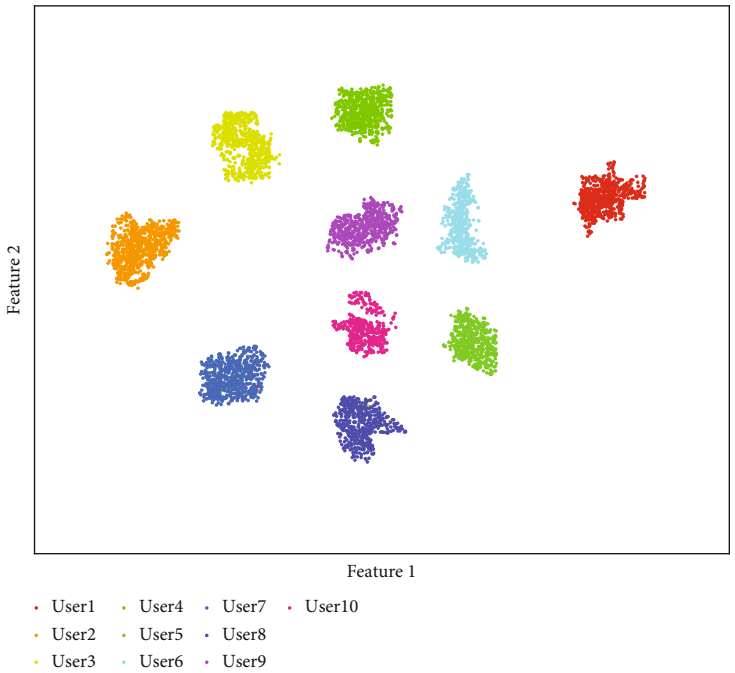


Figure 7: The T-SNE visualization of MVMD-based features for 10 persons.

2.3. System Overview. The framework of RF-Gait is shown in Figure 4. To perform person identification based on gait patterns, we deploy several RFID tags to form a tag array for comprehensive gait perception. For the received RF signals from the tag array, the main processing steps consist of three main components: (1) *Data preprocessing*: with the RF signals received from the tag array, RF-Gait first filters data with interpolation and smoothing and then analyzes for walking detection and direction. (2) *Feature extraction*: based on the processed signal series, RF-Gait extracts feature vectors via MVMD algorithm and intrinsic mode sifting criterion. (3) *Person identification*: our method finally identifies the person who is walking in sensing area according to the gait features stored in the trained model.

3. Method Design

In this section, we detail the method design of RF-Gait including data collection, data preprocessing, feature extraction, and person identification.

3.1. Data Collection. The first step of our method is to collect the RFID signal measurements. We deploy the reader antenna and multiple tags on both sides of the walking path. When a person passes between the reader antenna and tags in different ways, we collect the raw RSSI measurements $R(t)$ and phase measurements $\theta(t)$.

3.2. Preprocessing

3.2.1. Data Sanitizing. The commercial reader uses the slot-ALOHA mechanism to read the channel information, resulting in the nonuniform distribution of the received data in the time domain. In order to facilitate the subsequent processing, we carry on the linear interpolation with 80 sampling points per second to the measurements. Furthermore, in order to solve equipment noise and measurement errors, we consider using the Savitzky-Golay filter [13] for data smoothing. The key insight of the filter is to perform weighted filtering in the predefined window. Moreover, it can more effectively retain the change information of the measurements while smoothing the data. The raw RSSI and phase measurements and the data after sanitizing are shown in Figure 5, and the sanitized signal composed of sanitized RSSI $\widehat{R}(t)$ and phase $\widehat{\theta}(t)$ can be denoted as

$$\widehat{S}(t) = \sum_{i=1}^{I} \widehat{S}_i(t). \tag{5}$$

3.2.2. Walking Detection. To better extract gait features from walking patterns, we need to perform motion detection and segment the time series of the sanitized measurements. Therefore, in this section, we first analyze the short cumulative energy of signal series as follows:

$$E(t) = \sum_{l=1}^{L} A^2(t+l) = \sum_{l=1}^{L} \sqrt{10^{R_i^2(t+l)/10^{-3}}}, \tag{6}$$

Figure 8: The topology of the tag array in real scenarios.

where $A(t)$ denotes the tth amplitude value, and L denotes the length of the sliding window.

As a result, we calculate the signal energy in each window and compare it with a predefined moving detection threshold. As shown in Figure 6, we utilize this scheme to judge the starting and ending points of the movement.

The next step is aimed at distinguishing the walking direction. For this problem, because different directions have different effects on labels, we can consider using the RSSI difference of two columns of labels to solve it. The difference can be represented as

$$Diff_R = R_{T_{left}} - R_{T_{right}}, \tag{7}$$

where R_{T_l} and R_{T_r} mean the tags' RSSI of two different columns in the same row. When $Diff_R \geq 0$, this means that the person walks from the tag in the right column to the tag in the left column, and vice versa.

3.3. Feature Extraction

3.3.1. Multivariate Variational Mode Decomposition. Since the dynamic component of the signal is nonstationary and superimposes with the static component in a nonlinear way, thus we resort to an adaptive method to extract the intrinsic modes from the signals consisting of multichannel tag measurements automatically. The basic idea of MVMD is to decompose the original multichannel signals into several simple and high quality signals (i.e., intrinsic mode functions which are termed as IMFs). As a result, the preprocessing multichannel signals $\widehat{S}(t)$ are decomposed into

Figure 9: Experimental scenario.

Table 1: Confusion matrix of walking detection.

Classified as	Walking	Nonwalking
Walking	0.998	0.02
Nonwalking	0.03	0.997

Table 2: Verification time.

Scene	Gateway	Walkway	Office
Time/(s)	0.241 s	0.312 s	0.518 s

predefined $K \times I$ number IMFs and the residual ε:

$$\widehat{S}(t) = \sum_{k=1}^{K} C_k(t) + \varepsilon, \tag{8}$$

where $C_k(t) = [c_{k,1}(t), c_{k,2}(t), \cdots, c_{k,I}(t)]$.

The goal is to extract simple oscillatory mode IMFs involved in the signals $\widehat{S}(t)$, which are the informative information for realizing person identification. The result cost function of our problem Equation (8) is given by

$$\text{minimize}_{\{c_{k,i}\},\{\omega_k\}} \left\{ \sum_k \sum_i \left\| \partial_t \left[c_+^{k,i}(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\}. \tag{9}$$

In the above proposed variational model of decomposition task, we aim to minimize the sum of bandwidths of all IMFs. In order to resolve the multiobjective optimization problem (9), we adopt the Alternating Direction Method of Multiplier- (ADMM-) based optimization strategy [6]. First of all, we transform the original constrained optimization problem into the form of unconstrained optimization by Lagrange multiplier. In the following, two alternate update steps are given for the mode $c_{k,i}$ and the center frequency ω_k.

Step 1. Updating mode $c_{k,i}$: The other variables are fixed, and $c_{k,i}$ can be updated by solving the problem:

$$c_{k,i}^{n+1} = \underset{\{c_{k,i}\}}{\text{argmin}} \left\{ \alpha \left\| \partial_t \left[c_+^{k,i}(t) e^{-j\omega_k t} \right] \right\|_2^2 + \left\| s_i(t) - \sum_K c_{k,i}(t) + \frac{\lambda_i(t)}{2} \right\|_2^2 \right\}, \tag{10}$$

where α and λ are the parameters of Lagrange multiplier method.

Step 2. Updating center frequency ω_k: In this step, ω_k is updated by solving the following problem:

$$\omega_k^{n+1} = \underset{\{\omega_k\}}{\text{argmin}} \left\{ \sum_i \left\| \partial_t \left[c_+^{k,i}(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\}. \tag{11}$$

The detailed optimization process is given in Algorithm 1.

After MVMD processing, we get $K \times I$ IMFs, but there is only a small number involving rich gait information. Hereinafter, we rank the IMFs' quality of each tag via the correlation coefficients, and for the *ith* tag, the correlation of the *kth* IMF $C_{i,k}$ is defined as

$$\zeta_{i,k} = \frac{C_{i,k}(t)\widehat{S}_i(t)}{\sqrt{C_{i,k}(t)^2 \widehat{S}_i(t)^2}}. \tag{12}$$

According to the above Equation (12), we sift IMFs of each tag with the highest correlation $\sum_{i=1}^{I} \widehat{C}_i$.

3.3.2. Feature Construction. Considering the real-time requirements for person identification, we choose a lightweight and efficient scheme to extract features based on statistics theory. Specifically, for the above selected set of IMFs $\{\sum_{i=1}^{I} \widehat{C}_i\}$, we leverage 6 features to portray the time and frequency profile of the walking pattern. The features are the (1) mean, (2) standard deviation, (3) skewness, (4) kurtosis, (5) form factor, and (6) crest factor.

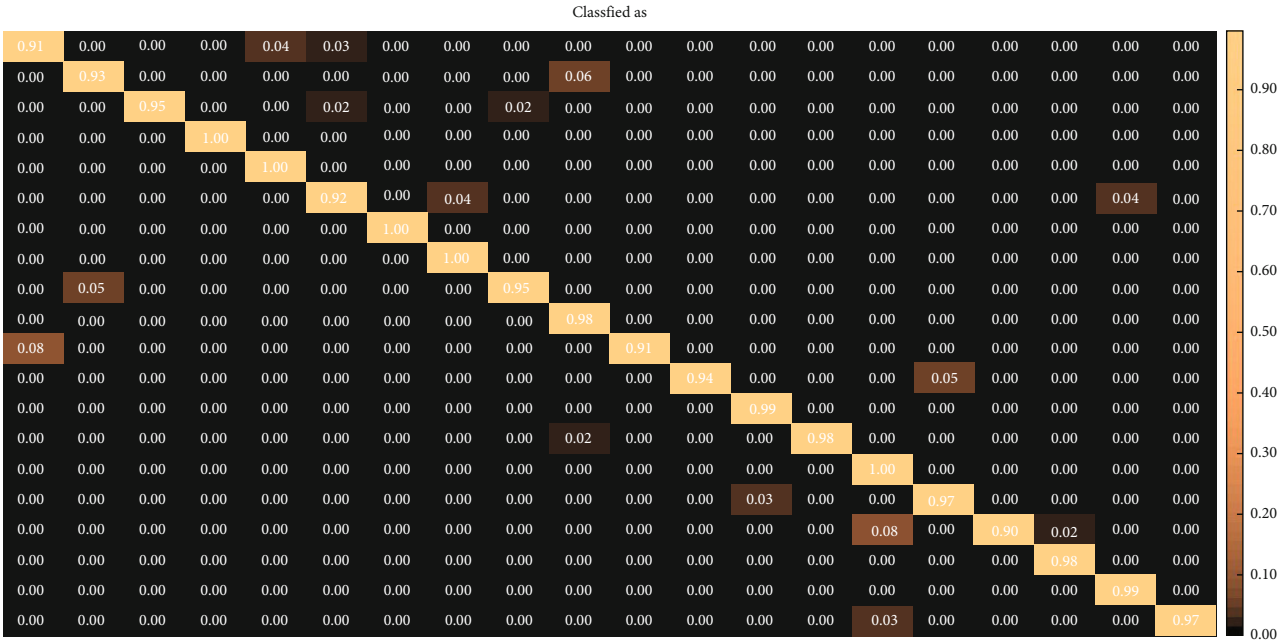


Figure 10: The confusion matrix in our experiment.

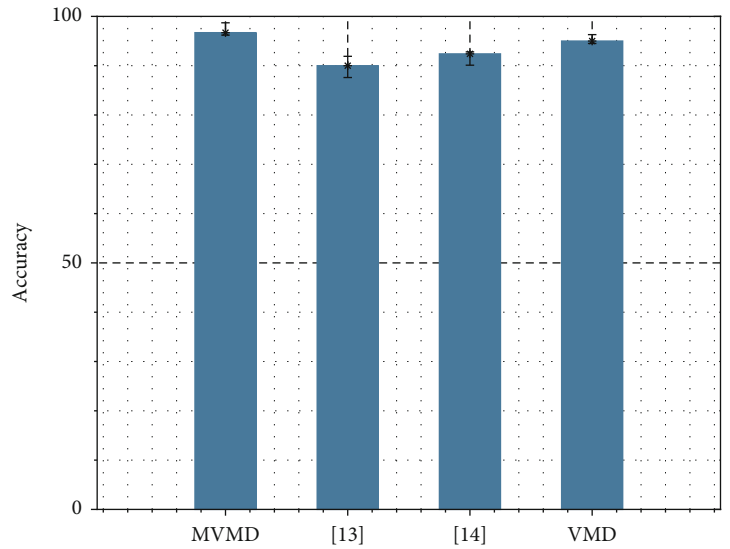


Figure 11: Performance comparison of different gait-based identification methods.

In this way, we can finally form the feature matrix F as

$$F = [F_1, F_2, \cdots, F_i, \cdots, F_I], \tag{13}$$

where F_i contains 6 time domain features and 6 frequency features.

To visualize the learned features, we use t-SNE [14] to project them into a two-dimensional feature space. Figure 7 illustrates the gait features of 10 subjects, which have excellent discrimination capability.

3.4. Person Identification. In this step, we introduce how to perform person identification based on extracted features. Due to the multiclassification problem and the need to ensure higher recognition accuracy in a short period of time, we choose support vector machine (SVM) [15] for identification. Specifically, RF-Gait feeds the extracted IMFs' feature vectors into an SVM model with a radial basis function kernel function in the model training step. Then, when a person is walking in the sensing area, RF-Gait determines the identity of this person via the trained model.

4. Performance Evaluation

4.1. Experiment Settings. In this paper, we perform extended experiments to verify the proposed method RF-Gait in a

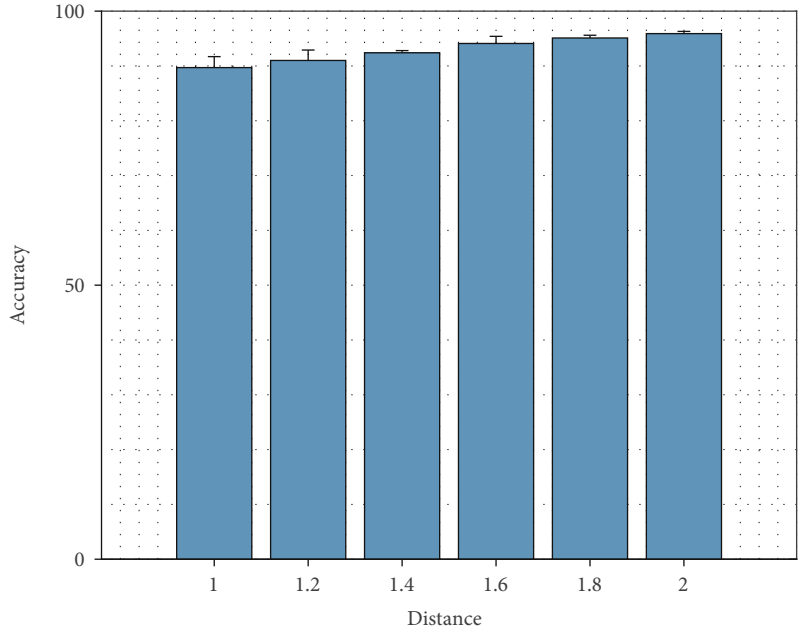


Figure 12: The accuracy varies with the distance.

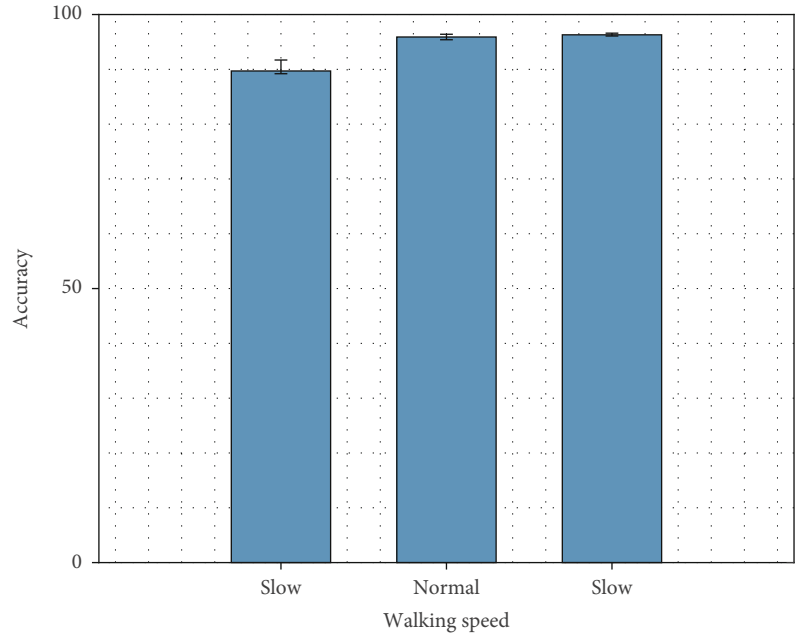


Figure 13: The accuracy varies with walking speed.

typical indoor environment. The experiment contains an Impinj Speedway R420 reader with a circular polarization antenna (9*dBi* gain) and Impinj-H47 tags that form a tag array. The reader operates at the stable frequency of 920.625 MHz, and the tags are deployed to form a uniform linear tag array which is shown in Figure 8. The tag array is $2\,m$ away from the reader antenna, and the height of the reader antenna is $1.5\,m$. The plan of the experimental scene is shown in Figure 9. As shown in Figure 9, multiple places in the scene are used to collect walking data. We recruit 10-20 volunteers, everyone with 60 walking times on each place. From the data set, 80% of the data is randomly selected for model training, and the remaining 20% is used for testing our proposed method RF-Gait.

4.2. System Performance Analysis

4.2.1. Walking Detection Accuracy. In this section, we first evaluate the accuracy of the walking detection scheme. To verify the effectiveness of our scheme, we asked volunteers to walk normally, intermittently, and pause in the sensing

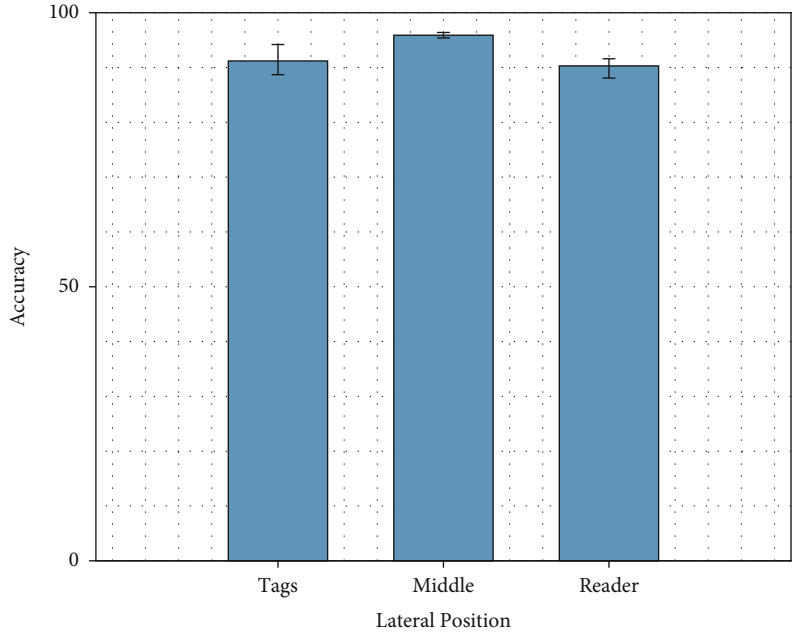


Figure 14: The accuracy varies with lateral position.

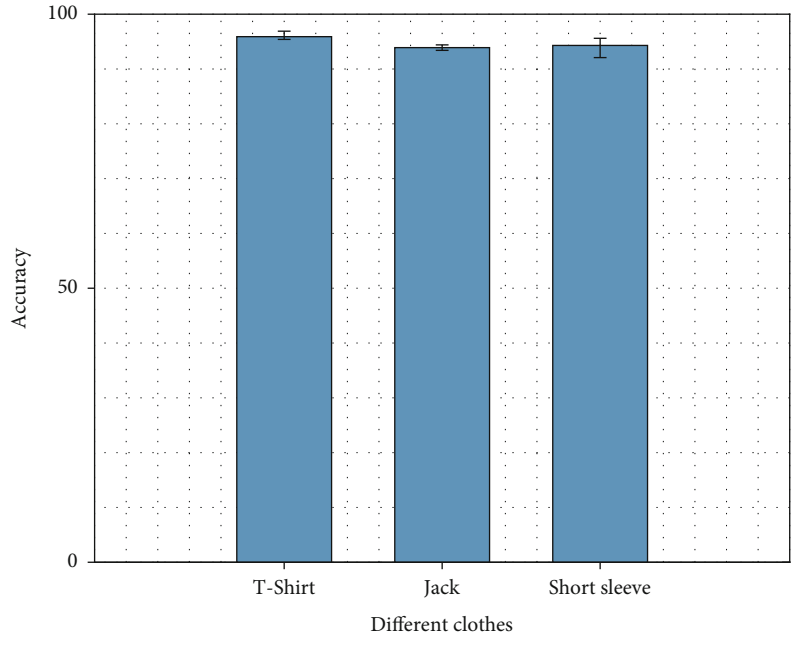


Figure 15: The accuracy varies with different clothes.

area and we used the camera to record the results as the ground truth. As shown in Table 1, our scheme achieves a good result.

4.2.2. Time Required for Verification. Many gait recognition methods based on video image processing need to run complex image processing algorithms. The method proposed in this paper is very lightweight, and the time required for authentication is very short. As shown in Table 2, the average time required for signal processing, feature extraction, and comparison to authenticate a person is only 0. 357 s.

4.2.3. Person Identification Accuracy. To evaluate the accuracy of RF-Gait, we utilize a confusion matrix to describe the overall performance of RF-Gait as shown in Figure 10. The diagonal elements of the matrix represent the probability of correct recognition using the extracted feature matrix, and other elements represent the probability of the wrong classification into others. Figure 10 shows that our system achieves an average accuracy of 96.7% in the typical complex indoor environment.

Furthermore, we compare RF-Gait with variational mode decomposition and two state-of-art gait-based person

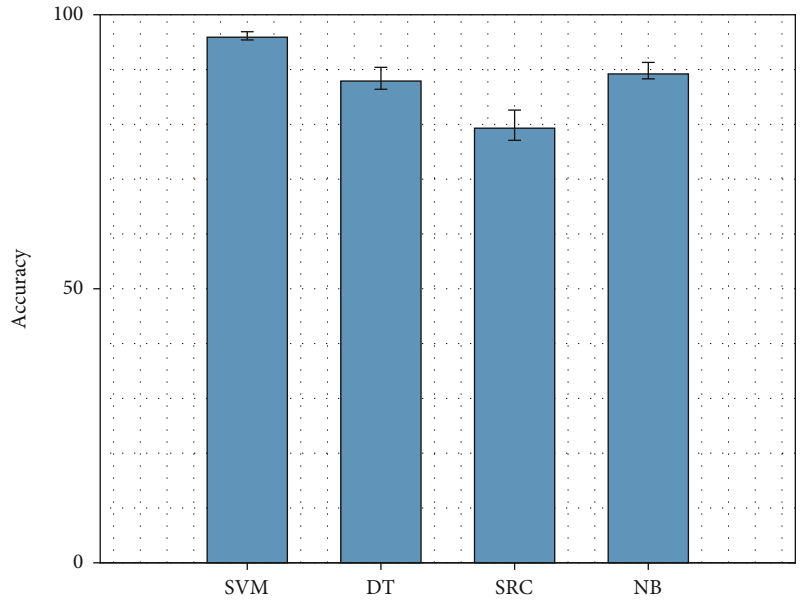


Figure 16: The accuracy varies with different classifiers.

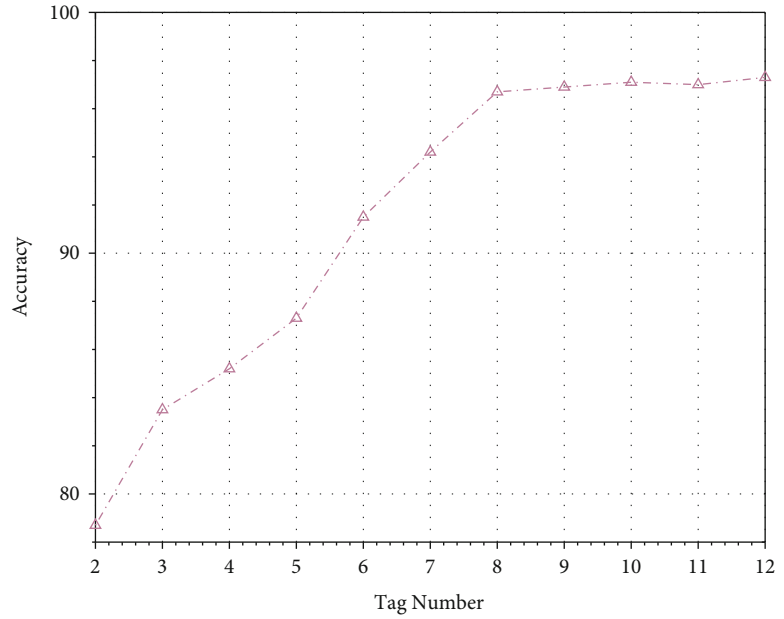


Figure 17: The accuracy varies with the tag number.

identification methods [16, 17] which are empirical mode decomposition-based methods in Figure 11. The accuracy of our proposed method is higher than the other three methods. Our method is 3.7% higher than the traditional VMD scheme, 6.7% higher than the literature [16], and 4.3 % higher than the literature [17], respectively.

4.2.4. The Impact of the Horizontal Distance between the Reader Antenna and Tag Array. In our proposed method, the width of an entrance determines the distance between the reader and the tag array. We evaluate the effect of entrance width on the system performance. The distance varies from $1m$ to $2m$ at intervals of $0.2m$. As shown in Figure 12, we find that the recognition accuracy tends to get better as the distance from the entrance becomes larger. Through literature research, we believe that the dynamic path, that is, personnel walking, has a more significant influence than the static RF signal component [18].

4.2.5. The Impact of Walking Difference. In this part, we mainly evaluate the influence of walking speed and user lateral positions on the recognition effect. First, we evaluate the effects of different walking speed by letting the same 10 subjects walk through the sensing area at different velocities. As shown in Figure 13, except for the fast walking which achieves a decreased accuracy of 89.7%, RF-Gait maintains

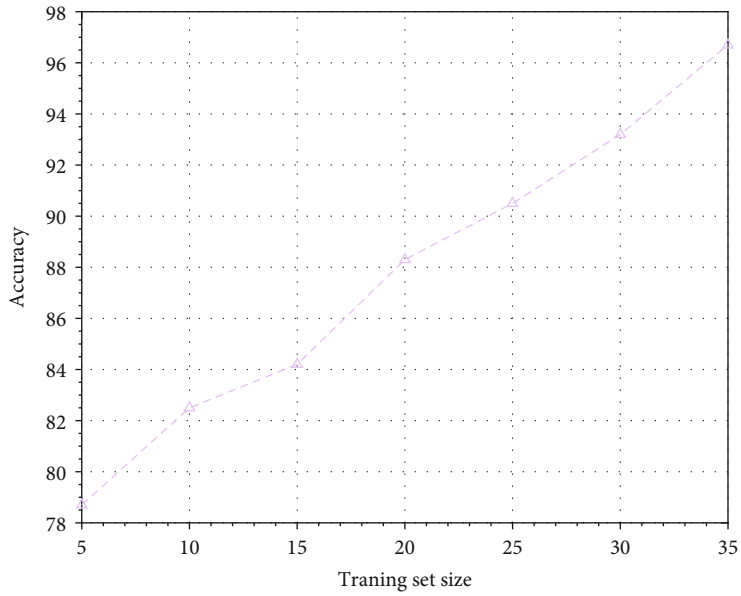


Figure 18: The accuracy varies with the training set.

high recognition accuracy when walking slowly and normally. Second, we set three kinds of lateral positions, one is close to the tags, one is close to the reader, and the other is walking in the middle. Then, we evaluate the influence of this condition on the recognition accuracy. As shown in Figure 14, when people walk in the middle, the recognition effect is the best. We speculate that when close to the reader and tags, the normal reading of the RF signal will be affected [19], so the recognition effect will be reduced.

4.2.6. The Impact of Different Clothes. In the actual scene, the same person will wear different clothes, so we need to evaluate the impact of different clothes on RF gait. As shown in Figure 15, there is no significant impact on the recognition effect when the clothes of walkers change. We believe that this is because normal clothing materials have little effect on the RF signal of UHF band [20].

4.2.7. The Impact of Different Classifiers. In order to analyze the recognition accuracy of different classifiers on selected features, we select four commonly used multiobjective classifiers, support vector machine with RBF (SVM), decision tree (DT), sparse representation classification (SRC), and naive Bayes (NB). As shown in Figure 16, we find that the result of SVM is the best, so we choose it as the classifier in our paper.

4.2.8. The Impact of Tags' Number. As more tags can acquire richer spatial gain for person identification while increasing the person identification delay, we study the impact of the number of tags on the performance. In the scenario shown as Figure 17, we change the number of tags from 2 to 12 and show the results in Figure 17. We find that with the increase of the number of tags, the recognition accuracy grows significantly, but the growth is not obvious after the number is 8. Therefore, we make a trade-off between recognition accuracy and system delay and select eight tags to collect gait information for person identification.

4.2.9. The Impact of Training Set Size. In practical application, the less the number of training samples, the less the cost of actual deployment. In order to investigate the influence of the number of training samples on the results, this paper changes the number of training samples from 5 to 35 and uses the remaining samples as the test. The certification results are shown in Figure 18.

5. Related Work

This section reviews the related literature in RFID-based sensing and gait recognition techniques.

5.1. RFID-Based Sensing. As a mature automatic identification technology, RFID technology is widely used in industrial manufacturing [21], warehousing [22], and logistics [23]. RFID-based sensing technology utilizes signal characteristics reflected back from RFID tags for indoor contextual sensing. The emerging RFID-based sensing applications are widely applied to user authentication [12, 24, 25], indoor localization [26, 27], activity recognition [7, 28], and so on [29]. RF-Mehndi [25] is an RFID-based user authentication system. The authors find that the coupling effect of a tag array is distinctive when different users touch the tags so as to achieve biometric acquisition. 3DLRA [27] leverages RFID tgas to develop a 3D indoor localization system and analyzes the variation characteristic of signal indicators using deep learning. In [28], the authors quantify the correlation between RF phase values and human activities by modeling the signal reflection of RFID tags in contact-free scenarios. Furthermore, RFID is also be used for material sensing, vibration sensing, and so on. In [29], the impedance-related phase change is utilized for material sensing and finally the authors can detect the category of the material.

5.2. Gait Recognition. Existing gait recognition methods mainly focus on three categories: methods based on video image processing [30], methods based on sensors [31, 32], and methods based on radios [33–35]. Among them, the method based on video image processing performs gait feature extraction and identification by extracting the contour image of the user's movement. In [31], the authors use Gait Energy Image (GEI) as a template for human identification by gait. The sensor-based method extracts the user's acceleration and other information when walking through various sensors carried by the user and analyzes the user's gait behavior. In [36], inertial measurement units (IMU) are consider for recognizing gait and the authors design a deep convolutional neural network to extract discriminative gait features. The radio-based methods identify human gait by analyzing the wireless signal and establishing the mapping relationship with human walking. In [35], the authors apply weighted multidimensional dynamic time warping to compute the similarity of two walking profiles which are collected from the RFID tags.

6. Summary and Future Work

This paper introduces a walking-induced method RF-Gait which is capable of identifying persons using COTS RFID. RF-Gait employs multiple spatially distributed tags to obtain fine-grained profiles of the human gait and enables human identification via the MVMD algorithm and SVM-based identification model. Experimental results show that the average identification rate of this method can reach 96.3%, and it has good robustness and stability. However, this paper only considers the gait perception of a single person, which is not well performing in multiperson situations. An in-depth study is needed to identify the gaits of more than one person at a time.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant no. XDC02040300).

References

[1] I. Rida, N. Almaadeed, and S. Almaadeed, "Robust gait recognition: a comprehensive survey," *IET Biometrics*, vol. 8, no. 1, pp. 14–28, 2019.

[2] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: a survey," *IEEE Communication Surveys and Tutorials*, pp. 1629–1645, 2019.

[3] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," in *UbiComp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 363–373, Heidelberg Germany, 2016.

[4] Y. Zhang, Y. Zheng, G. Zhang, K. Qian, C. Qian, and Z. Yang, "Gaitsense: towards ubiquitous gait-based human identification with wi-fi," *ACM Transactions on Sensor Network (TOSN)*, vol. 1, no. 1, 2022.

[5] Q. Zhang, D. Li, R. Zhao, D. Wang, Y. Deng, and B. Chen, "RFree-ID: an unobtrusive human identification system irrespective of walking cofactors using COTS RFID," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–10, Athens, Greece, 2018.

[6] N. u. Rehman and H. Aftab, "Multivariate variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 6039–6052, 2019.

[7] H. Ding, L. Shangguan, Z. Yang et al., "Femo: a platform for free-weight exercise monitoring with rfids," in *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pp. 141–154, Seoul South Korea, 2015.

[8] C. Wang, J. Liu, Y. Chen et al., "Multi-touch in the air: device-free finger tracking and gesture recognition via COTS RFID," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 1691–1699, Honolulu, HI, USA, 2018.

[9] J. D. Griffin and G. D. Durgin, "Complete link budgets for backscatter-radio and RFID systems," *IEEE Antennas and Propagation Magazine*, vol. 51, no. 2, pp. 11–25, 2009.

[10] M. P. Murray, A. B. Drought, and R. C. Kory, "Walking patterns of normal men," *The Journal of Bone and Joint Surgery*, vol. 46, no. 2, pp. 335–360, 1964.

[11] M. Ghaddar, L. Talbi, T. A. Denidni, and A. Sebak, "A conducting cylinder for modeling human body presence in indoor propagation channel," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 11, pp. 3099–3103, 2007.

[12] J. Ning, L. Xie, C. Wang et al., "RF-badge: vital sign-based authentication via RFID tag array on badges," *IEEE Transactions on Mobile Computing*, p. 1, 2021.

[13] R. W. Schafer, "What is a Savitzky-Golay filter? [Lecture notes]," *IEEE Signal Processing Magazine*, vol. 28, no. 4, pp. 111–117, 2011.

[14] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 53–65, 2008.

[15] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and Their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

[16] J. Wang, Y. Zhao, X. Fan, Q. Gao, X. Ma, and H. Wang, "Device-free identification using intrinsic CSI features," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8571–8581, 2018.

[17] X. Wang, F. Li, Y. Xie, S. Yang, and Y. Wang, "Gait and respiration based user identification using Wi-Fi signal," *IEEE Internet of Things Journal*, vol. 9, 2022.

[18] T. Xin, B. Guo, Z. Wang et al., "FreeSense: a robust approach for indoor human detection using Wi-Fi signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 2, no. 3, pp. 1–23, 2018.

[19] S. Ahmed, S. M. Rehman, L. Ukkonen, and T. Bjorninen, "Glove-integrated slotted patch antenna for wearable UHF RFID reader," in *2018 IEEE International Conference on RFID Technology & Application (RFID-TA)*, pp. 1–5, Macau, Macao, 2018.

[20] H. Li, C. Ye, and A. P. Sample, "IDSense: a human object interaction detection system based on passive UHF RFID," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2555–2564, Seoul Republic of Korea, 2015.

[21] L. Zhekun, R. Gadh, and B. Prabhu, "Applications of RFID technology and smart parts in manufacturing," *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 46970, pp. 123–129, 2004.

[22] B. Vijayaraman and B. A. Osyk, "An empirical study of RFID implementation in the warehousing industry," *The International Journal of Logistics Management*, vol. 17, no. 1, pp. 6–20, 2006.

[23] C. Sun, "Application of RFID technology for logistics on Internet of Things," *AASRI Procedia*, vol. 1, pp. 106–111, 2012.

[24] N. Peng, X. Liu, and S. Zhang, "RF-Ubia: user biometric information authentication based on RFID," in *International Conference on Wireless Algorithms, Systems, and Applications*, pp. 135–146, Nanjing, China, 2021.

[25] C. Zhao, Z. Li, T. Liu et al., "RF-Mehndi: a fingertip profiled RF identifier," in *IEEE INFOCOM 2019- IEEE Conference on Computer Communications*, pp. 1513–1521, Paris, France, 2019.

[26] C. Duan, J. Liu, X. Ding, Z. Li, and Y. Liu, "Full-dimension relative positioning for rfid-enabled self-checkout services," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 5, no. 1, pp. 1–23, 2021.

[27] S. Cheng, S. Wang, W. Guan, H. Xu, and P. Li, "3DLRA: an RFID 3D indoor localization method based on deep learning," *Sensors*, vol. 20, no. 9, p. 2731, 2020.

[28] Y. Wang and Y. Zheng, "Modeling RFID signal reflection for contact-free activity recognition," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 2, no. 4, pp. 1–22, 2018.

[29] B. Xie, J. Xiong, X. Chen et al., "Tagtag: material sensing with commodity RFID," in *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*, pp. 338–350, New York New York, 2019.

[30] M. Hu, Y. Wang, Z. Zhang, D. Zhang, and J. J. Little, "Incremental learning for video-based gait recognition with LBP flow," *IEEE Transactions on Cybernetics*, vol. 43, no. 1, pp. 77–89, 2012.

[31] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan, "Human identification using temporal information preserving gait template," *IEEE transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 34, no. 11.

[32] W. Xu, C. Javali, G. Revadigar, C. Luo, N. Bergmann, and W. Hu, "Gait-Key: a gait-based shared secret key generation protocol for wearable devices," *ACM Transactions on Sensor Networks (TOSN)*, vol. 13, no. 1, pp. 1–27, 2017.

[33] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "FreeSense: indoor human identification with Wi-Fi signals," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, Washington, DC, USA, 2016.

[34] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 1–12, Vienna, Austria, 2016.

[35] Q. Zhang, R. Zhao, D. Li, and D. Wang, "Unobtrusive and robust human identification using COTS RFID," *Computer Networks*, vol. 166, article 106818, 2020.

[36] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-based gait recognition using convolutional neural networks and multi-sensor fusion," *Sensors*, vol. 17, no. 12, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3306661, 9 pages
https://doi.org/10.1155/2022/3306661

Research Article

Research and Application of Key Technologies of Ocean Virtual Scene Display Based on Digital Image

Xiaonan Ren, Jie Ning, and Joung Hyung Cho

Department of Marine Design Convergence Engineering, Pukyong National University, Busan 48513, Republic of Korea Xiaonan

Correspondence should be addressed to Xiaonan Ren; xiaonan77@pukyong.ac.kr
and Joung Hyung Cho; 202056741@pukyong.ac.kr

Received 5 January 2022; Revised 11 February 2022; Accepted 4 April 2022; Published 3 June 2022

Academic Editor: Liqin Shi

Digital imaging technology originated from scientific and technological competitions in the military field, allowing continuous high-intensity digital photography work. The ocean virtual scene display of digital images is more flexible and more tolerant. The ocean virtual scene can be used to model the structure of different regions, create a diversified ocean experience pavilion as much as possible, and create a multi-dimensional visual product. Through this research, we get: 1. The best frame rate is 100.607 when High Dynamic Range (HDR) is off and 98.47 when HDR is on. When A =0.0161, Me =26, K1 = 0.6488, G =28, K2 = -0.0072, α =58, Omega =2.5229, β =78. The simulation experiment of ocean scene construction shows that it only needs to calculate the ocean virtual scene of LONC area, and repeatedly set and splice the ocean virtual scene to form a large modeling scene. 2. When the FFT model is optimized, the parameters are set as follows: Meta Flight =0.946, Open Flight =0.441, α =0.828, Omega =089, β =0.754; The optimal parameters of multiple kernel learning model in ocean virtual scene are: Meta Flight =0.757, Pen Flight =0.818, α =0.781, Omega =0.157, β =1.739; The best parameters of the ocean virtual scene of Vega Prime model are: Meta Flight =0.285, Open Flight =0.803, α =0.701, Omega =0.725, and β =0.757. And the constructed ocean virtual scene can achieve the best effect. 3. In the FFT model, when the related technical parameters are set optimally, the two-dimensional animation is integrated into the digital image as a visual special effect element or an animation subtitle, so that the digital image is more interesting, creative and propagable.

1. Introduction

Digital imaging technology originated from scientific and technological competitions in the military field. In the initial development period of digital imaging, this technology was mainly used in military fields such as aerospace and navigation. Digital imaging is significantly different from traditional film. The former is based on the photosensitive component Complementary Metal Oxide Semiconductor (CMOS), which can carry out continuous and high-intensity digital photography. Compared with traditional film imaging, the technical means are relatively backward and require a lot of manual operation. The requirements are high, and high-throughput work cannot be achieved. The development of digital imaging benefits from the development of Charge-coupled Device (CCD). Charge-coupled devices are used as storage devices [1–3]. Later, it was developed and applied in the field of digital imaging combined with photoelectric effect for storage of video effects. CMOS and CCD are continuously polished and updated to perfectly match the digital image. Digital imaging technology can help people understand the aesthetics of digital art, starting from the digital imaging features of human interaction. Digital imaging technology is a digital imaging art with digitization as the core. Its development may have an impact on the form of traditional imaging. However, the audience for digital imaging and traditional imaging is different and may reduce the huge contrast brought about by the impact. Digital photography technology has accelerated the human society from the writing age into the prosperous ocean virtual scene display, which greatly broadens the human visual experience. The creative form and scene construction are also more exaggerated and bold. The marine virtual scene display combines people and the machine has changed the traditional aesthetic feeling and way of aesthetics. Have a stronger sense of participation [4–8]. In the field of

computer and information science, influence interaction can be understood as "human-computer interaction" or the transmission of information back and forth. Interaction can also be interpreted as a two-way communication between the media and users, which can realize information interaction in the true sense. In a diversified network platform, digital images are the main form of information dissemination today, and digital images are used to promote the company's image and brand. At the beginning of its release, the Mi 11 Ultra reached the top of the professional imaging evaluation agency DXOMARK list. Mi 11 Ultra uses a 1/1.12-inch GN2 sensor with an ultra-outsole, which has been appreciated by users in both photography and video shooting. In addition to improving the picture quality, mobile phones must also consider the user's experience and interactive experience. The application of AI algorithms in photography and the needs of beauty are all user needs. For digital photography, anti-shake technology and high-resolution image quality The increase in rate, combined with technologies such as optical image stabilization, can significantly improve the digital camera experience of mobile phones. The 5G era and short video complement each other and promote the development of digital photography technology. 5G technology is of great help to the cloud platform technology of sports cameras. The video transmission quality of drones can reach high frame rates and 8 K images. Virtual technology combines 3D scanners, thermal imagers, and surveying instruments to form a framework of digital influence. Digital photography technology accelerates the human society from the writing age into the present prosperous ocean virtual scene display, which greatly broadens the visual experience and has a revolutionary change. With the great development of the marine virtual scene field [9, 10]. More and more people prefer virtual technology to experience some scenes, and virtual technology that can be close to the real experience is more popular. With its unique visual effects, the combination of ocean virtual scenes and real shot videos still has a lot of space and more forms of expression under different carriers and network communication methods. The advancement of virtual scene technology has made the production of 3D animation and digital special effects more and more realistic. The international and domestic markets are generally pursuing the production of 3D animation and film and television special effects. The traditional form of ocean scenes combined with video has encountered bottlenecks in the development and innovation of film forms, and it is difficult to make new breakthroughs. The ocean virtual scene display of digital images is more flexible and tolerant. The ocean virtual scene can be used to model and construct different regions, build a diversified ocean experience pavilion as much as possible, and tap more ocean dynamic elements to apply to reality. Take a video. The art form of inserting the ocean virtual scene into the real shot video is unique [11–15]. The flexible drawing style and the freely changing visual space, if used properly, will form a unique ocean virtual scene display, creating a multi-dimensional visual product. Insert some two-dimensional animations into the construction of ocean virtual scenes, and integrate two-dimensional animations as a kind of visual special effects elements or animated subtitles into digital images, making digital images more interesting, creative, and spreading wider. Although real ocean scenes and digital photography have more perfect viewing visual effects, the production of complex underwater worlds and large aquariums requires a long period of time for animal needs, and the investment in money and time is still large. Video works in the era of digital imaging pay more attention to the speed of dissemination and the evaluation and influence of audience groups than in the past. However, the combination of traditional animation films has lacked freshness for today's audiences. Combine two-dimensional animation with real shot video, explore more ways to combine the two and the effect of combining different real shot video content, both in terms of visual performance of the video and production cost, it is a win-win situation s Choice.

2. Digital Imaging Technology and Virtual Scene Construction Technology

2.1. Digital Imaging Technology. Digital video is the use of a video camera to transform the real world under the ocean into electrical signals, which are images recorded in digital form. The development of imaging technology is divided into traditional imaging technology and digital imaging technology. With the help of the fermentation of today's network environment, digital images have formed digital images of production methods, digital images of storage methods, and digital images of broadcast methods. The carrier of digital images is digital photography equipment, which can record and output optical and electrical signals. Figure 1 shows the main framework of imaging technology [16].

2.2. Storage and Transmission of Digital Images. As shown in Figure 2, the process of storing and transmitting digital images is visualized. Digital image storage devices include digital tapes, digital P2 cards, digital Blu-ray discs and digital disks, which are spread through network video streams, and finally form a network pandemic to achieve the final flow realization purpose. In the creation of digital video, two-dimensional animation can be used for creation, which brings together the advantages of animation software, and the blessing of a computer can definitely contribute to the effect and quality of the film. The digital camera shoots the scene as an entity and can record the truest side of the world. Two-dimensional animations are relatively weak, and sometimes they cannot achieve perfect presentation effects due to different perspectives. Digital imaging has interactive functions in the field of digital art, and can be used to build ocean virtual scenes by combining computer and information science, communication, aesthetics, and psychology. The interaction of the ocean virtual scene is "to realize the expression of artistic concepts through the artistic expression method of ocean virtual and the way of virtual scene construction mode.

2.3. Digital Interaction. The interaction of the ocean virtual scene includes four levels, including: viewing on the spot, browsing on the network, using virtual equipment for experience, and on-site control. As an art form supported by digital technology, the ocean virtual scene art of digital imaging is a powerful expansion of interaction. After all, pure VR viewing cannot achieve the immersive experience. Only when the experiencer decides the direction of things in the

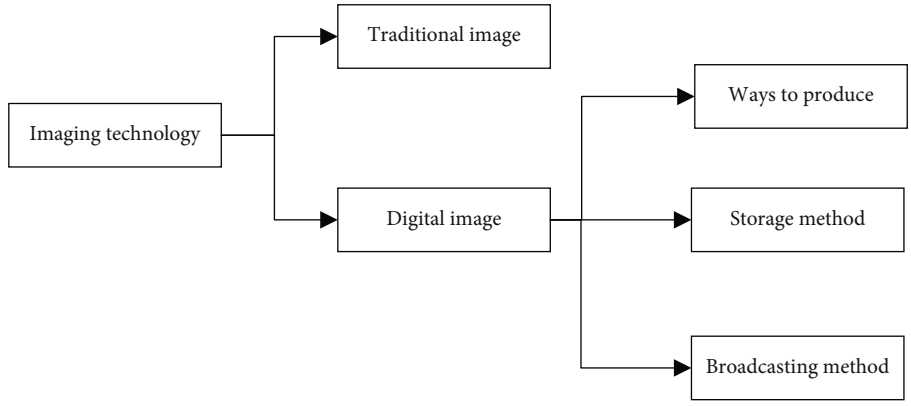


Figure 1: Image technology.

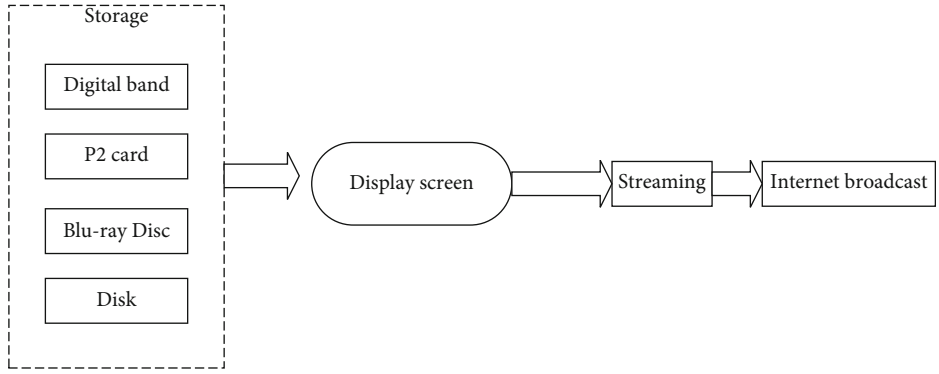


Figure 2: Storage and transmission of digital images.

virtual world, can they control and change the direction of things in virtual modeling. This lies in the digital interactive magic of ocean virtual scenes, which can truly be expressed as aesthetic activities. In digital art, its essence has not changed, because the law of communication of interaction is the same, and the difference is only the innovation and breakthrough in the degree of interaction between content and form. The interactivity of digital art is a way of making use of specific information transmission equipment and feedback systems in the virtual environment of the Internet platform to carry out the individualized participation of people and people, people and works, and people and systems.

2.4. Target Detection Algorithm. R-CNN (Regions with CNN features or Region-based Convolution Neural Networks) is an A region-based convolutional neural network algorithm. Traditional art cannot provide audiences with a platform for interactive aesthetic manipulation of virtual ocean scenes due to the limitations of technical conditions, appreciation experience, and form. Customers who visit the aquarium will not have a sense of participation in art. The spread of digital ocean virtual scenes has linear characteristics, so that customers of the aquarium can personally participate in the process of art formation. Under the conditions of rapid development of digital technology, the strategic phenomenon of regional recommendation has been changed, and mastering digital technology has become a target positioning model. In the formation of a bottom-up art acceptance process, the audience broke through the single, one-way passive aesthetic method of traditional target detection algorithms, and instead focused on the selection and control of multi-scale sliding, and mastered the target. The initiative of regional artistic aesthetics. R-CNN abandons the traditional ocean virtual scene idea, creatively combines the ocean virtual scene with CNN, and finally makes the construction speed of the ocean virtual scene and the speed and accuracy of target detection have been significantly improved.

Table 1: FFT transform method.

Frame rate	HDR off	HDR on	Lonc	Latc
Frame rate at that time	93.569	81.267	95.429	88.32
Average frame rate	95.429	83.591	95.769	89.65
Worst frame rate	90.769	74.629	82.32	96.65
Optimal frame rate	100.607	98.476	96.23	84.65

3. Application of Digital Images in Ocean Virtual Scenes

(a) FFT [17–20]

Digital image

$$\alpha^T = [a_0, a_1, \cdots, a_{n-1}] \quad (1)$$

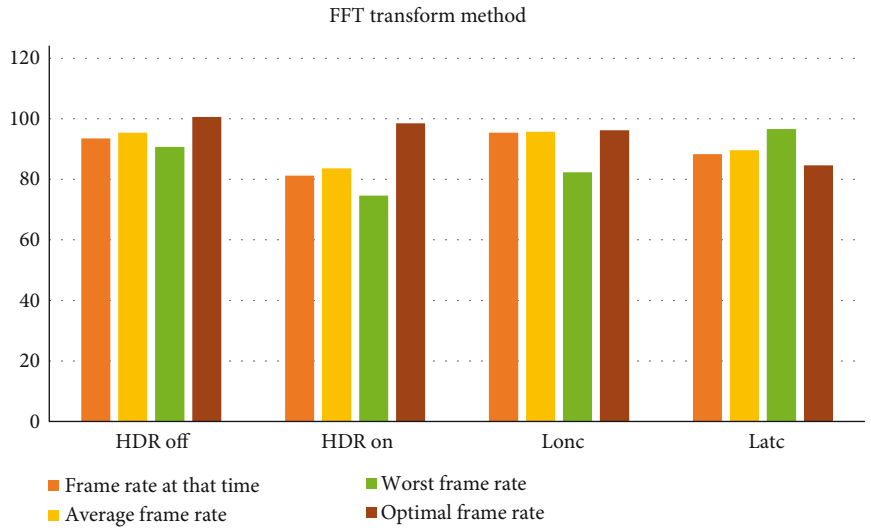


Figure 3: FFT transform method.

$$P_a(x) = \sum_{i=0}^{n-1} a_i x^i \tag{2}$$

Digital image storage method

$$P_a(x) = Xa^T \tag{3}$$

$$X^T = \left[1, x, x^2, \cdots, x^{n-1}\right] \tag{4}$$

Digital video broadcast method

$$P_a = \Omega a \tag{5}$$

Ocean virtual scene

$$P(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{n-1} x^{n-1} \tag{6}$$

MKL [21–23]

$$P_e\left(x^2\right) = a_0 + a_2 x^2 + a_4 \left(x^2\right)^2 + \cdots + a_{n-2} x^{n-2} \tag{7}$$

Equipment for experience

$$xP_e\left(x^2\right) = x\left[a_0 + a_2 x^2 + a_4 \left(x^2\right)^2 + \cdots + a_{n-2} x^{n-2}\right] \tag{8}$$

On-site control

$$P(x) = P_e\left(x^2\right) + xP_e\left(x^2\right) \tag{9}$$

$$P_e(x) = a_0 + a_2 x + a_4 x^2 + \cdots + a_{n-2} x^{n-2/2} \tag{10}$$

$$P_{ee}(x) = a_0 + a_2 x + a_8 x^4 + \cdots + a_{n-4} x^{n-2/2-1} \tag{11}$$

$$xP_{ee}(x) = x\left[a_0 + a_2 x + a_8 x^4 + \cdots + a_{n-4} x^{n-2/2-1}\right] \tag{12}$$

Table 2: Scene simulation.

NO	Wind speed	Fetch length
a	5 m/s	1500 km
b	10 m/s	50 km
c	10 m/s	500 km
d	15 m/s	1500 km
e	20 m/s	50 km
f	20 m/s	750 km
g	20 m/s	1500 km

Table 3: Construction of ocean scene.

NO	A	Me	K1	G	K2	α	Omega	β
1	0.0161	26	0.6488	28	-0.0072	58	2.5229	78
2	0.0064	89	2.1143	85	-0.0092	66	4.5544	45
3	0.0161	26	0.648	89	-0.0317	83	2.5229	78
4	0.0196	39	0.8359	95	0.03301	51	2.8648	81
5	0.0104	35	0.4084	78	-0.0528	72	2.0101	24
6	0.0104	35	0.4011	65	0.09335	27	2.0101	24
7	0.0033	9	0.6375	48	-0.1206	75	2.5229	78
8	0.0027	77	4.3809	62	-0.1211	18	6.5574	53
9	0.0011	35	7.0132	27	11.2821	93	11.423	35
10	0.0008	94	10.506	58	-12.154	21	12.566	71
11	0.0008	94	-10.06	56	-12.522	26	12.566	71
12	0.0011	35	-3.257	49	12.8787	35	11.423	35
13	0.0011	35	-2.383	69	13.0687	22	11.423	35
14	0.0011	35	0.4647	67	13.2761	58	11.423	35
15	0.0008	94	6.5805	79	14.6564	24	12.566	31
16	0.0008	94	-5.686	83	15.0257	75	12.566	71
17	0.0008	94	1.8399	31	15.9602	27	12.566	31

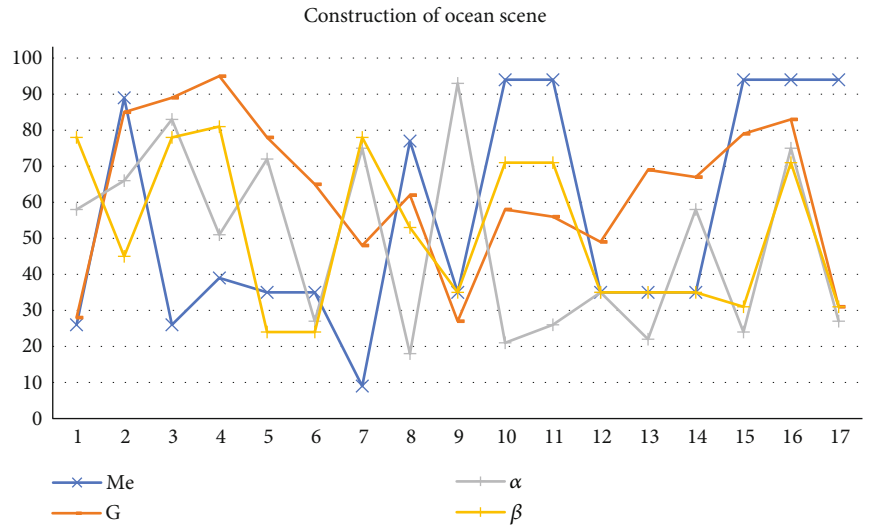


Figure 4: Construction of ocean scene.

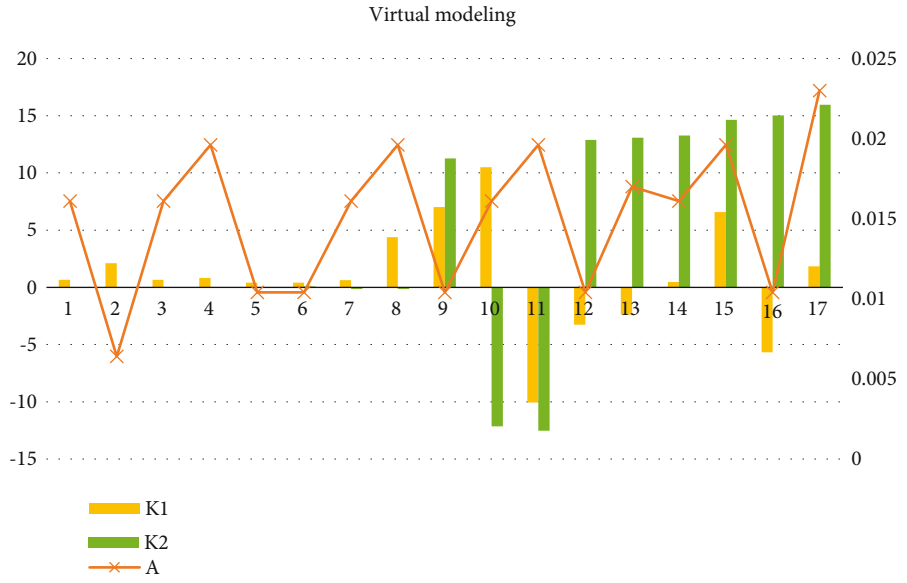


Figure 5: Virtual modeling.

Sea wave crest

$$P_e(x) = P_{ee}(x^2) + xP_{eo}(x^2) \tag{13}$$

$$P_o(x) = P_{oe}(x^2) + xP_{oo}(x^2) \tag{14}$$

$$n \in \{\omega_n^0, \omega_n^1, \cdots, \omega_n^{n-1}\} \tag{15}$$

The wave crest is sharper. Database operation based on XML data scene

$$P(\omega_n^k) = P_e(\omega_n^{2k}) + \omega_n^k P_0(\omega_n^{2k}) \tag{16}$$

Where $k = 0, 1, \cdots, n-1$.

Table 4: Comparison of model parameters.

FFT	MKL	Vega prime	Meta flight	Open flight	α	Omega	β
0.256	0.443	0.192	0.946	0.441	0.828	0.089	0.754
0.807	0.353	0.913	0.823	0.893	0.598	0.365	0.348
0.557	0.701	0.725	0.757	0.818	0.781	0.157	0.739
0.222	0.976	0.764	0.78	0.724	0.866	0.744	0.751
0.841	0.776	0.557	0.256	0.782	0.353	0.913	0.823
0.876	0.235	0.222	0.285	0.803	0.701	0.725	0.757
0.247	0.369	0.773	0.634	0.589	0.752	0.88	0.461

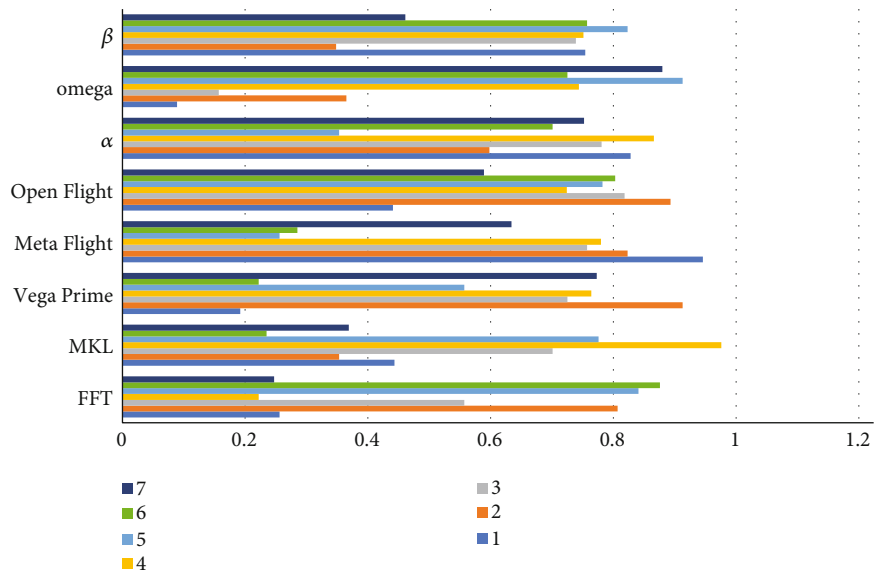


Figure 6: Comparison of model parameters.

C.Vega Prime [24, 25]

$$P\left(\omega_n^k\right) = P_e\left(\omega_{n/2}^{2k}\right) + \omega_n^k P_0\left(\omega_{n/2}^{2k}\right) \tag{17}$$

Where $k = 0, 1, \cdots, n/2 - 1$.

$$d = \frac{\sqrt{3}}{2}a \tag{18}$$

$$d \le Z_{far} \tag{19}$$

Flat trough

$$r_{x_1} = \frac{(x_1/d \cdot \tan(fov/2)) + 1}{2} \cdot r_{width} \tag{20}$$

$$r_{x_2} = \frac{(x_2/d \cdot \tan(fov/2)) + 1}{2} \cdot r_{width} \tag{21}$$

$$r_{x_2} - r_{x_1} = 1 \tag{22}$$

Ocean virtual scene wave Open Flight application range

$$\frac{(x_1/d \cdot \tan(fov/2)) + 1}{2} - \frac{(x_2/d \cdot \tan(fov/2)) + 1}{2} \cdot r_{width} = 1 \tag{23}$$

$$x_2 - x_1 = 2d\frac{\tan(fov/2)}{r_{width}} \tag{24}$$

$$T_{x_1} = \frac{(x_1/d) + 1}{2} \cdot T_{size} \tag{25}$$

Table 5: Marine virtual scene modeling.

Parameter	FFT	MKL	Vega prime
Me	52.08	25.12	15.76
K1	52.33	33.81	17.52
G	55.01	35.41	10.34
K2	40.22	26.88	36.08
α	31.02	38.20	12.66
Omega	34.61	32.71	10.31
β	45.93	44.31	21.91
Me2	38.00	37.68	15.94
K3	34.20	35.32	17.61
G2	45.42	24.83	32.68
K4	35.84	33.15	19.87
J	46.23	40.79	18.01
omega2	37.73	32.06	38.01
β2	39.35	31.81	12.02
M2	40.96	25.70	36.26
P2	42.47	21.38	13.43
t2	45.02	23.60	21.27

Model simulation under different wind speeds

$$T_{x_2} - T_{x_1} = 1 \tag{26}$$

$$x_2 - x_1 = \frac{2d}{T_{size}} \tag{27}$$

$$T_{size} = \frac{R_{width}}{\tan(fov/2)} \tag{28}$$

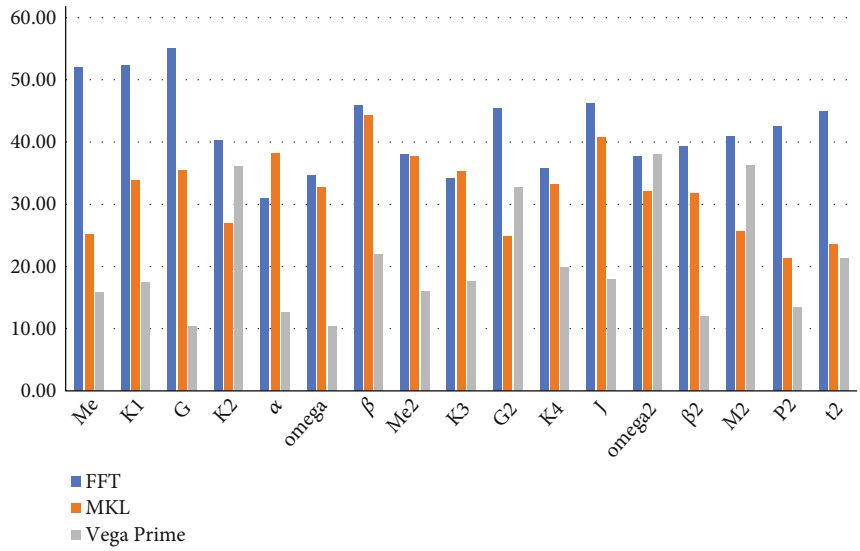


Figure 7: Comparison of model architecture.

4. Simulation Experiment

4.1. FFT Transform Method. The FFT transform method (shown in Table 1 and Figure 3) for modeling and simulation of ocean virtual scenes can improve the image presentation effect of the superposition method based on ocean waves. For the evaluation of modeling and simulation of ocean virtual scenes, The dispersibility of the plane ripple and the continuous sine wave are brought into the evaluation index. The FFT transform method can expand the virtual scene of the ocean. When HDR is turned off, the frame rate at that time =93.569, the average frame rate =95.429, the worst frame rate =90.769, and the optimal frame rate =100.607. When HDR is turned on, the current frame rate =81.267, the average frame rate = 81.26, the worst frame rate = 83.59, and the optimal frame rate = 98.47. Latc increased the frame rate of the simulation to 95.42 and the rendering speed to 89.65. The FFT average frame rate transformation has a periodicity of 96.65, so the worst frame rate is 84.65. There is no need to increase the calculation of grid nodes. Only the ocean virtual scene in the lonc area needs to be calculated, and this ocean virtual scene is repeatedly set and spliced to form a large modeling scene.

4.2. Vega Prime 3D Scene Simulation. Vega Prime has a powerful horizontal correction feature, users can control the correction factor on multiple platforms, and realize Meta Flight cross-platform operation. Meta Flight can increase the sea surface wave crest, the wave crest is sharper and the database operation based on XML data scene, the flatter shape of the wave trough can better play the operation of the database, ocean virtual scene, etc. The wave greatly expands the Open when the sea condition is high. Flight application range. The model simulation under different wind speeds is shown in Table 2, which can simulate the roll of sea waves.

4.3. Construction of Ocean Scene. Set the scene parameter A to be tidal force, M to wave crest, K1 to zp112 test, G to surface heat, K2 to groundwater, α to represent ground wind, omega to sea velocity, and β to surface wind. In the simulation experiment, the first group showed A =0.0161, Me =26, K1 = 0.6488, G =28, K2 = -0.0072, α =58, omega =2.5229, β =78. In the fifth optimization, A = 0.0104, Me = 35, K1 = 0.4084, G = 78, K2 = -0.0528, α = 72, omega = 2.0101, β = 24; when N = 15, A = 0.0008, Me = 94, K1 = 10.506, G =58, K2 = -12.154, α = 21, omega = 12.566, β = 71; N = 17 is the best A = 0.0008, Me = 94, K1 = 1.8399, G = 31, K2 = 15.9602, A = 27, omega = 12.566, β = 31. As shown in Table 3 and Figures 4 and 5.

4.4. Comparison of Model Parameters. The FFT model, MKL model, and Vega Prime model are simulated and compared, and the results are shown in Table 4 and Figure 6. In the FFT model, Meta Flight, Open Flight, α, mega, and β parameters are used for evaluation. The best parameters of the FFT model are Meta Flight =0.946, Open Flight =0.441, α =0.828, omega = 0.089, β =0.754; the best parameters of MKL are: Meta Flight =0.757, pen Flight =0.818, α =0.781, omega =0.157, β = 0.739; the best parameters of the Vega Prime model are Meta Flight =0.285, Open Flight =0.803, α =0.701, omega =0.725, β =0.757.

Insert some two-dimensional animations into the construction of the ocean virtual scene, as shown in Table 5, Figure 7, and Figure 8. In the FFT model, Me = 52, K1 = 52, G = 55, K2 = 40, α = 31, omega = 35, β = 46, Me2 = 38, K3 = 34, G2 = 45, K4 = 36, J = 46, omega2 = 38, β2 = 39, M2 = 41, P2 = 42, t2 = 45. Two-dimensional animation is integrated into digital images as a kind of visual special effects elements or animated subtitles, making digital images more interesting, creative, and spreading wider. In the MKL model, Me = 25, K1 = 34, G = 35, K2 = 27, α = 38, omega = 33, β = 44, Me2 = 38, K3 = 35, G2 = 25, K4 = 33, J = 41, omega2 = 32, β2 = 32, M2 =26, P2 = 21, t2 = 24; in the Vega Prime model, Me = 16, K1 = 18, G = 10, K2 = 36, α = 13,

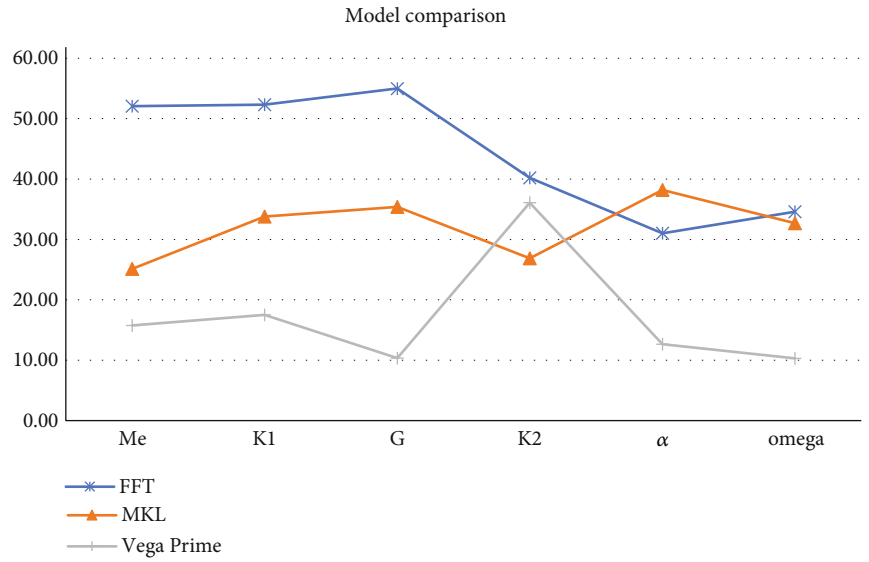


Figure 8: Model comparison.

omega = 10, $\beta = 22$, Me2 = 16, K3 = 18, G2 = 33, K4 = 20, J = 18, omega2 = 38, $\beta 2 = 12$, M2 = 36, P2 = 13, t2 = 21.

From Figure 8, FFT has high execution time and poor performance under different indexes. While MKL is at a general level, it has certain advantages over FFT as a whole. The Vega Prime model has good performance, can achieve lower execution time, for the other two models, the advantage is more obvious.

5. Conclusion

The ocean virtual scene display of digital images is more flexible and more tolerant. The ocean virtual scene can be used to model the structure of different regions, create a diversified ocean experience pavilion as much as possible, and create a multi-dimensional visual product. Two-dimensional animation is integrated into digital images as a kind of visual special effects elements or animated subtitles, making digital images more interesting, creative, and spreading wider. Digital image of the ocean virtual scene has a good application scene, can be applied to multi-dimensional visual effects. Using big data parallel processing technology to quickly analyze related features and improve analysis accuracy. It can be applied comprehensively from different feature attributes for different scenes, and reflects the complex application of digital images in marine virtual scenes.

Data Availability

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

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by a grant from Brain Korea 21 Program for Leading Universities and Students (BK21 FOUR) MADEC Marine Designeering Education Research Group.

References

[1] M. D. Presnar, A. D. Raisanen, D. R. Pogorzala, J. P. Kerekes, and A. C. Rice, "Dynamic scene generation, multimodal sensor design, and target tracking demonstration for hyperspectral/Polarimetric performance-driven sensing," *Proceedings of SPIE-The International Society for Optical Engineering*, vol. 7672, article 76720T, 2010.

[2] J. A. Rupkalvis and R. Gillen, *A New Method for Combining Live Action and Computer Graphics in Stereoscopic 3D[C]// Engineering Reality of Virtual Reality*, International Society for Optics and Photonics, San Jose, CA(US), 2008.

[3] G. Zhang, C. Huang, J. Li, and X. Zhang, "Constrained coordinated path-following control for underactuated surface vessels with the disturbance rejection mechanism," *Ocean Engineering*, vol. 196, p. 106725, 2020.

[4] J. Liu, D. Chen, Y. Wu, R. Chen, P. Yang, and H. Zhang, "Image edge recognition of virtual reality scene based on multi-operator dynamic weight detection," *IEEE Access*, vol. 8, pp. 111289–111302, 2020.

[5] S. Alkahtani, A. Eisa, J. Kannas, and G. Shamlan, "Effect of acute high-intensity interval cycling while viewing a virtual natural scene on mood and eating behavior in men: A randomized pilot trial," *Clinical Nutrition Experimental*, vol. 28, pp. 92–101, 2019.

[6] K. Zhbanova, "Ocean underwater scene dioramas of first graders with submarine porthole views," *Journal of STEM Arts, Crafts, and Constructions*, vol. 4, no. 1, article 4, 2019.

[7] T. Sieberth, A. Dobay, R. Affolter, and L. C. Ebert, "Applying virtual reality in forensics - a virtual scene walkthrough," *Forensic Science Medicine & Pathology*, vol. 15, no. 1, pp. 41–47, 2019.

[8] M. Liao, B. Song, S. Long, H. E. Minghang, C. Yao, and X. Bai, "SynthText3D: synthesizing scene text images from 3D virtual worlds," *SCIENCE CHINA Information Sciences*, vol. 63, no. 2, pp. 1–14, 2020.

[9] X. Ning, P. Duan, W. Li, and S. Zhang, "Real-time 3D face alignment using an encoder-decoder network with an efficient deconvolution layer," *IEEE Signal Processing Letters*, vol. 27, pp. 1944–1948, 2020.

[10] A. Bayramova, T. Mane, T. Ogunleye, S. C. Taylor, and E. Bernardis, "Photographing alopecia: how many pixels are needed for clinical Evaluation?," *Journal of Digital Imaging*, vol. 33, no. 6, pp. 1404–1409, 2020.

[11] N. P. Dang, K. Chandelon, I. Barthélémy, L. Devoize, and A. Bartoli, "A proof-of-concept augmented reality system in oral and maxillofacial surgery," *Journal of stomatology Oral and Maxillofacial Surgery*, vol. 19, no. 1, pp. 338–342, 2021.

[12] G. B. Chen, Z. Sun, and L. Zhang, "Road identification algorithm for remote sensing images based on wavelet transform and recursive operator," *IEEE Access*, vol. 8, pp. 141824–141837, 2020.

[13] T. Fromenteze, O. Yurduseven, P. D. Hougne, and D. R. Smith, "Lowering latency and processing burden in computational imaging through dimensionality reduction of the sensing matrix," *Scientific Reports*, vol. 11, no. 1, pp. 1–14, 2021.

[14] G. Chen, L. Wang, and M. M. Kamruzzaman, "Spectral classification of ecological spatial polarization SAR image based on target decomposition algorithm and machine learning," *Neural Computing and Applications*, vol. 32, no. 10, pp. 5449–5460, 2020.

[15] B. Huber, D. Mcduff, C. Brockett, M. Galley, and B. Dolan, "Emotional dialogue generation using image-grounded language models," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal, QC, Canada., April 2019.

[16] J. Wang, Z. Li, W. Hu, Y. Shao, and Y. Chen, "Virtual reality and integrated crime scene scanning for immersive and heterogeneous crime scene reconstruction," *Forensic Science International*, vol. 303, article 109943, 2019.

[17] W. Li, L. Liu, and J. Zhang, "Fusion of SAR and optical image for sea ice Extraction," *Journal of Ocean University of China*, vol. 20, no. 6, pp. 1440–1450, 2021.

[18] K. Rahimi, C. Banigan, and E. D. Ragan, "Scene Transitions and Teleportation in Virtual Reality and the Implications for Spatial Awareness and Sickness," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 6, pp. 2273–2287, 2018.

[19] T. C. Bybee and S. E. Budge, "Method for 3-D scene reconstruction using fused LiDAR and imagery from a Texel Camera," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8879–8889, 2019.

[20] Y. Chen, D. Wang, and G. Bi, "An image edge recognition approach based on multi-operator dynamic weight detection in virtual reality scenario," *Cluster Computing*, vol. 22, pp. 8069–8077, 2019.

[21] N. Sarhangnejad, N. Katic, Z. Xia, M. Wei, and R. Genov, "Dual-tap computational photography image sensor with per-pixel pipelined digital memory for intra-frame coded multi-exposure," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 11, pp. 3191–3202, 2019.

[22] M. Awad, A. Elliethy, and H. A. Aly, "Adaptive near-infrared and visible fusion for fast image enhancement," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 408–418, 2020.

[23] M. A. Guang-Ming, C. H. Wang, and L. I. Li-Ning, "Based on CityEngine 3D virtual campus scene design and implementation," *Computer Engineering & Software*, vol. 63, no. 2, pp. 1–14, 2019.

[24] Z. Sunm, Y. Q. Liu, C. R. Zhang, J. Shi, and Y. Y. Chen, "A scene-distributed interactive rendering system," *Journal of Graphics*, vol. 3, no. 2, pp. 1–14, 2019.

[25] Y. Liu, L. Shen, G. Zhang, and F. Liu, "Design on virtual simulation experiment for digital teaching video shooting and evaluation in complex scene," *Experimental Technology and Management*, vol. 6, pp. 1–14, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6144741, 10 pages
https://doi.org/10.1155/2022/6144741

Research Article

Energy-Efficient Resource Allocation for Backscatter-Assisted Wireless Powered Communication Networks in Twin Workshop

Yujian Li and Xinxing Zhang

Department of Mechanical and Electrical Engineering, Quzhou College of Technology, China

Correspondence should be addressed to Xinxing Zhang; redrobot@zjut.edu.cn

Received 1 March 2022; Revised 14 April 2022; Accepted 22 April 2022; Published 19 May 2022

Academic Editor: Yinghui Ye

The problem of energy shortage in sensor nodes caused by frequent data interactions is one of the major constraints on the development of twin workshops. A backscatter-assisted wireless powered communication network (BAWPCN) has been deemed a potential solution for addressing the problem of energy shortage in twin workshops. How to effectively ensure the link energy efficiency (EE) while satisfying the quality of servers for each user has been of high interest, while it has not been well studied in previous works. Inspired by this, in this paper, we propose a resource allocation scheme based on the max–min criterion, considering the user quality of service and energy-causality constraints. The optimization problem is formulated as a mixed-integer nonconvex fractional planning problem which is aimed at maximizing the minimum EE of each link. The generalized fractional theory is used to transform the nonconvex fractional planning problem into an equivalent mixed-integer nonconvex subtraction optimization problem, and then, the mixed-integer nonconvex subtractive optimization problem is transformed into an equivalent nonconvex optimization problem by introducing relaxation variables to eliminate the integer programming problem arising from the maximum-minimum function. Based on this, the block coordinate descent method is used to decompose the transformation problem into two convex subproblems, and an iterative algorithm is proposed to solve the transformation problem. Simulation results verify the fast convergence of the proposed iterative algorithm and show that the proposed resource allocation method can effectively guarantee the fairness of the energy efficiency of the system in twin workshops.

1. Introduction

Industry 4.0 refers to using the Cyberphysical System (CPS) to digitize and intellectualize the supply, manufacturing, and sales information in production and finally achieves fast, effective, and personalized product supply. The core of Industry 4.0 is the interconnection and intelligent operation of the physical and information worlds of manufacturing. The combination of intelligent automation (e.g., robotics) and a new generation of information technology (e.g., Internet of Things and artificial intelligence) produces the digital twin workshop, which is one of the current trends in manufacturing in the context of Industry 4.0 [1]. Digital twin (DT) is driven by the multidimensional virtual model and fused data and realizes monitoring, simulation, prediction, optimization, and other actual functional services and application requirements through virtual and real closed-loop interaction [2], as shown in Figure 1. Digital twin workshops (DTW), as an important enabling way to realize digital transformation, promote intelligent upgrading, and accelerate Industry 4.0, have moved from theoretical research to the practical application stage, among which the development of Internet of Things technology is the key driving force.

With the development of Internet of Things (IoT) technology and its wide application in the manufacturing industry, it greatly enhances the real-time data acquisition ability and the transmission ability of production factors in twin workshops. The realization of ubiquitous IoT requires the deployment of numerous low-power sensors, but the frequent data interaction

will greatly consume sensor energy [4]. One way to solve the problem is replacing the battery, but the deployment of sensors in an industrial production environment cannot meet the demand to frequently replace the battery. At the same time, as the power consumption and volume of wireless sensor nodes become smaller and smaller, the advantages of small and flexible wireless sensor nodes will be limited if they only rely on their own power supply. Therefore, the energy shortage of sensor nodes caused by frequent data interaction is one of the important factors restricting the development of twin workshops. In recent years, thanks to the progress of science and technology, scholars have proposed wireless power communication network (WPCN) and backscatter communication technologies to solve the energy limitation problem of sensor nodes. The WPCN deploys dedicated energy stations to provide the energy resource to the sensor nodes through the wireless energy transfer (WET) technology, and the nodes leverage the harvested energy to transmit information. Therefore, the core of WPCN design lies in the joint allocation of energy and time resource. The authors in [5] proposed the harvest-then-transmit (HTT) protocol by incorporating WET into wireless communication networks, which is the main basis for the operation of current WPCNs. In [6], the half-duplex mode is extended to the full-duplex mode, and the base station contains two antennas so that it can simultaneously transmit energy signals and receive transmission data. However, the transmitter performs active transmission (AT) by the HTT protocol, which leads to high power consumptions. Thus, the transmitter has to allocate a long period to harvest sufficient energy but leaves a short period to transmit information. Backscatter communication (BackCom) technology adopts a relatively simple modulation method to load its information to external RF signals for transmission. Thus, BackCom avoids power-consuming components, providing a transmission strategy with low power consumption and low transmission rate for IoT scenarios [7]. The long period for HTT to harvest energy is available for BackCom to increase the throughput. Thus, BackCom can be combined with HTT to achieve different tradeoffs between energy harvesting (EH) and data transmission.

The authors of [8] take the wireless digital TV signal as an example to show that the sensor can maintain communication by using the existing wireless signal without batteries. In [9, 10], the authors combined BackCom with WPCN and proposed a backscatter-assisted wireless powered communication network (BAWPCN). For the BAWPCNs, the authors of [11] considered a network combining multimode backscatter communication and HTT and proposed a time slot resource allocation scheme to maximize subuser link capacity. In [12], the authors studied the optimal time allocation of energy harvesting, backscatter, and wireless transmission to maximize the throughput. The authors of [13] proposed a resource allocation scheme to maximize the throughput by jointly optimizing the transmission time, reflection coefficient, and transmit power in the full-duplex BackCom. The authors of [14] considered the scenario of nonorthogonal multiple access and maximized the throughput by jointly optimizing the backscatter time and reflection coefficient, subject to the constraints of energy collection threshold and signal-to-noise ratio. Energy efficiency (EE) is also one of the important indicators of wireless communication networks; the authors in [15] studied the resource allocation method to maximize the user EE in the BAWPCNs. Then, users' EE was also studied in an unmanned aerial vehicle-based WPCN with backscatter communications [16]. Based on the analysis of existing literature, it can be seen that there are many researches on BAWPCNs from the evaluation of network outage capacity or spectral efficiency, while there are few researches from the perspective of energy efficiency. Meanwhile, the following three problems still exist in the research process: (1) there are only single nodes in BackCom; (2) only the time dimension is considered in the design of resource allocation, while the power resource allocation is ignored; and (3) the energy harvesting model adopts the linear mode. However, in the case of a multiuser network in the twin workshop, the communication network composed of multiple nodes should be considered. A large number of nonlinear components such as photoelectric sensors and Hall sensors are used in twin workshops, and the actual energy harvesting circuit presents nonlinear characteristics. At the same time, power resources also affect the level of energy efficiency. Inspired by the above three problems, aiming at the energy shortage caused by frequent data exchange of each sensor node in the twin workshop, this paper introduces a BAWPCN to study the time-power two-dimension resource allocation method for multiple users and optimize the energy efficiency of the communication link. Our main contributions are summarized as follows.

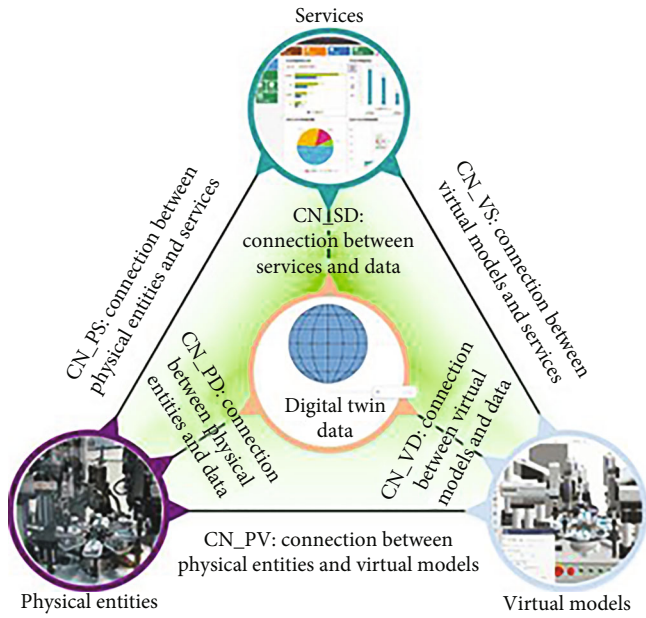


Figure 1: Digital twin model [3].

(i) A joint energy efficiency/fairness time-power two-dimensional resource optimization model for the system based on the max–min criterion is proposed. The proposed optimization model not only ensures fair access to communication resources for users but also considers the optimization of the time slot, transmit power of the power beacon, and nodes simultaneously

(ii) The joint optimization of time slot and power leads to coupling of multiple variables and a fractional form of the objective function. In addition, there is a max–min function of the objective function. The proposed optimization model is therefore a mixed-integer nonconvex fractional programming problem and cannot be solved directly using the existing tools such as CVX

(iii) The proposed iterative algorithm is verified to converge quickly through simulation and is shown to achieve better fairness of the link EE when compared with similar algorithms

The rest of this paper is organized as follows. In Section 2, the system model for BAWPCN is introduced. In Section 3, we formulate the problem to maximize the minimum links' EE and solve it for the EE fairness resource allocation scheme. In Section 4, the simulations are conducted to evaluate the performance of the proposed scheme. Section 5 concludes this paper.

2. System Model

Consider a BAWPCN as shown in Figure 2, the network consists of a gateway (GW) to receive information, a dedicated power beacon (PB) to provide energy, and M power-limited users (EUs).

The EUs need to upload their data to the gateway within time slot T. The PB was deployed to provide radio frequency (RF) signals to EUs who can use the received RF signals for backscatter communication and energy harvesting.

In order to avoid interference between EUs, time division multiple access is adopted to decompose the whole time slot T into multiple small time slots, as shown in Figure 2. In αT, PB broadcasts the unmodulated RF signals, and all EUs can use the received RF signals for energy harvesting or backscatter information. Specifically, in τ_0, all EUs work in the energy harvesting mode. At the time slot τ_m $(m = 1, 2, \cdots, M)$, EU m transmits data to the GW through backscatter technology while the other EUs continue to harvest energy. In $(1-\alpha)T$, the PB remains silent and EU m transmits its data to the GW in slot τ_m by active communication. The signal received by the mth EU in αT can be expressed as

$$y_m^{\tau_0} = w_m + \sqrt{P_0} g_m x, \tag{1}$$

where P_0 represents the transmitted power of the PB, x denotes RF signals transmitted by PB and $E[|x|^2] = 1$, w_m represents the receiving noise of mth EU and follows a Gaussian distribution with the mean of 0 and variance of σ^2, and g_m represents the channel coefficient from PB to mth EU. Using the nonlinear energy harvesting model proposed in [17], the energy harvested by EU m within time slot αT can be expressed as follows:

$$\Phi_m^{\text{total}}(P_0, \alpha, \tau_m) = (\alpha T - \tau_m) E_{\max} \frac{1 - \exp\left(-cP_0|g_m|^2\right)}{1 + \exp\left(-cP_0|g_m|^2 + cd\right)}, \tag{2}$$

where $E_{\max}$ is the maximum harvested power of the energy collector, c and d are parameters of the nonlinear energy model, and their values can be obtained by fitting the actual measured data. Since the value of harvested energy is positive, the value of c must meet the requirement $1 - \exp\left(-cP_0|g_m|^2\right) > 0$. It should be noted that the time for EU m to harvest energy is $(\alpha T - \tau_m)$, not αT, because EU m operates in the backscattering communication mode within the τ_m time slot. It should be pointed out that the energy collected in formula (2) will be used for mth EU's backscatter communication and HTT energy dissipation in τ_m and t_m time slots. In τ_m, EU m backscatters information to the gateway via the backscattering communication technique, where the instantaneous power of the reflected signal received by the gateway from mth EU can be expressed as

$$P_{\tau_m}(P_0) = \frac{4P_0|g_m|^2|h_m|^2\varepsilon^2(\Gamma_0 - \Gamma_1)^2}{\pi^2}, \tag{3}$$

where ε is the scattering efficiency of the backscattering communication module, h_m denotes the channel coefficient from mth EU to the gateway, and Γ_0 and Γ_1 denote the reflection coefficient. The Fries transfer formula is used to model the channel gain, $|g_m|^2 = G_p G_h \lambda^2/(4\pi d_{0m})^2$, $|h_m|^2 = G_h G_r \lambda^2/(4\pi d_{1m})^2$, where λ denotes the wavelength; G_p, G_h, G_r denote the antenna gain of the PB, the EUs, and the gateway, respectively; and d_{0m} and d_{1m} denote the distance from the PB to m th EU and mth EU to the gateway, respectively. According to Equation (3) and Shannon's capacity theorem, the achievable throughput of mth EU at τ_m can be calculated as follows:

$$C_{\tau_m}^{\text{Back}}(\tau_m, P_0) = B_0 \tau_m \log_2\left(1 + \frac{\xi P_{\tau_m}(P_0)}{\sigma^2}\right), \tag{4}$$

where B_0 denotes bandwidth. Since backscatter communication uses simple modulation, its channel capacity is smaller than that of conventional communication modes. In this paper, we use the same approach as in [18] to characterize this difference in channel capacity, i.e., multiplying the signal reception signal-to-noise ratio by a nonnegative real number ζ with a factor less than 1 $(0 < \zeta < 1)$. In $(1-\alpha)T$, mth EU transmits data to the gateway in time slot t_m in the traditional communication mode, so the throughput that mth EU can accomplish is expressed as

$$C_{t_m}^{\text{HTT}}(t_m, P_{t_m}) = B_0 t_m \log_2\left(1 + \frac{P_{t_m}|h_m|^2}{\sigma^2}\right), \tag{5}$$

where P_{t_m} denotes the transmit power of mth EU during time slot t_m. Thus, the total throughput by mth EU during the entire time slot T can be expressed as

$$C_m^{\text{total}}(\tau_m, t_m, P_{t_m}, P_0) = B_0\left(\tau_m \log_2\left(\left(1 + \frac{\xi P_{\tau_m}(P_0)}{\sigma^2}\right) + t_m \log_2\left(1 + \frac{P_{t_m}|h_m|^2}{\sigma^2}\right)\right)\right). \tag{6}$$

According to Equation (18) in [19], the EU EE is the ratio of

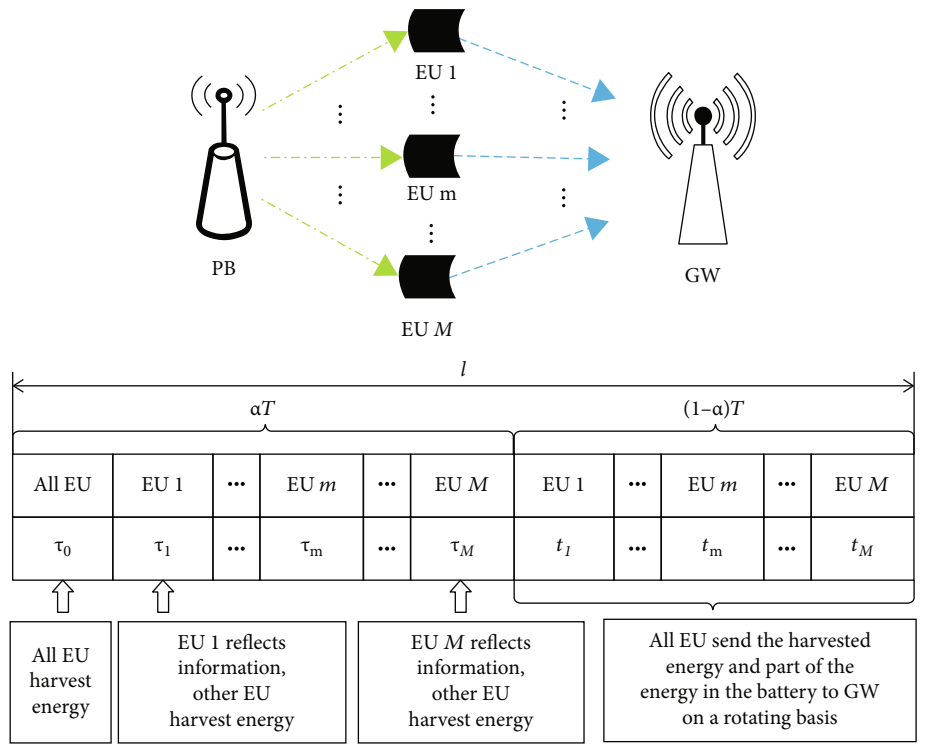


Figure 2: System model.

the total throughput achieved by the user to the consumed energy. Equation (6) already gives the total throughput that can be completed by user m throughout the time slot. Next, this section analyses the energy consumed by mth EU throughout the time slot and the expression of EE.

During time slots τ_0 and τ_m, mth EU collects energy from a dedicated energy station. During time slots τ_m and t_m, the mth EU needs to consume energy. During time slot τ_m, there is only circuit loss as the mth EU modulates its own information to the received RF signal (i.e., the RF signal emitted by the energy source) and does not need to generate its own carrier. During time slot t_m, the mth EU uses a conventional communication mode to transmit data, so its energy consumption consists of two components: the energy consumed by the transmit power and the circuit losses. In summary, the total system energy consumed by mth EU is

$$E_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0\right) = P_0(\tau_0 + \tau_m) + \left(p_{m,c}^{\text{HTT}} + P_{t_m}\right)t_m + p_{m,c}^{\text{Back}}\tau_m, \tag{7}$$

where $p_{m,c}^{\text{Back}}$ denotes the circuit loss when mth EU operates in the backscatter communication mode, $p_{m,c}^{\text{HTT}}$ denotes the circuit loss when mth EU is operating in the conventional communication mode, and P_{t_m} denotes the transmit power of m th EU in time slot t_m.

Combining Equations (6) and (7), the link EE of mth EU can be expressed as

$$\eta_m\left(\tau_m, t_m, P_{t_m}, P_0\right) = \frac{C_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0\right)}{E_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}\right)}. \tag{8}$$

3. Resource Allocation Method for Energy Efficiency Fairness of the System

3.1. Problem Formulation. In this subsection, we propose a resource allocation scheme that can effectively guarantee the EE fairness of communication links. The max–min criterion is an effective way to guarantee user fairness [20], so that the resource allocation method for guaranteeing energy-efficient fairness of links among EUs can be achieved by solving the following optimization problem, namely,

$$\begin{aligned}
&\max_{\{\tau_m\},\{t_m\},\left\{P_{t_m}\right\},P_0,\beta,\tau_0} \min_m \eta_m\left(\tau_m, t_m, P_{t_m}, P_0\right),\\
&\text{C1}: \tau_0 + \sum_{m=1}^{M}\tau_m \le \alpha T,\\
&\text{C2}: \sum_{m=1}^{M} t_m \le (1-\alpha)T,\\
&\text{C3}: p_{m,c}^{\text{Back}}\tau_m + \left(p_{m,c}^{\text{HTT}} + P_{t_m}\right)t_m \le \Phi_m^{\text{total}}(P_0, \alpha, \tau_m), \forall m,\\
&\text{C4}: 0 \le P_{t_m} \le P_m^{\max}, \forall m,\\
&\text{C5}: C_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0\right) \ge C_m^{\min}, \forall m,\\
&\text{C6}: 0 \le P_0 \le P_0^{\max},\\
&\text{C7}: \tau_0 > 0, \tau_m \ge 0, t_m \ge 0, \forall m.
\end{aligned} \tag{9}$$

C3 ensures that the harvested energy is greater than the energy consumed by user m. C4 constrains the maximum

transmit power of the mth EU when working in the traditional information transmission mode. C5 indicates that the throughput achieved by mth EU in the whole time slot T cannot be less than the given minimum value, i.e., the communication service quality of user m is guaranteed. C6 constrains the maximum transmit power of the PB.

As can be seen from the objective function, the EE of m th link η_m is a fraction function, which contains multiple coupled variables, such as the variables P_{tm} and t_m, P_0, τ_0, and τ_m. Secondly, since the max–min criterion is adopted, the integer variable M needs to be optimized. Therefore, the optimization problem (9) is a mixed-integer nonconvex fractional programming optimization problem. Therefore, the existing optimization tools such as CVX cannot be directly used to solve the original problem.

3.2. Problem Transformation. For the mixed-integer nonconvex fractional programming problem in Equation (9), we divide three steps to obtain its optimal solution. Firstly, the nonconvex fractional programming problem is transformed into an equivalent mixed-integer nonconvex subtractive optimization problem by using the generalized fractional programming theory. Secondly, the integer programming problem caused by the max–min function is eliminated by introducing relaxation variables; that is, the mixed-integer nonconvex subtraction optimization problem is transformed into an equivalent nonconvex optimization problem. Finally, the block coordinate descent (BCD) technology was used to decompose the transformation problem into two convex subproblems [21], and then, an iterative algorithm was designed to solve the transformation problem.

The BCD algorithm encompasses a wider range of traditional alternating optimization algorithms as well as coordinate descent algorithms. The BCD algorithm is usually applied to optimization problems with nonconvex objective functions or feasible domains, but when the optimization variables are blocked, the objective functions and feasible domains of such optimization problems are convex with respect to the blocks of variables. In addition, the scope of the BCD algorithm can be extended, for example, by using approximate optimization for nonconvex objective functions and feasible domains after blocking the variables, where the BCD algorithm is also feasible [22].

The specific processing steps are described below.

(1) Let the variable Q denote the value of the objective function of problem (9), i.e., the max–min EE. According to the generalized fractional programming theory, the sufficient condition for obtaining the optimal solution to the optimization problem (9) is that Equation (10) holds when the C1-C7 constraints are satisfied:

$$\max_{\{\tau_m\},\{t_m\},\{P_{t_m}\},P_0,\beta,\tau_0} \min_m C_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0\right) - Q * E_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}\right) = \min_m \left[C_m^{\text{total}}\left(\tau_m^*, t_m^*, P_{t_m}^*, P_0^*\right) - Q * E_m^{\text{total}}\left(\tau_m^*, t_m^*, P_{t_m}^*\right)\right] = 0. \tag{10}$$

We can obtain the optimal solution to the original problem by solving the optimization problem (10). However, in practice, Q^* is often unknown, but we can obtain Q by continuously updating the value of Q^*, which is shown in Algorithm 1.

According to Algorithm 1, the core of solving the original problem (9) is to solve the optimization problem. Compared to the original problem (9), the objective function of optimization problem does not have a fractional form, but it is still a mixed-integer nonconvex optimization problem.

(2) A relaxation variable is introduced to transform the mixed-integer nonconvex optimization problem into the following optimization problem, namely,

$$\begin{aligned} &\max_{\{\tau_m\},\{t_m\},\{P_{t_m}\},P_0,\alpha,\tau_0} \theta \\ &\text{s.t.C1} - \text{C7}, \\ &\text{C8}: C_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0\right) - QE_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}\right) \geq \theta, \forall m. \end{aligned} \tag{11}$$

In the optimization problem (11), the optimization objective is a linear function and the constraints C1, C2, C4, C6, and C7 are linear constraints, but the remaining constraints C3, C5, and C8 are nonconvex, so the optimization problem (11) is a nonconvex problem.

(3) The optimization problem (11) is transformed into two subproblems using the BCD technique as follows:

(1) Given $P_0^{(l)}$, solve for $\tau_m^{(l)}$, $t_m^{(l)}$, $P_{t_m}^{(l)}$, $\tau_0^{(l)}$, $\alpha^{(l)}$ by the following problem:

$$\begin{aligned} &\max_{\{\tau_m\},\{t_m\},\{P_{t_m}\},\alpha,\tau_0} \theta \\ &\text{s.t.C1, C2, C4, C7}, \\ &\text{C3}: p_{m,c}^{\text{Back}}\tau_m + \left(p_{m,c}^{\text{HTT}} + P_{t_m}\right)t_m \leq \Phi_m^{\text{total}}\left(P_0^{(l)}, \alpha, \tau_m\right), \forall m, \\ &\text{C5}: C_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0^{(l)}\right) \geq C_m^{\min}, \forall m, \\ &\text{C6}: 0 \leq P_0^{(l)} \leq P_0^{\max}, \\ &\text{C8}: C_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0^{(l)}\right) - QE_m^{\text{total}}\left(\tau_m, t_m, P_{t_m}, P_0^{(l)}\right) \geq \theta, \forall m. \end{aligned} \tag{12}$$

In the optimization problem (12), constraints C3, C5, and C8 still exist in the coupled variable and are jointly nonconvex. Specifically, in C3 and C8, the variable P_{t_m} and t_m are coupled; in C5 and C8, the joint nonconvexity is present in $t_m \log_2(1 + P_{t_m}|h_m|^2/\sigma^2)$. To solve the above problem, we introduce auxiliary variable $x_m = P_{t_m} t_m$ and bring them into

$$\begin{aligned}
&\max_{\{\tau_m\},\{t_m\},\{x_m\},\alpha,\tau_0} \theta \\
&\text{s.t.C1, C2, C7,} \\
&\text{C3}: p_{m,c}^{\text{Back}}\tau_m + p_{m,c}^{\text{HTT}}t_m + x_m \le \Phi_m^{\text{total}}\left(P_0^{(l)}, \alpha, \tau_m\right), \forall m, \\
&\text{C4}: 0 \le x_m \le P_m^{\max}t_m, \\
&\text{C5}: B_0\tau_m \log_2\left(\left(1 + \frac{\xi P_{\tau_m}\left(P_0^{(l)}\right)}{\sigma^2}\right)\left(1 + \frac{x_m|h_m|^2}{t_m\sigma^2}\right)\right) \ge C_m^{\min}, \forall m, \\
&\text{C6}: 0 \le P_0^{(l)} \le P_0^{\max}, \\
&\text{C8}: B_0\tau_m \log_2\left(\left(1 + \frac{\xi P_{\tau_m}\left(P_0^{(l)}\right)}{\sigma^2}\right)\left(1 + \frac{x_m|h_m|^2}{t_m\sigma^2}\right)\right) - Q\left(P_0^{(l)}(\tau_0 + \tau_m) + p_{m,c}^{\text{Back}}\tau_m + p_{m,c}^{\text{HTT}}t_m + x_m\right) \ge \theta, \forall m.
\end{aligned} \tag{13}$$

the optimization problem (13) to obtain the equivalent optimization problem as follows, i.e.,

(2) Given $\tau_m^{(l)}$, $t_{m^{(l)}}$, $P_{t_m}^{(l)}$, $\tau_0^{(l)}$, $\alpha^{(l)}$, we can obtain $P_0^{(l)}$ by solving the following optimization problem:

$$\begin{aligned}
&\max_{P_0} \theta \\
&\text{s.t.C1}: \tau_0^{(l)} + \sum_{m=1}^{M} \boxtimes\tau_m^{(l)} \le \alpha^{(l)}T, \\
&\text{C2}: \sum_{m=1}^{M} \boxtimes t_m^{(l)} \le \left(1 - \alpha^{(l)}\right)T, \\
&\text{C3}: p_{m,c}^{\text{Back}}\tau_m^{(l)} + p_{m,c}^{\text{HTT}}t_m^{(l)} + x_m^{(l)} \le \Phi_m^{\text{total}}\left(P_0, \alpha^{(l)}, \tau_m^{(l)}\right), \forall m, \\
&\text{C4}: 0 \le x_m^{(l)} \le P_m^{\max}t_m^{(l)}, \\
&\text{C5}: B_0\tau_m^{(l)} \log_2\left(\left(1 + \frac{\xi P_{\tau_m}(P_0)}{\sigma^2}\right)\left(1 + \frac{x_m^{(l)}|h_m|^2}{t_m^{(l)}\sigma^2}\right)\right) \ge C_m^{\min}, \forall m, \\
&\text{C6}: 0 \le P_0 \le P_0^{\max}, \\
&\text{C7}: \tau_0^{(l)} > 0, \tau_m^{(l)} \ge 0, t_m^{(l)} \ge 0, \forall m, \\
&\text{C8}: B_0\tau_m^{(l)} \log_2\left(\left(1 + \frac{\xi P_{\tau_m}(P_0)}{\sigma^2}\right)\left(1 + \frac{x_m^{(l)}|h_m|^2}{t_m^{(l)}\sigma^2}\right)\right) - Q\left(P_0\left(\tau_0^{(l)} + \tau_m^{(l)}\right) + p_{m,c}^{\text{Back}}\tau_m^{(l)} + p_{m,c}^{\text{HTT}}t_m^{(l)} + x_m^{(l)}\right) \ge \theta, \forall m.
\end{aligned} \tag{14}$$

Lemma 1. *The optimization problem (13) is a convex problem.*

Proof. Please see the appendix.

According to Lemma 1, we can use the CVX tool to obtain the optimal solution to solve the optimization problem (13). □

In the optimization problem (14), the objective function is linear and C1, C2, C4, and C7 are independent of the optimization variables. C6 is a linear constraint. Meanwhile, C3 and C5 are convex constraints. The left-hand side of C8 is shaped as $f(x) = C\log_2(1 + Dx) - ax + b$, where a, b, C, and D are all constants and greater than zero. The second-

1: Solve the optimal solution of the optimization problem given any Q greater than zero
2: Substitute the optimal solution obtained in the first step into the objective function of the optimization problem (9) to update $Q(0)$
3: If Q does not converge, let $Q = Q(0)$ and repeat the first step; if not, let $Q^* = Q$, and the optimal solution in the first step is the optimal solution of the original problem (9):

$$\max_{\{\tau_m\},\{t_m\},\{P_{t_m}^{(2)}\},P_0,\alpha,\tau_0} \min_m C_m^{\text{total}}(\tau_m, t_m, P_{t_m}, P_0) - QE_m^{\text{total}}(\tau_m, t_m, P_{t_m})$$

s.t.C1 − C7

Algorithm 1: The link EE fairness resource allocation algorithm.

1: Initialize the system parameters, given $P_0^{(l)}$, the convergence accuracy δ, and the maximum iteration number L
2: Repeat
3: For given $P_0^{(l)}$, solve the problem (13) by using CVX to obtain $\tau_m^{(l+1)}$, $t_m^{(l+1)}$, $x_m^{(l+1)}$, $\tau_0^{(l+1)}$, α^{l+1}
4: For given $\tau_m^{(l+1)}$, $t_m^{(l+1)}$, $x_m^{(l+1)}$, $\tau_0^{(l+1)}$, α^{l+1}, solve the problem (14) by using CVX to obtain $P_0^{(l+1)}$
5: Until $\theta^{(l+1)} - \theta^{(l)} \le \delta$ or $l = L$
6: Return the optimal solutions: $\tau_m^* = \tau_m^{(l+1)}$, $t_m^* = t_m^{(l+1)}$, $\tau_0^* = \tau_0^{(l+1)}$, $\alpha^* = \alpha^{(l+1)}$, $P_{t_m}^* = P_{t_m}^{(l+1)}$, $P_0^* = P_0^{(l+1)}$

Algorithm 2: Iterative algorithm for solving optimization problem.

1: Given any $Q^{(0)}$ greater than zero, use Algorithm 2 to obtain an optimal solution to the optimization problem (14)
2: Update Q by bringing the optimal solution obtained in the first step into the objective function of the optimization problem (9)
3: If Q does not converge, let $Q^{(0)} = Q$ and repeat the first step; if not, let $Q^* = Q$, and the solution obtained in the first step is the solution to the original problem (9)

Algorithm 3: The overall algorithm of link EE fairness resource allocation scheme.

order derivative of $f(x)$ can be calculated as $f''(x) = -CD^2/(1+Dx)^2 \ln 2$. $f''(x) < 0$, so C8 is a convex constraint. Therefore, the optimization problem (14) is convex, and we can use the CVX tool to obtain the optimal solution to the optimization problem (14).

Based on the above analysis, the optimization problem can be solved by an iterative BCD-based algorithm, as described in Algorithm 2.

As the optimization problems (13) and (14) are convex, the convergence of Algorithm 2 is guaranteed [23]. The overall algorithm of the link energy efficiency fairness resource allocation scheme can be described as Algorithm 3.

4. Simulation Results and Analysis

In this section, we provide simulation results to evaluate the performance of the proposed iterative algorithm. In addition, we illustrate the advantages of the proposed resource allocation scheme based on the max–min criterion to guarantee energy efficiency fairness by comparing with the total energy efficiency maximization scheme. Unless otherwise stated, the parameters listed in Table 1 are used in this section. In addition, the distance between the gateway and the PB is assumed to be 52 m; the distance between the three EUs and the PB is 1.8 m, 1.6 m, and 1.4 m, respectively. Based on the simulated channel gains $|g_m|^2$ and $|h_m|^2$ and the energy harvesting circuits, dedicated power stations produced by the powercast company, the antenna gains of the PB and gateways assume as 5 dBi, and the antenna gain of each EU is 2 dBi in this paper. Assume a carrier frequency of 915 MHz.

The convergence performance of the proposed resource allocation algorithm is shown in Figure 3. It can be seen that the proposed iterative algorithm converges to a certain constant after roughly three to four iterations, which validates the fast convergence of Algorithm 1. Secondly, we can see that ζ has an important influence on the energy efficiency of the system link. As in Equation (4), the max–min EE should increase with the ζ. And we can see that the larger the ζ, the greater the max–min EE from the figure.

Figure 4 shows the achieved EE versus the performance gap ζ. We compare the max–min link EE achieved by the proposed transmission strategy and the transmission strategies of the HTT mode and backscatter communication mode. As can be seen in Figure 4, the proposed transmission strategy achieves the max–min link EE no worse than these

Table 1: Simulation parameter settings.

Parameters	Symbols	Value
The number of EUs	M	3
The maximum transmit power of PB	P_0	3 W
The entire transmission time slot	T	1 s
Bandwidth	B_0	10 MHz
Circuit loss in backscatter communication mode	$p_{m,c}^{\text{Back}}$	400 μW
The maximum transmit power of mth EU	P_{t_m}	10 mW
Circuit loss of transmitting information in HTT mode	$p_{m,c}^{\text{HTT}}$	1 mW
Reflection coefficient	Γ_0	1
Reflection coefficient	Γ_1	-1
Scattering efficiency of the backscatter communication module	ε	-1.1 dB

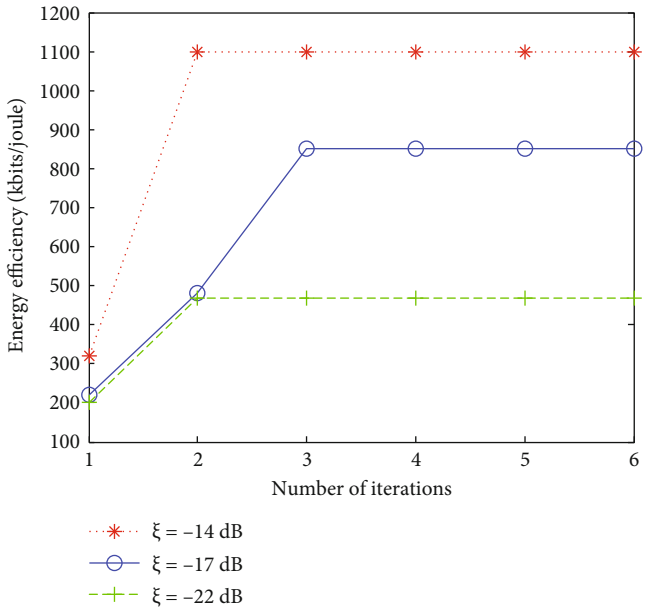


Figure 3: Convergence diagram for Algorithm 1.

transmission strategies regardless of the value of ζ. It is noted that the proposed hybrid transmission strategy can be degraded to these two transmission modes by adjusting parameters. It is consistent that the two transmission modes are special cases of the transmission strategy, and there is a tradeoff between the BackCom and HTT mode in the proposed resource allocation scheme as in theoretical analysis. Specifically, when ζ is less than or equal to -18 dB, the max–min link EE achieved by the proposed transmission strategy is the same as the max–min link EE in the HTT mode. When ζ is greater than -18 dB and less than -14 dB, the proposed transmission strategy makes EUs work in both backscatter and HTT modes in a full transmission time slot; when ζ is greater than -14 dB, it can be seen that the max–min EE accomplished by the proposed transmission strategy is the same as the max–min link EE in the backscatter communication mode, which means that when ζ is larger, the proposed transmission strategy degrades to the backscattered communication mode. In addition, compared to the throughput maximization resource allocation scheme, the EE by the proposed max–min fairness scheme presents better performance. The reason is that the throughput maximization scheme maximizing UE throughput in a BAWPCN does not take into account the factor of EE; the achieved EE is lower than that of the proposed max–min EE fairness scheme. The above analysis shows that the BAWPCNs can indeed combine the advantages of both backscatter and WPCNs, allowing them to adaptively adjust their parameters to meet the different communication goals in complex communication demands.

In Figure 5, we compare the total EE maximization resource allocation scheme with the proposed max–min link EE resource allocation scheme, under the case of large and small interuser channel differences. The major difference between the two resource allocation schemes is the objective function. According to Equation (9), the objective function is the minimum EE for the proposed max–min resource allocation scheme. For the total EE maximization resource allocation scheme, the objective function is the sum rates divided by the sum energy consumptions of M EUs calculated as $\sum_{m=1}^{M}$throughput of the mth SN/$\sum_{m=1}^{M}$energy of the mth SN. It can be seen that the difference in EE between the best and worst EUs under the total EE maximization resource allocation scheme is significantly greater than that of the proposed max–min resource allocation scheme. When the channel difference between users is small, the average EE of users under the total EE maximization resource allocation scheme is slightly higher than that under the proposed max–min resource allocation scheme, but the difference between the EE of the best and worst users is significantly higher than the proposed resource allocation scheme. When the channel difference between users is large, it can be seen that the EE by the total EE maximization scheme results in the energy efficiency of the best user being 2.2 times that of the worst link, but the EE of the best link is about 1.4 times that of the worst link by the proposed resource allocation scheme, which effectively ensures fair access to resources between EUs. This is because the total EE maximization resource allocation scheme inclines the EUs with good channel status to maximize the EE,

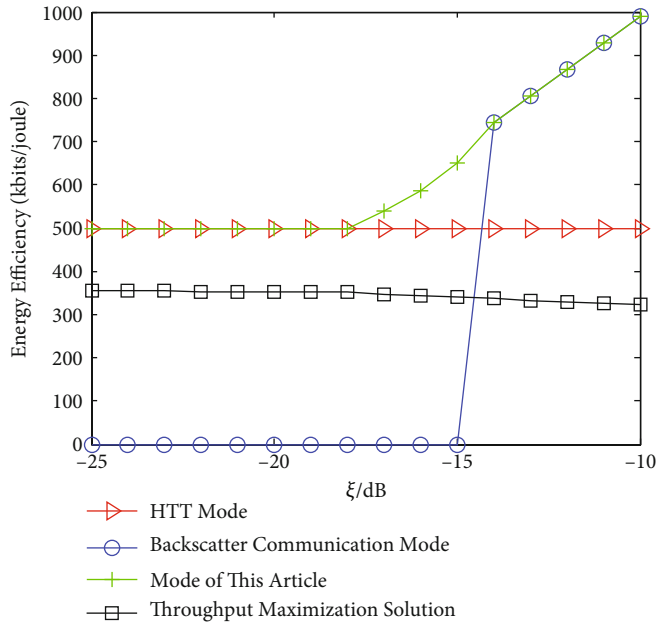


Figure 4: Performance comparison of the four resource allocation schemes.

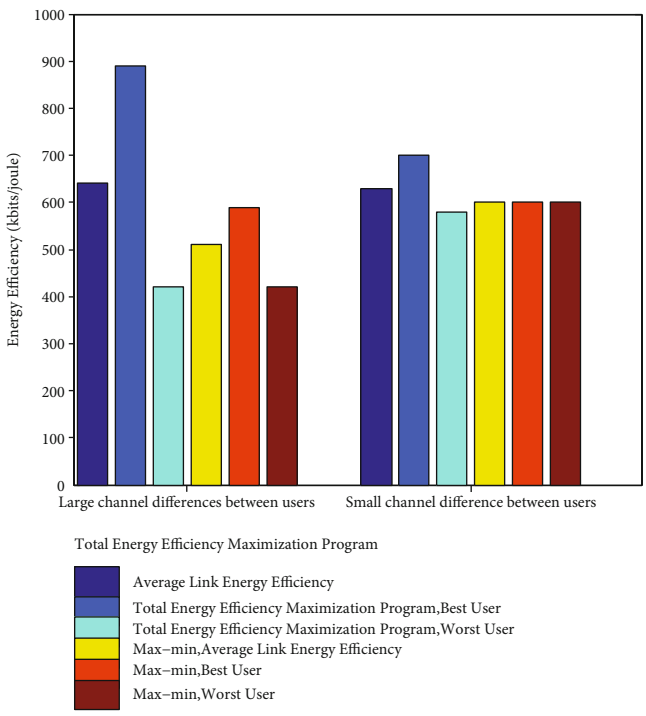


Figure 5: Comparison of the fairness of the two resource allocation options.

while the max–min scheme maximizes the EE of EUs with the worst link status to ensure fairness. As a result, the max–min scheme achieves much fairness at the sacrifice of less total EE.

5. Conclusion

This paper introduces a BAWPCN to address the energy shortage of sensor nodes due to frequent data interactions in digital twin workshops and proposes a resource allocation scheme to guarantee fairness in link EE. Considering the EUs' minimum rate requirement and energy causality constraints, the max–min system link EE optimization problem was formulated as a mixed-integer nonconvex fractional programming problem with time-power two-dimension resource joint optimization. By introducing generalized fractional theory, relaxation variables, BCD, and auxiliary variables, an iterative algorithm is designed to solve the transformation problem to obtain the resource allocation scheme. Finally, the following three conclusions are verified by simulation: (1) the proposed iterative algorithm can quickly converge and (2) the proposed resource allocation scheme can effectively guarantee the fairness of link EE.

Appendix

In the optimization problem (14), the objective function, constraints C1-C4 and C7 are linear, so we only need to prove that constraints C5 and C8 are convex constraints. In C5 and C8, $B_0\tau_m \log_2(1+\xi P_{\tau_m}(P_0^{(l)})/\sigma^2)$ and $p_{m,c}^{\text{Back}}\tau_m + p_{m,c}^{\text{HTT}} t_m + x_m$ are linear, so that a sufficient condition for the optimization problem (14) to be convex is that the Hessian matrix is $t_m \log_2(1+x_m|h_m|^2/t_m\sigma^2)$ seminegative definite. Construct the function $f = t_m \log_2(1+x_m|h_m|^2/t_m\sigma^2)$, and its Hessian matrix can be expressed as follows:

$$\nabla f^2 = \begin{bmatrix} -\frac{t_m|h_m|^4}{\sigma^4 \ln 2}\left(t_m + \frac{x_m|h_m|^2}{\sigma^2}\right)^{-2} & \frac{x_m|h_m|^4}{\sigma^4 \ln 2}\left(t_m + \frac{|h_m|^2 x_m}{\sigma^2}\right)^{-2} \\ \frac{x_m|h_m|^4}{\sigma^4 \ln 2}\left(t_m + \frac{|h_m|^2 x_m}{\sigma^2}\right)^{-2} & -\frac{x_m|h_m|^4}{t_m\sigma^4 \ln 2}\left(t_m + \frac{x_m|h_m|^2}{\sigma^2}\right)^{-2} \end{bmatrix}. \tag{A.1}$$

The first-order determinant of the Hesse matrix shown in formula (A.1) is less than 0, and the second-order determinant is equal to 0, so the Hesse matrix is a seminegative definite matrix. Therefore, both constraints C5 and C8 are convex constraints. Based on the above analysis, Lemma 1 is proved to be right.

Data Availability

The simulation data used to support the findings of this study are included within the article. The code 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.

References

[1] F. Tao, Q. Qi, L. Wang, and A. Y. C. Nee, "Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: correlation and comparison," *Engineering*, vol. 5, no. 4, pp. 653–661, 2019.

[2] L. Wright and S. Davidson, "How to tell the difference between a model and a digital twin," *Advanced Modeling and Simulation in Engineering Sciences*, vol. 7, no. 1, pp. 1–13, 2020.

[3] T. H. Q. Qi, F. Tao, T. Hu et al., "Enabling technologies and tools for digital twin," *Journal of Manufacturing Systems*, vol. 58, pp. 3–21, 2021.

[4] Y. Ye, L. Shi, X. Chu, and G. Lu, "On the outage performance of ambient backscatter communications," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7265–7278, 2020.

[5] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 418–428, 2014.

[6] X. Kang, C. K. Ho, and S. Sun, "Full-duplex wireless-powered communication network with energy causality," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5539–5551, 2015.

[7] D. Kuester and Z. Popovic, "How good is your tag?: Rfid backscatter metrics and measurements," *IEEE Microwave Magazine*, vol. 14, no. 5, pp. 47–55, 2013.

[8] R. J. Vyas, B. B. Cook, Y. Kawahara, and M. M. Tentzeris, "E-WEHP: a batteryless embedded sensor-platform wirelessly powered from ambient digital-tv signals," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 6, pp. 2491–2505, 2013.

[9] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: a contemporary survey," *IEEE Communication Surveys and Tutorials*, vol. 20, no. 4, pp. 2889–2922, 2018.

[10] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Ambient backscatter: a new approach to improve network performance for rf-powered cognitive radio networks," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3659–3674, 2017.

[11] B. Lyu, H. Guo, Z. Yang, and G. Gui, "Throughput maximization for hybrid backscatter assisted cognitive wireless powered radio networks," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2015–2024, 2018.

[12] P. Ramezani and A. Jamalipour, "Optimal resource allocation in backscatter assisted wpcn with practical energy harvesting model," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12406–12410, 2019.

[13] S. Xiao, H. Guo, and Y.-C. Liang, "Resource allocation for full-duplex-enabled cognitive backscatter networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3222–3235, 2019.

[14] G. Yang, X. Xu, and Y.-C. Liang, "Resource allocation in noma-enhanced backscatter communication networks for wireless powered iot," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 117–120, 2020.

[15] L. Shi, R. Q. Hu, J. Gunther, Y. Ye, and H. Zhang, "Energy efficiency for rf-powered backscatter networks using htt protocol," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13932–13936, 2020.

[16] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for uav-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 2876–2891, 2022.

[17] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for swipt systems," *IEEE Communications Letters*, vol. 19, no. 12, pp. 2082–2085, 2015.

[18] S. H. Kim and D. I. Kim, "Hybrid backscatter communication for wireless-powered heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6557–6570, 2017.

[19] M. Ismail, W. Zhuang, E. Serpedin, and K. Qaraqe, "A survey on green mobile networking: from the perspectives of network operators and mobile users," *IEEE Communication Surveys and Tutorials*, vol. 17, no. 3, pp. 1535–1556, 2015.

[20] H. Yang, Y. Ye, and X. Chu, "Max-min energy-efficient resource allocation for wireless powered backscatter networks," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 688–692, 2020.

[21] M. Hong, M. Razaviyayn, Z.-Q. Luo, and J.-S. Pang, "A unified algorithmic framework for block-structured optimization involving big data: with applications in machine learning and signal processing," *IEEE Signal Processing Magazine*, vol. 33, no. 1, pp. 57–77, 2016.

[22] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1758–1789, 2015.

[23] M. Razaviyayn, M. Hong, and Z. Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2012.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4502430, 9 pages
https://doi.org/10.1155/2022/4502430

Research Article

Analysis of Supply Chain Optimization Method and Management Intelligent Decision under Green Economy

Minyi Li and Yi Zhou

School of Economics and Management, Xinyu College, Xinyu 338004, China

Correspondence should be addressed to Minyi Li; liminyi@xyc.edu.cn

Received 22 January 2022; Revised 16 February 2022; Accepted 11 April 2022; Published 5 May 2022

Academic Editor: Yinghui Ye

As environmental issues become the focus of global attention, low-carbon economy based on the concept of low energy consumption, low pollution, and sustainable development is becoming the focus of global attention. This new trend brings new challenges and opportunities to supply chain management. In view of the current new trend, each node enterprise in the supply chain should have the dual compatibility of economy and environment. How to balance the profits of supply chain nodes is an important issue for traditional enterprises under the guarantee of environmental benefits and green supply chain management. This paper takes the green transformation of enterprises as the breakthrough point, combined with the comparison of three green supply chain models. Considering the different preferences of enterprises for environmentally friendly goals, a green supply chain model is constructed when manufacturers and retailers consider different goals. This paper discusses the impact of environmental preferences of manufacturers and retailers on the supply chain system. It is found that when the manufacturer's environmental preference is 1, the supply chain profit of the biobjective model is up to 1901. The results show that the model can achieve the dual optimization of the profit target and environmental friendly target and achieve the effect of green supply chain optimization.

1. Introduction

With the development of market economy, the supply chain system has gradually matured, and at the same time, new problems have arisen. The supply chain needs to pay more attention to the concept of low carbon and environmental protection in its development, and the optimization of the supply chain network of green economy has become an opportunity for enterprises to make rapid progress. In that optimization of green supply, more and more people begin to pay attention to the environmental factors of suppliers and manufacturers. People are also looking for suitable models in order to balance the relationship between them and weigh more environmental factors. Most of them put forward mathematical programming models for green supply chain structure. For example, Wu et al. [1] put forward an integer mixed model to solve the cost problem. The construction of mathematical model also has limitations in the actual supply chain, and the simulation results usually need to be corrected artificially. For the complexity of the problem, we can refer to the traditional accurate methods, software simulation, or heuristic algorithms proposed in reference [2]. The stochastic programming model proved by Xu and Nozick [3] is also significant. The model weighs the relationship between cost and environment. According to the actual situation, the same model can solve different problems, such as the same stochastic programming model. Soleimani et al. [4] considered using environmental risk value (CVaR) as an environmental risk evaluator and optimized the environmental impact caused by the supply chain path, and changing expectations in the supply chain in the secondary supply market mentioned in reference [5], fuzzy algorithm can be used to positively stimulate the demand of suppliers to improve the green degree of the supply chain network. Ahi and Searcy [6] add the customer factor of random variable in the supply chain and also found that the customer will lead to the environmental coordination of the supply chain. Reference [7] puts forward a value model that

can optimize risk to reflect the uncertainty of supply chain flow process. Literature [8] analyzes the main behavior of green supply chain management from 2006 to 2016 and expounds the practice of green supply chain management from a comprehensive perspective. According to Xu et al. [9], a large number of country supply chain approaches are categorized, based on 32 different stress scenarios. Luthra et al. [10] proposed that internal management and competitive green supply chain management are the key to achieve green supply chain management performance. In analyzing the indicators of enterprises, the benefits of green environment can also be realized, and the purpose of green economy can be achieved [11, 12]. References [13–15] take multiobjective analysis, supply chain structure analysis, the relationship between suppliers and manufacturers, and the environmental value of green supply chain as analysis factors to seek the optimal solution.

This paper will analyze and compare the multiobjective model with the basic model and single-objective model and explore the influence of the factors such as cost, price, and profit of manufacturers and retailers on the degree of environmental preference to determine the supply chain optimization method under the green economy.

2. Green Supply Chain and Its Basic Model

2.1. Green Supply Chain. In the process of supply chain and circulation, the products are transferred to upstream and downstream enterprises and customers and passed to consumers through a certain route [16]. Supply chain is a network structure built around suppliers, manufacturers, distributors, and retailers, and its theoretical basis has formed a systematic and rich one. If enterprises connected with supply are represented by nodes and links between enterprises are represented by line segments, in a word, the supply chain can realize the interaction between raw materials and consumers through activities such as planning, acquisition, storage, sales, and service and meet people's production and living needs [17].

In 1996, the American Manufacturing Research Association introduced the concept of environmental protection and environmental awareness in the supply chain. The concept of green supply chain is preliminarily put forward [18]. After years of development and application, many scholars have studied the green supply chain. It is considered that environmental protection should be fully considered in supply chain management, improve the utilization rate of supply chain resources and strengthen the energy-saving management of supply chain, integrate supply chain resources according to the green energy-saving mode, including a series of supply chain links such as suppliers, logistics and transportation, warehousing, product design, manufacturing, and consumption recovery, and further improve the environmental protection ability of supply chain. Green supply chain can effectively enhance the competitiveness of enterprises, realize the sustainable development and improvement of enterprise resources, and strengthen the scientific and normative supply operation of enterprises with green production and environmental protection as the biggest goal [19]. Green supply chain management includes five key parts: green procurement, green design, green production, green logistics, and green recycling. This paper expounds the main contents of green supply chain in detail from the aspects of suppliers and manufacturers.

2.2. Basic Model of Green Supply Chain. Green supply chain model is composed of raw material acquisition, production, assembly, distribution, and sales of specific products.

It can be abstracted as a network structure composed of a series of node sets and edge sets. Let$G_i = [N_i, L_i]$ and I denote a supply chain N_i composed of a node set L_i and an edge setG_i, which denotes a supply chain network composed of all competing supply chains. Let S represent the set of all potential market chains. Let X and $\forall a$ represent the market chain and nonnegative product flow, respectively. The side flow is the sum of the flow of the market chain in which it participates, that is,

$$X = \sum \delta_{as} X_s, \forall a \in L. \tag{1}$$

Each edge has a certain capacity constraint, so that $u_a \geq 0$ represents the nonnegative capacity constraint on edge a, and the capacity of each edge is the upper limit of product flow on that edge. Thus, the following inequality constraint in formula (2) holds

$$0 \leq X_a \leq u_a, \forall a \in L. \tag{2}$$

The edge cost is related to the product flow through the edge that is shown in

$$C_a = C_a(X_a), \forall a \in L. \tag{3}$$

Generally speaking, one edge is allowed to participate in multiple market chains in the model:

$$C_{as}(X_a) = \frac{C_a(X_a)}{X_a}. \tag{4}$$

Let p_{ij} denote the retail price of product I in market J, which depends on market demand; set market demand d_{ij}, that is,

$$p_{ij} = p_{ij}(d_{ij}), \forall i, j. \tag{5}$$

3. Green Supply Chain Model Based on Multiobjective Optimization

3.1. Overview of Multiobjective Optimization Theory. Multiobjective optimization generally studies the optimization of multiple objective functions in a given region, also known as multiobjective programming. In many fields, such as economy, management, military affairs, science, and engineering design, it is often difficult for people to measure the implementation quality of the whole plan with one index. Therefore, it is often necessary to compare multiple indicators, even if there are contradictory and complex

relationships among multiple targets. As early as the end of the 19th century, some foreign scholars studied it, and French economist Pareto and mathematicians such as Neumann, Kuhn, and Tucker took the lead in exploring multiple target problems [20]. Experts can often form the following methodologies on multiobjective problems: first, the main objective method, linear weighting method, and ideal point method are taken as examples to simplify multiobjective into single-objective and double-objective solutions as much as possible; the second is the hierarchical sequence method of solving the optimal solution of the second important goal on the basis of the optimal set of the important goal every time by assigning the goal value; third, it is solved by simplex method, analytic hierarchy process combining qualitative and quantitative methods and other multiobjective decision-making methods [21–23]. By constructing a two-level green supply chain model composed of a manufacturer and a retailer, the model is simulated as Figure 1. In the green supply chain, there is a positive correlation between the manufacturer's greenness and the wholesale price; that is, with the increase of greenness, the wholesale price will often increase.

3.2. Manufacturer's Single-Objective Model considering Profit and Environmental Friendliness. When the manufacturer considers the single objective of profit and environmental friendliness, the manufacturer takes the maximization of profit and environmental friendliness as the decision objective, and the retailer takes the maximization of its own profit as the decision objective. The optimization functions are formulas (6) and (7), respectively:

$$\operatorname{Max} \pi_m(w, g) = (w - c_m)(a - bp + kg) - \frac{1}{2} z g^2, \tag{6}$$

$$\operatorname{Max} \pi_r(p) = (p - w - c_r)(a - bp + kg). \tag{7}$$

Firstly, the wholesale price and greenness of products are defined by the manufacturers; secondly, retailers depend on the manufacturer's decision to set the price of products; finally, retailers sell their products to consumers to meet market demand. According to the above game order, the reverse induction method is used to solve the problem. For formula (7), the response function (8) of retailers is obtained by the first-order optimality condition:

$$p = \frac{a + kg + b(w + c_r)}{2b}. \tag{8}$$

Substituting formula (8) into formula (6), the Hessian matrix of π_m is calculated as

$$H = \begin{bmatrix} -b & \frac{k}{2} \\ \frac{k}{2} & -z \end{bmatrix}. \tag{9}$$

When the Hessian matrix satisfies $4bz - k^2 > 0$, there is a unique optimal solution for the manufacturer's profit function. Combined with the formula, the most suitable wholesale price (10) and product greenness (11) are solved:

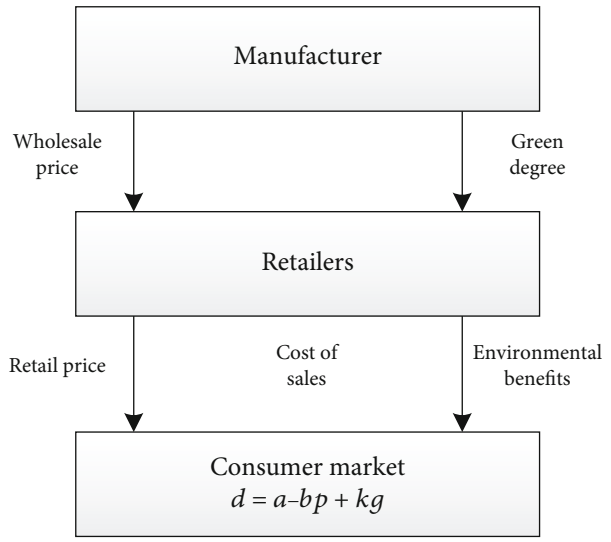


Figure 1: Supply chain model of manufacturer and retailer.

$$W = \frac{2[a + b(c_m - c_r)]Z - c_m k^2}{4bz - k^2}, \tag{10}$$

$$g = \frac{[a - b(c_m + c_r)]}{4bz - k^2}. \tag{11}$$

Substituting formula (10) and formula (11) into formula (8), the retailer's optimal retail price is obtained as

$$p = \frac{\left[3a + b(c_m + c_r)z - (c_m + c_r)k^2\right]}{4bz - k^2 ee}. \tag{12}$$

By substituting formulas (10) and (11) into formulas (6) and (8), respectively, the optimal product demand is obtained, and the optimal profits of manufacturers and retailers are as follows:

$$d = \frac{[a - b(c_m + c_r)]bz}{4bz - k^2}, \tag{13}$$

$$\pi_m = \frac{[a - b(c_m + c_r)]^2 z}{2\left(4bz - k^2\right)}, \tag{14}$$

$$\pi_r = \frac{[a - b(c_m + c_r)]^2 bz^2}{2\left(4bz - k^2\right)^2}. \tag{15}$$

From formulas (13) and (14), the maximum profit of the whole supply chain can be obtained as follows:

$$\pi = \frac{[a - b(c_m + c_r)]^2 \left(6bz - k^2\right) z}{2\left(4bz - k^2\right)^2}. \tag{16}$$

3.3. Two-Objective Model considering Both Profit and Environmental Friendliness for Manufacturers and Retailers. In addition to manufacturers' consideration of environmental friendliness, retailers, as downstream enterprises in the supply chain, are more susceptible to the influence of consumers' green preferences. The green supervision of the government, the public opinion guidance of the media, and the social services of non-profit organizations all urge retailers to implement green supply chain management [24]. I think the most likely application scenario of the dual-objective model is the automobile supply chain, because the vigorous development of new energy vehicles at present well reflects the environmental friendliness and the double harvest of supply chain profits. When both manufacturers and retailers consider profit and environmental friendliness at the same time, both manufacturers and retailers take profit and environmental friendliness maximization as their decision objectives. In this paper, superscript *mr* is introduced to represent the decision under *MR* model. The manufacturer's multiobjective optimization function is

$$\text{Max } F_m^{mr} = (\pi_m^{mr}, f^{mr}). \tag{17}$$

Of which, $\pi_m^{mr} = (w - c_m)(a - bp + kg) - 1/2zg^2$ and $f^{mr} = (a - bp + kg)g$. The retailer's multiobjective optimization function is

$$\text{Max } F_r^{mr} = (\pi_r^{mr}, f^{mr}). \tag{18}$$

Of which, $\pi_r^{mr} = (p - w - c_r)(a - bp + kg)$ and $f^{mr} = (a - bp + kg)g$. The decision-making goal of manufacturers and retailers is to achieve the synergistic optimization of profit and environmental friendliness, which is solved by linear weighting method. According to the target weight coefficients λ_m and λ_r of manufacturer and retailer, the multiobjective linear weighting function U_m of manufacturer is constructed as follows:

$$U_m^{mr} = \pi_m + \lambda_m f^{mr}. \tag{19}$$

The multiobjective linear weighting function of retailer is constructed as follows:

$$U_r^{mr} = \pi_r + \lambda_r f^{mr}. \tag{20}$$

λ_m and λ_r reflect the importance of environmental friendliness objectives λ_m and λ_r, and the greater the importance, the environmental friendliness objectives the more important the mark is. Especially, when $\lambda_m = 0$ and $\lambda_r = 0$, the optimization goal of manufacturers and retailers degenerates into profit single target case. Formula (19) can be converted to

$$\text{Max } U_m^{mr}(w, g) = (w - c_m)(a - bp + kg) - \frac{1}{2}zg^2. \tag{21}$$

Equation (20) is converted to

$$\text{Max } U_r^{mr}(p) = (p - w - c_r)(a - bp + kg). \tag{22}$$

The game sequence of the model is as follows: firstly, the manufacturer determines the wholesale price and greenness of products with the goal of optimizing profit and environmental friendliness; secondly, after observing the manufacturer's decision, retailers optimize the retail price of products with the goal of profit and environmental friendliness. In this paper, we use the inverse induction method to find the first derivative of P for formula (22) and make the first derivative equal to zero. We can see that the reaction function of retail price is

$$P^{mr} = \frac{a + kg + b(w + c_r) - \lambda_r bg}{2b}. \tag{23}$$

Substituting formula (23) into formula (21), the Hessian matrix of U_m is calculated as

$$H = \begin{bmatrix} -b & \dfrac{k - \lambda_m b + \lambda_r b}{2} \\ \dfrac{k - \lambda_m b + \lambda_r b}{2} & \lambda_m(k + \lambda_r b) - z \end{bmatrix}. \tag{24}$$

When $4bz - [(\lambda_m + \lambda_r)b + k]^2 > 0$satisfied, Hessian matrix is negatively definite, and formula (24) has a unique optimal solution. On this basis, the first-order partial derivatives of W and G for U_m are obtained and made equal to zero, and the optimal wholesale price and product greenness of manufacturers and retailers considering economic profit and environmental friendliness at the same time are as follows:

$$w^{mr} = \frac{2[a + b(c_m - c_r)] - \lambda_r bc_m[((\lambda_m + \lambda_r)b + 2\text{k})]}{4bz - [(\lambda_m + \lambda_r)b + k]^2}, \tag{25}$$

$$g^{mr} = \frac{[((\lambda_m + \lambda_r)b + \text{k})][a - b(c_m + c_r)]}{4bz - [(\lambda_m + \lambda_r)b + k]^2}. \tag{26}$$

Formula (25) and formula (26) are substituted into the formula (23), and the optimal retail price of the retailer is obtained as follows:

$$p^{mr} = \frac{[3a + b(c_m + c_r)]z - (c_m + c_r)k - (\lambda_m + \lambda_r)[(\lambda_m + \lambda_r)ab + ak + b(c_m + c_r)k)]}{4bz - [(\lambda_m + \lambda_r)b + k]^2}. \tag{27}$$

Table 1: Parameter settings.

Parameter name	Parameter content	Weight coefficient
Demand function	$d^m = 200 - 2p + g$	1
R&D input cost	$I^m = 3g^2$	0.8
Manufacturer's production cost	$c_m = 10$	0.7
Retailer cost of sales	$c_r = 6{,}0 \le \lambda_m \le 1.09$	0.9

By substituting formulas (25) and (27) into the corresponding equations, the optimal product demand is obtained as follows:

$$d^{mr} = \frac{[a - b(c_m + c_r)]bz}{4bz - [(\lambda_m + \lambda_r)b + k]^2}. \tag{28}$$

The optimal profit of the whole supply chain is

$$\pi^{mr} = \frac{\left\{6bz - \left[3(\lambda_m + \lambda_r)^2 b^2 + 4(\lambda_m + \lambda_r)bk + k\right]^2\right\}[a - b(c_m + c_r)]^2 z}{2\left\{4bz - [(\lambda_m + \lambda_r)b + k]^2\right\}^2}. \tag{29}$$

4. Experimental Simulation Analysis

4.1. Data Preparation. In order to more intuitively verify the above conclusions and theorems, but also in order to better reflect the impact of supply chain considering different objectives on its decision-making results, this section uses Maple software to do numerical simulation analysis to explore the changes of environmental preference in different models. Set the relevant parameters of the supply chain as shown in Table 1; it mainly simulates the ideal environment, that is, the balance between supply and demand. Reducing the weight will affect the cost of sales coefficient of retailers, resulting in an increase in the weight of cost of sales.

The above values are substituted into three models, namely, basic model, single-objective model, and double-objective model, and the following rules are obtained.

Under three different conditions, the greenness of products will increase with the improvement of manufacturers' environmental preference, and it will always be $g^m > g^m(\mu)$ as shown in Figure 2. Although the single-objective model does not consider the green degree, the single-objective model considers the manufacturer's profit. When the manufacturer's environmental preference degree increases, it will inevitably lead to the increase of production cost and indirectly improve the product green degree:

As shown in Figure 3, by analyzing the changing trend of enterprise environmental friendliness under the three models, it is found that in three different cases, environmental friendliness will increase with the improvement of manufacturers' environmental preference; that is, environmental friendliness is the largest in the two-objective model and the smallest in the basic model. Under this contract, the environmental friendliness goal of enterprises has been optimized and improved to some extent, but it has not achieved the perfect coordination towards centralization.

As shown in Figure 4, analyze the wholesale of products under the two models before and after coordination the trend of price change shows that the wholesale price of products always decreases with the increase of manufacturers' environmental preference, and the wholesale price of products in the double-objective model will always be less than that in the single-objective model. It can be seen that the manufacturer in a new sum can share the retailer's profits, and manufacturers can make profits by lowering the wholesale price, so as to promote the coordination of the whole supply chain decision-making and realize the optimization and improvement of the two objectives of profit and environmental friendliness.

As shown in Figure 5, the retail sales of products under the single-objective model and the double-objective model are analyzed. The change trend of price shows that the retail price of the products will first increase and then decrease with the increase of manufacturers' environmental preference. In addition, the increase in retail prices is relatively large in the overall trend; the decline is relatively large; that is, after manufacturers consider the goal of environmental friendliness, retailers follow the manufacturer's ring. With the improvement of environmental preference, more consumers are attracted by reducing retail prices and adopting the strategy of small profits but quick turnover market demand, so as to promote the Pareto improvement of the overall profit and environmental friendliness of the supply chain.

As shown in Figure 6, by analyzing the change trend of manufacturers' profits under the three models, it is found that retailers' profits will always increase with the increase of manufacturers' environmental preference, and retailers' profits will be optimized and improved after constant coordination. Therefore, although manufacturers share part of their own profits to retailers, manufacturers can still weaken the double marginal effect among supply chain members by adopting the strategy of small profits but quick turnover with lower product purchase price and realize the optimization and improvement of their own profits.

As shown in Figure 7, the change trend of total profit of supply chain under three models is analyzed. It is found that under the basic model, the profit of supply chain will always decrease with the increase of manufacturer's environmental preference, while under the single-objective model, the profit of supply chain will always increase first and then decrease slowly with the increase of manufacturer's environmental preference. When the degree of manufacturer's environmental preference is small, the degree of manufacturer's environmental preference cannot increase indefinitely under the dual model, and the high green input cost will inhibit the green transformation and upgrading of supply chain to a certain extent. Obviously, when the goal is to maximize the overall green benefit of the supply chain, the optimal decision-making of the supply chain fails to maximize the

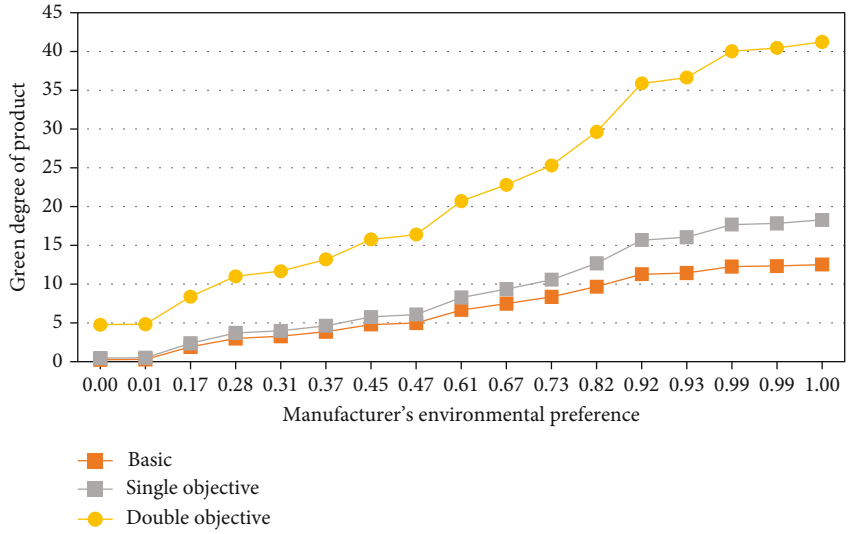


Figure 2: Change curve of product greenness before and after coordination.

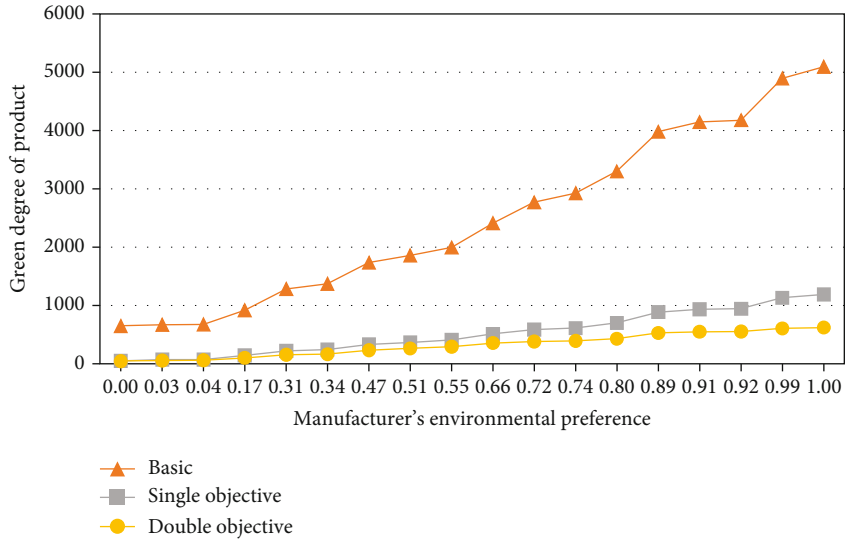


Figure 3: Change curve of environmental friendliness before and after coordination.

profit goal, which is because the supply chain will sacrifice economic profit and increase green R&D investment after considering the environmental friendliness goal, which will better meet its own expectation for multiobjective benefit. On the other hand, the total profit of the supply chain under the dual-objective model will always be higher than that of the single-objective model. It is concluded that the dual-objective model can achieve the dual optimization of profit goal and environmental friendliness goal and make it further achieve the desired effect.

5. Conclusion

When the current environment continues to deteriorate and resources are increasingly scarce, green supply chain has gradually aroused widespread concern, and improving the environmental friendliness in the process of product production and circulation has become a hot issue of universal concern all over the world. On the one hand, enterprises attach great importance to their own economic profits. On the other hand, the green transformation of enterprises is a realistic demand facing today's society. Therefore, on the premise of maintaining the profit distribution between the upper and lower members of the supply chain, it has become an important issue in the field of supply chain management to achieve the environmental goal of green supply chain and promote the sustainable management of the whole channel. Based on the comparison of multiobjective, single-objective, and basic model optimization, this paper focuses on the different value cognition background of enterprises to environmental objectives, constructs a green supply chain model when manufacturers and retailers consider different objectives, discusses the influence of manufacturers' environmental preference on supply chain system, and designs an

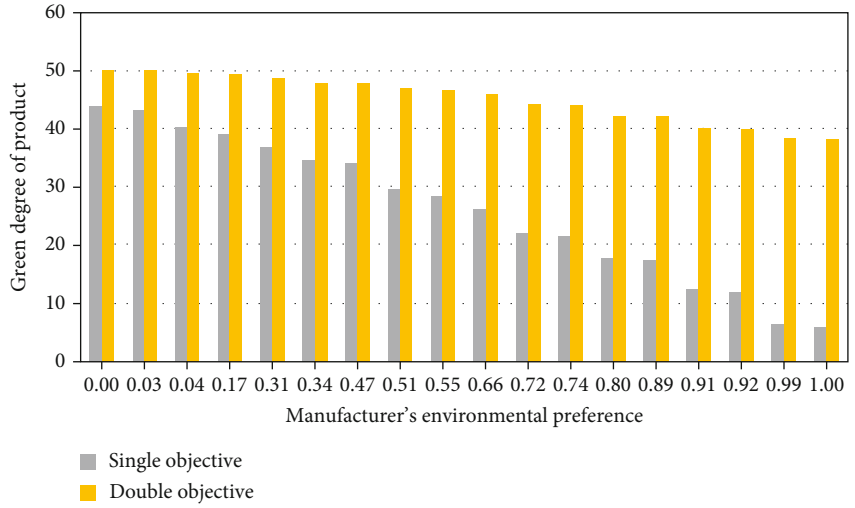


Figure 4: Changes of wholesale price before and after coordination.

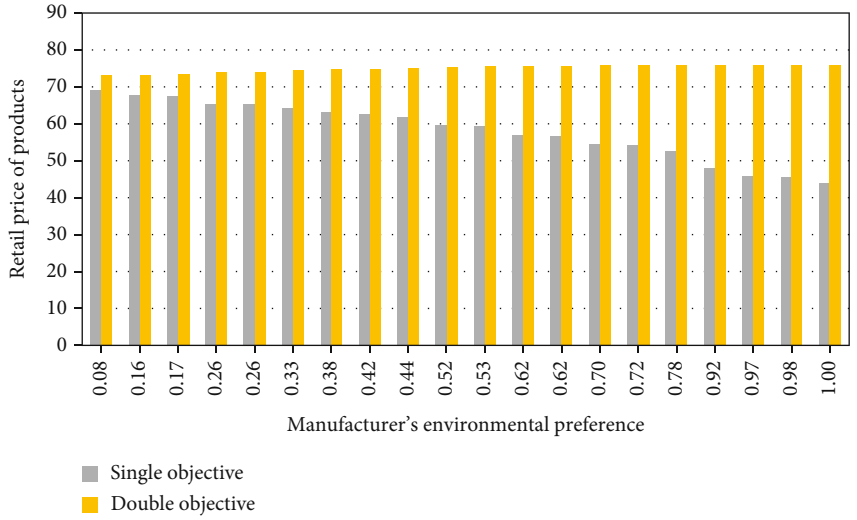


Figure 5: Changes of retail price of products before and after coordination.

effective optimization model. The main conclusions are as follows:

The basic supply chain optimization model is constructed. The supply chain model is based on the single-objective model of manufacturer only considering profit and environment and the dual-objective model of manufacturer and retailer considering profit and environment at the same time. The optimal decision of the three models is compared and analyzed, and the influence of environmental preference degree of manufacturer and retailer on supply chain is studied. The results show the following:

(1) Considering the goal of environmental friendliness by manufacturers and retailers can improve the green degree of products, the environmental friendliness of enterprises, and the total demand of products. With the improvement of environmental preference of manufacturers and retailers, the greenness and environmental friendliness of products are increasing; that is, the greenness, environmental friendliness, and product demand are the largest when manufacturers and retailers consider environmental friendliness at the same time and the smallest in the basic supply chain model. Obviously, considering the goal of environmental friendliness will significantly improve the environmental protection level of enterprises and occupy an advantage in the green consumption market

(2) When manufacturers consider profit and environmental friendliness at the same time, the higher the degree of environmental preference of manufacturers, the higher the retailer's profit and the lower the manufacturer's profit. When both manufacturers and retailers consider profit and environment-friendly objectives, manufacturers have lower environmental preference and retailers have higher environmental preference, while manufacturers have higher environmental preference and retailers properly consider environment-friendly objectives, which

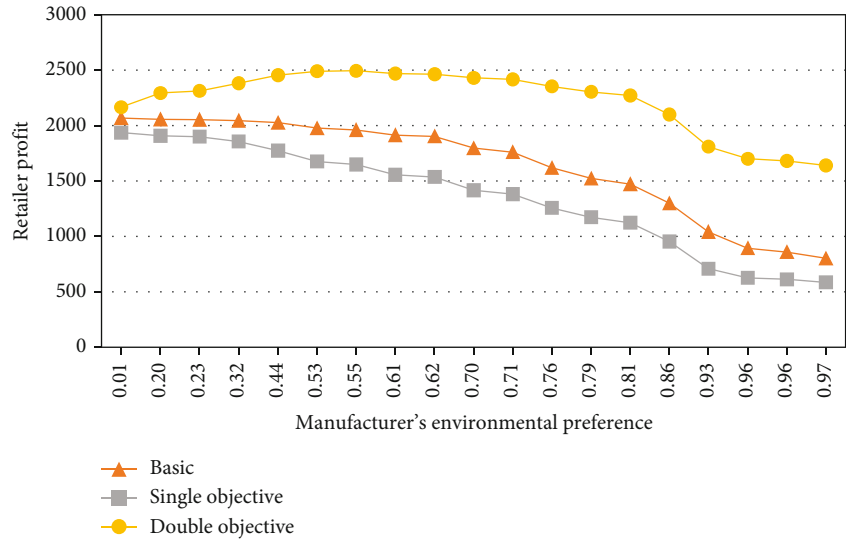


Figure 6: Curves of manufacturers' profits before and after coordination of three models.

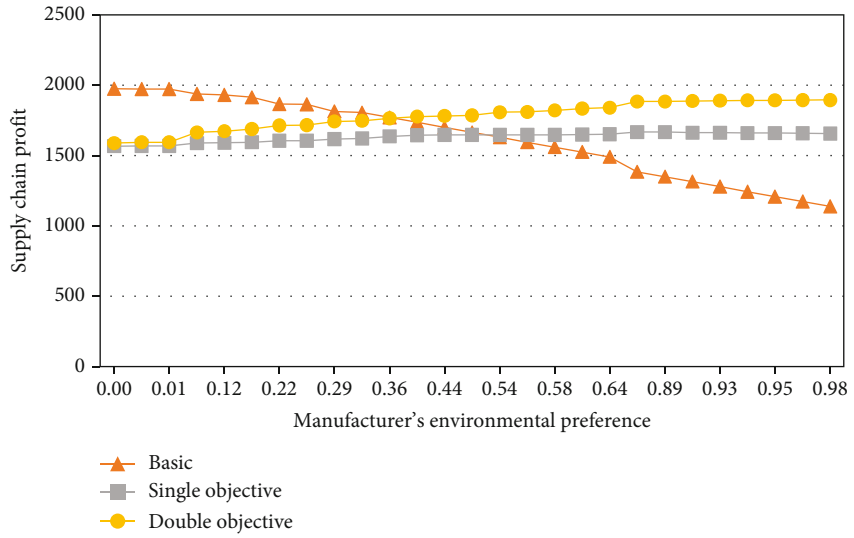


Figure 7: Change curves of supply chain profit before and after coordination of three models.

is beneficial to retailers' profits. The overall profit of supply chain increases at first and then decreases with the improvement of environmental preference of manufacturers and retailers. That is to say, under the premise that consumers have green preference, manufacturers and retailers properly consider the goal of environmental friendliness, which can not only improve the green degree of products but also promote the growth of supply chain profits and achieve a win-win situation between enterprise profits and ecological environment. Blind investment in green environmental protection will lead to serious damage to corporate profits

(3) Wholesale and retail prices increase first and then decrease with the increase of environmental preference of manufacturers or retailers, and the prices reach the highest when both manufacturers and retailers consider profit and environmental friendliness at the same time. At the initial stage when manufacturers and retailers consider environmental objectives, the whole supply chain can adopt high price strategy and maximize its own benefits by increasing the income per unit product. However, when the degree of environmental preference is high, the cost of green R&D is high, and the whole supply chain adopts the price reduction strategy to promote the increase of product demand, thus realizing the overall optimization situation of "small profits but quick turnover"

Data Availability

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

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] G. H. Wu, C. K. Chang, and L. M. Hsu, "Comparisons of interactive fuzzy programming approaches for closed-loop supply chain network design under uncertainty," *Computers & Industrial Engineering*, vol. 125, pp. 500–513, 2018.

[2] A. Amiri, "Designing a distribution network in a supply chain system: formulation and efficient solution procedure," *European Journal of Operational Research*, vol. 171, no. 2, pp. 567–576, 2006.

[3] N. X. Xu and L. Nozick, "Modeling supplier selection and the use of option contracts for global supply chain design," *Computers Operations Research*, vol. 36, no. 10, pp. 2786–2800, 2009.

[4] H. Soleimani, M. Seyed-Eafahani, and M. A. Shirazi, "A new multi-criteria scenario-based solution approach for stochastic forward/reverse supply chain network design," *Annals of Operations Research*, vol. 242, no. 2, pp. 399–421, 2016.

[5] G. P. Cachon and M. Fisher, "Supply chain inventory management and the value of shared information," *Management Science*, vol. 46, no. 8, pp. 1032–1048, 2000.

[6] P. Ahi and C. Searcy, "A comparative literature analysis of definitions for green and sustainable supply chain management," *Journal of Cleaner Production*, vol. 52, pp. 329–341, 2013.

[7] B. M. Beamon, "Supply chain design and analysis: models and methods," *International Journal of Production Economics*, vol. 55, no. 3, pp. 281–294, 1998.

[8] U. R. de Oliveira, L. S. Espindola, I. R. da Silva, I. N. da Silva, and H. M. Rocha, "A systematic literature review on green supply chain management: research implications and future perspectives," *Journal of Cleaner Production*, vol. 187, pp. 537–561, 2018.

[9] L. Xu, K. Mathiyazhagan, K. Govindan, A. N. Haq, N. V. Ramachandran, and A. Ashokkumar, "Multiple comparative studies of green supply chain management: pressures analysis," *Resources, Conservation and Recycling*, vol. 78, pp. 26–35, 2013.

[10] S. Luthra, D. Garg, and A. Haleem, "The impacts of critical success factors for implementing green supply chain management towards sustainability: an empirical investigation of Indian automobile industry," *Journal of Cleaner Production*, vol. 121, pp. 142–158, 2016.

[11] J. Su, C. Li, Q. Zeng, J. Yang, and J. Zhang, "A green closed-loop supply chain coordination mechanism based on third-party recycling," *Sustainability*, vol. 11, no. 19, p. 5335, 2019.

[12] J. Yang, J. Su, and L. Song, "Selection of manufacturing enterprise innovation design project based on consumer's green preferences," *Sustainability*, vol. 11, no. 5, p. 1375, 2019.

[13] J. Jian, Y. Guo, L. Jiang, Y. An, and J. Su, "A multi-objective optimization model for green supply chain considering environmental benefits," *Sustainability*, vol. 11, no. 21, p. 5911, 2019.

[14] D. Ghosh and J. Shah, "A comparative analysis of greening policies across supply chain structures," *International Journal of Production Economics*, vol. 135, no. 2, pp. 568–583, 2012.

[15] G. B. Chen and S. Li, "Network on chip for enterprise information management and integration in intelligent physical systems," *Enterprise Information Systems*, vol. 15, no. 7, pp. 935–950, 2021.

[16] M. V. Tatikonda and G. N. Stock, "Product Technology Transfer in the Upstream Supply Chain," *Journal of Product Innovation Management*, vol. 20, no. 6, pp. 444–467, 2003.

[17] C. T. Zhang and L. P. Liu, "Research on coordination mechanism in three-level green supply chain under non-cooperative game," *Applied Mathematical Modelling*, vol. 37, no. 5, pp. 3369–3379, 2013.

[18] H. You, L. Yu, S. Tian et al., "MC-Net: multiple max-pooling integration module and cross multi-scale deconvolution network," *Knowledge-Based Systems*, vol. 231, article 107456, 2021.

[19] B. Li, M. Zhu, Y. Jiang, and Z. Li, "Pricing policies of a competitive dual-channel green supply chain," *Journal of Cleaner Production*, vol. 112, no. 20, pp. 2029–2042, 2016.

[20] W. Yu and R. Han, "Coordinating a two-echelon supply chain under carbon tax," *Sustainability*, vol. 9, no. 12, p. 2360, 2017.

[21] J. Zhao and J. Wei, "The coordinating contracts for a fuzzy supply chain with effort and price dependent demand," *Applied Mathematical Modelling*, vol. 38, no. 9-10, pp. 2476–2489, 2014.

[22] I. Stanimirovic, M. Zlatanovic, and M. Petkovic, "On the linear weighted sum method for multi-objective optimization," *Facta universitatis - series: Mathematics and Informatics*, vol. 26, no. 4, pp. 49–63, 2011.

[23] G. Lou, H. Xia, J. Zhang, and T. Fan, "Investment strategy of emission-reduction technology in a supply chain," *Sustainability*, vol. 7, no. 8, pp. 10684–10708, 2015.

[24] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 9, pp. 3391–3402, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2806423, 12 pages
https://doi.org/10.1155/2022/2806423

Research Article

Energy Efficiency Maximization in the Wireless-Powered Backscatter Communication Networks with DF Relaying

Chuangming Zheng,[1] **Wengang Zhou**,[2] **and Xinxin Lu**[1]

[1]*School of Computer Science and Technology, Zhoukou Normal University, Zhoukou 466001, China*
[2]*City University of Macau, Macao, China*

Correspondence should be addressed to Chuangming Zheng; 20191050@zknu.edu.cn

Received 6 January 2022; Revised 7 March 2022; Accepted 18 March 2022; Published 16 April 2022

Academic Editor: Yinghui Ye

This paper focuses on the design of an optimal resource allocation scheme to maximize the energy efficiency (EE) in the wireless-powered backscatter communication networks (WPBCN) with decode and forward (DF) relaying. The two different devices are supported to operate in different modes, the harvest-then-transmit (HTT) mode and backscatter communication (BackCom) mode, respectively. In particular, we formulate an optimization problem to maximize system EE by jointly optimizing the transmit power of hybrid access point (H-AP) and the system time resource allocation. To deal with the nonconvex problem, we investigate the characteristics of the EE expression and a variable substitution approach. Then, the optimal power allocation scheme and iterative optimization algorithm were derived for achieving maximum EE. Extensive simulation results have demonstrated that the system EE can be improved about 10% because the proposed scheme provides more flexibility to utilize the resource efficiently by employing the proposed scheme.

1. Introduction

With the development of the Internet of Things (IoT), numerous wireless devices (WDs) have been deployed widely in the world to provide ubiquitous connectivity, which improves human being's life greatly in all aspects [1, 2]. However, the sustainable energy supply is a major challenge for the evolvable IoT networks. Fortunately, a new technology named as wireless power transfer (WPT) has been deemed as an attractive promising approach to power WDs conveniently and steadily [3–8]. Wireless-powered communication network (WPCN) is a novel communication solution for IoT that using WPT techniques solves the problem of sustainable energy supply for the tiny WDs. Consequently, WPCN has been widely studied in recent years, especially in [5–9]. A well-known protocol for WPCN is named as "harvest-then-transmit (HTT)" protocol proposed in [8]. Following the HTT protocol, hybrid access point (H-AP) first broadcasts the wireless energy to its served users for energy harvesting, and then, users transmit their information actively to H-AP by utilizing the harvested energy.

Recently, ambient backscatter communication (BackCom) [10] has been emerging as a promising technology for low-energy communication systems that was designed to communicate with WDs nearby without resorting to any existing energy supply or storage device. Different from the devices using the HTT mode, the BackCom devices transmit information by modulating and reflecting the instantaneous incident signals passively [11, 12]. Hence, the active RF components are not required at all, which can significantly decrease the circuit energy consumption and meet the low-power requirement of IoT devices. Furthermore, the dedicated energy harvesting (EH) time is also not necessary, and then, the information transmission time can be extended.

Essentially, there exist some different tradeoffs between EH time and data transmission time for HTT protocol and BackCom. They may complement each other for increasing the data rates and energy utilization while they are used in IoT. Hence, applying BackCom in WPCNs and wireless-powered backscatter communication networks (WPBCNs) is a very efficient way to exploit the advantages of the HTT mode and the BackCom mode, which have gained extensive attention in academic field, especially in [12–16]. The authors in [12] presented an optimal time resource allocation method to achieve the maximal throughput for WPBCNs. The literatures [13–16] played an emphasis on optimizing the EE performance of WPBCNs.

It is to be noted that EE is a crucial performance metric in IoT to achieve a better tradeoff between data rates and the overall energy consumption. Due to the characteristics of WPT, the energy cost of WPCNs, especially the energy consumption of the dedicated energy source devices, has drawn a great deal of attention [14, 15]. In [17], the authors proposed an EE resource allocation scheme which employed a Dinkelbach-based iterative algorithm to obtain the optimal time allocation, reflection coefficient, and transmit power of the dedicated RF energy source in BackCom networks. The literature [18] utilized energy beamforming communication and ambient BackCom to overcome the energy problem of network, and a EE cooperative communication scheme was presented. The EE maximum problem was investigated in cooperative sensor networks with bidirectional wireless information and power transfer [19]. By studying the derivative properties of the objective function, the authors derived the optimal power allocation and time allocation. The authors in [20] studied the EE performances for an unmanned aerial vehicle- (UAV-) assisted backscatter communication network. They maximized the EE of the network by jointly optimizing the UAV trajectory, the backscatter device scheduling, and the carrier emitter transmit power. The literature [14] studied the EE of WPBCNs which consist of two types of WDs that operate in different modes, the HTT mode and BackCom mode, respectively. The authors in [14] study the network which includes two different types WDs. One device named as HD, operating in the HTT mode, transmits its information directly to the access point (AP), and the other device named as BD, operating in the BackCom mode, backscatters its information directly to AP. By jointly optimizing the energy beamforming vector, power allocation, and time allocation, the system maximum EE is achieved. Obviously, the researchers have done many works in improving the EE of WPCNs and a lot of valuable achievements have been obtained. However, the EE optimization of WPBCNs that consists of two types of WDs has not been sufficiently exploited by the existing works.

The authors in [12] proposed a new network structure in which HTT mode is adopted in the last hop and BackCom mode is used in the other hops. They assumed that the devices could only consume the energy harvested within the current time block. Therefore, based on the above assumption, the nodes do not harvest the energy after they finish the information forwarding in one time block, which leads to the reduction of harvested energy and the inefficient energy utilization. Both in [19, 21], the authors think that the devices can store the harvested energy in their battery and do not consume all the energy harvested in the current time block. Especially in [19], the remote device is designed to harvest the energy from the relay during the phase of relay transmitting information to the base station. Inspired by the literatures [19, 21], we develop a new network model and the corresponding communication protocol which improve the EE and throughput of new model.

In this paper, we investigate a new WPBCN with DF relaying (WPB-R-CN) which is more suitable to the realistic IoT application scenario. WPB-R-CN consists of one hybrid access point (H-AP), one relay, and one backscatter user (SU) which is shown in Figure 1. Following the literature [12], SU operates in BackCom mode and relay transmits the information by operating in HTT mode. Different from [12], SU is required to harvest energy from RF signal transmitted by a relay device during the phase of relay transmitting information to H-AP. The energy harvested during this time by SU will be stored in its battery for use in the next time block. Obviously, the proposed new model can improve system EE further.

Obviously, WPB-R-CN is a new network model which is different with the existing models, such as the models of literatures [12, 14], which are more similar to our proposed model. In [12, 14], BackCom mode and HTT mode are used simultaneously in the models. However, in [14], the device named as HD, operating in the HTT mode, transmits its information directly to access point (AP) and the other device named as BD, operating in the BackCom mode, backscatters its information directly to AP. The authors in [14] only study one-hop link performane for two different types WDs. Different from the model of [14], the authors in [12] study multihop hybrid backscatter communication network. However, the other nodes do not harvest the energy transmitted by the node operating in HTT mode which leads to the reduction of harvested energy and the inefficient energy utilization. Therefore, the model we studied in this paper is a new one. No existing algorithms can achieve the maximum EE of WPB-R-CN.

We aim at WPB-R-CN EE maximization by optimizing its resource allocation and H-AP transmit power. To obtain the optimal parameters, we formulate a nonconvex system EE maximization problem. To make the optimization problem tractable, we investigate the characteristics of the EE expression. Then, the optimal power allocation scheme and iterative optimization algorithm were derived for solving the nonconvex optimal problem and achieving maximum EE. Our contribution can provide a useful insight for optimizing system performance in both system throughput and EE in WPBCN.

The main contributions of our work are summarized as follows:

(i) We propose a new WPB-R-CN network model which consists of one hybrid access point (H-AP), one relay, and one backscatter user (SU). To improve the system EE performance, SU required to harvest energy from RF signal transmitted by

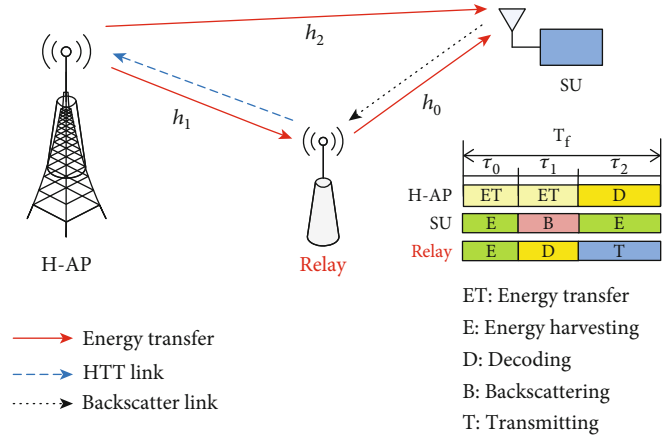


Figure 1: System model.

relay during the phase of relay transmitting information to H-AP

(ii) We develop a new protocol for the WPB-R-CN network to achieve the maximum EE. In the new protocol, SU harvests wireless energy from the relay during relay transmitting the data to H-AP and stores its harvested energy in a battery. And the stored energy can be scheduled across different transmission blocks

(iii) We formulate a system EE maximization problem and obtain the joint power allocation scheme for the proposed model. By investigating the characteristics of the EE expression, the optimal power allocation scheme and iterative optimization algorithm were derived for achieving maximum EE

(iv) We investigate the performances of the optimal power allocation scheme and the proposed iterative optimization algorithm. Comparing their performances with the other two schemes, our proposed scheme is proved to be efficient in improving system EE performance

The remainder of the paper is organized as follows. System model and communication protocol are described in Section 2. The problem of jointly optimizing H-AP transmit power and time resource allocation scheme is formulated to maximize the system EE in Section 3. Section 4 gives the optimal time allocation and power allocation schemes for maximizing the system EE. Simulation experiments are conducted to verify the effectiveness of the proposed strategies in Section 5. Finally, we conclude the paper in Section 6.

2. System Model and Communication Protocol

2.1. System Model. We consider a WPB-R-CN illustrated in Figure 1, which consists of one H-AP, one relay, and one SU. Each device is equipped with one antenna. H-AP is assumed to have sustainable power supply acting as a wireless power provider and information receiver. Relay and SU operate in the HTT mode and the BackCom mode, respectively. They do not have embedded energy supplies. Both relay and SU can harvest the energy from RF signal using their EH circuit and store the energy into a rechargeable battery. The stored energy can be consumed while relay and SU transmit their information. Moreover, by using the existing energy in SU battery, the information transmission can be initialized before energy harvesting and SU can consume all the harvested energy during the frame [19]. Each channel among nodes is assumed to be undergoing independent identically distributed (i.i.d) quasistatic block fading [22]. And the channels are reciprocal in two directions [19, 22]. Assume that no direct link exists between nodes H-AP and SU due to severe path loss and/or shadowing. Therefore, relay node has the obligation to relay the data of SU to H-AP.

Let h_0, h_1, h_2 denote the channel response of the relay–SU link, the H-AP–relay link, and the H-AP–SU link, respectively. Similarly, d_0, d_1, d_2 are the distance of relay and SU, H-AP and relay, and H-AP and SU. Let n_1, n_2, n_3 denote the additive white Gaussian noise at the receiver of SU, relay, and H-AP, respectively, $n_i \sim \mathcal{CN}(0, \sigma^2)$, $i \in \{1, 2, 3\}$. Without loss of generality, relay is closer to H-AP than SU in the system which means $d_1 < d_2$, resulting in $|h_1|^2 > |h_2|^2$. The duration of each frame T_f is one second, and the system bandwidth is B Hz. P_{cr} is the circuit power consumed while the H-AP and relay work as the receiver. P_{ct} is denoted as the circuit power consumed while SU and relay work as the transmitter.

2.2. Proposed New Communication Protocol. As shown in Figure 2, one time-frame duration T_f is divided into three time phases that are denoted by τ_0, τ_1, and τ_2, respectively. During the first time phase τ_0, H-AP transfers its wireless energy to relay and SU by transmitting the modulated signal $s(t) = \sqrt{P_H}x(t)$, where P_H denotes the transmit power of H-AP and $x(t)$ is a known signal with $\mathbb{E}[|x(t)|^2] = 1$. Correspondingly, both relay and SU work as the energy harvester and harvest the energy from the ambient RF signal using their EH circuit and store the energy into a rechargeable

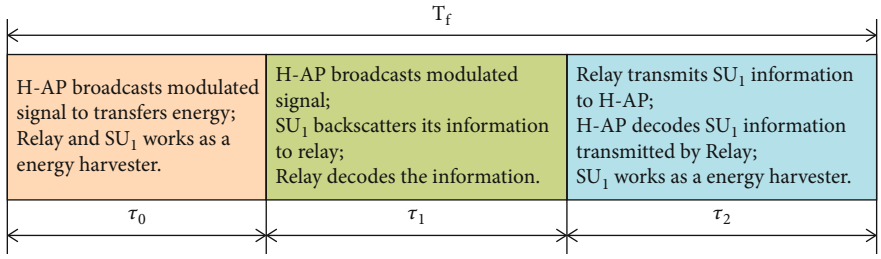


Figure 2: Framework of the communication protocol.

battery. The harvested energy E_{R_0} by relay and the harvested energy E_{U_0} by SU are expressed as

$$E_{R_0} = \eta P_H |h_1|^2 \tau_0, \tag{1}$$

$$E_{U_0} = \eta P_H |h_2|^2 \tau_0, \tag{2}$$

where $\eta \in (0, 1]$ denotes the energy harvesting efficiency. The harvested energy of relay is split into two parts. One part energy is used to maintain the circuit consumption of the relay during the time τ_1, and the other part is used to transmit information during the time τ_2. The harvested energy of SU is used to maintain its circuit consumption while SU backscatters the information during the time τ_1.

During the second time phase τ_1, H-AP continues to transmit the modulated signal $s(t)$. SU enters the active state to backscatter its information to relay by utilizing incident signals from H-AP, and relay also enters the active state to decode the information of SU. The received signal $y_1(t)$ by SU is given by

$$y_1(t) = \sqrt{P_H} h_2 x(t) + n_1(t). \tag{3}$$

Then, SU modulates its own signal $c_1(t)$ on the received signal $y_1(t)$ and $c_1(t)$ satisfies $\mathbb{E}[|c_1(t)|^2] = 1$. Therefore, the backscattered signal $y_2(t)$ by SU is written as

$$y_2(t) = \sqrt{P_H} h_2 x(t) c_1(t) + n_1(t) c_1(t). \tag{4}$$

The signal received by relay is denoted as $y_3(t)$ which is expressed as

$$\begin{aligned} y_3(t) = h_0 y_2(t) + h_1 s(t) + n_2(t) = \sqrt{P_H} h_2 h_0 x(t) c_1(t) \\ + h_0 n_1(t) c_1(t) + \sqrt{P_H} h_1 x(t) + n_2(t). \end{aligned} \tag{5}$$

Obviously, The first term of $y_3(t)$ is the received desired signal by relay. The second term of $y_3(t)$ is the noise caused by backscattering process. The third term of $y_3(t)$ is the interference from the H-AP, the power of which is typically larger than that of the desired signal. Following the literatures [23, 24], successive interference cancellation (SIC) technique is used to remove it from the $y_3(t)$ since relay has known the information $s(t)$ well. Thus, the signal-noise-ratio (SNR) at relay during the second time phase τ_1 is given by

$$\gamma_1 = \frac{P_H |h_0|^2 |h_2|^2}{\left(|h_0|^2 + 1\right)\sigma^2}. \tag{6}$$

While $|h_0|^2 << 1$, then equation (6) can be written as

$$\gamma_1 = \frac{P_H |h_0|^2 |h_2|^2}{\sigma^2}. \tag{7}$$

During the third time phase τ_2, H-AP stops transmitting the modulated signal $s(t)$ and acts as a information receiver. Relay operates in HTT mode and transmits the information decoded during the time τ_1. SU enters sleep state and acts as energy harvester. It harvests the RF energy transmitted by relay. Let $\mathscr{P}_R$ denote the transmit power of relay which is written as

$$\mathscr{P}_R = \frac{E_{R_0} - P_{cr}\tau_1 - P_{ct}\tau_2}{\tau_2}. \tag{8}$$

Relay transmits the information $c_2(t)$ to H-AP, and $c_2(t)$ satisfies $\mathbb{E}[|c_2(t)|^2] = 1$. The received signal of H-AP $y_4(t)$ and SNR γ_2 at H-AP is, respectively, given by

$$\begin{aligned} y_4(t) &= \sqrt{\mathscr{P}_R} h_1 c_2(t) + n_2(t), \\ \gamma_2 &= \frac{\mathscr{P}_R |h_1|^2}{\sigma^2}. \end{aligned} \tag{9}$$

The harvested energy E_{U_1} by SU is expressed as

$$E_{U_1} = \eta \mathscr{P}_R |h_0|^2 \tau_2. \tag{10}$$

Thus, the total energy harvested E_U by SU is written as

$$E_U = E_{U_0} + E_{U_1} = \eta P_H |h_2|^2 \tau_0 + \eta \mathscr{P}_R |h_0|^2 \tau_2. \tag{11}$$

Consequently, the SU–relay link and relay–H-AP link throughput are, respectively, calculated as

$$\mathscr{R}_1 = \tau_1 B \log_2(1 + \gamma_1) = \tau_1 B \log_2\left(1 + \frac{P_H |h_0|^2 |h_2|^2}{\sigma^2}\right), \tag{12}$$

$$\mathscr{R}_2 = \tau_2 B \log_2(1+\gamma_2) = \tau_2 B \log_2\left(1+\frac{\mathscr{P}_R|h_1|^2}{\sigma^2}\right). \quad (13)$$

Following the literature [19], the system throughput $\mathscr{R}$ is given by

$$\mathscr{R} = \min\{\mathscr{R}_1, \mathscr{R}_2\}. \quad (14)$$

Obviously, since both SU and relay are powered by harvested energy, only H-AP device consumes the energy in the system. Therefore, the total energy consumption of the whole system is also the energy consumption of H-AP which consists of two parts: the energy consumed in H-AP transmitting RF signal phase and the energy consumed in H-AP decoding information phase. Then, the total energy consumption of the whole system is written as

$$E_c = \left(\frac{P_H}{\zeta}\right)(\tau_0+\tau_1) + P_{\mathrm{cr}}\tau_2, \quad (15)$$

where $\zeta \in (0, 1]$ is the power amplifier efficiency.

3. Problem Formulation and Analysis

In this section, we formulate an optimization problem to maximize the system EE by jointly optimizing time resource allocation and H-AP transmit power. The system EE $\psi(P_H, \tau_0, \tau_1, \tau_2)$ is defined as the ratio of the achievable system throughput to the total energy consumption [14, 16], which is given by

$$\psi(P_H, \tau_0, \tau_1, \tau_2) = \frac{\mathscr{R}}{E_c} = \frac{\min\{\mathscr{R}_1, \mathscr{R}_2\}}{(P_H/\zeta)(\tau_0+\tau_1) + P_{\mathrm{cr}}\tau_2}. \quad (16)$$

Then, the optimization problem is formulated as

$$\begin{aligned} \mathbf{P1}: \max_{P_H,\tau_0,\tau_1,\tau_2} \quad & \frac{\min\{\mathscr{R}_1, \mathscr{R}_2\}}{(P_H/\zeta)(\tau_0+\tau_1) + P_{\mathrm{cr}}\tau_2} \\ \text{s.t.} \quad & \mathrm{C1}: \tau_0+\tau_1+\tau_2 = T_f \\ & \mathrm{C2}: \tau_0, \tau_1, \tau_2 \geq 0 \\ & \mathrm{C3}: 0 \leq P_H \leq P_{\max} \\ & \mathrm{C4}: E_U \geq P_{\mathrm{ct}}\tau_1 \\ & \mathrm{C5}: E_{R_0} \geq P_{\mathrm{cr}}\tau_1 \end{aligned} \quad (17)$$

In problem P1, C1 indicates that the summation of three time variables equals to the duration of one frame T_f. C2 limits that each time variables must be nonnegative. C3 constrains the transmit power range of H-AP. C4 and C5 guarantee that the total energy consumed does not exceed the total energy harvested for SU and relay, respectively. Notice that the WPB-R-CN is a new network which is distinguished from the conventional relaying and the main difference can be found in Section 2. These differences make the formulated EE problem noticeably different from that of the conventional relaying network.

Obviously, the above-formulated EE optimization problem is appealing in practice. For one thing, the time resource allocation can be exploited to satisfy the maximum system throughput requirement. For another, the EE can be further improved by optimizing the transmit power of H-AP. However, problem P1 cannot be solved directly for the following two main challenges. First, both the nominator and denominator of problem P1 include the variables $\{\tau_0, \tau_1, \tau_2\}, P_H$. Second, the optimization variables $\{\tau_0, \tau_1, \tau_2\}, P_H$ are coupled in both objective function and the constrains C4 and C5. Consequently, P1 is a nonconvex problem which cannot be solved directly. In general, there are no standard methods to solve the nonconvex optimization problems efficiently. Note that when P_H remains unchanged, the bigger nominator of objective function leads to the bigger EE. Then, the original problem P1 can be written as

$$\begin{aligned} \mathrm{P2}: \max_{P_H,\tau_0,\tau_1,\tau_2} \quad & \frac{\max\{\min\{\mathscr{R}_1, \mathscr{R}_2\}\}}{(P_H/\zeta)(\tau_0+\tau_1) + P_{\mathrm{cr}}\tau_2} \\ \text{s.t.} \quad & \mathrm{C1}-\mathrm{C5} \end{aligned} \quad (18)$$

Obviously, the problem P2 is also nonconvex optimization problem which is too difficult to obtain a globally optimal solution.

4. Energy Efficiency Maximization Resource Allocation

To solve problem P2 for obtaining its optimal solution $\{\tau_0^*, \tau_1^*, \tau_2^*, P_H^*\}$, we decompose the problem P2 into two subproblems to make it more tractable according to reference [19].

4.1. Optimal Resource Allocation Scheme. First, we formulate one subproblem to achieve the maximum system throughput by optimizing the resource allocation while P_H is considered as remaining unchanged. The optimal resource allocation is denoted as $\{\tau_0^*, \tau_1^*, \tau_2^*\}$. Let $\mathscr{R}^*$ denote the system throughput which corresponds to the optimal resource allocation $\{\tau_0^*, \tau_1^*, \tau_2^*\}$. It means that $\mathscr{R}^* > \mathscr{R}^+, \forall \{\tau_0^+, \tau_1^+, \tau_2^+\} \neq \{\tau_0^*, \tau_1^*, \tau_2^*\}$ when P_H remains unchanged, where $\mathscr{R}^+$ denotes the system throughput which corresponds to the resource allocation $\{\tau_0^+, \tau_1^+, \tau_2^+\}$.

Therefore, the first subproblem denoted as P2a is formulated as

$$\begin{aligned} \mathrm{P2a}: \max_{\tau_0,\tau_1,\tau_2} \quad & \{\min\{\mathscr{R}_1, \mathscr{R}_2\}\} \\ \text{s.t.} \quad & \mathrm{C1}-\mathrm{C5} \end{aligned} \quad (19)$$

Accordingly, P2a is still a nonconvex problem because there are coupling relationships among different optimization variables. In order to solve it, we present the following lemmas.

Lemma 1. *$\mathscr{R}_1 = \mathscr{R}_2$ is a necessary but insufficient condition for problem P2a obtaining the optimal solution.*

Proof. It is assumed that $\mathscr{R}_1 > \mathscr{R}_2$. Then, $\mathscr{R} = \mathscr{R}_2$ is the maximum system throughput. In order to cut down the SU–relay link throughput $\mathscr{R}_1$, we reduce SU backscattering time τ_1 to $\tau_{1'}$ for achieving $\mathscr{R}_1{}' = \mathscr{R}_2$. Let $\tau_1 = \tau_1{}' + \Delta$. Thus, we divide Δ into three parts $\{\Delta_{0'}, \Delta_{1'}, \Delta_{2'}\}$ in the ratio of $\tau_0 : \tau_{1'} : \tau_2$ and let $\tau_{0'} = \tau_0 + \Delta_{0'}$, $\tau_{1''} = \tau_{1'} + \Delta_{1'}$, and $\tau_{2'} = \tau_2 + \Delta_{2'}$. Obviously, the achieved system throughput $\mathscr{R}'$ is greater than $\mathscr{R}$ which is result from the new resource allocation scheme $\{\tau_{0'}, \tau_{1''}, \tau_{2'}\}$, which contradicts with the assumption. Then, Lemma 1 is proved. □

We apply Lemma 1 to the problem P2a. Then, the new optimization problem P3 is formulated as

$$\begin{aligned} \text{P3}: \max_{\tau_0, \tau_1, \tau_2} \quad & \{\mathscr{R}_1 = \mathscr{R}_2\} \\ \text{s.t.} \quad & \text{C1} - \text{C5} \end{aligned} \tag{20}$$

Similar to problem P2a, problem P3 is also a nonconvex problem. To tackle this problem, we first divide the problem P3 into two subproblems and solve two subproblems sequently. Finally, We use the optimal solutions of the two subproblems to obtain the optimal solution of problem P3 by iterative optimization and proportional compression algorithm.

At first, we relax the variable τ_2 and make $\tau_2 = 0$. Correspondingly, the constrain C1 becomes the new constrain $\tau_0 + \tau_1 = T_f$. Let $\tau_0 = \alpha T_f$ and $\tau_1 = (1-\alpha)T_f$, where $0 < \alpha < 1$ is the factor of SU transmission time. The first subproblem P3a is formulated as

$$\begin{aligned} \text{P3a}: \max_{\alpha} & \mathscr{R}_1 \\ \text{s.t.} \quad & \text{C3} - \text{C5}, \text{C6}: 0 < \alpha < 1 \end{aligned} \tag{21}$$

According to the references [4, 17, 22], $P_{\text{cr}} \le P_{\text{ct}}$ is always satisfied in WPCN. Obviously, $E_{R_0} > E_U$ is always satisfied according to the system model. Thus, the constrain C5 can be satisfied while the constrain C4 is satisfied. Based on the above conclusion, we formulate problem P4 as follows:

$$\begin{aligned} \text{P4}: \max_{\alpha} & \mathscr{R}_1 \\ \text{s.t.} \quad & \text{C3, C4, C6} \end{aligned} \tag{22}$$

Obviously, P4 is a more tractable problem. By use of the optimization method in [12, 17], P4 can obtain the maximum throughput while the constrain C4 satisfies $E_U = P_{\text{ct}}\tau_1$. Substituting $\tau_0 = \alpha T_f$, $\tau_1 = (1-\alpha)T_f$ and equation (2) into $E_U = P_{ct}\tau_1$, the following equation can be obtained shown as follows:

$$\eta P_H |h_2|^2 \alpha T_f = P_{\text{ct}}(1-\alpha)T_f. \tag{23}$$

The optimal solution α^* can be easily calculated from equation (16) which is given by

$$\alpha^* = \frac{P_{\text{ct}}}{\eta P_H |h_2|^2 + P_{\text{ct}}}. \tag{24}$$

Substituting the optimal solution α^* into the objective function, the maximum throughput $\widehat{R}_1$ of problem P3a is achieved which is written as

$$\widehat{R}_1 = \frac{\eta P_H |h_2|^2}{\eta P_H |h_2|^2 + P_{\text{ct}}} T_f B \log_2\left(1 + \frac{P_H |h_0|^2 |h_2|^2}{\sigma^2}\right). \tag{25}$$

Secondly, we relax the variable τ_1 to satisfy $\tau_1 = 0$ and relax the variable T_f to satisfy $T_f = T_R$, where $0 < T_R < T_f$ denotes the transmission cycle time of relay. Additionally, to simplify the optimization process, we make $\tau_0 = \beta T_R$ and $\tau_2 = (1-\beta)T_R$, where $0 < \beta < 1$ is the factor of relay transmission cycle time. Thus, another subproblem P3b of problem P3 is formulated as

$$\begin{aligned} \text{P3b}: \max_{\beta} & \mathscr{R}_2 \\ \text{s.t.} \quad & \text{C3, C7}: 0 < \beta < 1 \end{aligned} \tag{26}$$

Substituting $\tau_0 = \beta T_R$, $\tau_2 = (1-\beta)T_R$ and equation (7) into equation (12), $\mathscr{R}_2$ can be written as

$$\mathscr{R}_2 = (1-\beta)T_R B \log_2\left(1 + \frac{\eta P_H |h_1|^4 (\beta/(1-\beta)1-\beta) - P_{\text{ct}}|h_1|^2}{\sigma^2}\right). \tag{27}$$

Let $c = (\eta P_H |h_1|^4)/\sigma^2$ and $b = 1 - (P_{\text{ct}}|h_1|^2)/\sigma^2$. Then, the objective function $\mathscr{R}_2$ can be rewritten as

$$\mathscr{R}_2 = (1-\beta)T_R B \log_2\left(c\frac{\beta}{1-\beta} + b\right). \tag{28}$$

Lemma 2. *The optimization problem P3b is convex, and the optimal parameter $\beta^* = (x^* - b)/(x^* - b + c)$ and x^* is the solution of equation $cx\ln x - x + b - c = 0$.*

Proof. See the appendix.

Substituting the optimal solution β^* into the objective function of P3b, the maximum throughput $\widehat{R}_2$ of problem P3b is achieved which is given by

$$\widehat{R}_2 = (1-\beta^*)T_R B \log_2\left(c\frac{\beta^*}{1-\beta^*} + b\right). \tag{29}$$

In order to solve problem P3, we make $\widehat{R}_2 = \widehat{R}_1$ to obtain T_R which is expressed as

$$T_R = \frac{\widehat{R}_1}{(1-\beta^*)B \log_2(c(\beta^*/(1-\beta^*)1-\beta^*) + b)}. \tag{30}$$

1: Initialize the parameter ε
2: Divide the problem P3 into two subproblems by relaxing variable τ_2 and τ_1, respectively.
3: Solve problem P3a to obtain α^* and solve problem P3b to obtain β^*.
4: Calculate T_R by use of (29).
5: Let $\tau_{01} = \alpha^* T_f$, $\tau_{11} = (1-\alpha^*)T_f$, $\tau_{02} = \beta^* T_R$ and $\tau_{22} = (1-\beta^*)T_R$
6: **if** $\tau_{01} > \tau_{02}$ **then**
7: Let $\tau_0 = \tau_{01}$, $\tau_1 = \tau_{11}$ and $\tau_2 = \tau_{22}$
8: Calculate $\mathscr{P}_R$ and $\mathscr{R}_1$
9: Update $\tau_2 = \mathscr{R}_1 / B \log_2(1 + (\mathscr{P}_R |h_1|^2 / \sigma^2))$
10: **while** $|\tau_{22} - \tau_2| > \varepsilon$ **do**
11: $\tau_{22} = \tau_2$
12: Calculate $\widehat{\tau}_{01} = \tau_{01} - (\mathscr{P}_R |h_0|^2 / P_H |h_1|^2)\tau_{22}$
13: Let $\tau_0 = \max\{\widehat{\tau}_{01}, \tau_{02}\}$
14: Update $\mathscr{P}_R$
15: Calculate $\tau_2 = \mathscr{R}_1 / B \log_2(1 + (\mathscr{P}_R |h_1|^2 / \sigma^2))$
16: **end while**
17: **else**
18: Let $\tau_0 = \tau_{02}$, $\tau_1 = \tau_{11}$ and $\tau_2 = \tau_{22}$
19: **end if**
20: Then $\tau_0^* = \tau_0 T_f / \tau_0 + \tau_1 + \tau_2$, $\tau_1^* = \tau_1 T_f / \tau_0 + \tau_1 + \tau_2$, $\tau_2^* = \tau_2 T_f / \tau_0 + \tau_1 + \tau_2$.

Algorithm 1: Iterative optimization algorithm for solving problem P3.

According to our model, we can easily get $0 < T_R < T_f$ for the cause of $|h_1|^2 > |h_2|^2$. To find the optimal solution $\{\tau_0^*, \tau_1^*, \tau_2^*\}$, we define $\tau_{01} = \alpha^* T_f$, $\tau_{11} = (1-\alpha^*)T_f$, $\tau_{02} = \beta^* T_R$, and $\tau_{22} = (1-\beta^*)T_R$. Obviously, we select $\tau_{00} = \max\{\tau_{01}, \tau_{02}\}$ which can satisfy $\mathscr{R}_1 = \mathscr{R}_2$. Then, we devise a iterative optimization algorithm to obtain the optimal solution $\{\tau_0^*, \tau_1^*, \tau_2^*\}$ of problem P3 which is given in Algorithm 1 with an error ε. □

4.2. Optimal Transmit Power Allocation Scheme. The optimal power allocation scheme will be aimed at maximizing the energy efficiency in this subsection while the optimal resource allocation solution is obtained. Then, another subproblem P5 of problem P2 is formulated as

$$\mathbf{P5}: \max_{P_H} \psi(P_H) = \frac{\tau_1^* B \log_2\left(1 + \left(\left(P_H |h_0|^2 |h_2|^2\right) P_H |h_0|^2 |h_2|^2 / \sigma^2\right)\right)}{(P_H/\zeta)(\tau_0^* + \tau_1^*) + P_{\mathrm{cr}} \tau_2^*} \quad \text{s.t.} \quad \mathrm{C3} \tag{31}$$

Let $\lambda = \sigma^2 / |h_0|^2 |h_2|^2$ and $\varphi = 1/\zeta(\tau_0^* + \tau_1^*)$. By taking the derivative of $\psi(P_H)$ with respect to P_H, we get

$$\frac{\partial \psi(P_H)}{\partial P_H} = \frac{((\tau_1^* B \lambda (P_H \varphi + P_{\mathrm{cr}} \tau_2^*)) \tau_1^* B \lambda (P_H \varphi + P_{\mathrm{cr}} \tau_2^*) / (\ln 2(1 + P_H \lambda)) \ln 2(1 + P_H \lambda)) - \tau_1^* B \log_2(1 + P_H \lambda)\varphi}{(P_H \varphi + P_{\mathrm{cr}} \tau_2^*)^2}. \tag{32}$$

In equation (30), the denominator $(P_H \varphi + P_{\mathrm{cr}} \tau_2^*)^2$ is greater than zero. Therefore, the sign of $(\partial \psi(P_H))/\partial P_H$ depends on its numerator. We define the numerator as a new function $g(P_H)$ which is given by

$$g(P_H) = \frac{\tau_1^* B \lambda (P_H \varphi + P_{\mathrm{cr}} \tau_2^*)}{\ln 2(1 + P_H \lambda)} - \tau_1^* B \log_2(1 + P_H \lambda)\varphi. \tag{33}$$

Similarly, the derivative of $g(P_H)$ with respect to P_H can be written as

$$\frac{\partial g(P_H)}{\partial P_H} = \frac{-\tau_1^* B \lambda^2 (P_H \varphi + P_{\mathrm{cr}} \tau_2^*) \ln 2}{(\ln 2(1 + P_H \lambda))^2}. \tag{34}$$

It is evident that $(\partial g(P_H))/\partial P_H$ is less than zero, which means that $g(P_H)$ is a monotone decreasing function of P_H while $0 < P_H < P_{\max}$. In addition, $g(0)$ is greater than zero. Then, when $g(P_{\max}) \geq 0$, the optimal solution $P_H^* = P_{\max}$. The reason is that $(\partial \psi(P_H))/\partial P_H$ is always greater than zero while $0 < P_H < P_{\max}$. It means that $\psi(P_H)$ is a monotone increasing function while $0 < P_H < P_{\max}$. Therefore, the

1: Let $P_1 = 0$, $P_2 = P_{\max}$ and $\delta > 0$
2: Calculate $v = g(P_{\max})$ based on Eq.(32).
3: **if** $v < 0$ **then**
4: **while** $|P_2 - P_1| > \delta$ **do**
5: Let $P_H^+ = P_1 + P_2/2$
6: Calculate $u = g(P_H^+)$
7: **if** $u \geq 0$ **then**
8: Let $P_1 = P_1 + P_2/2$
9: **else**
10: Let $P_2 = P_1 + P_2/2$
11: **end if**
12: **end while**
13: $P_H^* = P_H^+$
14: **else**
15: $P_H^* = P_{\max}$.
16: **end if**

Algorithm 2: Power allocation algorithm.

maximum energy efficient can be obtained while $P_H^* = P_{\max}$. Otherwise, there exists P_H^+ satisfying $(\partial\psi(P_H))/\partial P_H = 0$. And the energy efficient $\psi(P_H)$ is a monotone increasing function while $0 < P_H < P_H^+$ and $\psi(P_H)$ is a monotone decreasing function while $P_H^+ < P_H < P_{\max}$. Therefore, we can solve the equation $g(P_H^*) = 0$ to obtain the optimal solution P_H^*. Thus, the optimal solution P_H^* is given by

$$P_H^* = \begin{cases} P_{\max} & g(P_{\max}) \geq 0 \\ P_H^+ & g(P_{\max}) < 0 \end{cases}, \tag{35}$$

where P_H^+ is the unique solution satisfying $g(P_H^+) = 0$.

To solve the optimal solution P_H^*, we devise a power allocation algorithm which is described in Algorithm 2 with the given error δ.

5. Simulation Results

In this section, the performance of WPB-R-CN is evaluated by the system-level simulation. The following parameters are used for simulation unless stated otherwise. The simulation duration is ten time frames, and the channel model between arbitrary nodes is modeled as $|h_i|^2 = |g_i|^2 d_i^{-3}$ [22], where $g_i \sim \mathcal{CN}(0, 1)$ denotes the channel coefficient between two nodes and is set as $g_i = 1$ for simplicity [14]; d_i is the distance between adjacent nodes. System bandwidth is set $B = 100$ Hz. The distance between relay and SU is set $d_0 = 20$ m, the distance between H-AP and relay is set $d_1 = 14$ m, and the distance between H-AP and SU is set $d_2 = 30$ m. $P_{\max} = 30$ dBm [14], $\sigma^2 = -70$ dBm, $\eta = 0.6$ [12], $T_f = 1.0$ s, $P_{\text{ct}} = -16.5$ dBm [12], and $P_{\text{cr}} = -20.5$ dBm [12].

Assume that $\{\tau_0^*, \tau_1^*, \tau_2^*, P_H^*\}$ is an optimal solution for achieving maximum EE of the system. According to Algorithm 1, if the energy E_{U_0} harvested by SU during the time τ_0^* is greater than the energy consumed by SU maintaining the circuit normal operation during the time τ_1^*, SU does not require to harvest the energy from the relay during the time τ_2^* . Therefore, it does not need to execute iteration. This scenario is not considered in this section. We mainly play an emphasis on the other scenario that includes $E_{U_0} < P_{\text{ct}}\tau_1^*$ and the energy E_{R_0} harvested by relay during the time τ_0^* being enough to transmit the required data to H-AP. Therefore, SU needs to harvest the energy from the relay during the time τ_2^* to store it into the rechargeable battery for use in the next frame.

Figure 3 depicts the EE of the proposed Algorithm 1 versus the number of iterations under different H-AP transmit powers P_H. It can be seen that Algorithm 1 converges after only three iterations, which demonstrate the efficiency of Algorithm 1. Simultaneously, Figure 3 shows that system EE increases while H-AP transmit power P_H increases from 19 dBm to 23 dBm. It shows that the optimal H-AP transmit power P_H^* is greater than 23 dBm.

Figure 4 shows the system EE performance versus H-AP maximum transmit power $P_{\max}$. Four different schemes have been simulated to verify the predominance of the proposed scheme. "Proposed scheme with iteration" denotes as "the proposed Algorithm 1" and the optimal H-AP transmit power P_H^* is obtained by use of Algorithm 2. "Proposed scheme without iteration" denotes as the "scheme 2" that the optimal τ_0^* equals to the maximal value of τ_{01} and τ_{02} in the proposed Algorithm 1, and the optimal H-AP transmit power P_H^* is obtained by the use of Algorithm 2. "Throughput maximization" denotes "scheme 3" that $\{\tau_0^*, \tau_1^*, \tau_2^*\}$ are solved by the use of the proposed Algorithm 1, and P_H^* is set to $P_{\max}$ for achieving the maximal system throughput. To show the superiority of the framework design and proposed scheme, we have simulated the performance of the existing scheme proposed in reference [12], which is denoted as "scheme 4." In "scheme 4," we aim at the maximum EE by solving for the optimal H-AP transmitting power. It can be seen from Figure 4 that the system EE of four schemes is increasing with the increasing $P_{\max}$ while $P_{\max} \leq P_H^*$. "The proposed Algorithm 1," "scheme 3," and "scheme 4" behave exactly the same in system EE when $P_{\max} \leq P_H^*$. However, the system EE of "scheme 3" decreases sharply and the other two schemes keep unchangeable while $P_{\max} \geq P_H^*$. We conclude that the throughput maximization scheme may achieve the lower EE than "the proposed Algorithm 1" due to the fact that the optimal resource allocation for throughput maximization is not energy efficient. It illustrates the importance of maximizing the EE. It also can be observed that "the proposed Algorithm 1" always achieves the better system EE than that of the other three schemes, because SU can harvest energy from the relay during the time τ_2^* and the τ_0^* is reduced by the iteration process.

In addition, we have also simulated the system throughput performance of four different schemes shown in Figure 5, which shows the system throughput performance versus H-AP maximum transmit power $P_{\max}$. "The proposed Algorithm 1," "scheme 2," and "scheme 4" behave similarly in the system throughput. Their system throughput is increasing when $P_{\max} \leq P_H^*$ and kept unchanged when $P_{\max} \geq P_H^*$, whereas the system throughput of "scheme 3" always increases with the increasing $P_{\max}$. The reason is that

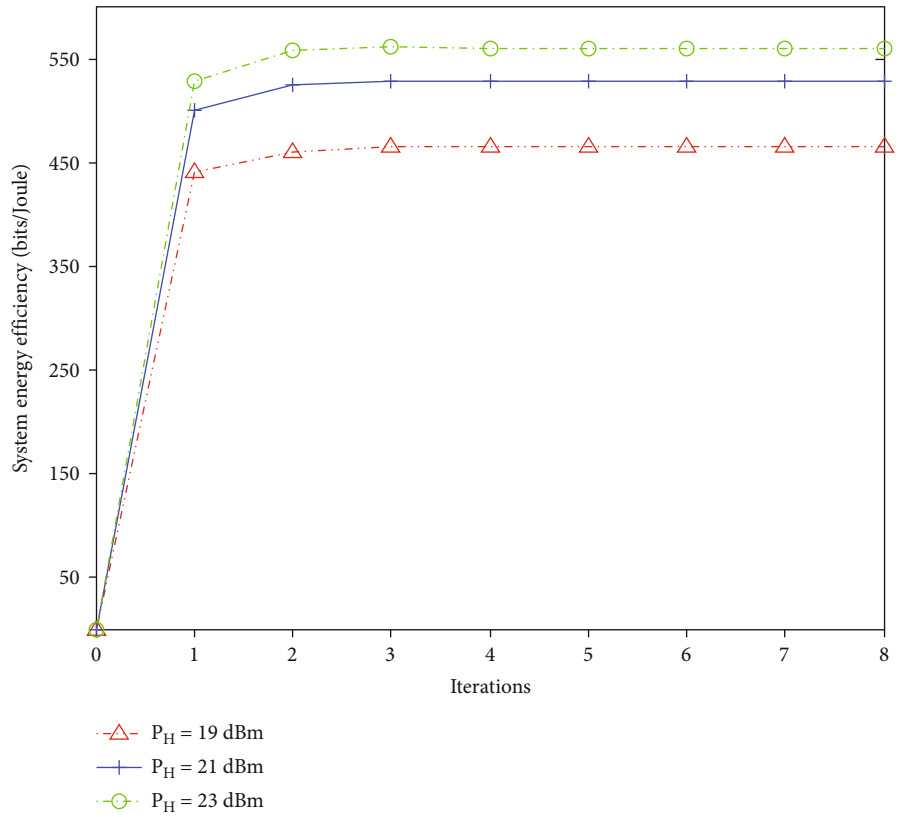


Figure 3: The convergence of Algorithm 1.

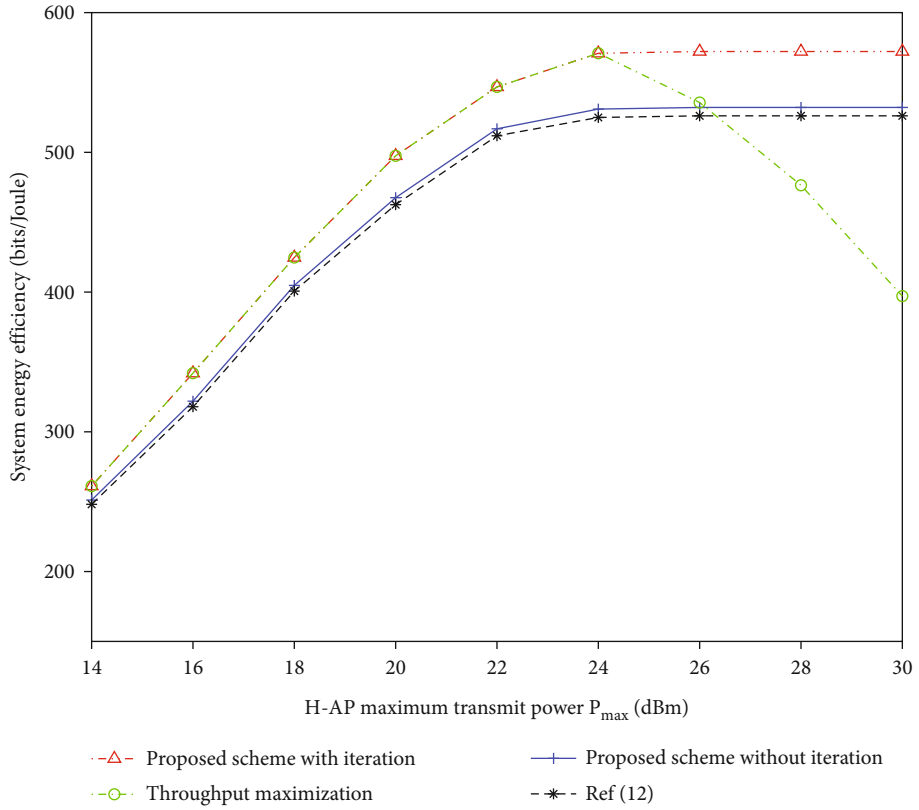


Figure 4: System EE versus H-AP maximum transmit power P_{max}.

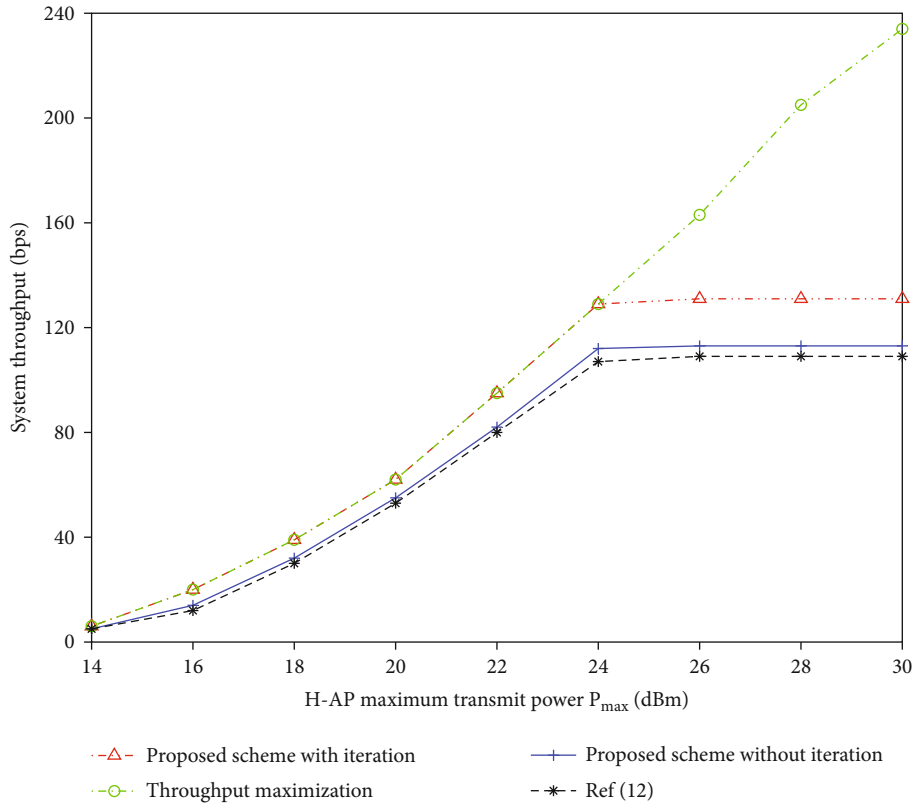


Figure 5: System throughout versus H-AP maximum transmit power $P_{\max}$.

"the proposed Algorithm 1," "scheme 2," and "scheme 4" are aimed at the maximum system EE. When $P_{\max} \geq P_H^*$, the algorithms always select $P_H = P_H^*$ as the optimal transmit power. However, "scheme 3" always sets $P_H = P_{\max}$ to achieve the maximum system throughput. Equations (11), (12), and (13) show that the system throughput is increasing function with the variant P_H. Therefore, When $P_{\max} \geq P_H^*$, H-AP transmit power is kept unchanged in Algorithm 1" and "scheme 2" and increasing in "scheme 3," which leads to the different throughput performances.

Simultaneously, Figures 4 and 5 tell us that our proposed scheme can always achieve the better performance both in EE performance and throughput performance. The reason is that the optimal parameters are obtained by searching algorithm while our scheme adopts more accurate iterative optimization method. Therefore, the proposed scheme can achieve the better performance compared with the scheme in reference [12]. Simulation results demonstrate that the proposed algorithm is efficient and achieves the more system EE due to adopting the proposed new communication protocol. Figures 4 and 5 illustrate the importance of considering the EE. Another observation is that our proposed scheme can achieve the highest EE among these schemes since the proposed scheme provides more flexibility to utilize the resource efficiently.

6. Conclusion

In this article, a new WPBCN with DF relaying has been studied which is more suitable to the realistic IoT application scenario. In order to achieve high system EE in this network, a new communication protocol was developed, in which SU harvests wireless energy from the relay during relay transmitting the data to H-AP and stores its harvested energy in a battery. And the stored energy can be scheduled across different transmission blocks. To maximize the system EE, we obtained the joint power allocation scheme for the proposed model. Then, by investigating the derivative of the EE expression, the optimal power allocation scheme and iterative optimization algorithm were derived for achieving maximum EE. Extensive simulation results have demonstrated that the system EE can be improved about 10% because the proposed scheme provides more flexibility to utilize the resource efficiently by employing the proposed scheme. Our contribution can provide a useful insight for optimizing system performance in both system throughput and EE in WPBCN.

Appendix

A. Convex Proof of Lemma 2

By taking the derivative of $\mathscr{R}_2$ with respect to β, we get

$$\frac{\partial \mathscr{R}_2}{\partial \beta} = \frac{T_R B}{\ln 2}\left(-\ln\left(c\frac{\beta}{1-\beta}+b\right)+\frac{1}{c\beta+b(1-\beta)}\right). \quad (A.1)$$

Then, the second-order derivative of $\mathscr{R}_2$ with respect to β is expressed as

$$\frac{\partial^2 \mathscr{R}_2}{\partial^2 \beta} = \frac{T_R B}{\ln 2(c\beta + b(1-\beta))}\left(-\frac{c}{1-\beta} - b(1-\beta) - \frac{c-b}{c\beta + b(1-\beta)}\right). \tag{A.2}$$

Since always hold with $0 < \beta < 1$, we conclude that P3b is a convex problem. The optimal solution β^* is obtained while $\partial \mathscr{R}_2 / \partial \beta = 0$. Then, we get the following equation:

$$\ln\left(c\frac{\beta}{1-\beta} + b\right) = \frac{1}{c\beta + b(1-\beta)}. \tag{A.3}$$

Let $x = c(\beta/(1-\beta)1-\beta) + b$ with $x > 1$. Then, we can get $\beta = (x-b)/(x-b+c)$. Substituting these into the above equation, we get the following equation:

$$\ln(x) = \frac{x-b+c}{cx}. \tag{A.4}$$

$cx \ln x - x + b - c = 0$ and x^* is the solution. Define $f(x) = cx \ln x - x + b - c$. It is easy to prove that $f(x)$ is a monotone-increasing function of x if $x > 1$. Therefore, we observe that x^* is unique. Thus, the optimal solution $\beta^* = (x^* - b)/(x^* - b + c)$ and x^* is the solution of equation (33).

Data Availability

All the data used to support the findings of this study are included within the article.

Conflicts of Interest

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

Acknowledgments

This work was supported in part by the Key Scientific and Technological Research Projects in Henan Province under Grant 212102210572 and in part by Key Scientific Research Projects of Colleges and Universities in Henan Province under Grant 22A510011.

References

[1] Y. Xu, Z. Qin, G. Gui, H. Gacanin, H. Sari, and F. Adachi, "Energy efficiency maximization in NOMA enabled backscatter communications with QoS guarantee," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 353–357, 2021.

[2] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: a contemporary survey," *IEEECommunications Surveys Tutorials*, vol. 20, no. 4, pp. 2889–2922, 2018.

[3] D. Song, W. Shin, J. Lee, and H. V. Poor, "Sum-throughput maximization in NOMA-based WPCN: a cluster-specific beamforming approach," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10543–10556, 2021.

[4] L. Shi, Y. Ye, X. Chu, and G. Lu, "Computation energy efficiency maximization for a NOMA-based WPT-MEC network," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10731–10744, 2021.

[5] D. Li and Y. Liang, "Adaptive ambient backscatter communication systems with MRC," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12352–12357, 2018.

[6] G. Lu, L. Shi, and Y. Ye, "Maximum throughput of TS/PS scheme in an AF relaying network with non-linear energy harvester," *IEEE Access*, vol. 6, pp. 26617–26625, 2018.

[7] T. Ruan, Z. J. Chew, and M. Zhu, "Energy-aware approaches for energy harvesting powered wireless sensor nodes," *IEEE Sensors Journal*, vol. 17, no. 7, pp. 2165–2173, 2017.

[8] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 418–428, 2014.

[9] Y. Xu, H. Sun, and Y. Ye, "Distributed resource allocation for SWIPT-based cognitive ad-hoc networks," *IEEE Transactions on CognitiveCommunications and Networking*, vol. 7, no. 4, pp. 1320–1332, 2021.

[10] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: wireless communication out of thin air," in *Proceedings of ACM SIGCOMM*, pp. 39–50, Hong Kong, China, 2013.

[11] S. Gong, X. Huang, J. Xu, W. Liu, P. Wang, and D. Niyato, "Backscatter relay communications powered by wireless energy beamforming," *IEEETransactions on Communications*, vol. 66, no. 7, pp. 3187–3200, 2018.

[12] S. H. Kim and D. I. Kim, "Hybrid backscatter communication forwireless-powered heterogeneous networks," *IEEE Transactions on WirelessCommunications*, vol. 16, no. 10, pp. 6557–6570, 2017.

[13] H. Yang, Y. Ye, and X. Chu, "Max-min energy-efficient resource allocation for wireless powered backscatter networks," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 688–692, 2020.

[14] B. Lyu, Z. Yang, F. Tian, and G. Gui, "Energy-efficient resourceallocation for wireless-powered backscatter communication networks," in *2018 IEEE International Conference on Communication Systems(ICCS)*, pp. 72–77, Chengdu, China, 2018.

[15] B. Gu, Y. Xu, C. Huang, and R. Q. Hu, "Energy-efficient resource allocationfor OFDMA-based wireless-powered backscatter communications," in *ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, Montreal, Canada, 2021.

[16] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, p. 1, 2021.

[17] Y. Ye, L. Shi, R. Qingyang Hu, and G. Lu, "Energy-efficient resourceallocation for wirelessly powered backscatter communications," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1418–1422, 2019.

[18] T. Liu, X. Qu, W. Tan, and Y. Cheng, "An energy efficient cooperative communication scheme in ambient RF powered sensor networks," *IEEE Access*, vol. 7, pp. 86545–86554, 2019.

[19] R. Chen, Y. Sun, Y. Chen, X. Zhang, S. Li, and Z. Sun, "Energy efficiency analysis of bidirectional wireless information and power transfer for cooperative sensor networks," *IEEE Access*, vol. 7, pp. 4905–4912, 2019.

[20] G. Yang, R. Dai, and Y.-C. Liang, "Energy-efficient UAV backscatter communication with joint trajectory design and resource optimization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 926–941, 2021.

[21] Y. Long, G. Huang, D. Tang, S. Zhao, and G. Liu, "Achieving high throughput in wireless networks with hybrid backscatter and wireless-powered communications," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10896–10910, 2021.

[22] L. Shi, Y. Ye, R. Q. Hu, and H. Zhang, "Energy efficiency maximization for SWIPT enabled two-way DF relaying," *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 755–759, 2019.

[23] X. Kang, Y. Liang, and J. Yang, "Riding on the primary: a new spectrum sharing paradigm for wireless-powered IoT devices," *IEEE Transactionson Wireless Communications*, vol. 17, no. 9, pp. 6335–6347, 2018.

[24] G. Yang, Q. Zhang, and Y. Liang, "Cooperative ambient backscatter communications for green Internet-of-Things," *IEEE Internet of ThingsJournal*, vol. 5, no. 2, pp. 1116–1130, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 5445676, 10 pages
https://doi.org/10.1155/2022/5445676

Research Article

Physical Layer Security of Two-Way Ambient Backscatter Communication Systems

Hao Wang[1], **Junjie Jiang**[1], **Gaojian Huang**[1], **Wenbin Wang**[2], **Dan Deng**[3], **Basem M. Elhalawany**[4] and **Xingwang Li**[1]

[1]*School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China*
[2]*Henan Chuitian Technology Co., Ltd., Hebi 458000, China*
[3]*School of Information Engineering, Guangzhou Panyu Polytechnic, Guangzhou 410630, China*
[4]*Benha University, Cairo 11672, Egypt*

Correspondence should be addressed to Gaojian Huang; g.huang@hpu.edu.cn

Received 25 December 2021; Revised 11 February 2022; Accepted 2 March 2022; Published 29 March 2022

Academic Editor: Yinghui Ye

Achieving high network traffic demand in limited spectrum resources is a technical challenge for the beyond 5G and 6G communication systems. To this end, ambient backscatter communication (AmBC) is proposed for Internet-of-Things (IoT), because the backscatter device (BD) can realize communication without occupying extra spectrum resources. Moreover, the spectral efficiency can be further improved by using two-way (TW) communication. However, secure communication is a great challenge for accessing massive IoT devices due to the broadcasting nature of wireless propagation environments. In light of this fact, this article proposed a two-way ambient backscatter communication (TW-AmBC) network with an eavesdropper. Specifically, the physical layer security (PLS) is studied through deriving the analytical/asymptotic expressions of the outage probability (OP) and intercept probability (IP). Moreover, the outage probabilities (OPs) in high signal-to-noise ratio (SNR) regions are studied for the asymptotic behavior and the intercept probabilities (IPs) in high main-to-eavesdropping ratio (MER) regions as well. Through analysis and evaluation of simulation performance, the results show that (i) when considering the target node, the OP of the BD decreases with increasing SNR, that is, enhancing the reliability; (ii) an optimal value of BD's reflection coefficient that maximize the reliability of backscatter link can be obtained; (iii) in high SNR regions, the OPs approach a constant; thus, the diversity orders are zero; (iv) when increasing the MER, the IP of target node decreases, suggesting the security enhances; (v) a trade-off exists between reliability and security which can be optimized by carefully designing the parameters.

1. Introduction

Due to faster network traffic demand and limited spectrum resources, the next-generation wireless communication techniques urgently need to improve the spectrum efficiency (SE) and reduce latency [1]. The ambient backscatter communication (AmBC) that could utilize the environmental wireless signals for both powering low-power devices and backscattering signals was used to address the problem of limited spectrum resources [2]. The main advantages of AmBC for Internet-of-Things (IoT) can be summarized as follows: (i) the SE can be improved since information is conveyed without consuming additional bandwidth [3]; (ii) the expensive and energy-consuming components such as oscillators, filters, and amplifiers are not required for the backscatter device (BD); thus, it enjoys low cost and low power [4]. As a new green paradigm, AmBC has aroused great research interest. However, due to the simplicity of the components, BD cannot separate interfering signals from received signals and is vulnerable to security attacks.

In particular, the performance of combining AmBC with other technologies was studied in much of the existing literature. For instance, in [5], the authors first studied the outage probability (OP) of symbiotic system incorporating the nonorthogonal multiple access (NOMA) and backscatter techniques. The authors in [6] acquired the analytical

expressions of the OP and studied the outage performance of AmBC system with considering in-phase and quadrature-phase imbalance (IQI). In [7], the backscatter technique was studied with intelligent reflecting surface (IRS) which is known as another low-power technique; the proposed scheme can achieve a balance between the average throughput and coverage probability according to practical condition. In [8], the multiantenna backscatter tag AmBC system was proposed, wherein the multiantenna technique was used to enhance the reliability. In [9], an adaptive reflection coefficient scheme was proposed to minimize the OP of backscatter link, which provided a guideline to design the optimal reflection coefficient for practical AmBC systems.

Meanwhile, in AmBC system, the power of BD is limited which is in contrast to the traditional key mechanism requiring high computing power; thus, some researchers focused on the security problems. Without consuming high computing power, the physical layer security (PLS) can be achieved by using the inherent randomness of physical media and the difference between legal channels and eavesdropping channels. Thus, it has attracted much attention in academia and industry and was studied in different scenarios [10–12]. For instance, in [13], the secrecy performance of AmBC systems considering IQI was investigated by deriving the analytical expressions of the OP and intercept probability (IP). In [14], the secrecy outage probability (SOP) of the multitag backscatter systems over the Rayleigh fading channels was given, where the channel correlation between the forward and backscatter links may exist. In [15], the authors studied the PLS for the relay selection schemes of NOMA systems, indicating that better security performance can be achieved when increasing the number of relays.

On the other hand, due to the network congestion caused by information explosion, it is imperative to study the techniques increasing the throughput which is defined as the amount of data successfully transferred per unit time of a communication channel. The two-way (TW) communication method enables the capability of improving throughput and SE since the relay in the TW communication system can receive information from both nodes in a single time slot. The TW communication system therefore has been extensively studied specifically on the PLS aspect. For example, in [16], the OP of a TW relay system can be minimized by an improved dynamic scheme, which both the power-splitting ratios and the power allocation ratio can be dynamically adjusted. In [17], the OP of TW full-duplex relay system considering self-interference was evaluated, and asymptotic OP was presented for more insight. Different from the above works, reconfigurable intelligent surface (RIS) technique with great channel capacity advantage was employed in wireless communication systems to improve the performance [18]. Similarly, in [19], the performance of the RIS-assisted TW communication systems was studied wherein the channels can either be reciprocal or nonreciprocal, and the closed-form expressions of OP were derived for single-element RIS. Moreover, in [20], an artificial noise-assisted opportunistic relay selection scheme was proposed to enhance the security of underlay cognitive TW relay network.

1.1. Motivation and Contributions. In the existing works, in [21], an optimization algorithm was proposed to maximize throughput of relaying system where an unmanned aerial vehicle (UAV) enabled TW relay assist communication. In [22], the secrecy of TW relay NOMA systems was evaluated in terms of the ergodic secrecy sum rate, which degraded when the distance between the eavesdropper and either user becomes closer. In [23], the outage performance of a TW model wherein BD is embedded in the relay for backscattering was investigated. In [24], a BD cooperative relay communication network was proposed and the system reliability performance was studied. It is noted that all the reported works established models with relays and for the case of without relay was neglected. In this article, we first consider a two-way ambient backscatter communication (TW-AmBC) system without relay and study the reliability and security of the proposed system, where sources communicate with each other and also via BD with an eavesdropper. The main contributions of the article can be listed as follows:

(i) We propose a novel TW-AmBC model without relay and study the PLS of target node, wherein BD and nodes transmit signals in the presence of an eavesdropper

(ii) We derive the analytical expressions of the outage probabilities (OPs) for the target node and BD and the analytical expressions of the intercept probabilities (IPs) for eavesdropper. It is found that the reflection coefficient and transmitted power have the opposite effect on OPs and IPs. The trade-off between security and reliability can be flexibly adjusted

(iii) We evaluate the performance of OPs and IPs in high signal-to-noise ratio (SNR) and main-to-eavesdropping ratio (MER) regions. Furthermore, we discuss the diversity orders of the target node and BD in high SNR regions. The diversity orders are zero owing to the OPs approach a constant, proving that the joint error-floor exists

(iv) We analyze the comprehensive influence factor on performance of both the reflection coefficient and transmitted power in the TW-AmBC network. When transmitted power is high, the system performance enjoys little fluctuation, and when the reflection coefficient increases, the security of the target node can be increased

1.2. Organization and Notations. The remainder of the article is composed as follows. In Section 2, the TW-AmBC system without relay is first constructed, followed by the elaboration of deriving expressions of OPs and IPs in Section 3. In Section 4, the correctness of the theoretical analysis is validated via numerical results. Finally, conclusions are given in Section 5.

$E(\bullet)$ is deemed as the expectation operation. $\Pr(\bullet)$ denotes the probability and $K_v(\bullet)$ denotes the v-th order

modified Bessel function of the second kind. $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex Gaussian random variable with mean μ and variance σ^2. Besides, $Ei(\bullet)$ is the exponential integral function and $W_{u,v}(\bullet)$ is the Whittaker function. n denotes the natural number, and $n!$ denotes the factorial operation. The cumulative distribution function (CDF) is expressed as $F(\bullet)$, and the probability dense function (PDF) is expressed as $f(\bullet)$.

2. System Model

As illustrated in Figure 1, we consider a TW-AmBC system which is composed of two source nodes (A and B), a BD, and an eavesdropper (E), wherein h_n, h_0, and h_e, respectively, denote the channel responses from A to BD, B, and E. g_n and f_e represent the channel responses from BD to B and E, respectively. In different time slots, B and A can directively transmit signals to each other, also via a BD. The eavesdropper can intercept signals from nodes and BD. We assume that (i) all nodes are equipped with a single antenna; (ii) all channel parameters are subjected to the independent Rayleigh fading. We have $h_0 \sim \mathcal{CN}(0, \lambda_{h0})$, $h_n \sim \mathcal{CN}(0, \lambda_{hn})$, $h_e \sim \mathcal{CN}(0, \lambda_{he})$, $g_n \sim \mathcal{CN}(0, \lambda_{gn})$, $g_e \sim \mathcal{CN}(0, \lambda_{ge})$, and $f_e \sim \mathcal{CN}(0, \lambda_{fe})$.

2.1. Received Signals at B. The received signals at B include signals from A and the backscattered signals from BD. The BD adds its own message $c(t)$ to signals from A or B and then backscatters, where $E(|c(t)|^2) = 1$. Thus, the received signals y_B at B can be expressed as

$$y_B = h_0\sqrt{P_s}x_A(t) + \beta h_n g_n \sqrt{P_s}x_A(t)c(t) + n_1(t), \quad (1)$$

where β is the complex reflection coefficient used to normalize $c(t)$ and $n_1 \sim \mathcal{CN}(0, \sigma^2)$ is the complex Gaussian noise with zero mean. $x_A(t)$ is the transmitted signal of A. P_s denotes the transmitted power. Furthermore, in order to extract the real-information from the scrambled signals, successive interference cancellation (SIC) technique is adopted at the receiver end [25]. Therefore, the received signal-to-interference-plus-noise ratio (SINR) extracting $x_A(t)$ at B can be written as

$$\gamma_B = \frac{\gamma|h_0|^2}{\gamma|\beta|^2|h_n|^2|g_n|^2 + 1}, \quad (2)$$

where $\gamma = P_s/\sigma^2$ is the transmit SNR. By eliminating the interference from A, the SNR at B can be written as

$$\gamma_{\mathrm{BD}} = \gamma|\beta|^2|h_n|^2|g_n|^2. \quad (3)$$

2.2. Wiretapped Signals. In the 1st time slot, the received signals at E can be expressed as

$$y_e(t) = g_e\sqrt{P_s}x_B(t) + \beta g_n f_e \sqrt{P_s}x_B(t)c(\mathrm{t}) + n_2(t), \quad (4)$$

where $n_2(t)$ is the complex Gaussian noise at E and follows $n_2 \sim \mathcal{CN}(0, \sigma^2)$. $x_B(t)$ is the transmitted signal of B. When only direct link signals are decoded at E, the SINR can be given by

$$\gamma_{e,B} = \frac{\gamma|g_e|^2}{\gamma|\beta|^2|g_n|^2|f_e|^2 + 1}. \quad (5)$$

It is noted that E can eliminate direct link signals and extract the signals from BD with aid of SIC technique. In this case, the SNR can be expressed as

$$\gamma_{e,\mathrm{BD}} = \gamma|\beta|^2|g_n|^2|f_e|^2. \quad (6)$$

3. Performance Analysis

In this section, we derive the analytical expressions of OPs and IPs to investigate the security and reliability of the TW-AmBC system. Furthermore, the asymptotic behaviors of OPs and IPs are studied by asymptotic expressions, and the diversity orders are derived in high SNR regions.

3.1. Outage Performance Analysis

3.1.1. Outage Probability Expressions for B. Considering the direct link only, the outage event would occur when signal $x_A(t)$ is unsuccessfully decoded at B. Thus, the OP at B can be given by

$$P_{\mathrm{out}}^B = 1 - \Pr\left(\gamma_B > \gamma_{\mathrm{th}}^B\right), \quad (7)$$

where γ_{th}^B is the SNR threshold at B.

Theorem 1. *For the direct link, we can get the OP analytical expression, which can be expressed as*

$$P_{out}^B = 1 + Q_1 e^{Q_1 - \gamma_{th}^B/\gamma\lambda_{h0}} Ei(-Q_1), \quad (8)$$

where $Q_1 = \lambda_{h0}/\gamma_{th}^B|\beta|^2\lambda_{hn}\lambda_{gn}$, $Ei(x)$ represents the exponential integral function and $Ei(x) = \int_{-\infty}^{x}(e^t/t)dt$, where t denotes the variable. $Ei(x)$ can be expanded to a series form which is given by

$$Ei(x) = \nu + \ln x + \sum_{n=1}^{\infty}\frac{x^n}{n \cdot n!}, \quad (9)$$

where ν is the Euler constant, $\nu \approx 0577$. n denotes the natural number, and $n!$ denotes the factorial operation for n.

Proof. See Appendix A. □

Corollary 1. *Under high signal-to-noise ratio (SNR) case, the OP asymptotic expression of direct link can be given as*

$$P_{out,\infty}^B = 1 + Q_1 e^{Q_1}\left(1 - \frac{\gamma_{th}^B}{\gamma\lambda_{h0}}\right)Ei(-Q_1). \quad (10)$$

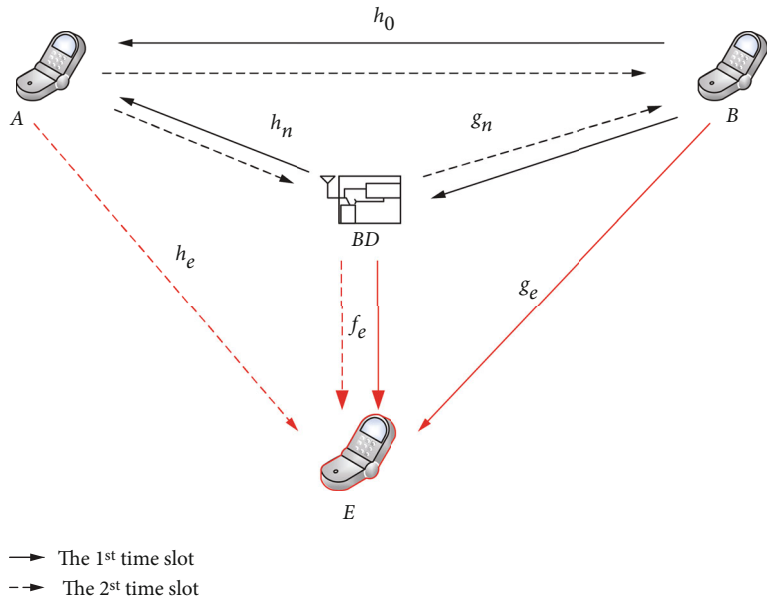


Figure 1: Two-way ambient backscatter communication system model.

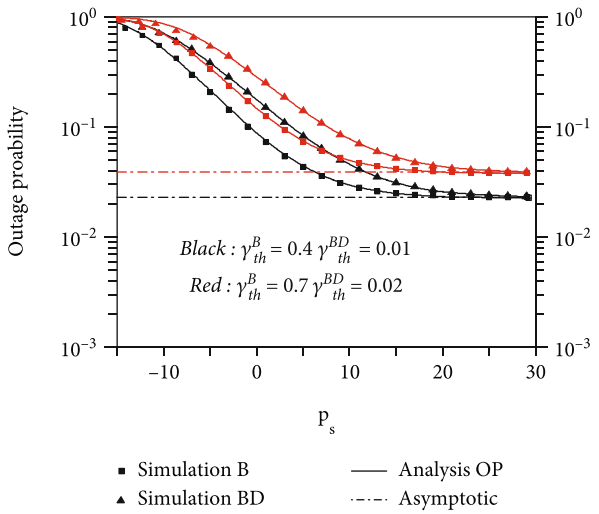


Figure 2: OPs versus P_s for different SNR threshold values.

Proof. According to asymptotic principle, Equation (10) can be derived from Equation (8) by using approximate equation $e^{-x} \approx 1 - x$ when $\gamma \longrightarrow \infty$. □

3.1.2. Outage Probability Expressions for BD. For the backscatter link, the outage event occurs when direct link signals are successfully decoded but failing with backscattered signals. Therefore, the OP of backscattered signals can be expressed as

$$P_{\text{out}}^{\text{BD}} = 1 - \Pr\left(\gamma_B > \gamma_{\text{th}}^B, \gamma_B > \gamma_{\text{th}}^{\text{BD}}\right), \tag{11}$$

where $\gamma_{\text{th}}^{\text{BD}}$ is the SNR threshold for backscattered signals.

Theorem 2. *For the backscatter link, we can obtain the OP analytical expression, which can be expressed as*

$$P_{out}^{BD} = 1 + Q_1 e^{Q_1 - \gamma_{th}^B/\gamma\lambda_{h0}} Ei(-Q_1) + \frac{\pi\Delta_2}{N}\sum_{i=1}^{N} e^{-\Delta_1} K_0\left(\sqrt{2\Delta_2(\delta_i + 1)}\right)\sqrt{1-\delta_i^2}, \tag{12}$$

where $\Delta_1 = (\gamma_{th}^B\gamma_{th}^{BD}(\delta_i + 1) + 2\gamma_{th}^B)/2\gamma\lambda_{h0}$, $\Delta_2 = \gamma_{th}^{BD}/\gamma|\beta|^2\lambda_{hn}$ λ_{gn}, $\delta_i = \cos\left(((2i-1)/2N)\pi\right)$, N *is a trade-off parameter between accuracy and complexity, and* $K_0(\bullet)$ *is the 0th order modified Bessel function of the second kind.*

Proof. See Appendix B. □

Corollary 2. *Under high SNRs case, the OP asymptotic expression for backscatter link in the proposed system can be given as*

$$P_{out,\infty}^{BD} = 1 + Q_1 e^{Q_1}\left(1 - \frac{\gamma_{th}^B}{\gamma\lambda_{h0}}\right)Ei(-Q_1) - \frac{\pi\Delta_2}{N}\sum_{i=1}^{N}(1-\Delta_1)\ln \cdot \left(\sqrt{\frac{\Delta_2(\delta_i+1)}{2}}\right)\sqrt{1-\delta_i^2}. \tag{13}$$

Proof. According to the asymptotic principle, through using two approximate equations, i.e., $K_0(x) \approx -\ln(x/2)$ and $e^x = 1 + x$, Equation (13) can be obtained.

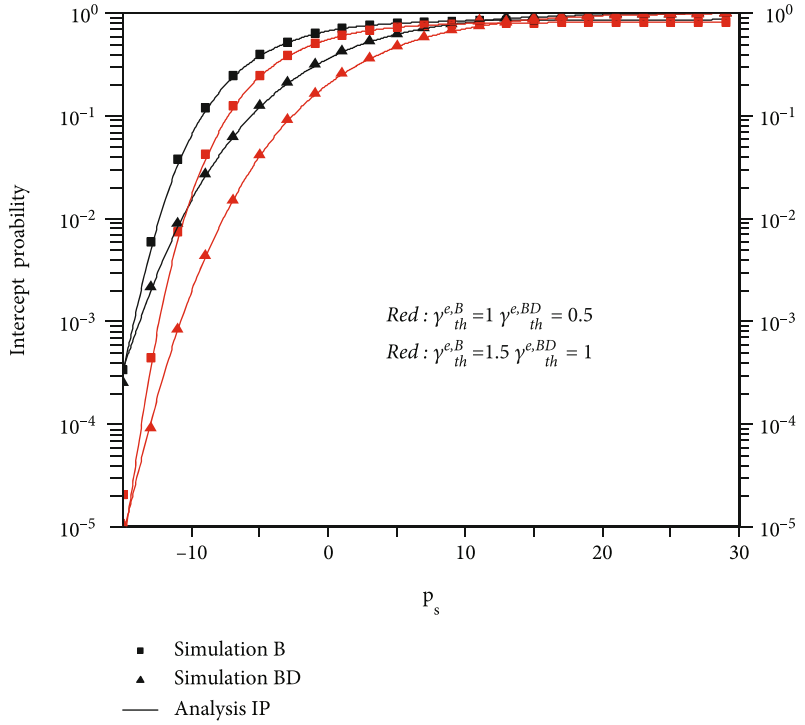


FIGURE 3: IPs versus P_s for different secrecy SNR threshold values.

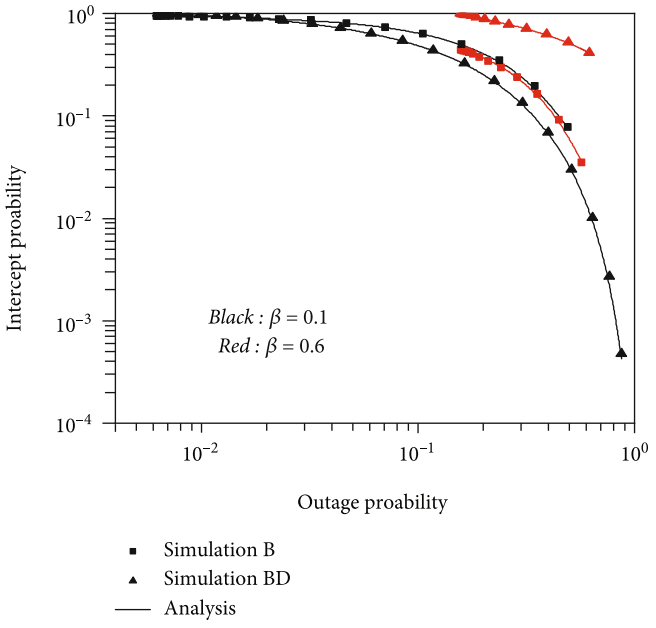


FIGURE 4: IPs versus OPs for different β values.

For further insights on the backscatter link and direct link, we also study the diversity order d_ψ, $\psi \in \{B, \mathrm{BD}\}$. The diversity order is given [26]:

$$d_\psi = -\lim_{\gamma \longrightarrow \infty} \frac{\log\left(P^{\psi}_{\mathrm{out},\infty}\right)}{\log \gamma}. \tag{14}$$

□

Corollary 3. *The diversity orders of the two links can be calculated as*

$$d_B = d_{BD} = 0. \tag{15}$$

Remark 1. From Corollaries 1–3 and Theorems 1 and 2, we can know that (1) when increasing P_s, the OPs decrease, enhancing the reliability of the system; (2) when increasing β, the OP of direct link increases; (3) when γ is large, the latter of Equation (12) is almost zero; thus, when $\gamma \longrightarrow \infty$, $P^{\mathrm{BD}}_{\mathrm{out},\infty}$ approximates $P^{B}_{\mathrm{out},\infty}$; (4) when $\gamma \longrightarrow \infty$, asymptotic OPs exist the joint error-floor, causing that the diversity orders are zero.

3.2. Security Performance Analysis. When the SINR or SNR at E surpasses the secrecy SNR threshold, direct link signals or backscattered signals would be intercepted. In further detail below, the expressions of IPs are derived to investigate the system security.

3.2.1. Intercept Probability Expressions for B. The IP expression of direct link signals can be written as

$$P^{B}_{\mathrm{int}} = \Pr\left(\gamma_{e,B} > \gamma^{e,B}_{\mathrm{th}}\right), \tag{16}$$

where $\gamma^{e,B}_{\mathrm{th}}$ is the secrecy SNR threshold for B.

Theorem 3. *For the direct link, we can get the IP analytical expression, which can be expressed as*

$$P^{B}_{\mathrm{int}} = -e^{-\gamma^{e,B}_{th}/\gamma\lambda_{ge}} Q_2 e^{Q_2} Ei(-Q_2), \tag{17}$$

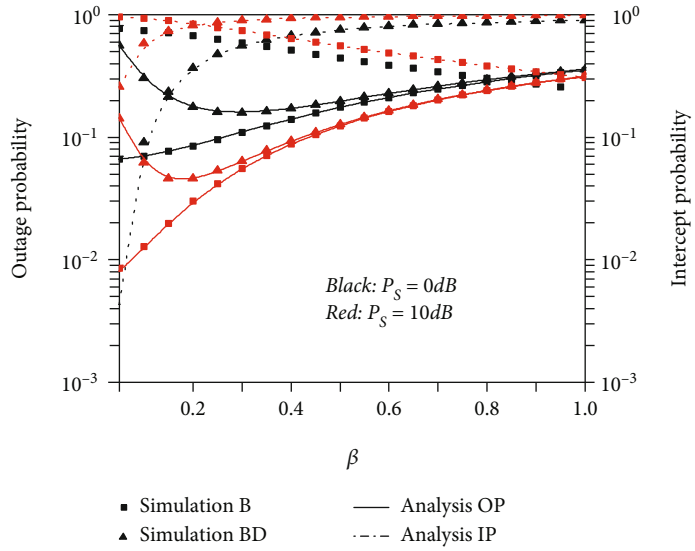


Figure 5: OPs and IPs versus β for different P_s values.

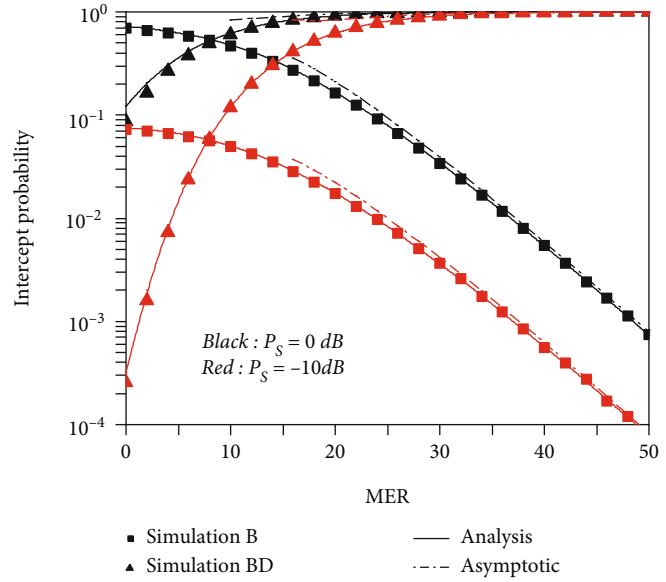


Figure 6: IPs versus MER for different P_s values.

where $Q_2 = \lambda_{ge}/\gamma_{th}^{e,B}|\beta|^2\lambda_{gn}\lambda_{fe}$.

Proof. Substituting Equation (5) into Equation (16), the IP of direct link signals can be expressed as

$$P_{\text{int}}^{B} = e^{-\gamma_{\text{th}}^{e,B}/\gamma\lambda_{ge}} \int_0^{\infty} e^{-\gamma_{\text{th}}^{e,B}|\beta|^2 y/\lambda_{ge}} f_{|gn|^2|fe|^2}(y)dy, \quad (18)$$

where $f_{|gn|^2|fe|^2}(y) = (2/\lambda_{gn}\lambda_{fe})K_0(2\sqrt{y/\lambda_{gn}\lambda_{fe}})$. The integral can be resolved by Appendix A, and Equation (17) can be derived. □

Here, we use the MER metric to study the asymptotic behavior of IPs, which depends on the ratio of the main link to eavesdropping link [27]. The main link and eavesdropping link are B to BD and BD to E, respectively. The MER can be expressed as $\lambda_{me} = \lambda_{gn}/\lambda_{fe}$. In the next corollary, the asymptotic analysis for IPs in high MER regions is derived.

Corollary 4. *Under high main-to-eavesdropping ratio (MER) case, the IP asymptotic expression can be given by*

$$P_{\text{int},\infty}^{B} = -e^{-\gamma_{th}^{e,B}/\gamma\lambda_{ge}} Q_2^*(1 + Q_2^*) \ln(-Q_2^*), \quad (19)$$

where $Q_2^* = \lambda_{ge}/\gamma_{th}^{e,B}|\beta|^2\lambda_{me}\lambda_{fe}^2$.

Proof. According to the asymptotic principle, Equation (19) can be obtained by substituting $e^x \approx 1 + x$ and $Ei(x) \approx \ln x + C$ into Equation (17), where C is a constant. □

3.2.2. Intercept Probability Expressions for BD. The IP expression of backscattered signals can be written as

$$P_{\text{int}}^{\text{BD}} = \Pr\left(\gamma_{e,\text{BD}} > \gamma_{\text{th}}^{e,\text{BD}}\right), \tag{20}$$

where $\gamma_{\text{th}}^{e,\text{BD}}$ is the secrecy SNR threshold of BD.

Theorem 4. *For the backscatter link, we can get the IP analytical expression, which can be expressed as*

$$P_{\text{int}}^{BD} = 2\sqrt{Q_3}K_1\left(2\sqrt{Q_3}\right), \tag{21}$$

where $Q_3 = \gamma_{th}^{e,BD}/\gamma|\beta|^2\lambda_{gn}\lambda_{fe}$ *and* $K_1(\bullet)$ *is the 1st order modified Bessel function of the second kind.*

Proof. Substituting Equation (6) into Equation (20), $P_{\text{int}}^{\text{BD}}$ can be expressed as

$$P_{\text{int}}^{\text{BD}} = \frac{1}{\lambda_{fe}}\int_0^{\infty} e^{-\gamma_{\text{th}}^{e,\text{BD}}/\gamma|\beta|^2\lambda_{gn}y - y/\lambda_{fe}}dy, \tag{22}$$

where the integral can be calculated by Equation (3.324.1) [28]. □

Corollary 5. *Under high MER case, the IP asymptotic expression for backscatter link can be written as*

$$P_{\text{int},\infty}^{BD} = 1 + 2Q_3^* \ln\left(\sqrt{Q_3^*}\right), \tag{23}$$

where $Q_3^* = \gamma_{th}^{e,BD}/\gamma|\beta|^2\lambda_{me}\lambda_{fe}^2$.

Proof. By using approximate equation $K_1(x) \approx (1/x) + (x/2)\ln(x/2)$, we can get Equation (23). □

Remark 2. From Theorems 3 and 4 and Corollaries 4 and 5, we can know that (1) with increasing reflection coefficient β, the security of B gets enhanced owing to P_{int}^{B} decreases; (2) the security of B and BD decrease when P_s increases; (3) when λ_{me} increases, the security of B is enhanced; (4) when $\lambda_{me} \longrightarrow \infty$, $P_{\text{int},\infty}^{\text{BD}}$ is close to 1, suggesting λ_{me} takes small value to ensure BD security.

4. Numerical Results

In this section, the correctness of previous analysis is validated via numerical simulation results. In these simulation examples, we consider 10^5 Monte Carlo trials. It is assumed that $\lambda_{h0} = 6$, $\lambda_{hn} = 4$, $\lambda_{gn} = 5$, $\lambda_{fe} = 6$, $\lambda_{ge} = 4$, $\beta = 0.2$, and $\sigma^2 = 1$.

The simulation results of OPs versus P_s for B and BD are plotted in Figure 2. It can be seen that in high P_s regions, the OPs are getting close to constant with increasing P_s, resulting in 0 diversity orders. From Remark 1, we can know that the value of the joint error-floor depends on the SNR threshold at B. Furthermore, it can be seen that the OPs increase when the SNR thresholds, i.e., γ_{th}^{B} and $\gamma_{\text{th}}^{\text{BD}}$, increase.

The IPs versus P_s for B and BD are plotted in Figure 3. The IPs decrease as the secrecy SNR thresholds increase. In low P_s regions, it is shown that the IP decreases significantly, indicating that the TW-AmBC system has higher security performance. In high P_s regions, the IPs are close to 1; thus, the security is impaired. Meanwhile, from Figures 2 and 3, it can be observed that when P_s increases, the OPs decrease but the IPs increase. Thus, it can be inferred that a trade-off between reliability and security exists in this system.

Figure 4 illustrates the IPs versus OPs for different β values. It can be observed that the system outage performance gets impaired when β changes from 0.1 to 0.6. This is because the backscattered signals make the interference stronger when B decodes signals. In addition, it can be observed that the secrecy performance of direct link is enhanced when β increases. It can be explained that the backscattered signals increase with increasing β; thus, it can project more interference when E decodes the signals from B. Meanwhile, when increasing β, the scopes of OP and IP are obviously reduced, which indicates that the trade-off between security and reliability degrades. In summary, high reliability and high security cannot be achieved simultaneously, which indicates that the reliability is compromised to get better security and vice versa.

The OPs and IPs versus β for different P_s values are plotted in Figure 5. In low β regions, the IP of BD decreases significantly, which greatly improves the system security performance. Thus, when β is small, the eavesdropper can easily intercept the messages from B, but it is difficult to decode the messages from BD. The IP trend curves of BD and B are opposite, indicating that high security of one link can be achieved with compromised security performance of the other link. Thus, a trade-off of security exists between the two links. Meanwhile, the curve of BD's OP firstly decreases and then increases, suggesting that we can get an optimal value β to minimize OP.

Figure 6 reveals the variations of IPs versus MER in conditions of different P_s values, with $\beta = 0.1$. We consider 10^6 Monte Carlo trials to reduce the impact of randomness on experimental accuracy. It can be observed that when MER increases, the IP of B decreases and that of BD increases, and in high MER regions, the IP of BD is close to 1, suggesting that the security of backscatter link can only be achieved when MER is low. When P_s decreases, the IPs decrease, indicating that the security gets enhanced. Beyond that, there is the strict approximation relationship between the theoretical value and the asymptotic value.

5. Conclusions

In this article, a novel TW-AmBC network structure was proposed, which integrated the benefits of both TW communication and AmBC. Through analyzing the PLS performance of the proposed TW-AmBC, it is found that the parameters can be designed to achieve an optimal trade-off between reliability and security, such as reflection coefficient β and the transmitted power P_s. Specifically, the secrecy performance of B can be enhanced when reflection coefficient β increases. When the transmitted power P_s is

high, the asymptotic OPs approach a joint error-floor, indicating that the reliability of two links is similar. These findings would be useful in instructing to apply BD in limited spectrum resource application scenarios.

Appendix

A. Proof of Theorem 1

Substituting Equation (2) into Equation (7), the OP of B can be given as

$$P_{\text{out}}^{B} = 1 - \underbrace{\Pr\left(\frac{\gamma|h_0|^2}{\gamma|\beta|^2|h_n|^2|g_n|^2+1} > \gamma_{\text{th}}^{B}\right)}_{I_1}, \tag{A.1}$$

where I_1 can be calculated as follows:

$$I_1 = \int_0^{\infty}\int_{\gamma_{\text{th}}^{B}|\beta|^2 y+\gamma_{\text{th}}^{B}/\gamma}^{\infty} \frac{1}{\lambda_{h0}} e^{-x/\lambda_{h0}} f_{|hn|^2|gn|^2}(y)dxdy, \tag{A.2}$$

where $f_{|hn|^2|gn|^2}(y)$ is the joint probability density function. The probability density functions of hn and gn are $f_{|hn|^2}(x) = (1/\lambda_{hn})e^{-x/\lambda_{hn}}$ and $f_{|gn|^2}(y) = (1/\lambda_{gn})e^{-y/\lambda_{gn}}$, respectively. We make $Z = X \bullet Y$; the joint probability density function $f_{|hn|^2|gn|^2}(z)$ can be expressed as

$$\begin{aligned} f_{|hn|^2|gn|^2}(z) &= \int_0^{\infty} \frac{1}{u} f_{|hn|^2}\left(\frac{z}{u}\right) f_{|gn|^2}(u)du \\ &= \frac{1}{\lambda_{hn}\lambda_{gn}} \int_0^{\infty} \frac{1}{u} e^{-z/u\lambda_{hn}-u/\lambda_{gn}} du \\ &= \frac{2}{\lambda_{hn}\lambda_{gn}} K_0\left(2\sqrt{\frac{z}{\lambda_{hn}\lambda_{gn}}}\right), \end{aligned} \tag{A.3}$$

where the integral can be derived by Equation (3.324.1) [28] and $K_0(\bullet)$ is the 0th modified Bessel function of the second kind. Substituting Equation (A.3) into Equation (A.2), I_1 can be expressed as

$$I_1 = \frac{2e^{-\gamma_{\text{th}}^{B}/\gamma\lambda_{h0}}}{\lambda_{hn}\lambda_{gn}} \int_0^{\infty} e^{-\gamma_{\text{th}}^{B}|\beta|^2 y/\lambda_{h0}} K_0\left(2\sqrt{\frac{y}{\lambda_{hn}\lambda_{gn}}}\right) dy, \tag{A.4}$$

where the integral can be resolved by Equation (6.643.3) [28], we have

$$\int_0^{\infty} e^{-\alpha x} K_0\left(2\theta\sqrt{x}\right) dx = \frac{e^{\theta^2/2\alpha}}{2\theta\sqrt{\alpha}} W_{-1/2,0}\left(\frac{\theta^2}{\alpha}\right), \tag{A.5}$$

where $\alpha = \gamma_{th}^{B}|\beta|^2/\lambda_{h0}$, $\theta = \sqrt{1/\lambda_{hn}\lambda_{gn}}$, and $W_{-1/2,0}(\theta^2/\alpha)$ is the Whittaker function, which can be substituted with the exponential function by Equation (9.222.1) [28], we have

$$W_{-1/2,0}\left(\frac{\theta^2}{\alpha}\right) = \left(\frac{\theta^2}{\alpha}\right)^{1/2} e^{-\theta^2/2\alpha} \int_0^{\infty} \frac{e^{-\left(\theta^2/\alpha\right)t}}{1+t} dt, \tag{A.6}$$

where the integral can be resolved by Equation (3.352.4) [28], and thus, $W_{-1/2,0}(\theta^2/\alpha) = -\sqrt{(\theta^2/\alpha)}e^{\theta^2/2\alpha}Ei(-\theta^2/\alpha)$. I_1 can be expressed as

$$I_1 = -Q_1 e^{Q_1 - \gamma_{\text{th}}^{B}/\gamma\lambda_{h0}} Ei(-Q_1), \tag{A.7}$$

where $Q_1 = \lambda_{h0}/\gamma_{\text{th}}^{B}|\beta|^2\lambda_{gn}\lambda_{hn}$.

Thus, Equation (8) can be given by substituting Equation (A.7) into Equation (A.1).

B. Proof of Theorem 2

The OP of BD can be denoted as

$$P_{\text{out}}^{\text{BD}} = 1 - \underbrace{\Pr\left(\gamma_B > \gamma_{\text{th}}^{B}, \gamma_{\text{BD}} > \gamma_{\text{th}}^{\text{BD}}\right)}_{I_2}, \tag{B.1}$$

where

$$I_2 = \int_{\gamma_{\text{th}}^{\text{BD}}/\gamma|\beta|^2}^{\infty}\int_{\gamma_{\text{th}}^{B}|\beta|^2 y+\gamma_{\text{th}}^{B}/\gamma}^{\infty} \frac{1}{\lambda_{h0}} e^{-x/\lambda_{h0}} f_{|hn|^2|gn|^2}(y)dxdy. \tag{B.2}$$

We can see from Equations (A.2) and (B.2) that there exists a difference between I_1 and I_2 on the lower limit of integration for x. Thus, we have

$$I_1 - I_2 = \int_0^{\gamma_{\text{th}}^{\text{BD}}/\gamma|\beta|^2}\int_{\gamma_{\text{th}}^{B}|\beta|^2 y+\gamma_{\text{th}}^{B}/\gamma}^{\infty} \frac{1}{\lambda_{h0}} e^{-x/\lambda_{h0}} f_{|hn|^2|gn|^2}(y)dxdy. \tag{B.3}$$

Meanwhile, it is challenging to calculate the integral I_2, thus changing I_2 into the following form.

$$I_2 = I_1 - \underbrace{\frac{2e^{-\gamma_{\text{th}}^{B}/\gamma\lambda_{h0}}}{\lambda_{hn}\lambda_{gn}} \int_0^{\gamma_{\text{th}}^{\text{BD}}/\gamma|\beta|^2} e^{-\gamma_{\text{th}}^{B}|\beta|^2 y/\lambda_{h0}} K_0\left(2\sqrt{\frac{y}{\lambda_{hn}\lambda_{gn}}}\right) dy}_{I_3}, \tag{B.4}$$

where can be given as Equation (B.5) by the Gaussian Chebyshev approximation Equation (20) [29].

$$I_3 = \frac{\pi\gamma_{\text{th}}^{\text{BD}}}{N\gamma|\beta|^2\lambda_{hn}\lambda_{gn}} \sum_{i=1}^{N} e^{-\gamma_{\text{th}}^{B}\gamma_{\text{th}}^{\text{BD}}(\delta_i+1)+2\gamma_{\text{th}}^{B}/2\gamma\lambda_{h0}} K_0\left(\sqrt{\frac{2\gamma_{\text{th}}^{\text{BD}}(\delta_i+1)}{\gamma|\beta|^2\lambda_{hn}\lambda_{gn}}}\right)\sqrt{1-\delta_i^2}. \tag{B.5}$$

Finally, Equation (12) can be obtained by substituting Equations (A.7), (B.4) and (B.5) into Equation (B.1).

Data Availability

The data is available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

This work was supported in part by the Programs of Henan Polytechnic University (No. B2017-57 and B2022-2), in part by the Fundamental Research Funds for the Universities of Henan Province (No. NSFRF200335), in part by the Natural Science Foundation of Guangdong Province with grant number 2022A030313736, in part by the Scientific Research Project of Education Department of Guangdong with grant number 2021KCXTD061, in part by the Science and Technology Program of Guangzhou with grant number 202207010389, Yangcheng Scholar, in part by the Scientific Research Project of Guangzhou Education Bureau with grant number 202032761, and in part by the Application Technology Collaborative Innovation Center of GZPYP with grant number 2020ZX01.

References

[1] X. Yue and Y. Liu, "Performance analysis of intelligent reflecting surface assisted NOMA networks," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.

[2] W. Zhang, Y. Qin, W. Zhao et al., "A green paradigm for Internet of things: ambient backscatter communications," *China Communications*, vol. 16, no. 7, pp. 109–119, 2019.

[3] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: a contemporary survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 2889–2922, 2018.

[4] Y. Ye, L. Shi, R. Qingyang Hu, and G. Lu, "Energy efficient resource allocation for wirelessly powered backscatter communications," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1418–1422, 2019.

[5] Q. Zhang, L. Zhang, Y.-C. Liang, and P.-Y. Kam, "Backscatter-NOMA: a symbiotic system of cellular and Internet-of-Things networks," *IEEE Access*, vol. 7, pp. 20000–20013, 2019.

[6] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[7] Z. Yang, L. Feng, F. Zhou, X. Qiu, and W. Li, "Analytical performance analysis of intelligent reflecting surface aided ambient backscatter communication network," *IEEE Wireless Communications Letters*, vol. 10, no. 12, pp. 2732–2736, 2021.

[8] Z. Niu, W. Ma, W. Wang, and T. Jiang, "Spatial modulation-based ambient backscatter: bringing energy self-sustainability to massive internet of everything in 6G," *China Communications*, vol. 17, no. 12, pp. 52–65, 2020.

[9] Y. Ye, L. Shi, X. Chu, and G. Lu, "On the outage performance of ambient backscatter communications," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7265–7278, 2020.

[10] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.

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

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

[13] X. Li, M. Zhao, Y. Liu, L. Li, Z. Ding, and A. Nallanathan, "Secrecy analysis of ambient backscatter NOMA systems under I/Q imbalance," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12286–12290, 2020.

[14] Y. Zhang, F. Gao, L. Fan, X. Lei, and G. K. Karagiannidis, "Secure communications for multi-tag backscatter systems," *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1146–1149, 2019.

[15] Z. Wang and Z. Peng, "Secrecy performance analysis of relay selection in cooperative NOMA systems," *IEEE Access*, vol. 7, pp. 86274–86287, 2019.

[16] L. Shi, Y. Ye, R. Q. Hu, and H. Zhang, "System outage performance for three-step two-way energy harvesting DF relaying," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3600–3612, 2019.

[17] Z. Wang, W. Shi, W. Liu, Y. Zhao, and K. Kang, "Performance analysis of two-way full-duplex relay mixed RF/FSO system with self-interference," *IEEE Communications Letters*, vol. 25, no. 1, pp. 209–213, 2021.

[18] G. Li, H. Liu, G. Huang, X. Li, B. Raj, and F. Kara, "Effective capacity analysis of reconfigurable intelligent surfaces aided NOMA network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, pp. 1–16, 2021.

[19] S. Atapattu, R. Fan, P. Dharmawansa, G. Wang, J. Evans, and T. A. Tsiftsis, "Reconfigurable intelligent surface assisted two-way communications: performance analysis and optimization," *IEEE Transactions on Communications*, vol. 68, no. 10, pp. 6552–6567, 2020.

[20] Z. Cao, X. Ji, J. Wang, S. Zhang, Y. Ji, and J. Wang, "Security-reliability tradeoff analysis for underlay cognitive two-way relay networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 6030–6042, 2019.

[21] X. Guo, B. Li, J. Cong, and R. Zhang, "Throughput maximization in a UAV-enabled two-way relaying system with multi-pair users," *IEEE Communications Letters*, vol. 25, no. 8, pp. 2693–2697, 2021.

[22] M. K. Shukla and H. H. Nguyen, "Ergodic secrecy sum rate analysis of a two-way relay NOMA system," *IEEE Systems Journal*, vol. 15, no. 2, pp. 2222–2225, 2021.

[23] Y. Liu, Y. Ye, G. Yan, and Y. Zhao, "Outage performance analysis for an opportunistic source selection based two-way cooperative ambient backscatter communication system," *IEEE Communications Letters*, vol. 25, no. 2, pp. 437–441, 2021.

[24] S. T. Shah, K. W. Choi, T.-J. Lee, and M. Y. Chung, "Outage probability and throughput analysis of SWIPT enabled

cognitive relay network with ambient backscatter," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3198–3208, 2018.

[25] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, 2016.

[26] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*, Cambridge university Press, 2007.

[27] S. Jacob, V. G. Menon, K. S. Fathima Shemim, B. Mahapatra, and M. Mukherjee, "Intelligent vehicle collision avoidance system using 5G-enabled drone swarms," in *In Proceedings of the 2nd ACM MobiCom Workshop on Drone Assisted Wireless Communications for 5G and Beyond*, pp. 91–96, New York, NY, USA, 2020.

[28] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, NY, USA, 2007.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 6216372, 11 pages
https://doi.org/10.1155/2022/6216372

Research Article

Computation Offloading in Multi-UAV-Enhanced Mobile Edge Networks: A Deep Reinforcement Learning Approach

Bin Li,[1,2] **Shiming Yu**,[1] **Jian Su**,[1] **Jianghong Ou**,[3] **and Dahua Fan**[3]

[1]*School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China*
[2]*Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications), Ministry of Education, Nanjing 210003, China*
[3]*Starway Communication, No. 31, Kefeng Road, Guangzhou Science City, Guangzhou 510663, China*

Correspondence should be addressed to Jian Su; sj890718@gmail.com

Received 5 December 2021; Revised 11 February 2022; Accepted 23 February 2022; Published 7 March 2022

Academic Editor: Shu Fu

In this paper, we investigate an unmanned aerial vehicle- (UAV-) enhanced mobile edge computing network (MUEMN), where multiple UAVs are deployed as aerial edge servers to provide computing services for ground moving equipment (GME). Each GME is trained to simulate movement by a Gauss-Markov random model in this MUEMN. Under the condition of limited energy cost, UAV dynamically plans its flight position according to the movement trend of GME. Our objective is to minimize the total energy consumption of GME by jointly optimizing the offloading decisions of GME and the flight positions of UAVs. More explicitly, we model the optimization problem as a Markov decision process and achieve real-time offloading decisions via deep reinforcement learning algorithm according to the dynamic system state, where the asynchronous advantage actor-critic (A3C) framework with asynchronous characteristics is leveraged to accelerate the learning process. Finally, numerical results confirm that our proposed A3C-based offloading strategy can effectively reduce the total of energy consumption of GME and ensure the continuous operation of the GME.

1. Introduction

Mobile users usually have limited computing capabilities and battery storages; it is challenging to provide a satisfactory computing service and achieve a low service delay when they face with the emerging applications with computation-intensive features [1–3]. In this context, mobile edge computing (MEC) is considered as a key technology to mitigate these issues [4]. With the help of MEC, mobile devices have the option to offload their computing tasks to nearby edge servers with powerful computing capabilities, enabling the demands for lower energy consumption [5, 6] and reduced latency. Nevertheless, the location of MEC server is usually fixed and cannot be changed flexibly according to the needs of mobile users, which restricts the extension of MEC [7, 8]. At present, frequent occurrence of natural disasters may destroy basic communication facilities on the ground, which makes it difficult for rescue communication efforts. Compared with the general communication infrastructure, unmanned aerial vehicles (UAVs) are highly flexibility and inexpensive, enabling reliable communication. UAVs equipped with MEC servers greatly enhance the application scalability of the traditional MEC model [9, 10].

With the development and maturity of UAV-related technologies, they have been paid much attention in disaster rescue, mineral mining, geological exploration and other wireless scenarios [11, 12]. On the one hand, in regions with incomplete or damaged basic communication facilities, where large-scale outdoor activities are required within a short period of time, UAVs can be deployed in the air on demand to enhance network connectivity and provide reliable communication services. On the other hand, in many civilian application scenarios, such as live broadcast and video shooting, the flow of people tends to be huge, and the offloading of various data generated by mobile devices in these areas to the cloud or base stations (BSs) can trigger high latency [13]. Fortunately, UAVs equipped with the computing resources can serve as the edge nodes to relieve

the pressure on computing resources and improve the user experience. As such, joint development of UAV technology and MEC model, i.e., adopting UAVs to enhance mobile edge computing capabilities, is a promising direction for MEC development.

The current phase of research works on UAV-assisted mobile edge computing is divided into two categories: single/multiple UAV deployment [14] and latency reduction or energy reduction [15, 16]. Note that the ideal layout of the UAV can optimize the total coverage of the UAV, thereby maximizing network advantages. Nevertheless, despite being interesting, the UAV has size and weight constraints, and limited energy profoundly affects sustainable operations. To do this, the flight state of the UAV must be studied to optimize the use of UAV energy. Guo and Liu in [17] designed a single UAV-assisted mobile edge computing network. Under the UAV energy consumption constraint, the authors derived a suboptimal UAV trajectory layout by introducing block coordinate descent and successive convex approximation methods. Distinguished from [17], Liu et al. in [18] employed the Gauss-Markov random model (GMRM) to simulate the mobility of ground moving equipment (GME) and continuously adapted the UAV flight trajectory in the light of the time-varying location of terminal users to promote the quality of service for each mobile terminal user. The performance of the UAV-enabled MEC network is quite limited when a single UAV is used as a computation server in the large-scale scenarios, which motivates the deployment of multiple UAVs. Unlike single UAV deployment, multi-UAV-assisted MEC has more complex trajectories. In [19], Wang et al. synthesized the inter-UAV collision problem and presented a differential evolution algorithm with an elimination operator to optimize the layout of multiple UAVs. Shang and Liu in [20] obtained the target of minimizing the sum energy consumption of users by jointly optimizing users' association, resource allocation, and UAV layout. They further recommended the coordinate descent algorithm to decompose the energy consumption minimization problem into several subproblems to explore the suboptimal solution. In [21], Guo et al. studied a UAV-assisted MEC network with the goal of minimizing the sum delay of all users, adopting the theories of successive convex approximation and difference of convex programming to obtain the suboptimal solution. However, most of the literature defaults to static ground users; the work on jointly optimizing multiple UAV positions and offloading decisions considering ground user movement remains relatively scarce.

Sparked by the above-mentioned observations, in this paper, we propose an MUEMN architecture to provide edge computing for GME. We optimize the task offloading decisions of GME and the flight locations of UAVs in the network to achieve the goal of minimizing the total energy consumption of all GME. The resultant optimization problem is a mixed-integer nonconvex problem, and we propose a deep reinforcement learning- (DRL-) based asynchronous advantage actor-critic (A3C) algorithm, which asynchronously trains optimal computational offloading decisions for all GME in different environments and then uniformly uploads the training parameters to the global network to update the parameters and continuously train them to finally obtain optimal network parameters.

Specifically, the main contributions of this paper can be summarized as follows:

(1) Considering the dilemma of the traditional MEC model, we propose a multi-UAV-enhanced MEC network. Different from the fixed setting of ground equipment in most work, the ground equipment in our network follows the GMRM and moves within a certain period of time. UAVs continuously optimize their flight position with reference to the movement trend of GME

(2) We comprehensively consider the issues of UAV signal coverage, collisions between UAVs, and UAV energy consumption in the multi-UAV scenario. Under the constraints of these background issues, we introduce the A3C algorithm to find the suboptimal solution that minimizes the total energy consumption of all GME and derive the optimal computing task offload decisions and flight positions of UAVs

(3) Numerical results show that under the constraint of calculation delay, as the size of the calculation task increases, the offloading strategy based on the traditional algorithm is difficult to effectively reduce the total energy consumption of GME. In this paper, the proposed A3C algorithm with asynchronous characteristics can generate an effective offloading strategy

2. System Model and Problem Formulation

We describe the network model, communication model, computation model, flying model, and problem formulation in this section.

2.1. Network Model. We consider a multi-UAV-enhanced mobile edge computing network (MUEMN), including M UAVs deployed with MEC servers, $\mathcal{M} = \{1, 2, 3, \cdots, M\}$, and K GME, $\mathcal{K} = \{1, 2, 3, \cdots, K\}$. The network model is shown in Figure 1. We assume that the UAVs with limited energy can provide task offloading service for K GME within a certain period. Without loss of generality, the GME k and the UAV m serve one-to-one during this period, and all tasks must be guaranteed to be completed within the specified time period L. To simulate the mobility of GME and UAVs, we divide the calculation time of nonexecuting tasks in the period L into T frames, and the time of each frame t is uniform, which is denoted as $t = \{0, 1, 2, \cdots, T\}$. In this paper, the UAVs are assumed to be flying at a constant altitude H without frequenting ups and downs and maintain communication with K GME in each frame through the periodic time division multiple access (TDMA) protocol. Similar to prior studies, we use 3D Cartesian coordinate system to simulate the position of each node, and the coordinate unit is meter. Note that the 3D position of the UAV m is $\mathbf{U}_m^{\mathrm{u}}(t) = (x_m^{\mathrm{u}}(t), y_m^{\mathrm{u}}(t), H)$, whereas the two

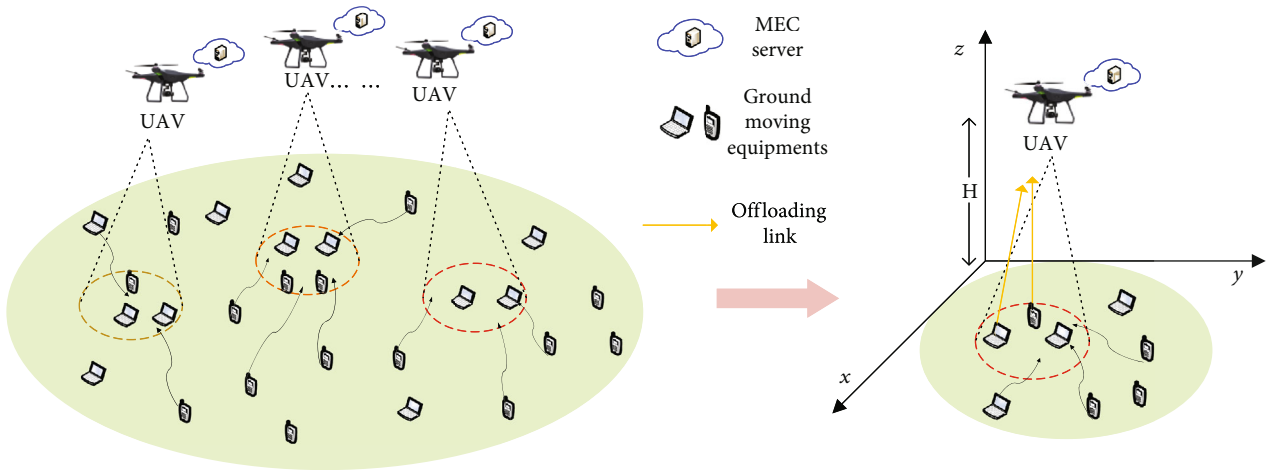


Figure 1: Task offloading network model for GME in multi-UAV scenarios.

UAVs need to meet the constraint $d_{\min}^{uu} \le d_{m_1 m_2}^{uu}(t)$, where $d_{\min}^{uu}$ represents the minimum allowable distance between two adjacent UAVs, and $d_{m_1 m_2}^{uu}(t) = \sqrt{(x_{m_1}^{u}(t) - x_{m_2}^{u}(t))^2 + (y_{m_1}^{u}(t) - y_{m_2}^{u}(t))^2}, \forall m_1, m_2 \in \mathcal{M}, m_1 \neq m_2$ represents the spacing between two adjacent UAVs.

In this MUEMN, we consider that all GME has random positions at $t = 0$ and do not change their positions within $\triangle_{t,t+1}$. Based on the GMRM [22], the movement speed and direction angle of the GME k at the tth ($t > 0$) frame are denoted as

$$v_k(t) = \tau_1 v_k(t-1) + (1-\tau_1)\bar{v}_k + \sqrt{1-\tau_1^2}\Omega_k, \quad (1)$$

$$\alpha_k(t) = \tau_2 \alpha_k(t-1) + (1-\tau_2)\bar{\alpha}_k + \sqrt{1-\tau_2^2}\Psi_k, \quad (2)$$

where $0 \le \tau_1, \tau_2 \le 1$ indicate the parameters for adjusting the state of the previous frame and $\bar{v}_k$ and $\bar{\alpha}_k$ stand for the average velocity and movement direction angle of the GME k, respectively. Also, Ω_k and Ψ_k follow two uncorrelated random Gaussian distributions with different mean-variance to simulate the random mobility of the GME k. From (1) and (2), the 3D X-coordinate and 3D Y-coordinate of the GME k at the tth frame can be deduced as

$$\begin{aligned} x_k(t) &= x_k(t-1) + v_k(t-1)\triangle_{t-1,t} \cos(\alpha_k(t-1)), \\ y_k(t) &= y_k(t-1) + v_k(t-1)\triangle_{t-1,t} \sin(\alpha_k(t-1)). \end{aligned} \quad (3)$$

To sum up, the 3D position coordinates of the GME k at the tth frame is $\mathbf{G}_k(t) = (x_k(t), y_k(t), 0)$. The visualized 3D model of the network unit can be referred to the right side of Figure 1.

2.2. Communication Model. In this paper, the line-of-sight wireless channels between GME and UAVs are more dominant than other channel impairments due to the high altitudes of UAVs. Therefore, the channel link between the GME k and the UAV m can be denoted by the free-space path loss model as follows:

$$h_{k,m}^{ul}(t) = \frac{\beta_0}{(x_m^u(t) - x_k(t))^2 + (y_m^u(t) - y_k(t))^2 + H^2}, \quad (4)$$

where β_0 is the channel power gain at a reference distance of 1 m.

Since each UAV can receive the offloaded task from at most one GME in each frame, the communication interference between channels can be neglected. As a result, the uplink transmission data rate between GME k and UAV m in a certain frame is calculated as

$$R_{k,m}^{ul}(t) = B \log_2\left(1 + \frac{h_{k,m}^{ul}(t)p_k}{\sigma^2}\right), \quad (5)$$

where B is the available channel bandwidth, p_k is the transmission power of the GME k, and σ^2 denotes the Gaussian noise power.

2.3. Computation Model. Considering that all GME distributed in the MUEMN generate a computationally intensive, latency-sensitive task $W_k = \{L_k, C_k, t_k^{\max}\}$, where L_k denotes the data size for calculating the offload task, C_k stands for the number of CPU cycles required to calculate each bit of task data, and $t_k^{\max}$ expresses the maximum tolerable task latency. The UAVs collaborate with each other to provide computing services to GME. Herein, $a_{k,m} \in \{0, 1\}$ is used to denote the GME k task offloading decision variable, where $a_{k,m} = 0$ indicates that the GME k chooses to perform local computation, and $a_{k,m} = 1$ expresses that the GME k chooses task offloading to the UAV m.

2.3.1. Local Computing. When the GME k decides to perform the calculation locally, the calculation execution time can be expressed as

$$t_k^{\mathrm{loc}} = \frac{L_k C_k}{f_k^{\mathrm{loc}}}, \tag{6}$$

where f_k^{loc} is the local computing power of the GME k. Correspondingly, the energy consumed by local calculation can be calculated as

$$E_k^{\mathrm{loc}} = \rho_k^{\mathrm{loc}} L_k C_k \left(f_k^{\mathrm{loc}}\right)^2, \tag{7}$$

where ρ_k^{loc} marked as the chip correlation coefficient of the GME k.

2.3.2. UAV Edge Computing. When the GME k moves into the coverage area of the UAV m, i.e., the constraint $d_{k,m}^{\mathrm{gu}}(t) \le R$ is satisfied and the UAV m becomes an option for the GME k to offload the computational task, where $d_{k,m}^{\mathrm{gu}}(t) = \sqrt{(x_k(t) - x_m^{\mathrm{u}}(t))^2 + (y_k(t) - y_m^{\mathrm{u}}(t))^2}, \forall k \in \mathcal{K}, \forall m \in \mathcal{M}$ denotes the horizontal distance between the GME k and the UAV m, $R = H \tan \vartheta$ indicates the coverage radius of each UAV, and ϑ is UAV antenna elevation angle [23]. When the GME k is in the coverage of multiple UAVs, the GME k randomly selects a UAV to offload the computational task. The process of offloading computational tasks from a GME to a UAV is divided into three main steps: (1) the GME offloads the computing task to a selected UAV; (2) the selected UAV receives the computational task and performs the calculation; (3) the selected UAV returns the results to the corresponding GME. As a result, the amount of data returned is small enough to be negligible. Therefore, the transmission time required for the GME k to offload the computational task to the UAV m, the energy consumption transmitted by the GME k, and the energy consumption received by the UAV m are expressed, respectively, as

$$\begin{aligned} t_{k,m}^{\mathrm{tr}}(t) &= \frac{L_k}{R_{k,m}^{\mathrm{ul}}(t)}, \\ E_{k,m}^{\mathrm{tr}}(t) &= p_k \frac{L_k}{R_{k,m}^{\mathrm{ul}}(t)}, \\ E_{k,m}^{\mathrm{re}}(t) &= p_m^{\mathrm{u}} \frac{L_k}{R_{k,m}^{\mathrm{ul}}(t)}, \end{aligned} \tag{8}$$

where p_m^{u} is the receiving power of the UAV m.

2.4. Flying Model

2.4.1. The Energy Consumption of Edge Computing. For a UAV with limited energy to work continuously, we need to constrain the UAV's energy. In this paper, the energy consumption of the UAV is divided into three main components: (1) reception energy consumption and calculated energy consumption (collectively known as edge computing energy consumption); (2) UAV flight energy consumption; (3) UAV hovering energy consumption. Let f_m^{u} and ρ_m^{u} be the computational power and the chip correlation coefficient of UAV m, respectively. Correspondingly, the time required and the energy consumed for the task calculation of the UAV m can be calculated as

$$t_{k,m}^{\mathrm{cal}} = \frac{L_k C_k}{f_m^{\mathrm{u}}}, \tag{9}$$

$$E_{k,m}^{\mathrm{cal}} = \rho_m^{\mathrm{u}} L_k C_k (f_m^{\mathrm{u}})^2. \tag{10}$$

According to (9) and (10), the edge computing energy consumption can be derived as

$$E_{k,m}^{\mathrm{edg}}(t) = E_{k,m}^{\mathrm{re}}(t) + E_{k,m}^{\mathrm{cal}} = p_m^{\mathrm{u}} \frac{L_k}{R_{k,m}^{\mathrm{ul}}(t)} + \rho_m^{\mathrm{u}} L_k C_k (f_m^{\mathrm{u}})^2. \tag{11}$$

2.4.2. The Energy Consumption of UAV Flying. Given that the UAV is flying at a constant altitude H, there is no change in the gravitational potential energy of the UAV in this paper. To this end, the UAV flight energy consumption only needs to consider kinetic energy, the flight speed, and energy consumption of the UAV m at the tth frame given by

$$\begin{aligned} v_m^{\mathrm{u}}(t) &= \frac{\mathbf{U}_m^{\mathrm{u}}(t) - \mathbf{U}_m^{\mathrm{u}}(t-1)}{\triangle}, \\ E_m^{\mathrm{f}}(t) &= \frac{1}{2} w \triangle \| v_m^{\mathrm{u}}(t) \|^2, \end{aligned} \tag{12}$$

where w is the effective weight of the UAV and $\triangle$ denotes the duration of each frame.

2.4.3. The Energy Consumption of UAV Hovering. The UAV receives a task offload request from a GME within the signal coverage area and will switch from flight state to hover state for the entire edge computing cycle. In this paper, the task offloading consists of two main phases: task transfer and execution of task calculation, and the calculation is reflected as

$$t_{k,m}^{\mathrm{edg}}(t) = t_{k,m}^{\mathrm{tr}}(t) + t_{k,m}^{\mathrm{cal}}. \tag{13}$$

To simplify the problem analysis, the energy consumed by the UAV m hovering E_m^{st} is considered as a constant.

By reason of the foregoing, under the premise that the total energy of the UAV mE_m^{u} is limited, the UAV m operation needs to satisfy the energy constraint

$$\sum_{k \in \mathcal{K}} a_{k,m} \left(E_{k,m}^{\mathrm{edg}}(t) + E_m^{\mathrm{st}} t_{k,m}^{\mathrm{edg}}(t)\right) + \sum_{t=1}^{T} E_m^{\mathrm{f}}(t) \le E_m^{\mathrm{u}}. \tag{14}$$

2.5. Problem Formulation. In this paper, we aim to minimize the total energy consumption of all GME for multi-UAV-enhanced MEC network by jointly optimizing the offloading decision variable $\mathbf{a} \triangleq \{a_{k,m}, \forall k \in \mathcal{K}, \forall m \in \mathcal{M}\}$ and UAV location $\{(x_m^{\mathrm{u}}, y_m^{\mathrm{u}})\}$. As such, the corresponding optimization problem can be formulated as

$$\begin{aligned}
\min_{\mathbf{a},\{(x_m^{\mathrm{u}},y_m^{\mathrm{u}})\}} \quad & \sum_{k=1}^{K}\sum_{m=1}^{M}\left((1-a_{k,m})E_k^{\mathrm{loc}}+a_{k,m}E_{k,m}^{\mathrm{tr}}\right),\\
\text{s.t.} \quad & \mathrm{C1}:\sum_{k\in K}a_{k,m}\left(E_{k,m}^{\mathrm{edg}}(t)+E_m^{\mathrm{st}}t_{k,m}^{\mathrm{edg}}(t)\right)+\sum_{t=1}^{T}E_m^{\mathrm{f}}(t)\le E_m^{\mathrm{u}},\forall k\in\mathscr{K},\forall m\in\mathscr{M},\\
& \mathrm{C2}:a_{k,m}d_{k,m}^{\mathrm{gu}}\le R,\forall k\in\mathscr{K},\forall m\in\mathscr{M},\\
& \mathrm{C3}:d_{\min}^{\mathrm{uu}}\le d_{m_1m_2}^{\mathrm{uu}},\forall m_1,m_2\in\mathscr{M},m_1\ne m_2,\\
& \mathrm{C4}:(1-a_{k,m})\frac{L_kC_k}{f_k^{\mathrm{loc}}}+a_{k,m}\left(t_{k,m}^{\mathrm{tr}}+t_{k,m}^{\mathrm{cal}}\right)\le t_k^{\max},\forall k\in\mathscr{K},\forall m\in\mathscr{M},\\
& \mathrm{C5}:a_{k,m}\in\{0,1\},\forall k\in\mathscr{K},\forall m\in\mathscr{M},\\
& \mathrm{C6}:\sum_{m=1}^{M}a_{k,m}=1,\forall k\in\mathscr{K},\forall m\in\mathscr{M},
\end{aligned}\tag{15}$$

where constraint C1 regulates the use of the UAV energy, constraint C2 indicates the coverage of the UAV signal, constraint C3 ensures the minimum distance between adjacent UAVs to prevent collisions, constraint C4 denotes the maximum latency allowed for the computing task, constraint C5 refers to the binary constraint, and C6 guarantees that each GME connects to at most one UAV. It can be clearly seen that Problem (15) is a mixed-integer nonlinear and nonconvex problem due to the nonconvex objective function and the constraint, which is challenging to solve and requires highly computational complexity to find a globally optimal solution utilizing classical mathematical tools. To this end, appropriate algorithms need to be designed for solving this type of problem efficiently [24]. In the following sections, we propose an A3C-based computational offloading algorithm to obtain suboptimal solution.

3. Proposed DRL-Based Approach: A3C

In this paper, we intend to use DRL-based A3C algorithm [25] to explore unknown environments, where GME goes through different task offloading decisions and UAVs learn from feedback by trying different moves. Continuously, the global network optimizes task offloading decisions and location moves until a suboptimal solution is obtained.

3.1. An Overview of A3C Algorithm. Compared with the traditional deep reinforcement learning algorithms, the A3C algorithm optimizes and improves the actor-critic (AC) algorithm [26]. Based on this, the A3C algorithm solves the problem that the AC algorithm is difficult to converge and achieves fast convergence, which can meet our needs. In detail, the AC algorithm uses an approximate value function to guide the policy parameter updates, and its single-step update can speed up the convergence. However, despite being effective, the AC algorithm requires a complete sequence of states, and iteratively updates the policy function separately, so that it is not easy to converge. As shown in Figure 2, the A3C algorithm utilizes its asynchronous feature to start multiple threads at the same time, while the agents learn by interacting with the environments in multiple threads separately. Each thread will complete the training independently and uploads the training data to the global model parameters in an asynchronous manner. At the same time, the model parameters of the threads are periodically synchronized with the global model parameters, and then, a new round of training is performed with the new parameters.

3.2. A3C-Based Offloading of Computing Task. In the MUEMN model, the GME with computational tasks may choose to compute locally or offload to UAVs in the current signal coverage area within each frame. Subject to the anti-collision constraint, energy constraint, and delay constraint, we aim to minimize the total energy consumption of all GME. The objective optimization problem can be modelled as an MDP by offloading GME tasks.

An MDP consists of a five-tuple: MDP $= \langle \mathscr{S}, \mathscr{A}, \mathscr{P}, \mathscr{R}, \gamma \rangle$, where $\mathscr{S}$ denotes the set of states of the environment, $\mathscr{A}$ describes the set of actions, $\mathscr{P}$ indicates the state transfer probability, $\mathscr{R}$ expresses the reward function, and γ is the decay coefficient. The MDP formulation of the MUEMN is as follows.

The state space in the MUEMN is described as

$$\mathscr{S}=\{s_t|s_t=\{\mathbf{U}_m^{\mathrm{u}}(t),\mathbf{G}_k(t),E_m^{\mathrm{u}}\}\},\quad \forall k\in\mathscr{K},\forall m\in\mathscr{M}.\tag{16}$$

The action space in the MUEMN consists of two kinds of actions, i.e., local computation and offloading to the UAV, expressed as follows

$$\mathscr{A}=\{a_k(t)|a_k(t)=\{0,1\}\},\quad \forall k\in\mathscr{K}.\tag{17}$$

The state transfer and action decision of the GME in the MUEMN is only related to the positions of GME and UAVs and the energy states of UAVs, so the state transfer probability can be expressed as

$$\mathscr{P}_{ss'}=\mathscr{P}\left(s_{t+1}=s'|s_{t=s}\right).\tag{18}$$

To minimize the total energy consumption of all GME, we consider designing a reward function, which assigns a negative reward if the action taken by the GME k in the state of the current frame satisfies constraints C1-C6. Briefly, the reward function can be calculated as

$$r(s_t,a_t)=-\left((1-a_{k,m})E_k^{\mathrm{loc}}+a_{k,m}E_{k,m}^{\mathrm{tr}}\right).\tag{19}$$

On the contrary, if the GME k violates the constraints, we will punish it. For instance, the GME k local calculation violates the delay constraint, we will do the following processing for its local calculation energy consumption

$$r(s_t,a_t)=-\frac{t_k^{\mathrm{loc}}}{t_k^{\max}}E_k^{\mathrm{loc}}.\tag{20}$$

With regard to the optimization Problem (15), it can be observed that the value sequence of the binary decision variables directly affects the suboptimal solution of the optimization problem. We pass the state of the environment to the local network to obtain a sequence of task offloading decisions and then accumulate the reward value adopting

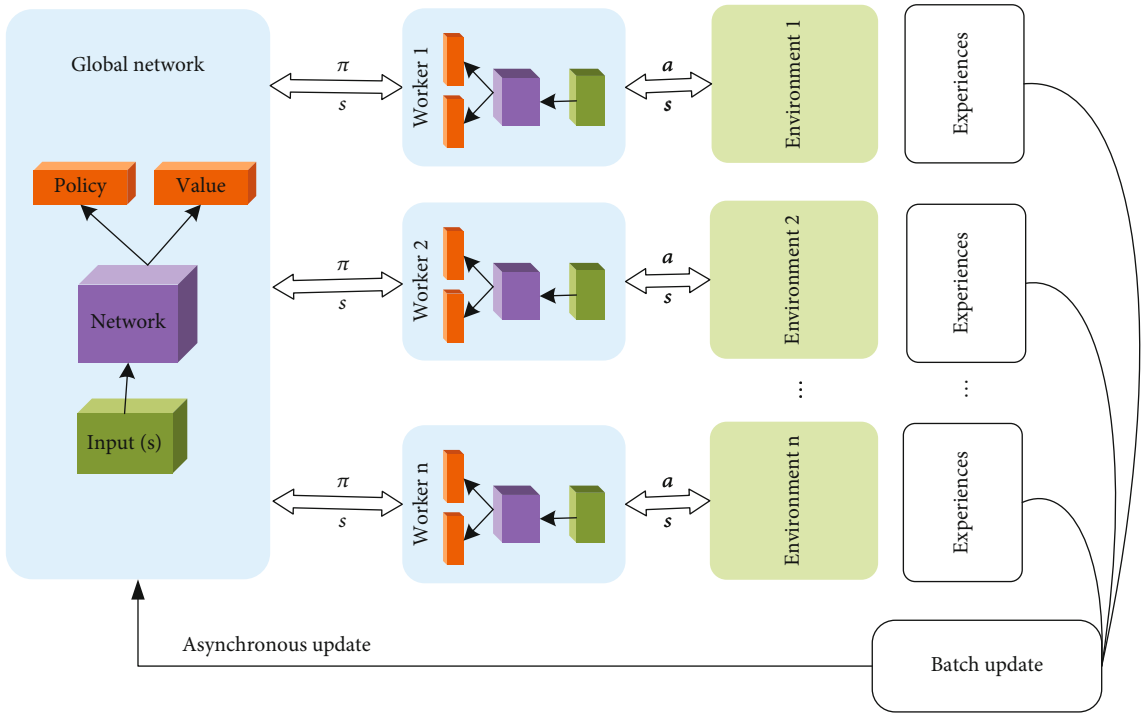


Figure 2: A3C algorithm asynchronous training framework.

the reward function. Multiple threads proceed asynchronously in this manner, leaving the training parameters to the global network for coordination. Ultimately, an optimal network parameter and a suboptimal reward value are derived. As shown in Algorithm 1, we give the detailed steps of the optimal network parameters for the A3C-based offloading strategy in the MUEMN.

3.3. Calculating Offloading Decision Generation. In particular, we introduce the interaction process of a certain thread's environment state sequence and action sequence in this subsection. For the computational task L_k generated by the GME k at tth frame, we consider the position of the GME k, the positions of the UAVs, and the UAVs' energy states as a set of state. Further, we input the state sequence $\mathcal{S}$ into the local network model of a thread, which is trained by the network to produce an action sequence $\mathcal{A}$, the elements of which correspond to the task offloading decisions of each of the K GME.

4. Numerical Results

4.1. Simulation Configurations. In this section, the simulation results are presented to evaluate the performance of our proposed A3C algorithm. We compare A3C with the following there commonly used baseline methods:

(1) Greedy: when the GME is in the coverage area of the UAV, the GME selects either local execution or UAV execution for the computation task depending on the magnitude of the local computation delay and transmission delay [27]

(2) Random: the GME within UAV signal coverage can randomly select the object of computational task execution, i.e., local execution or UAV execution

(3) DQN: the neural network accepts the environment state to calculate the value function and then uses the ε-greedy strategy to output the task offload decisions [28]

In the simulation, the software environment is Python 3.7 with TensorFlow and Visual Studio Code, and the hardware environment is a computer with Intel Core i5-9500 CPU and RAM 8.0 GB. Consider that the simulation scenario consists of M UAVs and K GME, and the area is a 300 m ◊ 300 m square single cell area. The horizontal plane flight altitude of the UAV $H = 80$ m. The effective weight of the UAV is set to 10 kg, the energy budget of the UAV E_m^{u} is set to 200 kJ, and the hovering energy consumption E_m^{st} is set as 200 W [29]. In addition, we set the total duration of each task completion cycle as $L = 10$ s, and the part of equipment that moves freely during this time can be divided into $T = 50$ frames; thus, the duration of each frame can be expressed as $\triangle = (L - \max\{a_{k,m} t_{k,m}^{\mathrm{cal}}, \forall k \in \mathcal{K}, \forall m \in \mathcal{M}\})/T$. Furthermore, we assume that the channel power gain β_0 at the reference distance of 1 m is set to -50 dB. The available bandwidth B is set to be 40 MHz and the noise power $\sigma^2 = 10^{-16}$ W. The coefficients related to the GME and the UAVs are set as $\rho_k^{\mathrm{loc}} = \rho_m^{\mathrm{u}} = 10^{-28}$.

Regarding the size of the computational tasks, we assume that they are randomly arranged in a certain interval. Meanwhile, the computing power of the GME k is set to $f_k^{\mathrm{loc}} = 0.5$ G cycles/s, and the computational capability of the UAV m

Input: The decay value of the reward γ, global shared count N, and global maximum shared count $N_{\max}$.
Output: Optimal network parameters θ and ω as well as the reward value $\mathscr{R}(s_n, a_n)$.
Initialization: Actor network parameter θ and critic network parameter ω in the global shared parameters, actor network parameter θ', and critic network parameter ω' in this thread;
Initialize local count $n = 1$;
repeat
  Reset gradient of local actor network and critic network: $d\theta \longleftarrow 0, d\omega \longleftarrow 0$;
  Synchronize parameters from the global network to this thread network: $\theta' = \theta, \omega' = \omega$;
  $n_{\text{start}} = n$;
  Initialize state s_n;
  repeat
    Based on the strategy $\pi(a_n|s_n;\theta')$ select out action a_n;
    Execute action a_n to get reward value r_n and new state s_{n+1};
    $N \longleftarrow N + 1, n \longleftarrow n + 1$;
  until s_n is the terminal state or $n - n_{\text{start}} == n_{\max}$;
  Calculate the value of $Q(s, n)$ for state s_n at the last count n:
  $$Q(s,n) = \begin{cases} 0, & s_n \text{ is the terminal state,} \\ V(s_n, \omega'), & \text{otherwise;} \end{cases}$$
  for $i \in (n-1, n-2, \cdots, n_{\text{start}})$ **do**
    $Q(s, i) = r_i + \gamma Q(s, i+1)$;
    Calculate the cumulative gradient of local actor parameter θ:
    $d\theta \longleftarrow d\theta + \nabla_{\theta'} \log \pi(a_n|s_n;\theta')(Q(s,i) - V(S_i, \omega'))$
    Calculate the cumulative gradient of local critic parameter ω:
    $d\omega \longrightarrow d\omega + (\partial(Q(s,i) - V(S_i, \omega'))^2 / \partial\omega'$;
  end for
  Update the global network model parameters θ and ω using the local cumulative gradient $d\theta$ and $d\omega$ asynchronously, respectively;
until $N > N_{\max}$

Algorithm 1: A3C-based offloading of computational tasks—arbitrary single-threaded execution process.

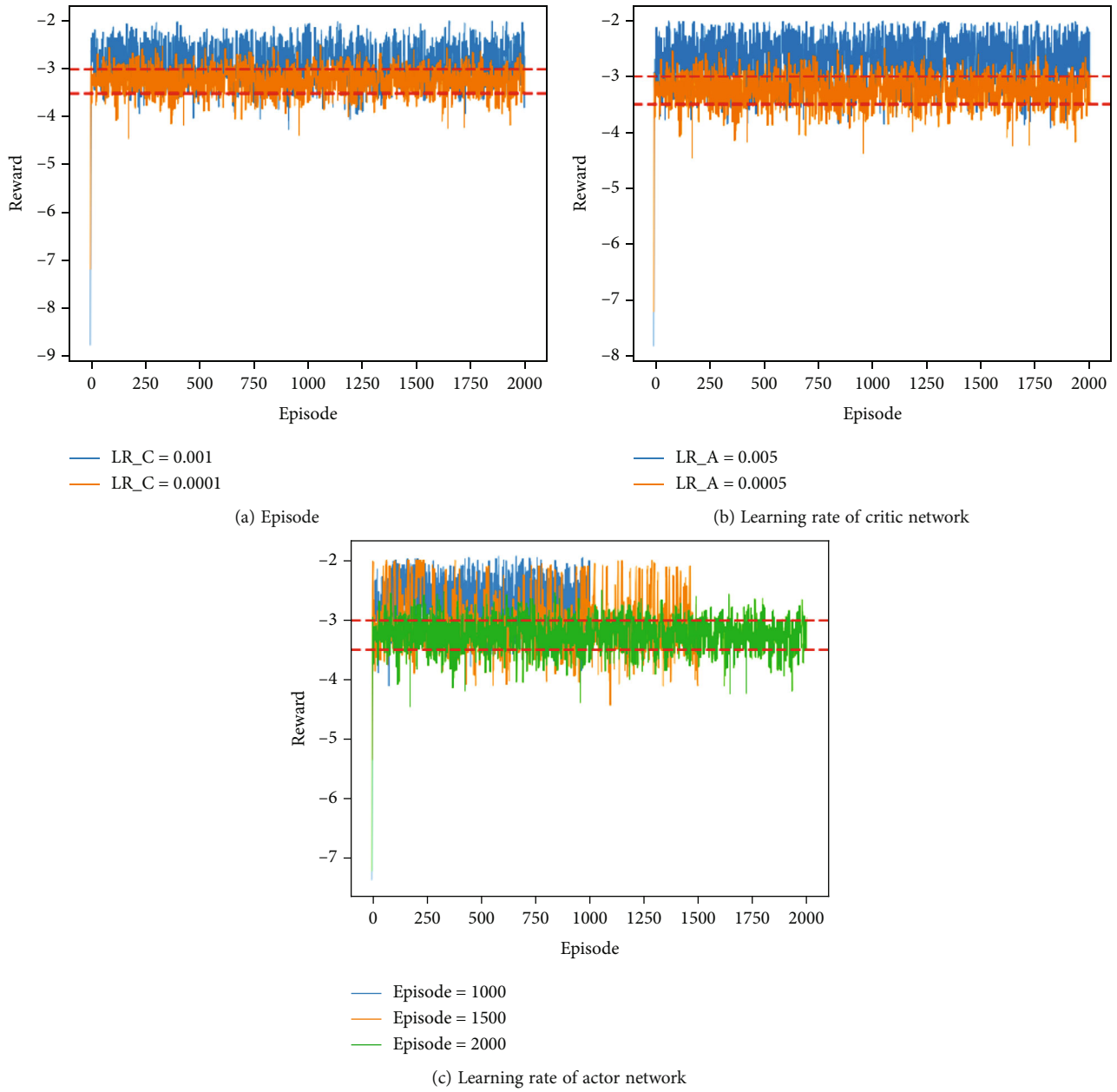
to each GME is set to $f_m^{\text{u}} = 5$ G cycles/s by reference to [30]. The specific parameter settings are shown in Table 1.

Table 1: Parameter setting.

Parameters	Values	Parameters	Values
β_0	-50 dB	ϑ	$\pi/4$
σ^2	10^{-16} W	p_k	50 mW
$\rho_k^{\text{loc}}, \rho_m^{\text{u}}$	10^{-28}	p_m^{u}	50 mW
f_k^{loc}	0.5 G cycles/sec	B	40 MHz
f_m^{u}	5 G cycles/sec	L_k	[5, 10] MB
$d_{\min}^{\text{uu}}$	4 m	C_k	[150, 200] cycles/bit

4.2. Performance Comparison. Assuming the number of GME is 20 and the number of UAVs is 3, i.e., $K = 20, M = 3$, we can clearly observe from Figure 3 that the total energy consumption of GME decreases rapidly within several iterations. The asynchronous nature of the A3C algorithm makes the reward value oscillate in an interval, and we need to reduce the oscillation interval as much as possible. When the scale of GME is large, it is acceptable for the reward value to fluctuate within 0.5. Figure 3(a) shows that as the number of episodes increases, the oscillation interval gradually decreases. At this point, we can regard it as the reward value gradually converging. Figure 3(b) shows that the oscillation interval of the reward value shrinks rapidly, indicating that the decrease of the critic network learning rate can reduce the oscillation interval of the reward value and accelerate the convergence of the reward value. Coincidentally, we reduce the learning rate of the actor network and obtain the goal of rapid convergence of the reward value in Figure 3(c). It is important that due to the characteristic that the reward value oscillates in a certain range, we use the average value of the upper and lower limits of the oscillating range as the final result of the reward value.

Figure 4 shows the minimum total energy consumption of GME as the number of GME increases. In this figure, the number of UAVs is 3, the size of task is set as 8 MB, and the number of CPU cycles to compute each bit is set as 160 cycles/bit, i.e., $M = 3$, $L_k = 8$ MB, and $C_k = 160$ cycles/bit. For different offloading strategies, the total energy consumption of GME also increases linearly with the increase of GME. When the UAVs' coverage is low and the number of GME is small, it is difficult to satisfy that all GME is within the UAV signal coverage. In the figure, the total energy consumption of GME under the four strategies is not much different at $K = 5$. But it can be seen that under the same computing task requirements, the greater the number of GME, the greater the total energy consumption of GME, and the offloading strategy based on A3C algorithm proposed by us is more advantageous.



(a) Episode

(b) Learning rate of critic network

(c) Learning rate of actor network

Figure 3: Comparison of total energy consumption with different number of GME.

Figure 5 compares the proposed offloading strategy based on A3C algorithm with other strategies in terms of all GME energy consumption versus different sizes of computation task. In this figure, the number of UAVs is 3, the number of GME is set as 20, and the number of cycles to compute each bit is set as 160 cycles/bit, i.e., $M = 3, K = 20$, and $C_k = 160$ cycles/bit. It can be seen that with the increase of the computational task size, the energy consumption gap of the four offloading strategies gradually increases. The reason is that with the increase of the data scale, due to the limitation of the calculation delay, the random strategy and the greedy strategy gradually lose their effect. By analysing the linear trend of Random algorithm, Greedy algorithm, DQN algorithm, and A3C algorithm in the graph, we can see that the larger the amount of data, the clearer the advantage of our proposed offloading strategy.

Figure 6 describes the sum energy consumption of all GME corresponding to the number of CPU cycles required for different calculations per bit of task data. In this figure, the size of task is set as 8 MB, the number of UAVs is 3 and the number of GME is set as 20, i.e., $L_k = 8$ MB, $M = 3$, and $K = 20$. As shown in Figure 6, it is interesting to note that there is a significant gap between the offloading strategy based on the DRL algorithm and the offloading strategy based on the random algorithm and the greedy algorithm. The reason is that the larger the number of cycles for calculating each bit, the higher the calculation delay

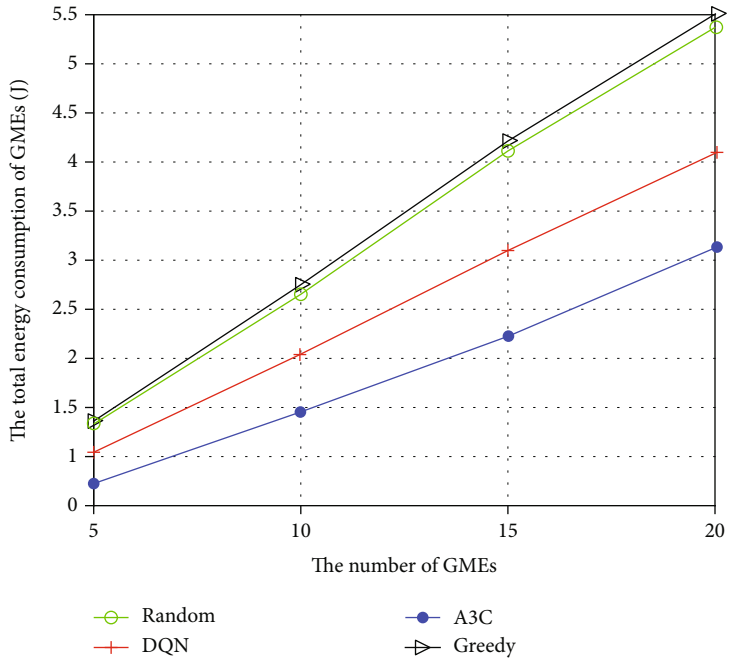


Figure 4: Comparison of total energy consumption with different number of GME.

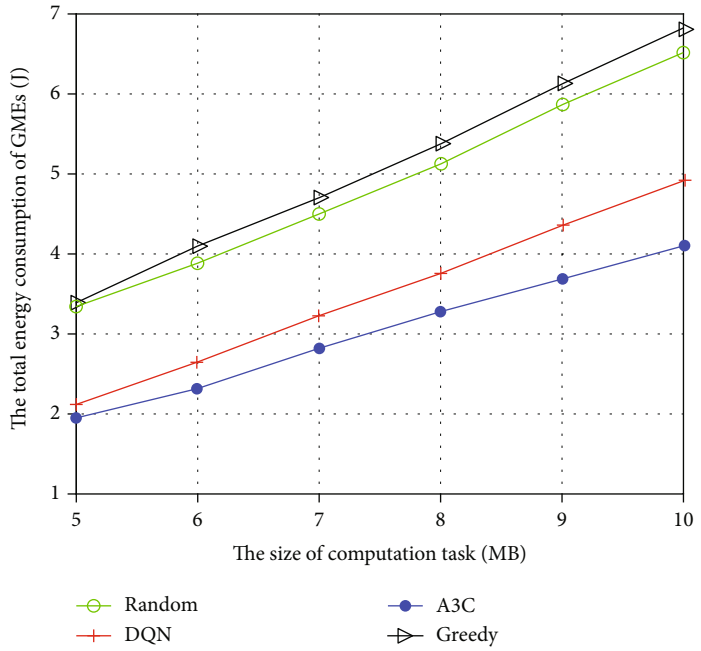


Figure 5: The total energy consumption of GME with different sizes of computation task.

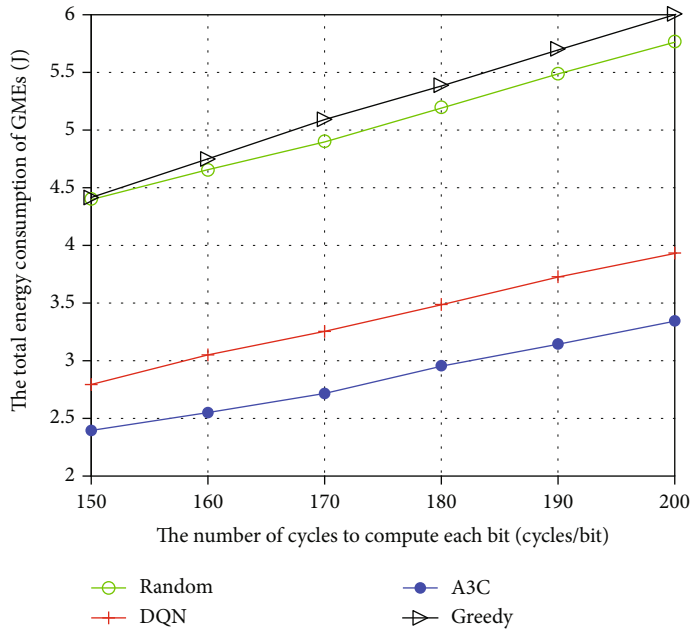


Figure 6: Performance comparison of different strategies versus the number of cycles to compute each bit.

requirements, and the traditional algorithms are difficult to meet such task offloading requirements.

5. Conclusions

In this paper, we researched the computational task offloading problem in an MUEMN and formulated a constrained optimization problem with the objective of minimizing the total energy consumption of all GME. We proposed a model-free DRL scheme with an asynchronous A3C algorithm to effectively generate offloading decisions. A large number of numerical results showed that the proposed A3C algorithm can accelerate the convergence speed of the algorithm and effectively reduce the total energy consumption of GME. In theory, the greater the number of UAVs, the task calculation delay can be greatly reduced, and the energy consumption of GME can also be reduced. However, too many UAVs can be a waste of resources when the limited space of application scenarios. In future work, we plan to study the optimal number of UAVs deployed in the MUEMN with limited space.

Data Availability

All the data used to support the findings of this study are included within the article.

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 62101277), by the Natural Science Foundation of Jiangsu Province (no. BK20200822), by the Natural Science Foundation of Jiangsu Higher Education Institutions of China (no. 20KJB510036), and by the open research fund of Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications) under grant JZNY202103, Ministry of Education.

References

[1] L. Zhao, G. Han, Z. Li, and L. Shu, "Intelligent digital twin-based software-defined vehicular networks," *IEEE Network*, vol. 34, no. 5, pp. 178–184, 2020.

[2] L. Zhao, W. Zhao, A. Y. Al-Dubai, G. Min, A. Y. Zomaya, and C. Gong, "Novel online sequential learning-based adaptive routing for edge software-defined vehicular networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 2991–3004, 2021.

[3] Z. Chang, L. Liu, X. Guo, and Q. Sheng, "Dynamic resource allocation and computation offloading for IoT fog computing system," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3348–3357, 2021.

[4] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—a key technology towards 5G," *ETSI White Paper*, vol. 11, no. 11, pp. 1–16, 2015.

[5] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*1.

[6] Z. Yang, C. Pan, K. Wang, and M. Shikh-Bahaei, "Energy efficient resource allocation in UAV-enabled mobile edge

computing networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4576–4589, 2019.

[7] L. Zhao, K. Yang, Z. Tan, X. Li, S. Sharma, and Z. Liu, "A novel cost optimization strategy for SDN-enabled UAV-assisted vehicular computation offloading," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3664–3674, 2021.

[8] Y. Ye, R. Q. Hu, G. Lu, and L. Shi, "Enhance latency-constrained computation in MEC networks using uplink NOMA," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2409–2425, 2020.

[9] F. Zhou, Y. Wu, R. Q. Hu, and Y. Qian, "Computation rate maximization in UAV-enabled wireless-powered mobile-edge computing systems," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 9, pp. 1927–1941, 2018.

[10] B. Li, Z. Fei, Y. Zhang, and M. Guizani, "Secure UAV communication networks over 5G," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 114–120, 2019.

[11] Y. Kawamoto, H. Nishiyama, N. Kato, F. Ono, and R. Miura, "Toward future unmanned aerial vehicle networks: architecture, resource allocation and field experiments," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 94–99, 2019.

[12] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: recent advances and future trends," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241–2263, 2019.

[13] Q. Tang, Z. Fei, B. Li, and Z. Han, "Computation offloading in LEO satellite networks with hybrid cloud and edge computing," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9164–9176, 2021.

[14] S. Shakoor, Z. Kaleem, D.-T. Do, O. A. Dobre, and A. Jamalipour, "Joint optimization of UAV 3-D placement and path-loss factor for energy-efficient maximal coverage," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9776–9786, 2021.

[15] L. Qian, Y. Wu, J. Ouyang, Z. Shi, B. Lin, and W. Jia, "Latency optimization for cellular assisted mobile edge computing via non-orthogonal multiple access," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5494–5507, 2020.

[16] Z. Wu, B. Li, Z. Fei, Z. Zheng, and Z. Han, "Energy-efficient robust computation offloading for fog-IoT systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4417–4425, 2020.

[17] H. Guo and J. Liu, "UAV-enhanced intelligent offloading for internet of things at the edge," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2737–2746, 2020.

[18] Q. Liu, L. Shi, L. Sun, J. Li, M. Ding, and F. Shu, "Path planning for UAV-mounted mobile edge computing with deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5723–5728, 2020.

[19] Y. Wang, Z.-Y. Ru, K. Wang, and P.-Q. Huang, "Joint deployment and task scheduling optimization for large-scale mobile users in multi-UAV-enabled mobile edge computing," *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3984–3997, 2020.

[20] B. Shang and L. Liu, "Mobile-edge computing in the sky: energy optimization for air-ground integrated networks," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7443–7456, 2020.

[21] F. Guo, H. Zhang, H. Ji, X. Li, and V. C. Leung, "Joint trajectory and computation offloading optimization for UAV-assisted MEC with NOMA," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, Paris, France, Apr-May 2019.

[22] S. Batabyal and P. Bhaumik, "Mobility models, traces and impact of mobility on opportunistic routing algorithms: a survey," *IEEE Communication Surveys and Tutorials*, vol. 17, no. 3, pp. 1679–1707, 2015.

[23] L. Shi, Z. Jiang, and S. Xu, "Throughput-aware path planning for UAVs in D2D 5G networks," *Ad Hoc Networks*, vol. 116, p. 102427, 2021.

[24] P. M. Pardalos and S. A. Vavasis, "Quadratic programming with one negative eigenvalue is NP-hard," *Journal of Global Optimization*, vol. 1, no. 1, pp. 15–22, 1991.

[25] W. Chen, X. Qiu, T. Cai, H.-N. Dai, Z. Zheng, and Y. Zhang, "Deep reinforcement learning for internet of things: a comprehensive survey," *IEEE Communication Surveys and Tutorials*, vol. 23, no. 3, pp. 1659–1692, 2021.

[26] N. Cheng, F. Lyu, W. Quan et al., "Space/aerial-assisted computing offloading for IoT applications: a learning-based approach," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 5, pp. 1117–1129, 2019.

[27] X. Chen, H. Zhang, C. Wu, S. Mao, Y. Ji, and M. Bennis, "Optimized computation offloading performance in virtual edge computing systems via deep reinforcement learning," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4005–4018, 2019.

[28] Z. Chen and X. Wang, "Decentralized computation offloading for multi-user mobile edge computing: a deep reinforcement learning approach," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, 21 pages, 2020.

[29] C. Di Franco and G. Buttazzo, "Energy-aware coverage path planning of UAVs," in *2015 IEEE International Conference on Autonomous Robot Systems and Competitions*, pp. 111–117, Vila Real, Portugal, Apr 2015.

[30] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 829–846, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1442745, 20 pages
https://doi.org/10.1155/2022/1442745

Research Article

A Robust Image Segmentation Framework Based on Nonlocal Total Variation Spectral Transform

Jianwei Zhang,[1] **Yue Shen**,[1] **Zhaohui Zheng**,[2] **and Le Sun**[3]

[1]*School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China*
[2]*Department of Clinical Immunology, Xijing Hospital, Fourth Military Medical University, No. 127 Changle West Rd., Xi'an 710032, China*
[3]*Jiangsu Engineering Center of Network Monitoring, School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China*

Correspondence should be addressed to Zhaohui Zheng; zhengzh@fmmu.edu.cn

Received 14 September 2021; Revised 27 November 2021; Accepted 22 January 2022; Published 24 February 2022

Academic Editor: Zheng Chu

Image segmentation plays an important role in various computer vision tasks. Nevertheless, noise always inevitably appears in images and brings a big challenge to image segmentation. To handle the problem, we study the nonlocal total variation (NLTV) spectral theory and build up an image segmentation framework with NLTV spectral transform to segment images with noise. Firstly, we decompose an image into the NLTV flow in the NLTV spectral transform, with which the max response time of each pixel is calculated. Secondly, a separation surface is constructed with the max response time to distinguish the objects and preserve the structure details in segmentation. Thirdly, the image is filtered by the surface in the NLTV spectral domain, and a rough segmentation result is obtained by means of an inverse transform. Finally, we use a binary process and morphological operations to refine the segmentation result. Experiments illustrate that our method can preserve edge structures effectively and has the ability to achieve competitive segmentation performance compared with the state-of-the-art approaches.

1. Introduction

Image segmentation refers to partitioning images into multiple homogeneous parts or objects. It plays a significant role in a broad range of computer vision applications, including scene understanding [1], image compression [2], and image retrieval [3, 4]. To date, two categories of segmentation methods have been widely proposed: data-driven methods [5–7] and model-driven methods [8–28].

Among data-driven methods, the common strategy is to extract the semantic features of images using deep convolutional neural networks, based on which each pixel can obtain a semantic label to realize segmentation. The popular deep neural networks for semantic segmentation consist of FCN [5], U-Net [6], SegNet [7], etc., which can obtain satisfying segmentation results without any postprocess techniques. However, deep neural networks often suffer from high computational resource consumption and need a great mass of labeled data. Moreover, the interpretability of neural networks is always an Achilles' heel. Therefore, model-driven methods are our research centrality.

According to different segmentation strategies, model-driven methods can be further categorized as boundary-based methods, region-based methods, hybrid methods, and transform-based methods. Boundary-based methods separate objects from the background by edge or shape. The representative methods include edge detection [8–10] and graph-cut methods [11, 12]. The former uses intensity discontinuity to segment an object. Common edge detection operators contain Prewitt [8], Sobel [9], Roberts [9], and Canny [10]. Compared with the edge detection approaches, graph-cut-based methods can achieve better segmentation accuracy. Nonetheless, the extraction of gradients is sensitive to noise, which makes the boundary-based models produce unsatisfying segmentation results for noisy images.

Region-based approaches recognize similar regions and complete segmentation by means of statistical techniques. The Chan-Vese model [13] and FCM [14] are representative

works. The Chan–Vese model makes the contour curve close to the object boundary by minimizing the energy on both sides of the evolution curve [15]. Nevertheless, the Chan–Vese model fails to obtain satisfying results because of the intensity inhomogeneity. FCM improves the tolerance to ambiguity and obtains more reasonable segmentation results by introducing a membership matrix. However, FCM is unrobust to noise because of the fact that it merely considers gray-level information. To solve the problem, many variants of FCM [16–19] have been developed, which bring good segmentation performance. Nonetheless, the improved methods are still sensitive to the complex background and intensity inhomogeneity.

Hybrid methods employ boundary information to detect the region of objects and then use region information to preserve the boundary structures. Recently, transition region (TR)-based image thresholding [20–23] has been proposed as a type of hybrid method. The method, firstly, uses edge detectors or statistical techniques to extract a transition region, which is a structure similar to the image edge, and then, it segments the image by a threshold, which is a gray level mean value of the transition region. TR-based image thresholding additionally exploits the spatial information to acquire more satisfying segmentation results. However, it is a global thresholding method, which is unrobust to intensity inhomogeneity.

The aforementioned model-driven methods segment the image using spatial features, which results in sensitivity to noise. Differently, transform-based approaches, firstly, transform the image to a specific domain according to mathematical theories, where noise and image details have different performances. Then, denoised images are obtained by filtering and inverse transformation, on which post-processing is performed to segment the image. As one of the popular transform approaches, wavelet transform is widely used in diverse computer vision tasks because of its ease of use and multiresolution processing ability. The common operation of the wavelet transform in image processing is to decompose the image to obtain multiscale sub-bands in the wavelet domain with the help of Mallat's pyramid algorithm [24]. Then, filter the image by low-pass, band-pass, or high-pass filter to obtain the required features. Finally, the processed image can be obtained by inverse transform. To get satisfying segmentation results, wavelet transform is often combined with other segmentation methods, such as watershed segmentation [25], clustering approaches [26], and image thresholding methods [27]. For instance, the method in [25], firstly, decomposes the original image into a multiscale pyramid representation in the wavelet transform domain. Secondly, the watershed algorithm is applied to segment every image of the multiscale pyramid into several regions, including objects and background. Thirdly, the reverse wavelet transform is conducted on the split regions to get the next higher resolution representation. Finally, the size of split regions gradually becomes the same as that of regions in the ground truth to achieve the segmentation result. Nonetheless, wavelet transform-based methods are sensitive to contrast, and the segmentation results are influenced by the selection of wavelet basis functions.

Recently, the NLTV spectral theory has been introduced [28] and has attracted people's attention. The NLTV spectral transform can transform the image from the spatial domain to the spectral domain, in which objects with different contrast, size, and detailed structures can be distinguished well. Additionally, the NLTV spectral transform can preserve image structures because of its nonlocal operators [28]. To this end, we further discuss the performance of NLTV spectral theory and attempt to further enhance the applicability of the NLTV spectral transform. Inspired by the work [29], we demonstrate the sensitivity of the NLTV spectral transform to size, contrast, and its detailed structures in images with or without noise. We also indicate that the spectral transform is invariance to rotation and translation. Besides, we are motivated to put forward a robust image segmentation framework with NLTV spectral transform. The main process is as follows: firstly, the NLTV flow is imposed on an image to acquire the NLTV spectral transform, by which spectral response and a salient time map of the image are calculated. The elements in the salient time map represent the max response time of each pixel of the image. Secondly, we filter the salient time map by a Gaussian filter to remove the isolated points and perform a least-squares regression using a polynomial on the filtered map to fit a separation surface. Thirdly, the image is filtered by the surface in the NLTV spectral domain, followed by the NLTV inverse transform to obtain a rough segmentation result. Finally, we use morphological operators and a binary process to refine the segmentation result.

It should be noticed that the total variation (TV) spectral transform-based method [30] has a similar idea in segmenting images with noise. However, the TV spectral transform used in [30] calculates the horizontal and vertical gradient of every pixel, which means only local information is selected to describe object features. In reference [30], the TV flow is obtained by iteratively solving the ROF model, and then the TV spectral transform is yielded. Considering that the edge detail of objects is lost for solving the ROF model, the guided filter is adopted to refine the object edge in [30]. In contrast to the spectral transform strategy in [30], our method pays more attention to the difference between one pixel and all other pixels in the image, termed nonlocal gradients, to achieve NLTV spectral transform. With the nonlocal information, the edge details can be effectively preserved when segmenting the object in a variety of noises. In addition, our segmentation framework does not introduce the guided filter, which may bring the noise from the original image to the segmentation result. We perform the experiments on synthetic, natural, and medical cell images, which demonstrate that the proposed method can achieve competitive segmentation performance compared with the state-of-the-art methods.

Overall, the contributions of this work are twofold, which are as follows:

(i) We illustrate the properties of NLTV spectral transform by theoretical proof and experiments. The analysis demonstrates that objects with varying size, contrast, and detailed structures can be

distinguished in the NLTV spectral domain. Additionally, the transform is invariant to rotation and translation. These properties indicate the feasibility of segmentation based on NLTV spectral transform.

(ii) We propose an image segmentation framework using NLTV spectral transform, which fits a separation surface to filter sub-bands in the NLTV spectral domain, and it obtains segmentation results by means of postprocessing. Our method can achieve satisfying results for images with diverse noise or complex texture.

The rest of the article is structured as follows: section 2 gives an overview of the NLTV spectral theory. Section 3 discusses the properties of NLTV transform and introduces our segmentation framework based on NLTV spectral transform. Section 4 illustrates the experimental results of the proposed method. At last, the paper is concluded in section 5.

2. Preliminaries

This section introduces the NLTV spectral transform framework [28]. The framework is made of several parts: nonlocal operators, NLTV flow, NLTV spectral transform, and spectral response.

2.1. Nonlocal Operators. According to continuous definitions on the graphs of nonlocal gradient and divergence [31], three nonlocal operators, namely nonlocal derivatives, nonlocal gradients, and nonlocal divergences, are defined as follows:

Let $\Omega \subset R^2$ be a bounded domain and $w(X,Y) \geq 0$ be non-negative weights between any two points, $X, Y \in \Omega$. In the view of graphs, these weights correspond to a certain relationship between these points. For simplicity, we assume that these weights are symmetric, which means $w(X,Y) = w(Y,X)$. Then, Gilboa and Osher [28] extended the local derivative to a nonlocal version by the following definition:

$$\partial_Y u(X) = (u(Y) - u(X))\sqrt{w(X,Y)}, \quad X, Y \in \Omega, \tag{1}$$

where $u(X)$ is a real function, u: $\Omega \longrightarrow R, 0 < w(X,Y) < \infty$, and $\partial_Y u(X)$ represents the partial derivatives of $u(X)$ in the direction of point X and Y.

Similar to local gradients derived from local partial derivatives, nonlocal gradient $\nabla_w u(X)$: $\Omega \longrightarrow \Omega \times \Omega$ is defined as the vector composed of all partial derivatives.

$$(\nabla_w u)(X,Y) = (u(Y) - u(X))\sqrt{w(X,Y)}, \quad X, Y \in \Omega. \tag{2}$$

Before introducing nonlocal divergence, the definition of inner product for vectors is shown as below. Denoting vectors as $\overrightarrow{\mathbf{v}_1} = v_1(X,Y)$, $\overrightarrow{\mathbf{v}_2} = v_2(X,Y) \in \Omega \times \Omega$, an inner product is defined as follows:

$$\langle \overrightarrow{\mathbf{v}_1}, \overrightarrow{\mathbf{v}_2} \rangle = \int_\Omega v_1(X,Y) v_2(X,Y) \mathrm{d}Y. \tag{3}$$

Then nonlocal divergence $(\mathrm{div}_w \overrightarrow{\mathbf{v}})(X)$: $\Omega \times \Omega \longrightarrow \Omega$ is defined as the adjoint of nonlocal gradient. $(\mathrm{div}_w \overrightarrow{\mathbf{v}})(X) = \int_\Omega (v(X,Y) - v(Y,X))\sqrt{w(X,Y)}\mathrm{d}Y$.

2.2. NLTV Flow. The weight matrix $\mathbf{W}$ depends on the patch similarity. For fixed point X and arbitrary point Y in the image, $\mathbf{W}(X,Y)$ represents the weight between the points X and Y, which is defined as follows:

$$\begin{cases} \mathbf{W}(X,Y) = \dfrac{E(X,Y)}{\sum_{Y \in \Omega} E(X,Y)}, \\ E(X,Y) = \exp\left(-\dfrac{\|P(X) - P(Y)\|_2^2}{\sigma^2}\right), \end{cases} \tag{4}$$

where $P(X)$ and $P(Y)$ represent the patches centered at points X and Y in the image, respectively. σ is a parameter to control the decay of the exponential function. $E(X,Y)$ describes the similarity between the points X and Y.

NLTV is divided into two types, including isotropic NLTV and anisotropic NLTV. The former is defined as follows:

$$J_{\mathrm{ISO-NLTV}}(u) = \int_\Omega \left(\int_\Omega (u(X) - u(Y))^2 w(X,Y) \mathrm{d}Y \right)^{1/2} \mathrm{d}X. \tag{5}$$

The latter is defined as follows:

$$J_{\mathrm{ANISO-NLTV}}(u) = \int_{\Omega \times \Omega} |u(X) - u(Y)| \sqrt{w(X,Y)} \mathrm{d}Y \mathrm{d}X. \tag{6}$$

In our work, the anisotropic nonlocal TV is applied to calculate NLTV flow.

$$\begin{cases} -\dfrac{\partial u}{\partial t} \in \partial_u J_{\mathrm{NLTV}}(u), \\ u(0,X) = u(X). \end{cases} \tag{7}$$

2.3. NLTV Transform. The sine and cosine functions are the basic functions in Fourier transform. These basic functions' amplitude forms impulses in the Fourier domain. The work [28] generalized this to NLTV domain. By examining the elementary structures disks for NLTV functional, the second derivative in the time of NLTV flow is considered the representation of the impulse of the elementary structure. Hence, the NLTV transform is defined by the following:

$$\phi(t) = u_{tt} t, \tag{8}$$

where $t \in (0, \infty)$ is a time parameter of the NLTV flow equation (7), and u_{tt} is the second derivative in the time of the NLTV flow.

For NLTV transform, the inverse transform reconstructs a signal or image from all $\phi(t)$ elements.

$$I(X) = \int_0^{\infty} \phi(t, X)\mathrm{d}t + \overline{u}, \tag{9}$$

where $\overline{u} = (1/\Omega)\int_\Omega u(X)\mathrm{d}X$ is the residual part of NLTV transform, and it is also the mean value of the initial condition.

2.4. NLTV Spectral Response. Corresponding to the amplitude of the response in Fourier domain, the NLTV spectral response is defined as follows:

$$S(t) = \int_\Omega |\phi(t, X)|\mathrm{d}X, \quad t \in (0, \infty). \tag{10}$$

The NLTV spectral response can roughly measure the importance of image information at different time scales in the NLTV spectral domain [28]. The main features of the image emerge at the time scale corresponding to the high response. Otherwise, the NLTV spectral transform could be considered negligible.

3. Proposed Method

This section discusses the properties of the NLTV spectral transform and displays a segmentation method for images with noise using the NLTV spectral transform. Firstly, the seminal works [29, 30], which demonstrate the properties of TV spectral transform in images with or without noise, are extended to the NLTV spectral transform in motivation. Secondly, a segmentation method using NLTV spectral transform for images with noise is introduced.

3.1. Motivation. The section tries to research the properties of NLTV spectrum transform in images with or without noise. Theories and experiments without noise are shown, firstly. Then, the properties are extended to the noise condition by experiments. As known to all, the typical noises in digital images are additive noise, multiplicative noise, and impulse noise. For this reason, we corrupt the images with Gaussian noise, Salt & Pepper noise, and Speckle noise.

3.1.1. Property 1: Sensitivity to Size. A short proof about the property is provided. For the sake of simplicity, we consider scaling with a gray level image $f(X)$, where $X = (x, y) \in \Omega$. Then, the image after scaling can be denoted as $f(aX)$. With the above notations, we explore why NLTV spectral transform values over the time scale of images before and after scaling satisfy the following relationship:

$$\widetilde{\phi}(t, X) = a\phi(at, aX), \tag{11}$$

where $\widetilde{\phi}(t, X)$ and $\phi(t, X)$ are NLTV spectral transforms corresponding to images before and after scaling, respectively. Notice that for the original image $f(X)$, the NLTV flow can be derived from the following partial differential equation:

$$\begin{cases} -\dfrac{\partial u}{\partial t} \in \partial_u J_{\mathrm{NL-TV}}(u), \\ \\ u(0, X) = f(X). \end{cases} \tag{12}$$

Inspired by the case of TV, we consider the elementary structures called nonlocal disks for the image $f(X)$. A set $\mathbf{A}$ can be used as a nonlocal disk when two conditions are satisfied [28]: 1) $\mathbf{A}$ is a nonlocal calibrable set. 2) The curvature is constant on the internal boundary of the set $\mathbf{A}$.

The characteristic function of $\mathbf{A}$ is $\chi_{\mathbf{A}}(X) = \begin{cases} 1, & X \in \mathbf{A} \\ 0, & X \notin \mathbf{A} \end{cases}$. Then, the explicit solution of problem (12) with $u(0, X) = \chi_{\mathbf{A}}(X)$ is expressed as follows:

$$u(t, X) = \begin{cases} (1 - t\lambda_{\mathbf{A}})\chi_{\mathbf{A}}(X), & X \in \mathbf{A}, \\ 0, & X \notin \mathbf{A}, \end{cases} \tag{13}$$

where $\lambda_{\mathbf{A}} = (\mathrm{Per}(\mathbf{A})/|\mathbf{A}|)$ and $\mathrm{Per}(\mathbf{A})$ and $|\mathbf{A}|$ are, respectively, perimeter and normal of $\mathbf{A}$. In the same way, the NLTV flow of nonlocal disk $\mathbf{A}'$ for the image $f(aX)$ is as follows:

$$\widetilde{u}(\widetilde{t}, \widetilde{X}) = \begin{cases} (1 - \widetilde{t}\lambda_{\mathbf{A}'})\chi_{\mathbf{A}'}(\widetilde{X}), & \widetilde{X} \in \mathbf{A}', \\ 0, & \widetilde{X} \notin \mathbf{A}'. \end{cases} \tag{14}$$

The energy of points in the image $f(X)$ and $f(aX)$ decreases with the average speed of $\lambda_{\mathbf{A}}$ and $\lambda_{\mathbf{A}'}$, respectively. It is worth noting that $\lambda_{\mathbf{A}}$ is equal to $\lambda_{\mathbf{A}'}$ because the object patterns before and after scaling are similar. Hence, we have $\widetilde{u}(\widetilde{t}, \widetilde{X}) = u(t, X)$ with $\widetilde{t} = t/a$ and $\widetilde{X} = X/a$. For time scaling is $\widetilde{t} = t/a$, $\widetilde{u}_{\widetilde{t}\,\widetilde{t}}(\widetilde{t}, \widetilde{X}) = a^2 u_{tt}(t, X)$ is obvious. Therefore, $\widetilde{\phi}(t/a, X/a) = t\widetilde{u}_{\widetilde{t}\,\widetilde{t}}(\widetilde{t}, \widetilde{X})/a = t \cdot a^2 u_{tt}(t, X)/a = a\phi(t, X)$.

Figure 1 is an example showing how the NLTV spectral transform separates different size objects. The multiscale NLTV spectral descriptions of the pixels are shown in Figure 1(b), which shows that there is a positive correlation between the size and the time to reach the max spectral response. In addition, we can find that the disappearance order of objects in Figure 1(c) is consistent with the order of reaching max spectral response time in Figure 1(b). Figure 1(d) shows the visualization of subbands in the NLTV spectral domain, and it is a more intuitive interpretation of figure 1(b). Moreover, Figure 2 shows the sensitivity of NLTV spectral transform to size and similar performance in different noises.

3.1.2. Property 2: Sensitivity to Local Contrast. Combing the work [29], we attempt to provide a short proof. The image after gray-scale transformation by factor a is denoted as $af(X)$. Then, we plan to prove that the NLTV spectral signatures of $f(X)$ and $af(X)$ satisfy the following relationship:

$$\widetilde{\phi}(at, X) = \phi(t, X). \tag{15}$$

It is noting that $\phi(t, X)$ is still related with characteristic function $\chi_{\mathbf{A}}(X)$ mentioned in property 1. Copying the analysis of property 1, the NLTV flows of $f(X)$ and $f(aX)$ are as follows:

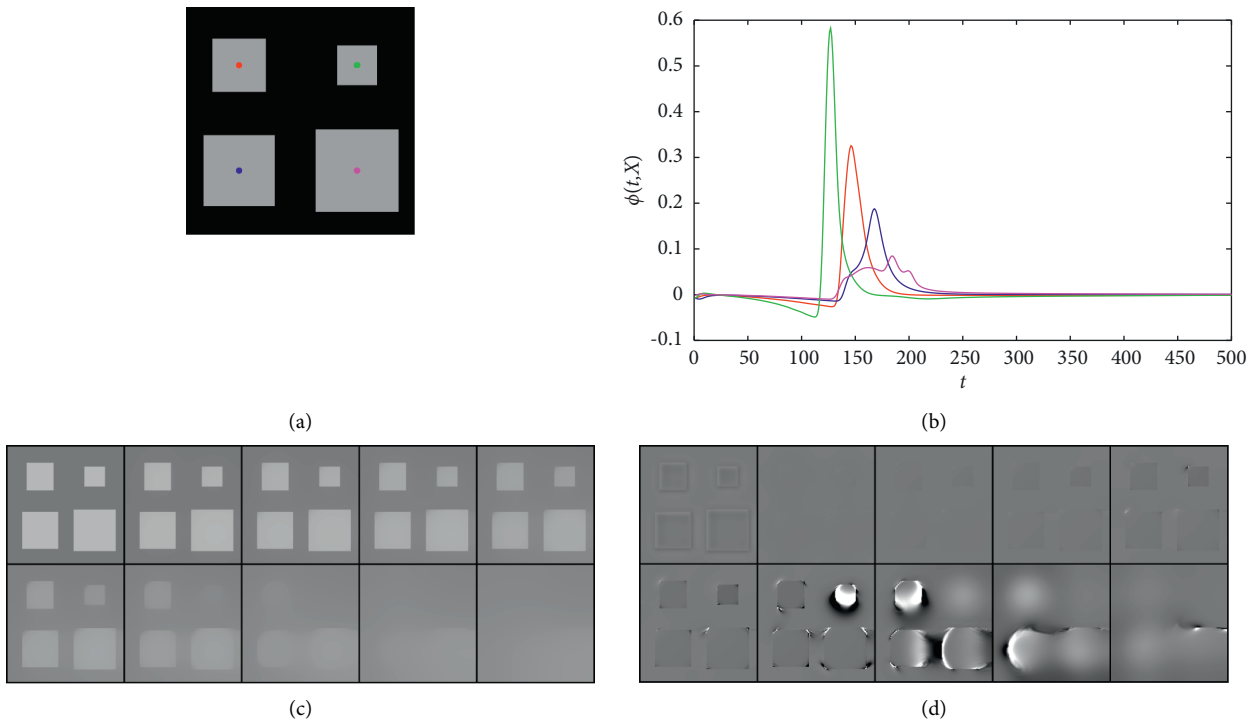


FIGURE 1: Demonstration of property 1. Signatures are distinguished because of their sensitivity to size. (a) Image f. (b) Multiscale NLTV spectral descriptions of different pixels. (c) Results of NLTV flow of f. (d) Multiscale NLTV spectral components.

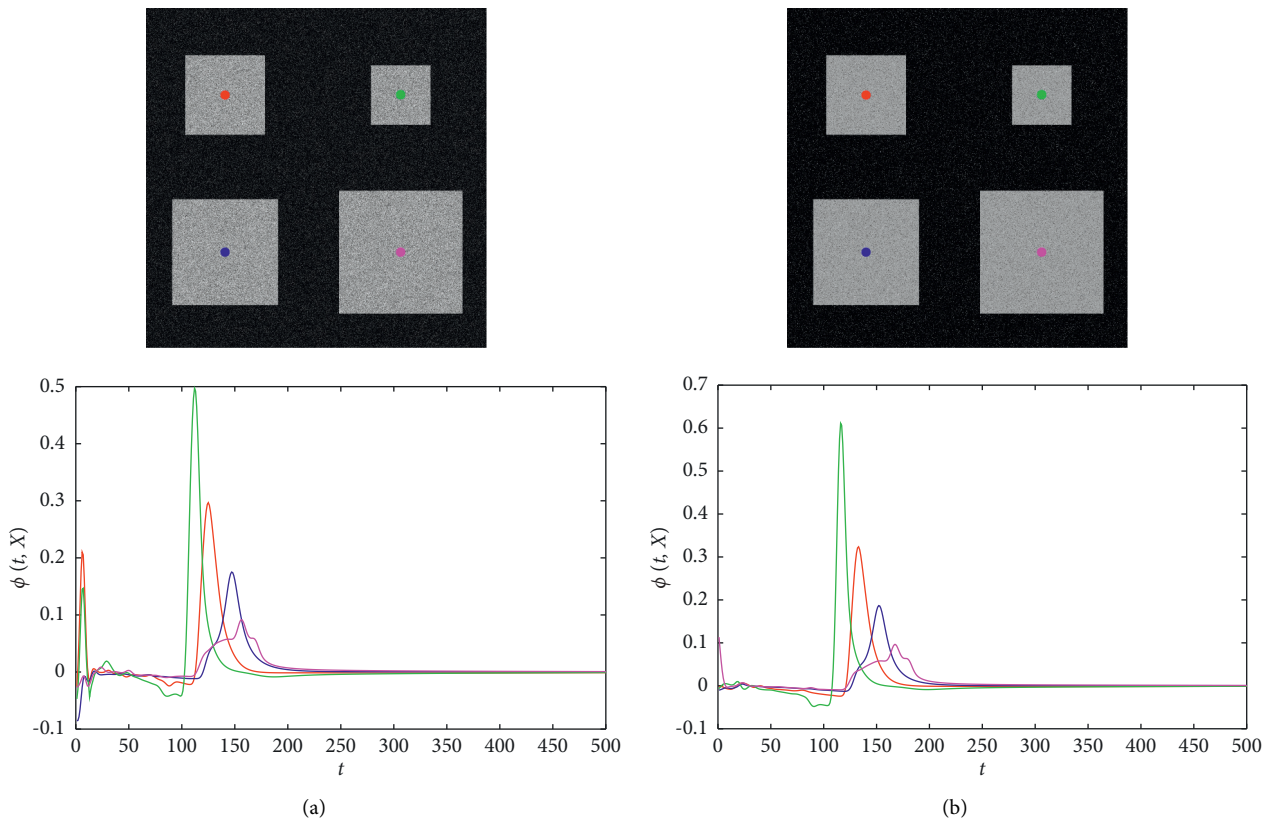
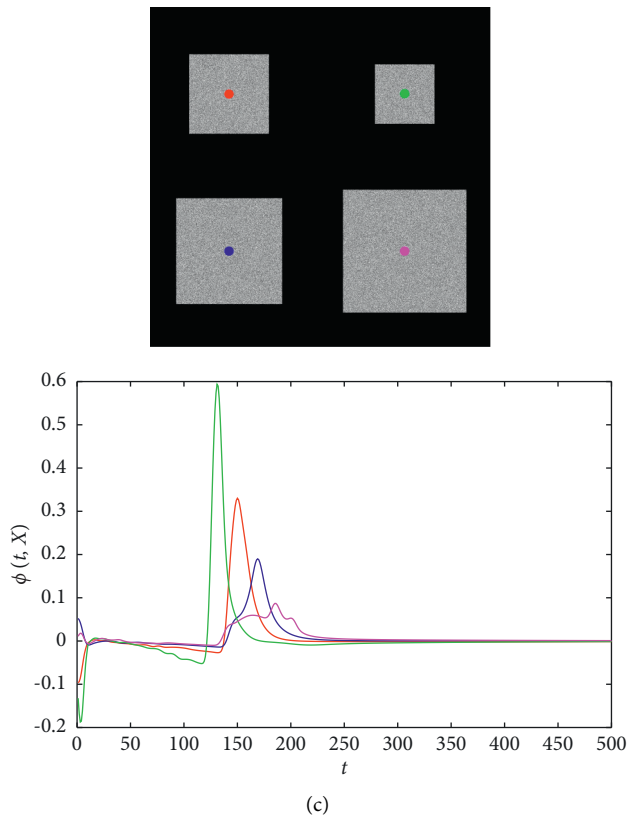


FIGURE 2: Continued.



(c)

FIGURE 2: NLTV spectral transform on different size objects corrupted with different noises. (a) is the image corrupted with Gaussian (10% variance), Salt & Pepper (10% density), and Speckle noise (10% variance), respectively. (b) is the multiscale NLTV spectral descriptions of different pixels corrupted with noises. (a) Gaussian noise. (b) Salt & Pepper noise. (c) Speckle noise.

$$u(t, X) = \begin{cases} (1 - t\lambda_{\mathbf{A}})\chi_{\mathbf{A}}(X), & X \in \mathbf{A}, \\ 0, & X \notin \mathbf{A}, \end{cases}$$
$$\widetilde{u}(\widetilde{t}, \widetilde{X}) = \begin{cases} a(1 - \widetilde{t}\lambda_{\mathbf{A}'})\chi_{\mathbf{A}'}(\widetilde{X}), & \widetilde{X} \in \mathbf{A}', \\ 0, & \widetilde{X} \notin \mathbf{A}', \end{cases} \tag{16}$$

where $\widetilde{t} = at$, $\mathbf{A} = \mathbf{A}'$, and $X = \widetilde{X}$. $\widetilde{u}(\widetilde{t}, \widetilde{X}) = au(t, X)$ and $\widetilde{u}_{\widetilde{t}\,\widetilde{t}}(a\widetilde{t}, \widetilde{X}) = u_{tt}(t, X)/a$ are given. Therefore, $\widetilde{\phi}(at, X) = at\widetilde{u}_{\widetilde{t}\,\widetilde{t}}(\widetilde{t}, \widetilde{X}) = at \cdot u_{tt}(t, X)/a = \phi(t, X)$.

An example is demonstrated on a synthetic image without noise, as shown in Figure 3. The image exhibited in figure 3(a) contains four different contrast squares with a black background. The NLTV spectral transform is calculated, and multiscale NLTV spectral descriptions of different pixels are shown in Figure 3(b). Figures 3(c) and 3(d) show more intuitive performance, which indicates that the low contrast squares disappear first. In addition, the NLTV spectral transform is implemented on different noises to verify its performance. As shown in Figure 4, except for small time scales, the NLTV spectral description has a similar performance, which demonstrates the sensitivity of the NLTV spectral transform to contrast images with noise.

3.1.3. Property 3: Sensitivity to Detailed Structures. Figure 5 shows objects with diverse structures. Figure 5(b) shows that different objects have different time scales when reaching the max spectral response. Figures 5(c) and 5(d) show an intuitive description. The center square with high contrast is decomposed, firstly. Then, the square ring to which the blue point belongs starts to be decomposed. The black square ring is decomposed finally. The experiment indicates the sensitivity of the NLTV spectral transform to detailed structures. The phenomena are caused by the nonlinear property of the NLTV spectral transform. Assuming that images f and g make up the image h, the response of these images satisfies the following:

$$\phi_h \neq \phi_f + \phi_g. \tag{17}$$

To observe the decomposition process of NLTV spectral transform within noise, examples are carried out on different noises. Figure 6 shows the decomposition results of different pixels in diverse noises. It can be seen that, except for small time scales, the NLTV spectral description is similar to the case shown in figure 5(b). The experiments demonstrate that the NLTV spectral transform has a sensitivity to detailed structures.

3.1.4. Property 4: Invariance to Rotation and Translation. Suppose the original image is denoted as $f(X)$, $X \in \Omega$. Then, the image after rotation by angle θ about the origin is $f(\mathbf{R}X)$, where $\mathbf{R}$ is the rotation matrix.

$$\mathbf{R} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$
$$f(X) = f(\mathbf{R}X). \tag{18}$$

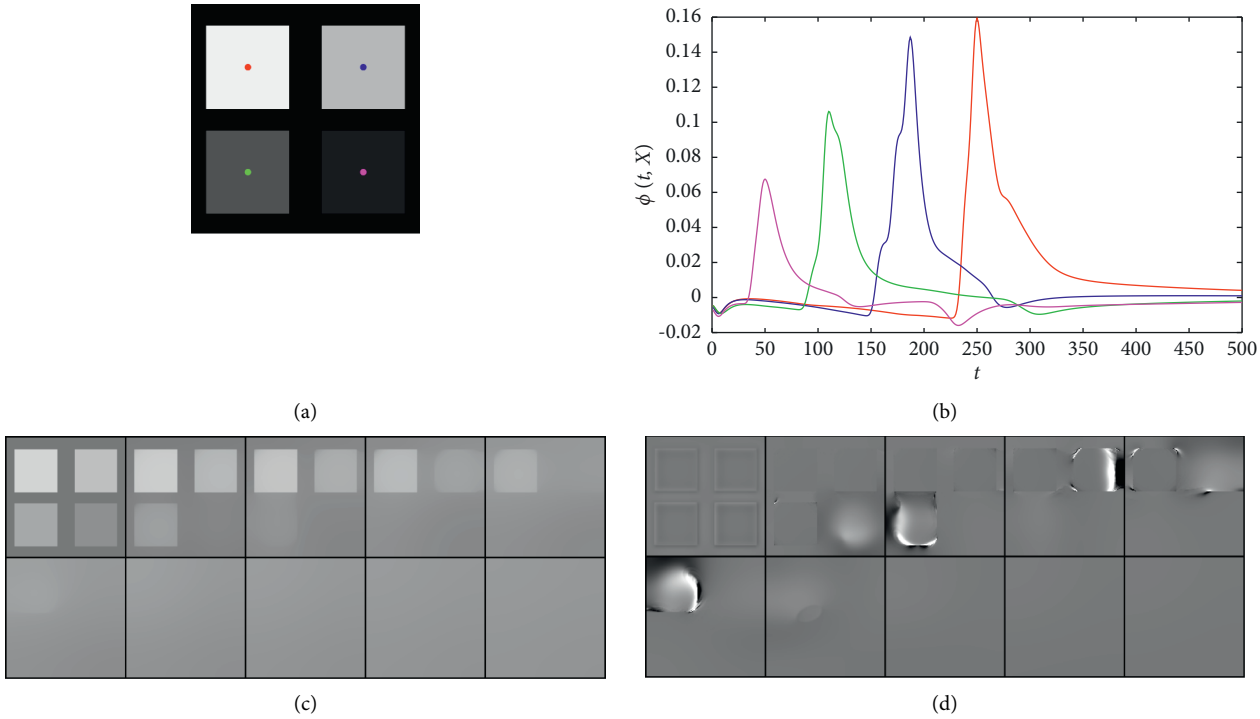


FIGURE 3: Demonstration of property 2. Signatures are distinguished because of their sensitivity to contrast. (a) Image f. (b) Multiscale NLTV spectral descriptions of different pixels. (c) Results of NLTV flow of f. (d) Multiscale NLTV spectral components.

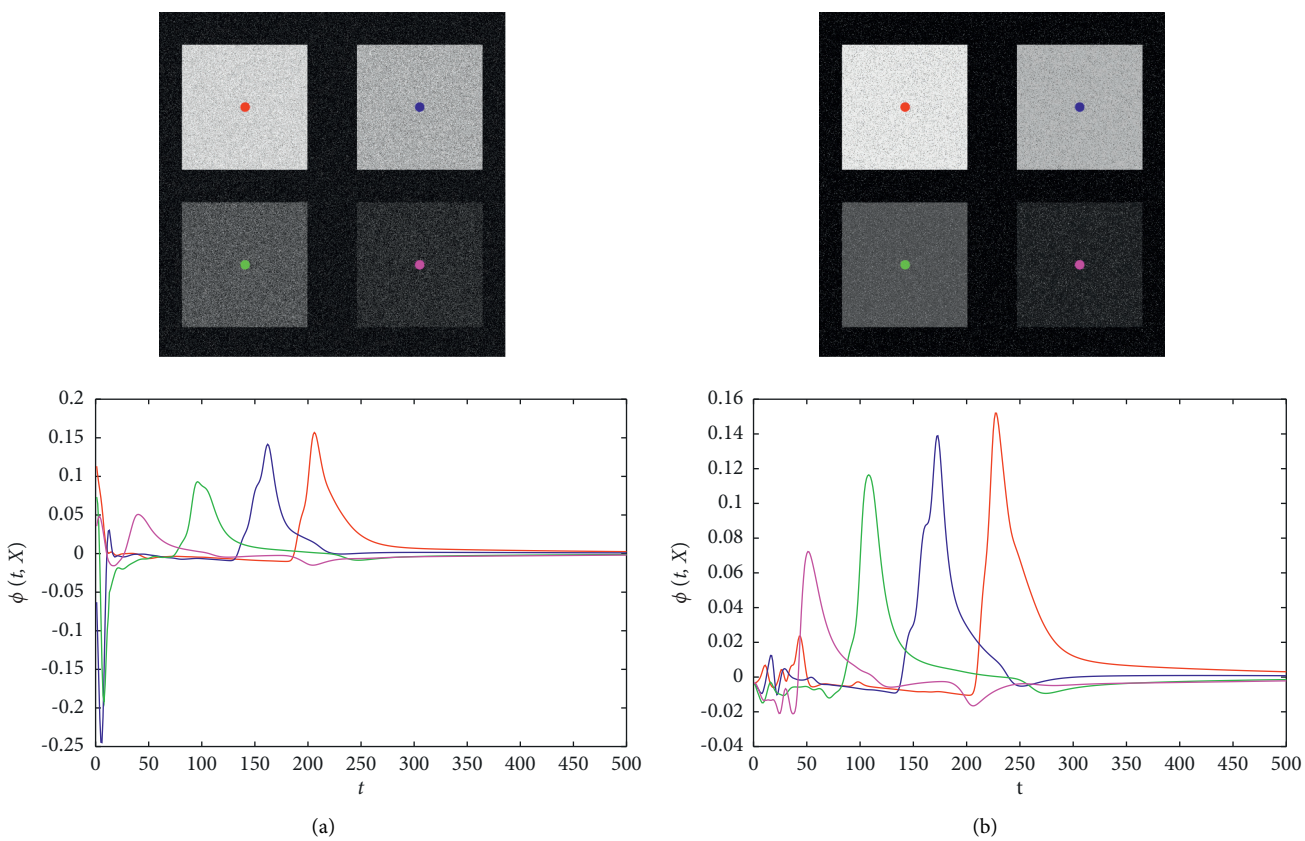
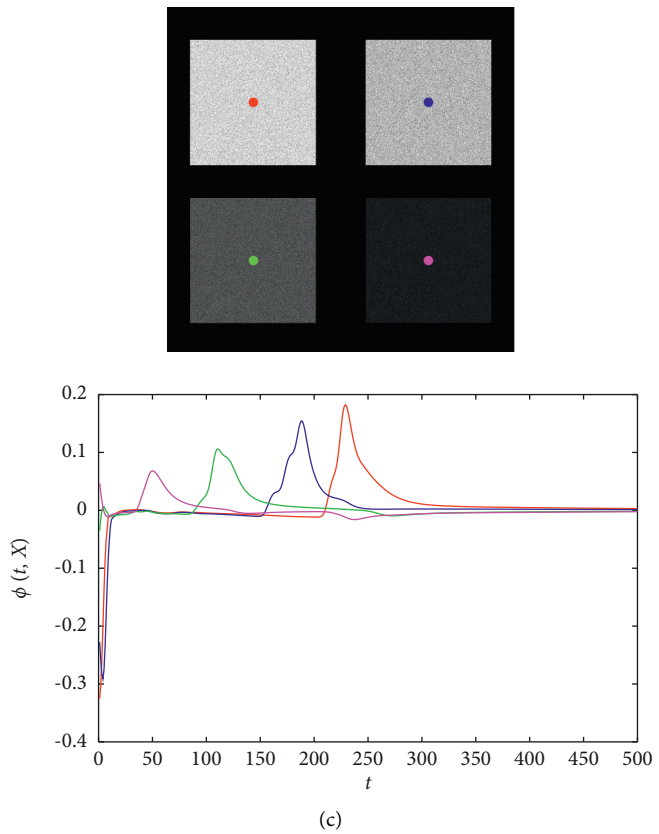
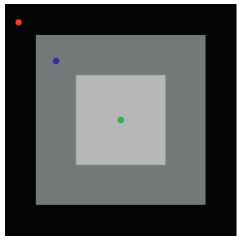


FIGURE 4: Continued.

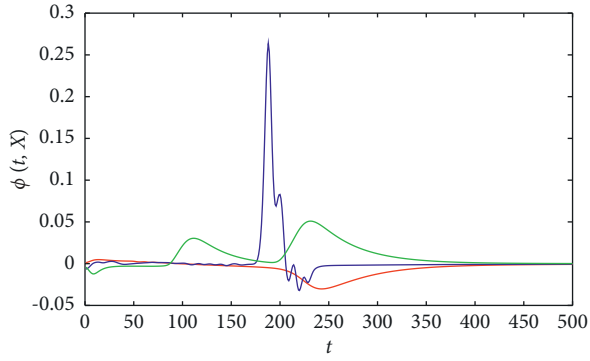


(c)

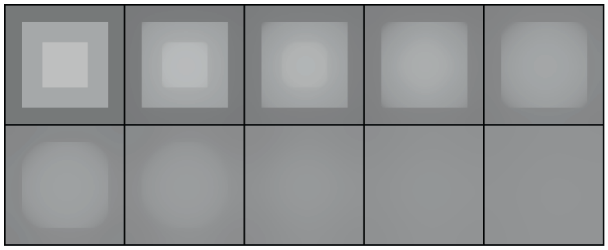
FIGURE 4: NLTV spectral transform on different local contrasts corrupted with different noises. (a) is the image corrupted with Gaussian (10% variance), Salt & Pepper (10% density), and Speckle noise (10% variance) respectively. (b) is the multiscale NLTV spectral descriptions of different pixels corrupted with noises. (a) Gaussian noise. (b) Salt & Pepper noise. (c) Speckle noise.



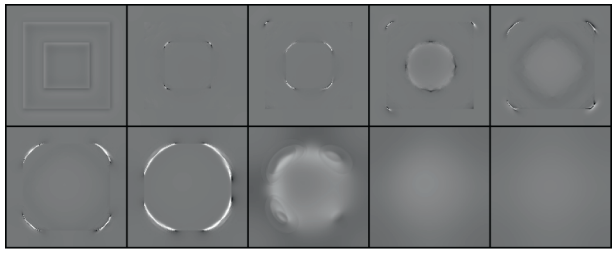
(a)



(b)



(c)



(d)

FIGURE 5: Demonstration of property 3. Signatures are distinguished because of their sensitivity to detailed structures. (a) Image f. (b) Multiscale NLTV spectral descriptions of different pixels. (c) Results of NLTV flow of f. (d) Multiscale NLTV spectral components.

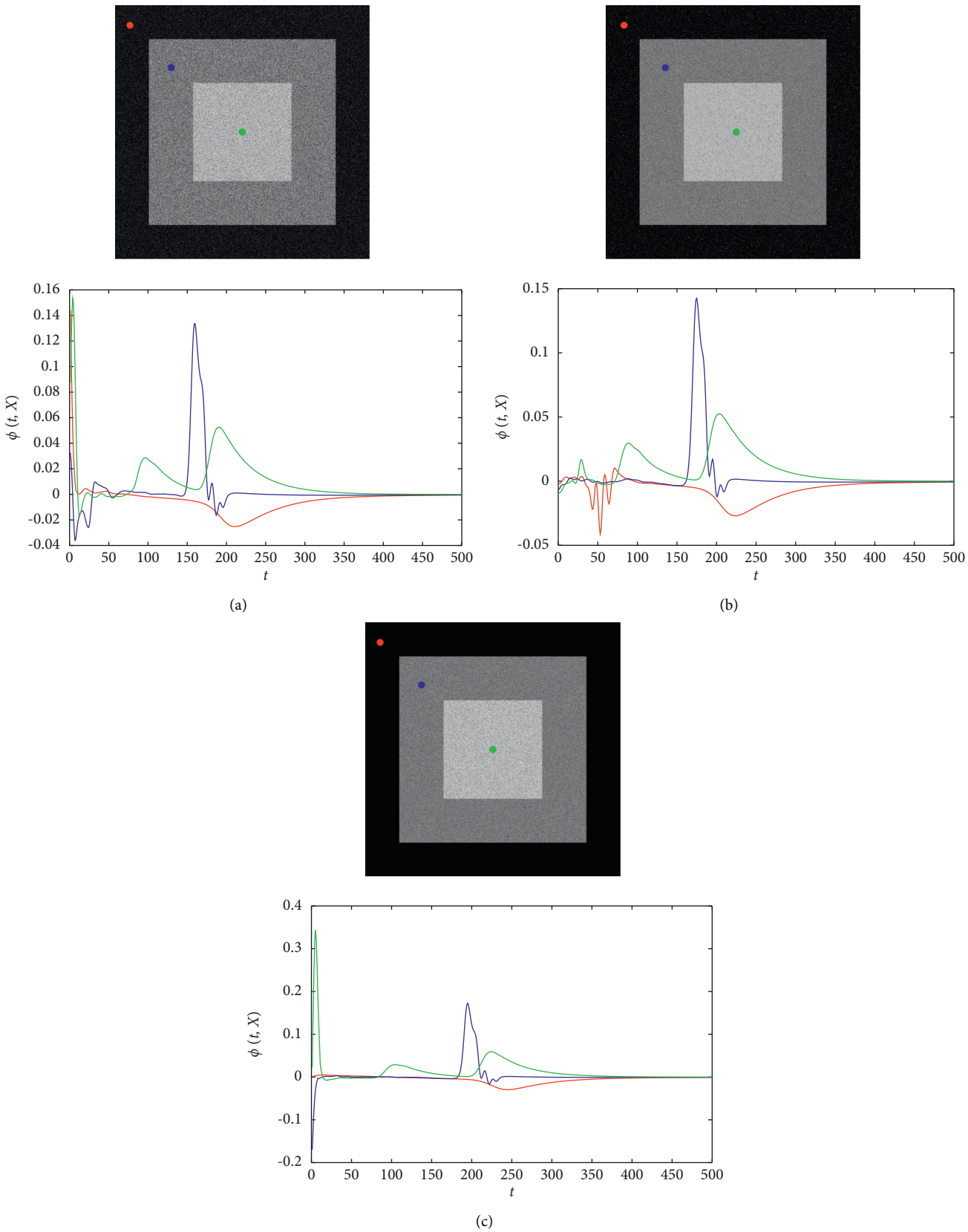


FIGURE 6: NLTV spectral transform on different structures corrupted with different noises. (a) is the image corrupted with Gaussian (10% variance), Salt & Pepper (10% density), and Speckle noise (10% variance), respectively. (b) is the multiscale NLTV spectral descriptions of different pixels corrupted with noises. (a) Gaussian noise. (b) Salt & Pepper noise. (c) Speckle noise.

Moreover, the image after translation by spatial shift on the original image is $f(X-d)$ and $f(X)=f(X-d)$. In essence, the rotation or translation of the image is equal to rotating or translating the coordinate system in the original image. On the other hand, the NLTV spectral transform is invariant to the coordinate system and sensitive to derivatives. Therefore, the NLTV spectral transform is invariant to rotation and translation, i.e.,

$$\begin{cases} \widetilde{\phi}_{f(\mathbf{R}X)}(t,X)=\phi_f(t,\mathbf{R}X), \\ \widetilde{\phi}_{f(X-a)}(t,X)=\phi_f(t,X-a). \end{cases} \tag{19}$$

There are three groups of objects with different shapes in figure 7(a). The objects in the same group have the same shape and contrast. Different objects have been translated in different positions and rotated at different angles. As figure 7(b) shows, the objects in the same group have a similar NLTV spectral description. More intuitive illustrations are displayed in figures 7(c) and 7(d), which present that the objects within the same group disappear simultaneously. Figure 8 shows the NLTV spectral descriptions of different pixels corrupted with noises. The bottom row of Figure 8 shows that the objects with the same shape have similar descriptions in large time scales, even though they have distinct rotations and translations.

3.2. NLTV Spectral Transform for Robust Image Segmentation

3.2.1. Overview of the Proposed Segmentation Flowchart. Figure 9 shows the flowchart of the proposed method. The method starts with the decomposition of an original image in the NLTV spectral domain. Then, the available information dimension of every pixel in the image increases from one to the number of time scales. To better get appropriate components, a soft threshold band-pass filter is selected to replace the traditional hard threshold band-pass filter. After obtaining the separation surface result, an inverse transform is used to get an abstract structure. The segmentation result is obtained with the help of the binary process and morphological operations.

3.2.2. NLTV Spectral Decomposition. In the subsection, the process of image decomposition using the NLTV spectral transform is illustrated in detail. Assuming that the number of decomposition components is N, the NLTV flow xxx can be calculated with the help of formulae (6) and (7). According to the definition of the NLTV spectral transform described in formula (8), the second derivative of the element $u(i)$ with respect to time scale needs to be computed. To speed up the calculation, the first and second derivatives are combined, expressed by formula (20).

$$u_{tt}(i,X)=\frac{(u(i+1,X)+u(i-1,X)-2\cdot u(i,X))}{\Delta t^2}, \tag{20}$$

where Δt is the time interval. NLTV transform is obtained based on u_{tt} by equation (21).

$$\phi(i,X)=u_{tt}(i,X)\cdot i\cdot \Delta t. \tag{21}$$

The NLTV spectral response can also be calculated using equation (10). The residual can be computed by equation (9). If the forward time difference $u_t(i)=(u(i+1)-u(i))/\Delta t$ is used to calculate the first derivatives, the residual part $\overline{f}$ can be transformed into formula (22).

$$\overline{f}=(N+1)\cdot u(N)-N\cdot u(N+1). \tag{22}$$

3.2.3. Object and Background Separation. After the decomposition of the original image in the NLTV spectral domain, the available information dimension of every pixel in the image increases from one to the number of time scales, i.e., the information used before decomposition is just pixel value. Inspired by the work [29], a separation surface is selected to effectively reduce the interference of noise on segmentation.

To better characterize the feature of objects in the image, time parameters t_1 and t_2 are chosen to construct a time range $[t_1,t_2]$. By the above analysis of the four properties of NLTV spectral transform, the max response time is computed to describe the image. The max response time here is different from the spectral response of equation (10). As equation (10) shows, the spectral response calculates the element ϕ(t) of the image in the NLTV spectral domain and can reflect the significant part of the image. The NLTV element $\phi(t)$ on the time scale t corresponding to the low response contains unimportant features, which can be discarded. However, formula (10) demonstrates that it fails to reflect the spatial information of the objects. To better analyze the performance of pixels in the NLTV spectral domain, the max response time is calculated. Specifically, the NLTV spectral transform, firstly, decomposes the image into several spectral components on a time scale, as shown in Figure 9. Then, every pixel in the image corresponds to a set of spectral responses. The time scale of the maximum spectral response is selected to indicate the performance of the local spatial information in the NLTV spectral domain. The maximum response time of pixels inside the same target tends to be close. Therefore, different objects of the image can be extracted by analyzing the max response time corresponding to each pixel. In other words, a salient time map $T(X)$ for each point X is calculated by equation (23).

$$T(X)=\arg\max_i \phi(i,X), \quad i\in[t_1,t_2]. \tag{23}$$

To extract more meaningful information about the segmentation target, we fit a separation surface whose role is a band-pass filter to separate the target from undesired information. Firstly, the filtered max response map $T_{\text{filter}}(X)$ is obtained by performing the Gaussian filtering on $T(X)$ to ensure the smoothness of separation surface. Then, the time scale corresponding to the maximum spectral response is stored as scatters, on which the least square regression is performed to finish fitting the surface $T_{\text{sur}}(X)$. The fitted surface can be regarded as a soft threshold in the range of

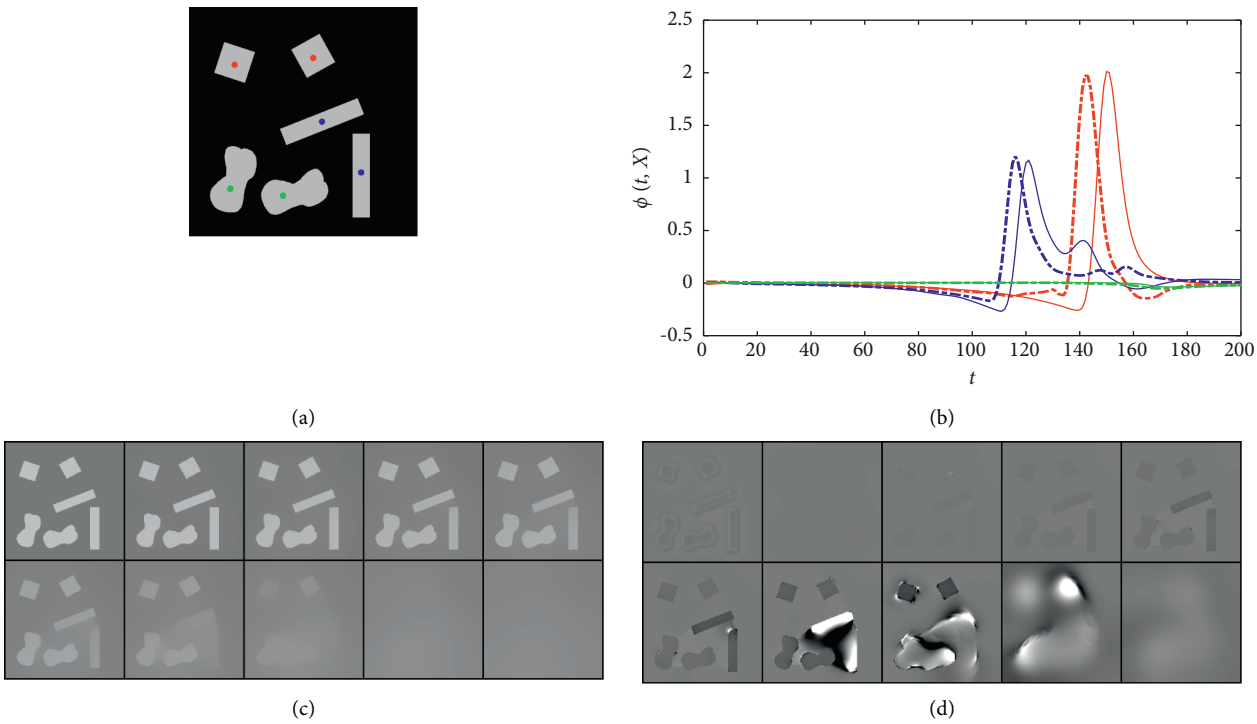


FIGURE 7: Demonstration of property 4. Signatures are similar to different rotations and translations. (a) Image f. (b) Multiscale NLTV spectral descriptions of different pixels. (c) Results of NLTV flow of f. (d) Multiscale NLTV spectral components.

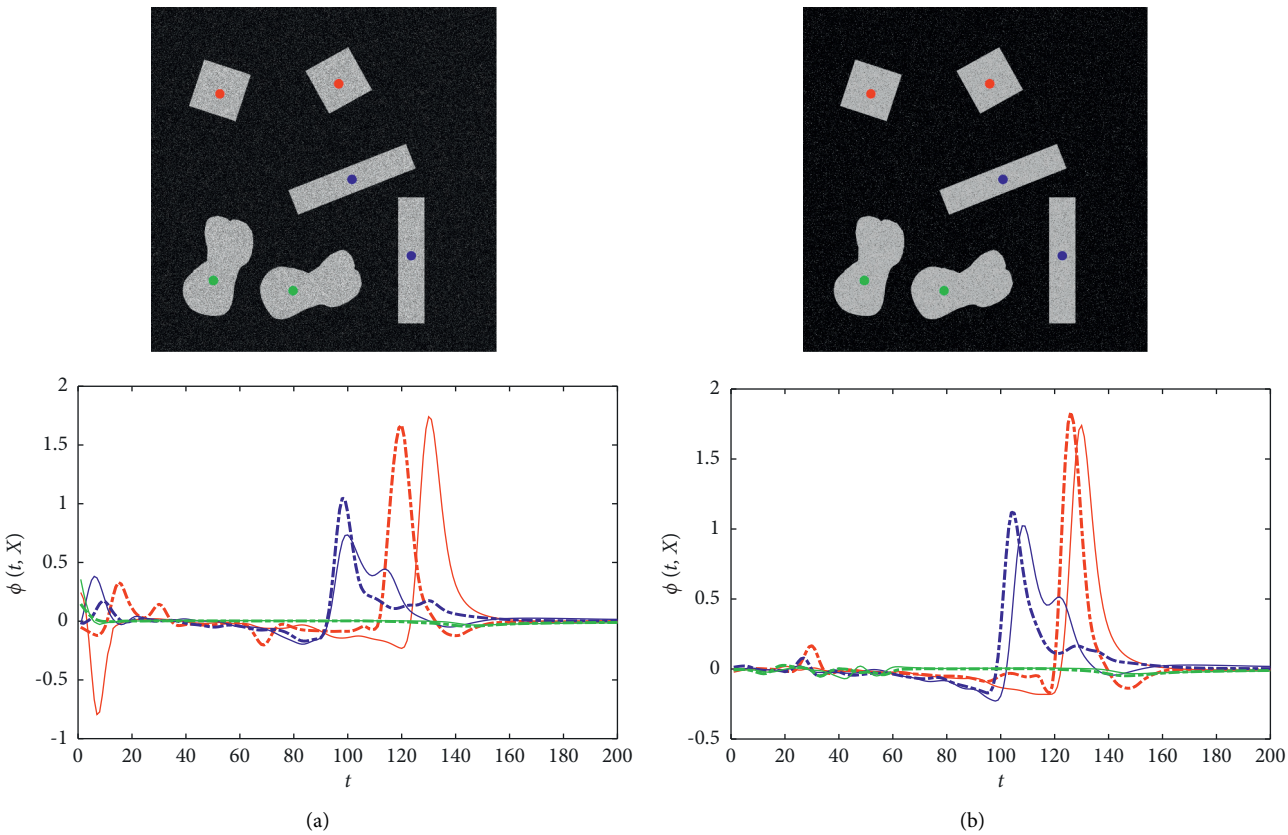


FIGURE 8: Continued.

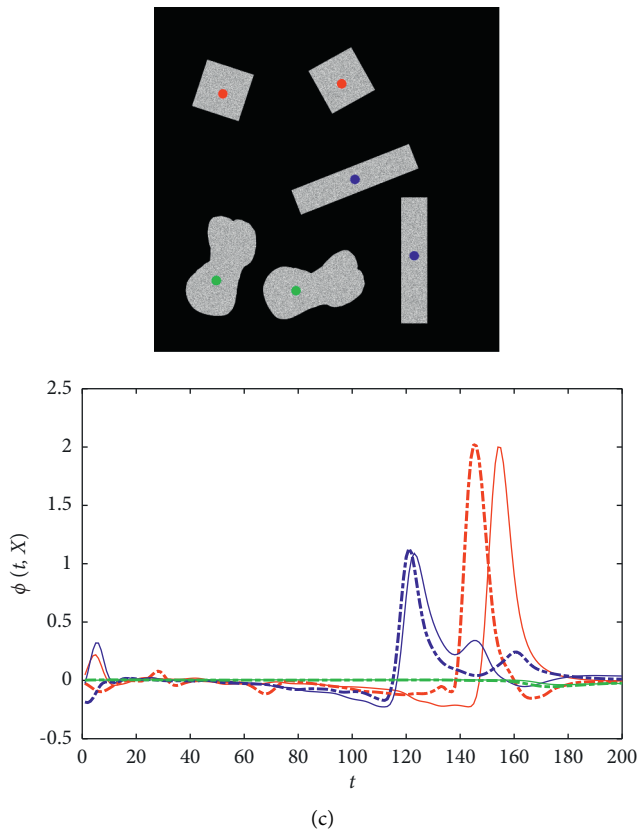


FIGURE 8: NLTV spectral transform on different groups corrupted with different noises. (a) is the image corrupted with Gaussian (10% variance), Salt & Pepper (10% density), and Speckle noise (10% variance), respectively. (b) is the multiscale NLTV spectral descriptions of different pixels corrupted with noises. (a) Gaussian noise. (b) Salt & Pepper noise. (c) Speckle noise.

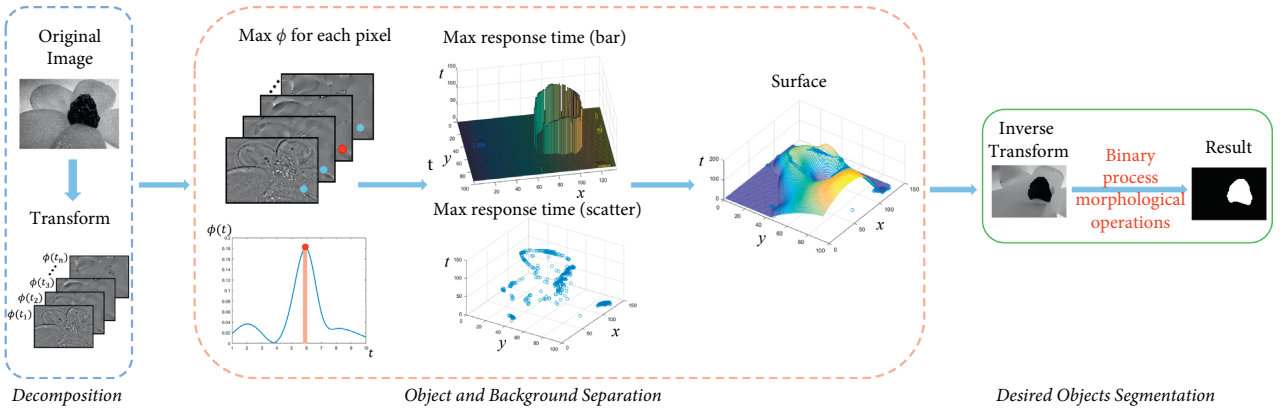


FIGURE 9: Flowchart of the proposed segmentation method using NLTV spectral transform.

$[t_1, t_2]$. For a certain point X, the surface divides it into two parts, $[t_1, T_{\text{sur}}(X)]$ and $[T_{\text{sur}}(X), t_2]$. The latter time range is usually chosen for image description to reduce the effect of noise. Therefore, the band-pass filter for each point X with time range $[t_1, t_2]$ can be denoted as follows:

$$HH_{\text{BPF},t_1,t_2}(i) = \begin{cases} 0, & 1 \le i < T_{\text{sur}}(X), \\ 1, & T_{\text{sur}}(X) \le i \le t_2, \\ 0, & t_2 < i \le N. \end{cases} \tag{24}$$

3.2.4. Desired Objects Segmentation. Image reconstruction, which is also called inverse transform, is implemented after surface fitting. The time scale band represents the integration times of each pixel for the object. The target in the original image is easily obtained by integrating over a specific time scale using reconstruction formula (25).

$$I(x) = \sum_{t=T_{\text{sur}}(X)}^{T} \phi(t, X) + \overline{f}. \tag{25}$$

Binary processing is performed after inverse transform to obtain the segmentation mask. Finally, morphological operations are used to refine the final mask. By the above operations, the desired segmentation mask f_{output} is obtained. To exhibit more details of the proposed method, Algorithm 1 shows the specific process of the NLTV spectral transform-based method for robust image segmentation.

4. Experiment Results

4.1. Data and Settings. To evaluate the performance of the proposed method, synthetic, natural, and medical images are used for experiments. 1) The first experiment contains 3 groups of synthetic images whose textures are taken from the Brodatz Textures dataset [32]. Speckle, Salt & Pepper, and Gaussian noises are added to each group of synthetic images separately. 2) The second experiment contains 3 groups of natural images taken from the MSRA-1000 dataset [33]. 3) The third experiment contains 1 group of cell images, which is taken from the Fluo-N2DH-SIM + dataset [34]. Three different types of noises are also added to natural and medical images.

We compare our segmentation method with four classical methods, i.e., the C-V model [13], FCM [14], FRFCM [19], and wavelet segmentation method (WSM) [27], which are used in the experiments. The experiments are implemented using the MATLAB R2020b platform and a PC with 16 GB RAM.

The parameter settings for the proposed method are as follows: experiments show that when the image is transformed into the NLTV domain, detailed information is located in a low time scale. Large scale, which is close to T, contains less important information. Objects are mostly distributed in the middle scale. Hence, a middle-scale time range $[t_1, t_2]$ is selected. In the following experiments, t_1 is set to $T/5$ and t_2 is set to $3T/5$. The parameters T and Δt are set to 9 and 0.03, respectively.

4.2. Quantitative Metrics. To quantitatively evaluate the performance of segmentation effect, four different metrics are chosen: FPR [21], FNR [21], dice similarity coefficient (DICE) [35], and segmentation accuracy (SA) [36].

To measure the difference between segmentation results and ground truths, FPR and FNR are chosen in the subsequent experiments. The former calculates the number of background pixels classified as object pixels relative to the total background pixels. FNR measures the number of object pixels classified as background pixels relative to the total object pixels. FPR and FNR are defined as follows:

$$\begin{aligned} \text{FPR} &= \frac{\left|B_R \cap O_G\right|}{\left|B_G\right|}, \\ \text{FNR} &= \frac{\left|O_R \cap B_G\right|}{\left|O_G\right|}, \end{aligned} \tag{26}$$

where B_R and B_G represent the number of background pixels in the segmentation results and ground truths, respectively. Additionally, O_R and O_G are the number of object pixels in the segmentation results and ground truths, respectively.

DICE measures segmentation accuracy by calculating the degree of spatial overlap. Specifically, for the result region $\mathbf{A}$ and target region $\mathbf{B}$,

$$\text{DICE}(\mathbf{A}, \mathbf{B}) = \frac{2(\mathbf{A} \cap \mathbf{B})}{\mathbf{A} + \mathbf{B}}, \tag{27}$$

where $\cap$ means the intersection of two sets. The value range of DICE is [0, 1]. The higher DICE indicates that the segmentation result is more precise. $\text{DICE}(\mathbf{A}, \mathbf{B}) = 1$ demonstrates that the segmentation result is the most complete, while $\text{DICE}(\mathbf{A}, \mathbf{B}) = 0$ shows that the segmentation result is the worst.

Another evaluation metric is SA, which can assess the number of well-classified pixels in the image. The definition of SA is given as follows:

$$\text{SA} = \frac{\sum_{i=1}^{N} f_i^{\text{truth}}}{N}, \tag{28}$$

where f_i^{truth} means the correctly segmented pixel and N denotes the total number of pixels in an image.

4.3. Parameter Analysis. This section analyzes the effects of Δt and T on the segmentation results of the proposed method through an experiment. The experiment was carried out on MSRA-1000, and the average SA was used as an indicator to show the influence of two parameters on the segmentation accuracy. The average SA was calculated by averaging the SA of all images on the dataset. The parameter Δt ranges from 0.01 to 0.1, and the step is 0.01. Additionally, the maximal time scale T ranges from 1 to 10, and the interval is 1. Figure 10 demonstrates the results for different Δt and T. The proposed method achieves the best performance when $\Delta t = 0.03$ and $T = 9$.

4.4. Synthetic Images. The first experiment was implemented on three synthetic images, which are shown in Figure 11. The first row shows a synthetic image containing multiple repeating structures and a dark grid-like background. A simple synthetic image, which has an irregular object, is arranged in the middle row. The object in the bottom row is complex and has a texture with inhomogeneous contrast. Moreover, three images are separately contaminated with Speckle (10% variance), Salt & Pepper (10% density), and Gaussian (10% variance) noise.

Table 1 lists the quantitative evaluations of different segmentation methods on various images. Combining with Figure 11 and Table 1, FCM got wrong segmentation results because of its sensitivity to noise. FRFCM achieved a good result on the first image and got a high DICE and SA value as shown in Table 1. However, it failed to distinguish the second and the third image because of the inhomogeneous contrast. WSM, which is based on spectral analysis, can remove the influence of noise. However, as Figure 11 shows, WSM oversmoothed the edge and damaged the edge details. Meanwhile, WSM was

Input: gray image f.
Output: segmentation mask f_{output}.
(1) Initialize: maximal time scale T, time step Δt.
(2) Calculate the number of decomposition components $N = T/\Delta t$.
(3) Compute NLTV flow $\{u(i)\}_{i=0}^{N+1}$ using equations (6) and (7).
(4) Calculate NLTV residual part $\overline{f}$ using equation (22).
(5) **for** $i = 1, 2, \ldots, N$ **do**
(6) Compute the second derivatives in time of flow for each pixel X by equation (20).
(7) Achieve NLTV transform by equation (21).
(8) Calculate NLTV spectral response using equation (10).
(9) **end for**
(10) Select time parameters t_1 and t_2 according to the NLTV spectral response.
(11) Compute the salient time map $T(X)$ by equation (23).
(12) Obtain $T_{\text{filter}}(X)$ by performing Gaussian filtering on $T(X)$.
(13) Get the fitted surface $T_{\text{sur}}(X)$ by performing least square regression on $T_{\text{filter}}(X)$.
(14) Reconstruct the result $I(X)$ using equation (25).
(15) Get the segmentation mask $f_{bw}(X)$ by thresholding segmentation on $I(X)$.
(16) Get the final mask $f_{\text{output}}(X)$ by performing morphological operations on $f_{bw}(X)$.

Algorithm 1: NLTV spectral transform-based method for robust image segmentation.

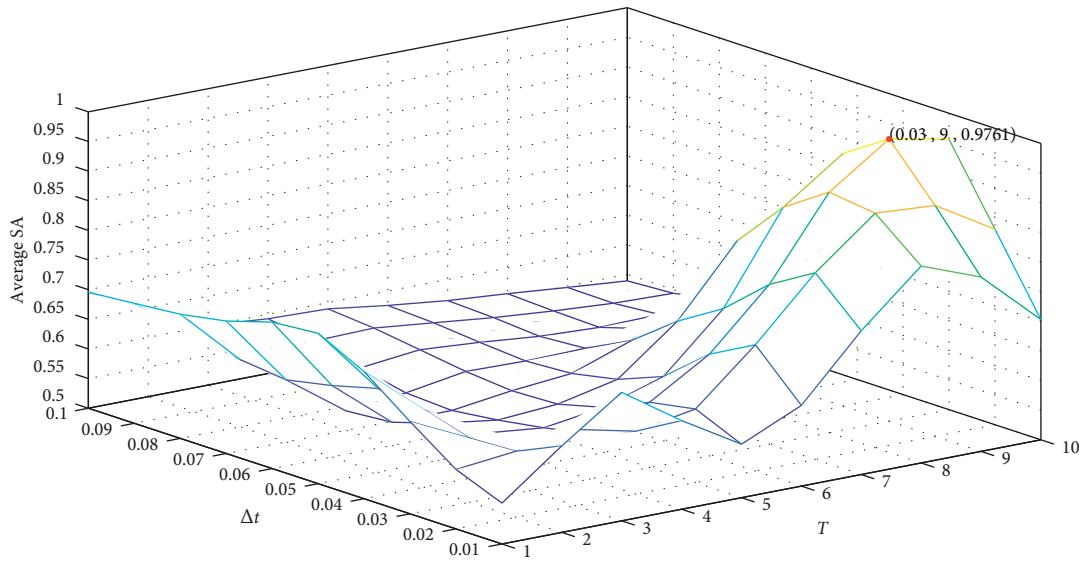


Figure 10: Effects of different Δt and T on the average SA.

unable to segment objects accurately on the second and third images. The reason is that WSM is sensitive to inhomogeneous contrast. The C-V model obtained the segmentation results of all images more correctly. One of the reasons was that the C-V model relies on an initial contour, which provides prior information about the approximate position of the object. Nevertheless, the C-V model was sensitive to noise. On the second and third images, the C-V model was unable to accurately segment the targets. The noises slowed down the convergence speed of the algorithm and made the method fall into the local minimum problem. However, the proposed method achieved the best results in all methods. The NLTV spectral transform-based method can segment the objects exactly and can reduce the influence of inhomogeneous contrast at the same time. The reason is that our method can segment objects, combining object size, contrast, and structures. As shown in Table 1, the proposed method got a high FNR on the second synthetic image, which intended an under-segmentation. The problem was caused by the morphological operators in the output of the proposed method, which may cause edge corrodes.

4.5. Nature Images. To further discuss the proposed method's segmentation ability for images with various noises, the second experiment was performed on three natural images, which are shown in Figures 12, 13, and 14. The object that has a similar contrast to the surroundings is shown in Figure 12. Figure 13 displays a complex scene that has lots of tiny structures in the background. The object in Figure 14 is a piece of paper containing words, and the

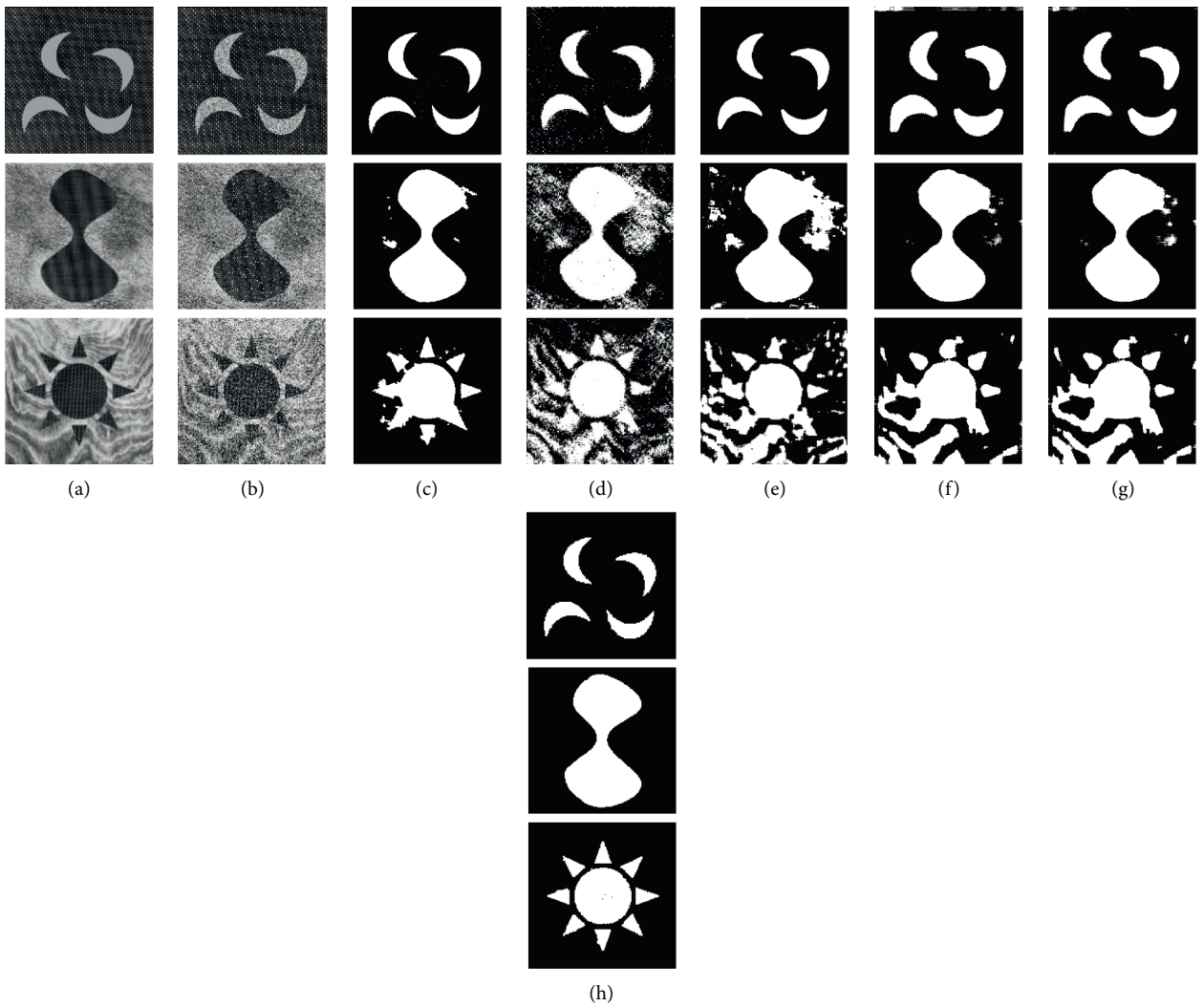


Figure 11: Segmentation results on synthetic images corrupted by different noises. (a) Original images. (b) Images (from top to bottom) that are corrupted with Speckle (10% variance), Salt & Pepper (10% density), and Gaussian noises (10% variance), respectively. (c) C-V Model. (d) FCM. (e) FRFCM. (f) WSM (db). (g) WSM (haar). (h) Proposed method.

Table 1: Evaluation metrics of compared methods for synthetic images, which are corrupted with Speckle (10% variance), Salt & Pepper (10% density), and Gaussian noise (10% variance), respectively.

Image	Metric	C-V model	FCM	FRFCM	WSM (db)	WSM (haar)	Proposed method
1	FPR	0.0025	0.0412	*0.0010*	0.0580	0.0402	**0.0009**
	FNR	0.0501	0.1977	0.0655	*0.0428*	0.0492	**0.0290**
	DICE	0.9596	0.7564	*0.9604*	0.7889	0.8328	**0.9821**
	SA	0.9907	0.9397	*0.9910*	0.9408	0.9558	**0.9956**
2	FPR	*0.0167*	0.2111	0.1355	0.0241	0.0326	**0.0038**
	FNR	*0.0231*	0.0743	**0.0187**	0.0389	0.0368	0.0273
	DICE	*0.9666*	0.7531	0.8427	0.9510	0.9415	**0.9798**
	SA	*0.9808*	0.8273	0.8959	0.9718	0.9659	**0.9884**
3	FPR	*0.0358*	0.3151	0.2352	0.1963	0.1984	**0.0235**
	FNR	0.0289	0.0901	**0.0217**	0.0253	0.0255	*0.0229*
	DICE	*0.9075*	0.5446	0.6402	0.6775	0.6751	**0.9333**
	SA	*0.9641*	0.7249	0.8015	0.8326	0.8308	**0.9748**

The best two results are highlighted in bold and italics fonts.

Table 2: Evaluation metrics of compared methods for "star," which are corrupted with Speckle (10%, 20%, and 30% variance) noise.

Noise level (%)	Metric	C-V model	FCM	FRFCM	WSM (db)	WSM (haar)	Proposed method
10	FPR	*0.0014*	0.3299	0.2172	0.2953	0.2934	**0.0010**
	FNR	0.0636	0.1170	0.1456	*0.0228*	**0.0044**	0.0544
	DICE	*0.9612*	0.2633	0.3553	0.3067	0.3117	**0.9667**
	SA	*0.9953*	0.6823	0.7910	0.7189	0.7213	**0.9957**
20	FPR	*0.0014*	0.3431	0.2677	0.2979	0.2935	**0.0011**
	FNR	0.0573	0.1296	0.0354	*0.0121*	**0.0107**	0.0578
	DICE	*0.9616*	0.2534	0.3294	0.3071	0.3109	**0.9649**
	SA	*0.9950*	0.6696	0.7461	0.7169	0.7212	**0.9955**
30	FPR	*0.0010*	0.3454	0.2668	0.3001	0.2913	**0.0010**
	FNR	0.0718	0.1155	0.0947	**0.0112**	*0.0170*	0.0584
	DICE	*0.9599*	0.2558	0.3207	0.3059	0.3113	**0.9652**
	SA	*0.9949*	0.6683	0.7460	0.7150	0.7231	**0.9950**

The best two results are highlighted in bold and italics fonts.

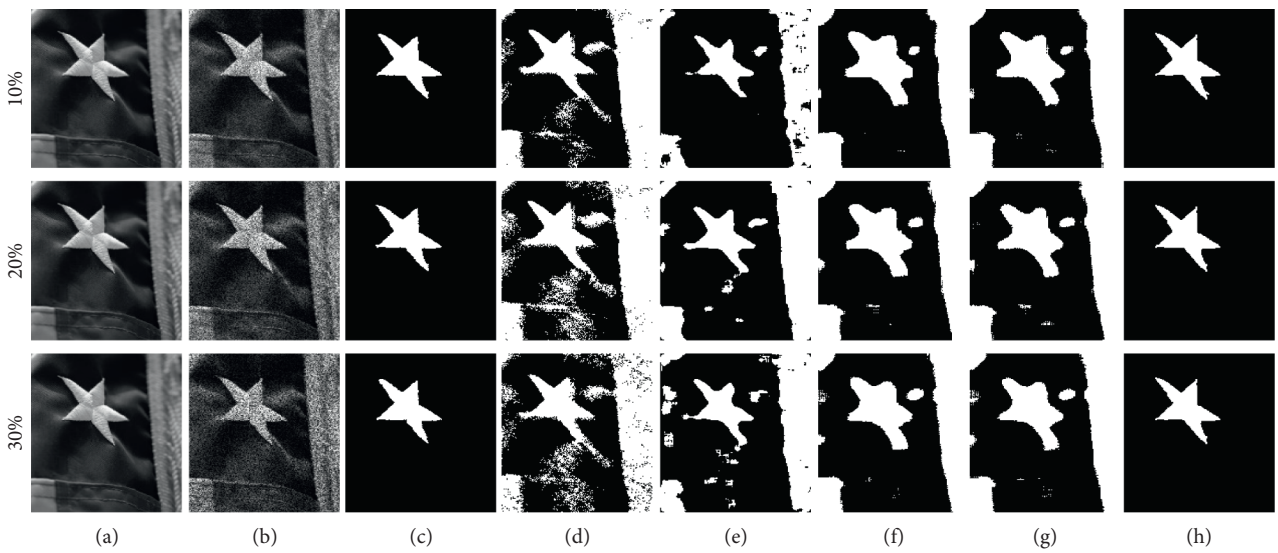


Figure 12: Segmentation results on image "star" corrupted by Speckle noise. (a) Original images. (b) Images (from top to bottom) that are corrupted with 10%, 20%, and 30% variance of Speckle noise, respectively (c) C-V model. (d) FCM. (e) FRFCM; (f) WSM (db); (g) WSM (haar); (h) Proposed method.

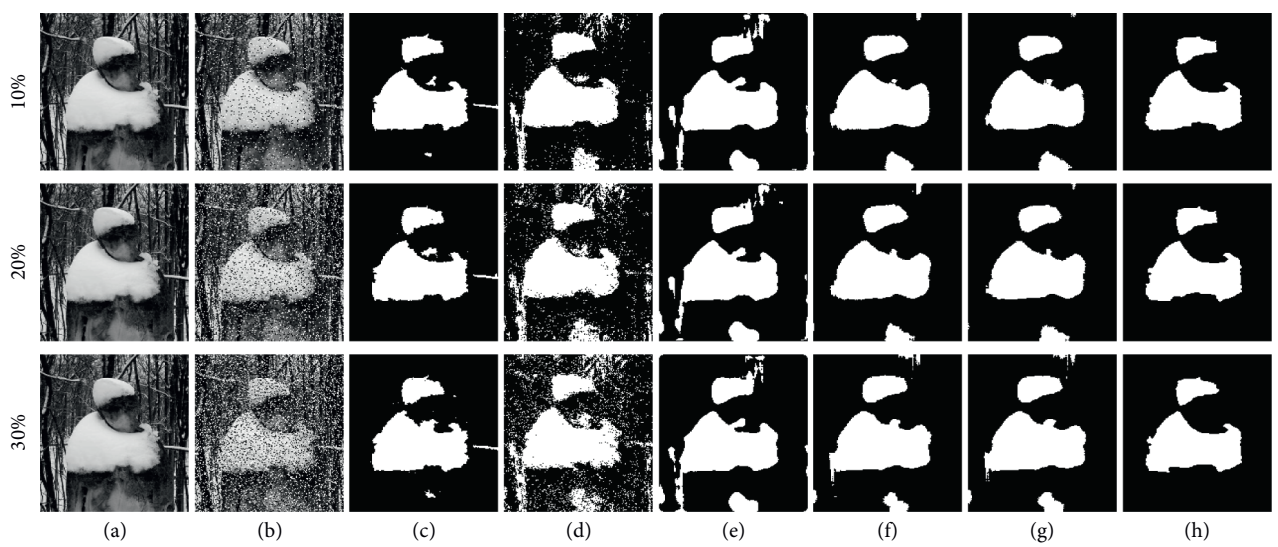


Figure 13: Segmentation results on image "snowman" corrupted by Salt & pepper noise. (a) Original images. (b) Images (from top to bottom) that are corrupted with 10%, 20%, and 30% variance of Speckle noise, respectively. (c) C-V Model. (d) FCM; (e) FRFCM. (f) WSM (db). (g) WSM (haar). (h) Proposed method.

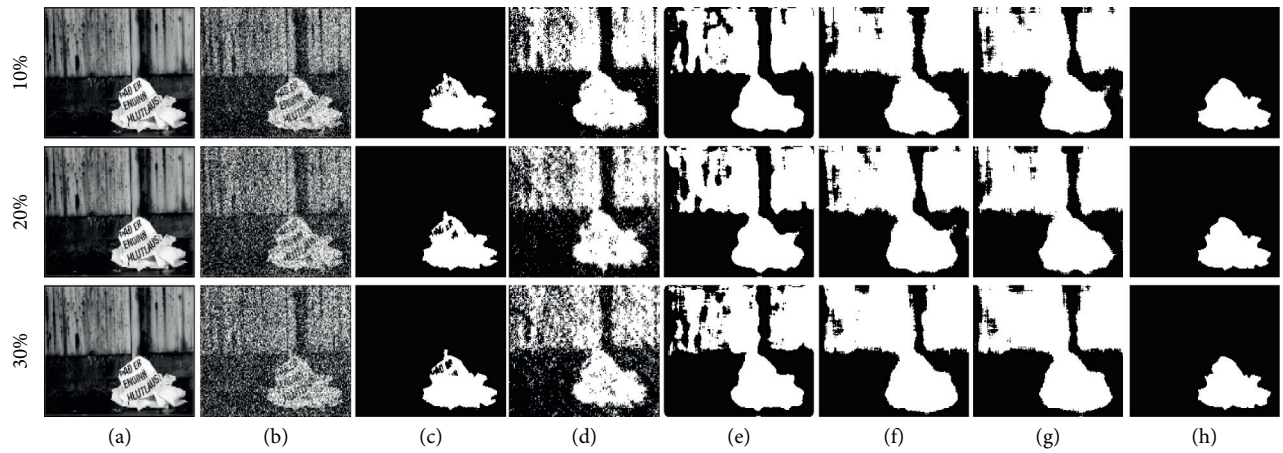


Figure 14: Segmentation results on image "paper scrap" corrupted by Gaussian noise. (a) Original images. (b) Images (from top to bottom) that are corrupted with 10%, 20%, and 30% variance of Speckle noise, respectively. (c) C-V Model. (d) FCM. (e) FRFCM. (f) WSM (db). (g) WSM (haar). (h) Proposed method.

words will interfere with segmentation methods. Moreover, three images are separately contaminated with Speckle (10%, 20%, and 30% variance), Salt & Pepper (10%, 20%, and 30% density), and Gaussian (10%, 20%, and 30% variance) noise.

As Figure 12 shows, the object has a similar contrast to the surrounding border. FCM separated the noise while segmenting the object because of its sensitivity to noise. FRFCM had better results than FCM, however, it still had wrong segmentation for noise. WSM can remove the influence of noise. However, WSM failed to remove the impact of inhomogeneous contrast. The C-V model achieved accurate segmentation results because of its initial contour. From Table 2, it can be seen that the C-V model had similar DICE and SA values with the proposed method. However, the C-V model was difficult to segment corner structure because of noises. The proposed method can better preserve structural information while segmenting.

Figure 13 shows the algorithms' performances on the natural images, which are corrupted by Salt & Pepper noise. Table 3 shows the corresponding quantitative metrics. In Figure 13, there are lots of small objects in the background, which have a similar contrast to the object. When these small targets are contaminated with Salt & Pepper noise, they cause serious interference with segmentation methods, which mainly rely on contrast. Figure 13 shows that WSM has a good result; however, it is unable to segment the areas surrounding the object correctly. Table 4 shows that the C-V model has better results than WSM; however, it still has an incorrect segmentation of the background. Because of the sensitivity of the NLTV spectral transform to contrast, size, and structures, the proposed method can still separate objects when the background has small size structures.

Figure 14 shows the segmentation results on the natural image when it is corrupted with different levels of Gaussian noise. Table 4 shows the corresponding quantitative metrics of algorithms. The natural image is difficult for segmentation methods because it has complex texture like words inside, which will affect the integrity of the segmentation results. The C-V model was capable of dealing with the background, however, it was unable to handle the interference of the internal texture of the object. FRFCM and WSM dealt with the effect of noise and internal texture but failed to remove the interference caused by contrast. Moreover, WSM cannot obtain accurate edge information of targets. As shown in Figure 14, WSM expanded the object and the edge details disappeared. However, our method can deal with the interference made by noise. The NLTV spectral transform was sensitive to local contrast and size. Hence, it can separate the low-contrast words on the paper scrap. Because of the contrast and structure difference between the paper scrap and the background, the proposed method can separate the object from the background and extract the object's edge details correctly. Table 4 shows that the proposed method has high FNR values. From Figure 14, the bottom edge in the results of the proposed method is a little expanded, and the left edge is obviously corroded. The main reason is that the morphological operator makes the segmentation result corroded.

4.6. Medical Image. The proposed method was evaluated on a medical image in this part. Because the medical image has a black background and the inference of speckle noise on the image is not obvious, the experiment was implemented on an image with Gaussian noise and Salt & Pepper noise. As Figure 15 shows, the top row is a cell image, which is contaminated with Gaussian noise, and the bottom row is the cell image corrupted with Salt & Pepper noise. On account of the noise, the initial contour of the C-V model generated a local minimum problem and was unable to be iteratively converged. As a result, the segmentation results of the C-V model can only be around the initial contour. FCM had wrong results because of its sensitivity to noise. FRFCM obtained the best result on the cell image corrupted with Salt & Pepper noise. However, Gaussian noise can cause FRFCM to generate an over-segmentation. WSM can

Table 3: Evaluation metrics of compared methods for "snowman," which are corrupted with Speckle (10%, 20%, and 30% variance) noise.

Noise level (%)	Metric	C-V model	FCM	FRFCM	WSM (db)	WSM (haar)	Proposed method
10	FPR	*0.0137*	0.1406	0.0737	0.0532	0.0467	**0.0049**
	FNR	*0.0577*	0.1756	**0.0408**	0.0848	0.0983	0.0890
	DICE	*0.9472*	0.6988	0.8556	0.8642	0.8675	**0.9524**
	SA	*0.9783*	0.8531	0.9332	0.9406	0.9431	**0.9810**
20	FPR	*0.0140*	0.1809	0.0763	0.0570	0.0497	**0.0053**
	FNR	*0.0553*	0.1768	**0.0402**	0.0764	0.0883	0.0979
	DICE	*0.9483*	0.6546	0.8520	0.8622	0.8682	**0.9505**
	SA	*0.9787*	0.8207	0.9312	0.9391	0.9428	**0.9804**
30	FPR	*0.0214*	0.2263	0.0785	0.0542	0.0548	**0.0175**
	FNR	*0.0529*	0.1651	**0.0380**	0.0880	0.0764	0.0997
	DICE	*0.9345*	0.6178	0.8505	0.8603	0.8656	**0.9479**
	SA	*0.9725*	0.7870	0.9301	0.9389	0.9408	**0.9732**

The best two results are highlighted in bold and italics fonts.

Table 4: Evaluation metrics of compared methods for "paper scrap," which are corrupted with Speckle (10%, 20%, and 30% variance) noise.

Noise level (%)	Metric	C-V model	FCM	FRFCM	WSM (db)	WSM (haar)	Proposed method
10	FPR	*0.0013*	0.4130	0.4138	0.5043	0.5063	**0.0002**
	FNR	0.1302	0.0962	0.0265	**0.0154**	**0.0230**	0.0688
	DICE	*0.9265*	0.3813	0.4023	0.3585	0.3557	**0.9754**
	SA	*0.9809*	0.6271	0.6340	0.5557	0.5533	**0.9939**
20	FPR	*0.0016*	0.4236	0.4158	0.5059	0.5077	**0.0005**
	FNR	0.1310	0.1144	*0.0202*	0.0225	0.0162	0.0743
	DICE	*0.9250*	0.3686	0.4037	0.3562	0.3568	**0.9725**
	SA	*0.9805*	0.6152	0.6333	0.5538	0.5527	**0.9932**
30	FPR	*0.0016*	0.4239	0.4164	0.5170	0.5137	**0.0011**
	FNR	0.1331	0.1129	*0.0230*	**0.0168**	0.0188	0.0893
	DICE	*0.9239*	0.3693	0.4034	0.3525	0.3534	**0.9661**
	SA	*0.9803*	0.6154	0.633	0.5446	0.5472	**0.9917**

The best two results are highlighted in bold and italics fonts.

Table 5: Evaluation metrics of compared methods for "cell" which are corrupted with gaussian (5% variance) and salt & pepper (5% density) noise.

Noise	Metric	C-V model	FCM	FRFCM	WSM (db)	WSM (haar)	Proposed method
Gaussian	FPR	0.4835	0.2775	0.1360	0.1170	*0.1020*	**0.0019**
	FNR	0.7958	0.1577	**0.0213**	*0.0277*	0.0325	0.3540
	DICE	0.0926	0.4688	0.6638	0.6892	*0.7112*	**0.7409**
	SA	0.4779	0.7332	0.8640	0.8796	*0.8920*	**0.9436**
Salt & Pepper	FPR	0.1256	0.1488	**0.0009**	0.2666	0.2491	*0.0180*
	FNR	0.2415	0.1140	0.3895	0.0376	*0.0502*	0.2780
	DICE	0.6000	0.6216	**0.8119**	0.5163	0.5299	*0.8026*
	SA	0.8640	0.8484	**0.9562**	0.7523	0.7672	*0.9459*

The best two results are highlighted in bold and italics fonts.

better remove the influence of Gaussian noise. However, WSM achieved high FPR values in the cell image corrupted with Salt & Pepper noise, which intends the over-segmentation. Nevertheless, our method obtained good segmentation performances in both noises. Our method achieved the best results on the image corrupted with Gaussian noise and obtained the second-best performance in Salt & Pepper noise. Table 5 shows that the proposed method achieves a high FNR value, which implies under-segmentation. As shown in the bottom row in Figure 15, the proposed method is difficult to segment the cells that have both small size and low contrast.

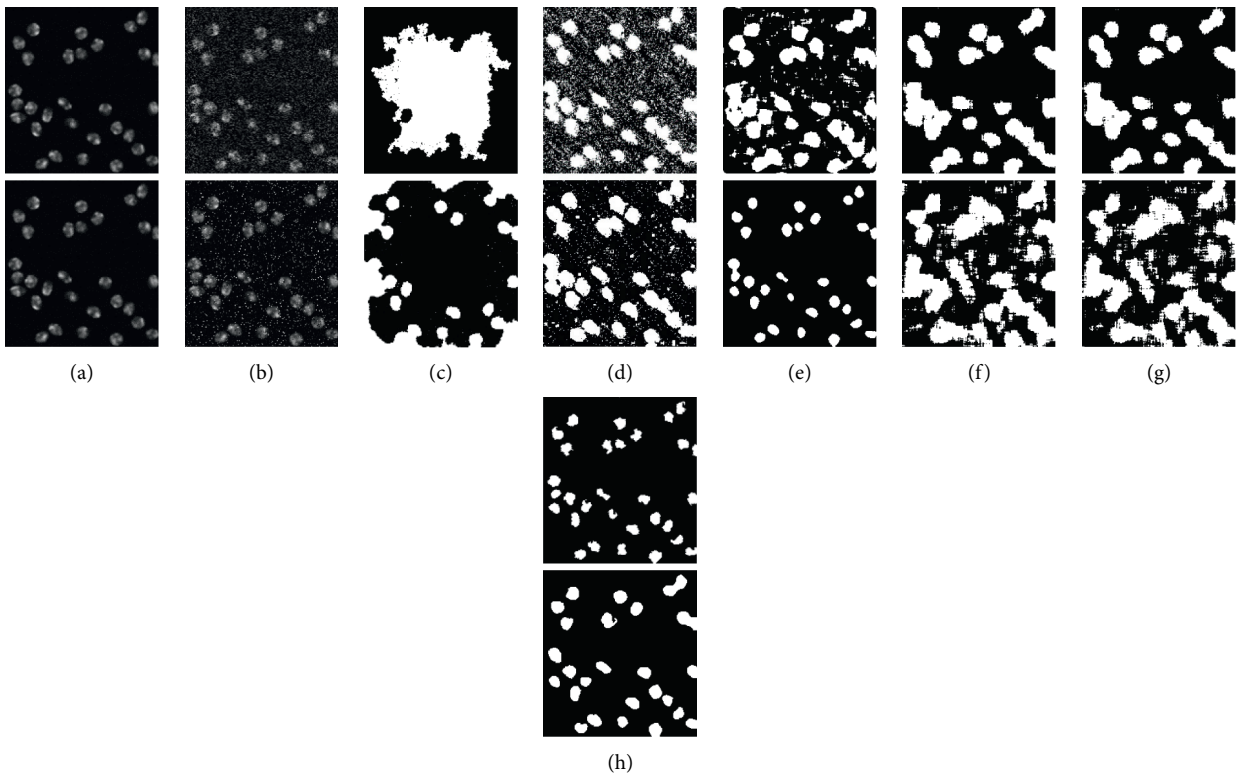


Figure 15: Segmentation results on image "cell" corrupted by Gaussian and Salt & Pepper noise. (a) Original images. (b) Images (from top to bottom) that are corrupted with Gaussian (5% variance) and Salt & Pepper (5% density) noise, respectively. (c) C-V Model. (d) FCM. (e) FRFCM. (f) WSM (db). (g) WSM (haar). (h) Proposed method.

5. Conclusion

We have analyzed the properties of NLTV spectral transform with the help of theoretical proof and experiments. Our analyses demonstrate that the object in an image corrupted with various noises can be separated its size, contrast, and detailed structure. The analyses also illustrate that the objects with same structures have similar descriptions in the NLTV spectral domain.

Furthermore, we have developed a novel transform-based method that segments images based on the NLTV spectral transform. The approach, firstly, decomposes an image into many sub-bands in the NLTV spectral domain and utilizes the max response time to represent the image features. Then, to better divide the object and background, the sub-bands in the NLTV spectral domain are filtered by fitting the separation surface, which is calculated based on maximum response time. Next, the filtered image is reconstructed by an inverse transform to obtain the rough segmentation result. Finally, the segmentation mask is calculated using postprocess methods. Subjective and objective evaluations show that the proposed method effectively protects the edge details while segmenting the object in a variety of noises.

However, one limitation of the proposed method is the high computational cost since the computation of nonlocal operators needs a long time and large memory storage. The other limitation of the method is the difficulty in fitting multiple separation surfaces accurately. We attempt to solve the aforementioned problems and develop a fast multiobject segmentation method in future work.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was partly supported by the National Natural Science Foundation of China (62076137, 61972206, 61971233, and 62011540407) and was also partly supported under the framework of international cooperation program managed by the National Research Foundation of Korea (NRF-2020K2A9A2A06036255 and FY2020).

References

[1] L. Sless, G. Cohen, S. B. El, and S. Oron, "Road scene understanding by occupancy grid learning from sparse radar clusters using semantic segmentation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, Seoul, Korea, October 2019.

[2] M. Akbari, J. Liang, and J. Han, "DSSLIC: deep semantic segmentation-based layered image compression," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2042–2046, Brighton, UK, May 2019.

[3] C. Bai, H. Li, J. Zhang, L. Huang, and L. Zhang, "Unsupervised adversarial instance-level image retrieval," *IEEE Transactions on Multimedia*, vol. 23, pp. 2199–2207, 2021.

[4] C. Bai, L. Huang, X. Pan, J. Zheng, and S. Chen, "Optimization of deep convolutional neural network for large scale image retrieval," *Neurocomputing*, vol. 303, pp. 60–67, 2018.

[5] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Massachusetts, MA, USA, June 2015.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Munich, Germany, October 2015.

[7] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[8] J. M. S. Prewitt, "Object enhancement and extraction," *Picture Processing and Psychopictorics*, vol. 10, pp. 15–19, 1970.

[9] R. Boyle and R. Thomas, *Computer Vision: A First Course*, pp. 48–51, Blackwell Scientific Publications, Hoboken, NJ, USA, 1988.

[10] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[11] M. N. Reza, I. S. Na, S. W. Baek, and K.-H. Lee, "Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images," *Biosystems Engineering*, vol. 177, pp. 109–121, 2019.

[12] D. Xiang, U. Bagci, C. Jin et al., "CorteXpert: a model-based method for automatic renal cortex segmentation," *Medical Image Analysis*, vol. 42, pp. 257–273, 2017.

[13] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, 266–277, 2001.

[14] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: the fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

[15] C. Liu, W. Liu, and W. Xing, "An improved edge-based level set method combining local regional fitting information for noisy image segmentation," *Signal Processing*, vol. 130, pp. 12–21, 2017.

[16] M. S. Yang and Y. Nataliani, "A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy," *IEEE Transaction Fuzzy System*, vol. 26, pp. 817–835, 2017.

[17] L. Guo, L. Chen, C. L. P. Chen, and J. Zhou, "Integrating guided filter into fuzzy clustering for noisy image segmentation," *Digital Signal Processing*, vol. 83, pp. 235–248, 2018.

[18] Y. Jiang, K. Zhao, K. Xia et al., "A novel distributed multitask fuzzy clustering algorithm for automatic MR brain image segmentation," *Journal of Medical Systems*, vol. 43, no. 5, pp. 118-119, 2019.

[19] N. Mahata, S. Kahali, S. K. Adhikari, and J. K. Sing, "Local contextual information and gaussian function induced fuzzy clustering algorithm for brain MR image segmentation and intensity inhomogeneity estimation," *Applied Soft Computing*, vol. 68, pp. 586–596, 2018.

[20] P. Parida and N. Bhoi, "2-D Gabor filter based transition region extraction and morphological operation for image segmentation," *Computers & Electrical Engineering*, vol. 62, pp. 119–134, 2017.

[21] P. Parida and N. Bhoi, "Wavelet based transition region extraction for image segmentation," *Future Computing and Informatics Journal*, vol. 2, no. 2, pp. 65–78, 2017.

[22] D. Palani and K. Venkatalakshmi, "An IoT based predictive modelling for predicting lung cancer using fuzzy cluster based segmentation and classification," *Journal of Medical Systems*, vol. 43, no. 2, pp. 1–12, 2019.

[23] Y. Wang, Q. Yuan, and C. He, "Indirect diffusion based level set evolution for image segmentation," *Applied Mathematical Modelling*, vol. 69, pp. 714–722, 2019.

[24] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, USA, 1998.

[25] H. Liu, Z. Chen, X. Chen, and Y. Chen, "Multiresolution medical image segmentation based on wavelet transform," in *Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 3418–3421, Shanghai, China, January 2006.

[26] H. Castillejos, V. Ponomaryov, L. Nino-de-Rivera, and V. Golikov, "Wavelet transform fuzzy algorithms for dermoscopic image segmentation," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 578721, 2012.

[27] J. Gao, B. Wang, Z. Wang, Y. Wang, and F. Kong, "A wavelet transform-based image segmentation method," *Optik*, vol. 208, Article ID 164123, 2020.

[28] J. F. Aujol, G. Gilboa, and N. Papadakis, "Fundamentals of non-local total variation spectral theory," in *Proceedings of the 5th International Conference on Scale Space and Variational Methods in Computer Vision*, pp. 66–77, Lège-Cap Ferret, France, June, 2015.

[29] G. Gilboa, "A total variation spectral framework for scale and texture analysis," *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 1937–1961, 2014.

[30] J. Zhang, J. Qi, Z. Zheng, and L. Sun, "A robust image segmentation framework based on total variation spectral transform," *Pattern Recognition Letters*, vol. 153, pp. 159–167, 2022.

[31] M. D'Elia, Q. Du, C. Glusa, M. Gunzburger, X. Tian, and Z. Zhou, "Numerical methods for nonlocal and fractional models," *Acta Numerica*, vol. 29, pp. 1–124, 2020.

[32] S. Abdelmounaime and H. Dong-Chen, *New Brodatz-Based Image Databases for Grayscale Color and Multiband Texture Analysis*, ISRN, 2013.

[33] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1597–1604, Miami Beach, FL, USA, June 2009.

[34] M. Maška, V. Ulman, D. Svoboda et al., "A benchmark for comparison of cell tracking algorithms," *Bioinformatics*, vol. 30, no. 11, pp. 1609–1617, 2014.

[35] S. Minaee, Y. Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: a survey," *IEEE Transaction Pattern Analysis Machine Intelligence*, 2021.

[36] S. Hao, Y. Zhou, and Y. Guo, "A brief survey on semantic segmentation with deep learning," *Neurocomputing*, vol. 406, pp. 302–321, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4392132, 11 pages
https://doi.org/10.1155/2022/4392132

Research Article

Throughput Fairness for Wireless Powered Cognitive Hybrid Active-Passive Communications

Shuang Fu, Chenyang Ding, and Peng Jiang

Institute of Information and Electrical Engineering, Heilongjiang Bayi Agricultural University, Daqing, China

Correspondence should be addressed to Shuang Fu; fushuang_dq@163.com

Received 24 November 2021; Revised 2 January 2022; Accepted 17 January 2022; Published 9 February 2022

Academic Editor: Liqin Shi

In this paper, we consider a backscatter communication (BackCom)-based cognitive network that consists of one primary transmitter, one primary receiver, multiple secondary transmitters (STs), and one secondary receiver (SR). Each ST operates in the BackCom or energy harvesting model. Our goal is to jointly optimize the energy harvesting and backscatter time, the transmit power of the primary transmitter, and the power reflection coefficient of each ST to maximize the sum throughput of all the STs under a nonlinear energy harvesting model while satisfying multiple constraints, i.e., the energy causality of each ST, the quality of service of the primary transmitter, etc. The formulated problem is nonconvex due to the coupled variables and is hard to solve. In order to address this problem, we decouple partially coupled variables by using the properties of the objective function and constructing auxiliary variables, and the remaining coupled variables are decoupled via successive convex approximation (SCA). On this basis, a SCA-based iterative algorithm is developed to solve the formulated problem. Simulation results are provided to support our work.

1. Introduction

Increasing demands for intelligent services have boosted the attention of Internet of Things (IoT), where massive tiny IoT devices should be deployed for connecting the physical environment and cyberspace seamlessly [1]. This poses an urgent need for developing a high spectrum efficient communication technology in the era of IoT networks. In this context, cognitive radio has been proposed, where the IoT nodes are allowed to share the spectrum with the primary users [2, 3]. Despite the improvement of spectrum efficiency for tiny IoT nodes, most of them still suffer from the short life span as they are powered by a limited battery capacity [4]. In order to address this challenge, wireless powered hybrid active-passive communication [5, 6] has been integrated into the cognitive radio, yielding a wireless powered cognitive hybrid active-passive communication network [7, 8].

Until now, there are considerable works on the design of resource allocation schemes for wireless powered cognitive hybrid active-passive communication network. In [9], the authors proposed to maximize the rate of an IoT node by jointly optimizing the energy harvesting time, the backscattering time, and the active communication time, while satisfying that the consumed energy of an IoT node does not exceed the harvested energy. Subsequently, this work was extended into a scenario with multiple IoT nodes [10], where the main focus was to find the optimal tradeoff among the energy harvesting time, the backscattering time, and each IoT node's active communication time. In the above two works [9, 10], the rate of the considered backscatter communication was assumed to be a constant. In [11], the authors considered another wireless powered cognitive hybrid active-passive communication network, where two different backscatter communications are introduced, and proposed an optimal time allocation scheme to maximize the throughput of the IoT node. Considering that energy efficiency is an important performance metric in wireless communications, the authors in [12] maximized the energy efficiency of an IoT node while considering the primary interference and imperfect spectrum sensing constraints. In [13], the authors considered multiple IoT nodes and

proposed to maximize the energy efficiency of all the IoT nodes by jointly optimizing the time and power resources, subject to the minimum throughput requirement of each IoT node and the energy causality constraint of each IoT node.

Although the above works have provided a solid foundation for understanding the resource allocation in wireless powered cognitive hybrid active-passive communication networks, some gaps still exist. First, the above works mainly ignored the interference introduced by the IoT node. More specifically, for the backscattering time, the interference from the IoT node to the primary was ignored. Second, fairness among IoT nodes has not been considered. Third, in the practical energy harvester, the output power is a nonlinear function with respect to the input power [14], while this nonlinearity has been ignored in most of the existing works. Motivated by the above observations, in this paper, we study the throughput fairness in a wireless powered cognitive hybrid active-passive communication network that consists of multiple backscatter devices (BDs) and backscatter receivers (BRs), one primary transmitter (PT), and one primary receiver (PR). The main contributions are summarized as follows:

(i) A throughput fairness problem is formulated. In particular, this problem maximizes the minimum throughput that achieved all IoT nodes by jointly optimizing the transmit power and time of the primary transmitter (PT), the BDs' time sharing among energy harvesting (EH), backscatter communication (BackCom) and active communications, and the power reflection coefficient and transmit power of each BD subject to the quality of service (QoS), energy causality, and transmit power constraints.

(ii) We develop an iterative algorithm to solve the formulated problem. More specifically, we first derive the optimal transmit power of the PT in a closed form via contradiction and then construct a series of auxiliary variables to decouple the coupled variables. Lastly, successive convex approximation (SCA) is leveraged to address the nonconvex QoS constraint. On this basis, an efficient iterative algorithm is proposed to solve the formulated problem.

(iii) We provide computer simulation to verify the quick convergence of the proposed algorithm and demonstrate that the fairness throughput can be ensured by our proposed scheme.

2. System Model

2.1. Basic Settings of the Considered Network. In this paper, the wireless powered cognitive hybrid active-passive communication and the wireless powered cognitive network with hybrid active-passive communications are used interchangeably, which is shown in Figure 1. Specifically, the whole network consists of one PT, one PR, K BDs, and their receivers, where the PT broadcasts its signals to the PR for

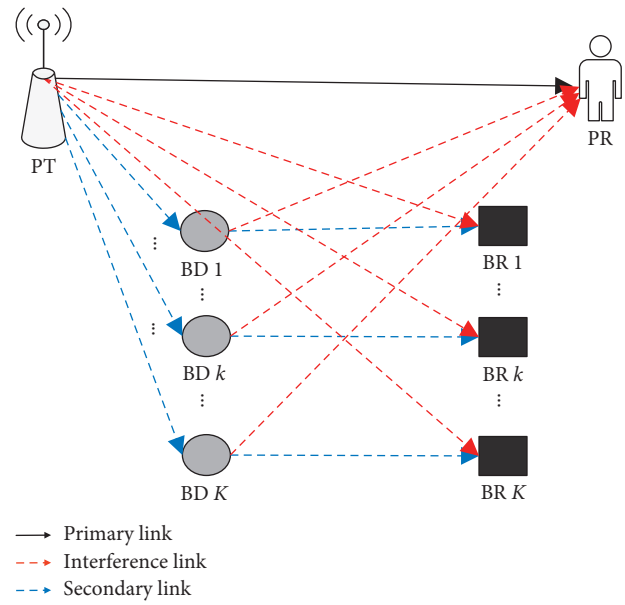


Figure 1: BackCom-based cognitive networks and its frame structure.

the primary transmission, and the PT's signals are also exploited by the K BDs for energy harvesting (EH) and BackCom. We note that when the PT is idle, each BD can also use its harvested energy to transmit its information by active communications. That is to say, each BD can backscatter its received signals for passive communications when PT is busy and use its harvested energy for active communications when PT is idle. We note that each BD only harvests energy from the energy signals from the PT since the energy harvested from other nodes during active and backscatter communications is very weak. All the devices are assumed to be equipped with a single antenna and always work in the half-duplex mode. The reasons are as follows: we note that both EH and BackCom are particularly applicable to wireless sensor networks for each sensor node's information transmission, where it may be difficult for low-cost small wireless sensor nodes to have multiple antennas and work at the full-duplex mode. Meanwhile, all devices with a single antenna have also been assumed in many related recent works. All the BDs are energy constrained, where each BD uses its harvested energy in each transmission block to support its energy consumption so that the operation time of each BD is prolonged. We assume that all channels are quasistatic fading. Let h_p and g_k $(k \in \mathcal{K} = \{1, 2, \ldots, K\})$ denote the channel gains from the PT to the PR and the k-th BD, respectively. We denote the channel gains from the k-th BD to its receiver and the PR as f_k and $f_{p,k}$. The channel gain from the PT to the k-th BD's receiver (BR) is expressed as h_k. In the beginning of each transmission block, the channel estimation is adopted by the PT so as to perfectly know the channel state information (CSI) of all links, and then the PT can determine the optimal resource allocation scheme according to the achieved CSI, and the optimal resource allocation scheme can be performed successively. In this work, we clarify how to obtain all the channel gains of

all links as follows: the channel gain from the PT to the PR (or the k-th BR) can be obtained by performing the existing advanced channel estimation methods, e.g., least-square estimation, etc. The value of the product of the forward channel gain and the backward channel gain from the k-th BD to the PT (or the k-th BR, the PR) can be obtained by performing least-square estimation. Due to the channel reciprocity, the forward channel gain equals the backward channel gain, and hence, the channel gain from the k-th BD to the PT (or the k-th BR, the PR) can be obtained.

2.2. Frame Structure. As shown in Figure 2, let T denote the duration of each transmission block. For the primary transmission, the whole block can be divided into two phases. In the first phase, the PT performs information transmission. Let β with $0 \le \beta \le T$ denote the transmission time of the primary transmission. In this phase, K BDs first work in the EH mode, where all the received signals are used to harvest energy and then take turns to work in the BackCom mode. In particular, let t_e be the EH time for all BDs and t_k be the backscattering time for the k-th BD. Then, we have $t_e + \sum_{k=1}^{K} t_k \le \beta$. We note that in the subphase t_k, the k-th BD performs BackCom, while the other BDs still work in the EH mode in order to improve its harvested energy. In the second phase with duration $T - \beta$, the PT stops its information transmission, and all the BDs can use their harvested energy to transmit their information. In order to avoid co-channel interference among BDs, all the BDs take turns to perform information transmission. Let τ_k denote the transmit time of the k-th BD in this phase. Then, we have $\sum_{k=1}^{K} \tau_k \le T - \beta$.

In the following part, we will clarify how the system works from both the primary transmission and BDs' transmissions. In the subphase t_e, we denote P_0 as the transmit power of the PT. Then, the received signals at the k-th ($k \in \mathcal{K} = \{1, 2, \ldots, K\}$) BD can be given by

$$y_{\mathrm{BD}}^{k} = \sqrt{P_0 g_k} x_p + N_{\mathrm{BD}}, \tag{1}$$

where x_p with $\mathbb{E}[|x_p|^2] = 1$ expresses the information transmitted by the PT to the PR, and N_{BD} is the additive white Gaussian noise (AWGN) at the k-th BD with mean zero and variance σ^2.

In this work, a nonlinear EH model proposed in [14] is considered to characterize the nonlinearity of practical EH circuits accurately. The reason of considering this nonlinear EH model is as follows: firstly, according to [14], the nonlinear EH model proposed in [14] is very accurate, even more accurate than the existing nonlinear EH models. Secondly, the use of the nonlinear EH model proposed in [14] removes the difficulty caused by the nonlinear EH model since we can prove that the nonlinear EH model proposed in [14] is concave by using the properties of practical EH circuits, which greatly reduces the difficulty of solving the formulated optimization problem. Then, the harvested energy at the k-th BD in this subphase can be calculated as

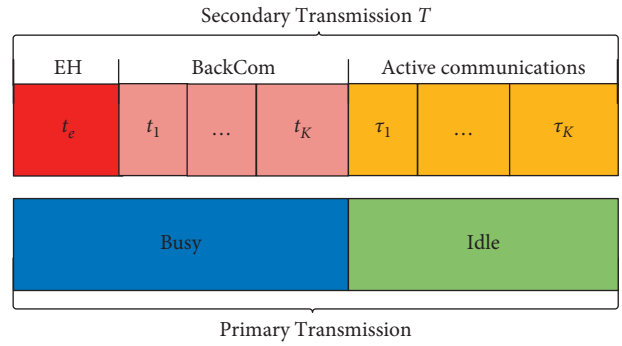


Figure 2: Frame structure for the considered network.

$$E_e^k = t_e \phi_k [P_0 g_k], \tag{2}$$

where $\phi_k[x] = ((a_k x + d_k)/(x + v_k)) - (d_k/v_k)$, a_k, d_k, and v_k are the given parameters of the considered nonlinear EH model at the k-th BD and may be different for different BDs.

For the primary transmission, the received signals at the PR can be expressed as

$$y_{\mathrm{PR}}^{e} = \sqrt{P_0 h_p} x_p + N_{\mathrm{PR}}, \tag{3}$$

where N_{PR} is the AWGN at the PR with mean zero and variance σ^2. Correspondingly, the achievable throughput at the PR in this subphase can be calculated as

$$C_e^p = t_e W \log_2\left(1 + \left(\frac{P_0 h_p}{\sigma^2}\right)\right), \tag{4}$$

where W is the bandwidth of the whole system.

In the subphase t_k, let α_k with $0 \le \alpha_k \le 1$ denote the power reflection coefficient of the k-th BD, which decides how many signals are received at the k-th BD to be backscattered. We note that the rest signals will be flowed into the EH circuit of the k-th BD for EH. Accordingly, the received signals at the BR in the subphase t_k can be represented as

$$y_{\mathrm{BR}}^{k} = \sqrt{\alpha_k P_0 g_k f_k} x_p x_{b,k} + \sqrt{P_0 h_k} x_p + N_{\mathrm{BR}}, \tag{5}$$

where $x_{b,k}$ with $\mathbb{E}[|x_{b,k}|^2] = 1$ is the information transmitted by the k-th BD, and N_{BR} is the AWGN at the BR with mean zero and variance σ^2.

It can be observed from (5) that each BD's transmission suffers from the co-channel interference to the primary transmission, leading to a reduction in the achievable throughput of the k-th BD via BackCom. Besides, owing to the double path loss fading of the BD's transmission, the signal power from the PT is always higher than that from the BD. In order to decode $x_{b,k}$ successfully, the successive interference cancellation (SIC) technology is performed at the k-th BR. Specifically, the BR should decode the PT's transmitted information x_p first by treating $\sqrt{\alpha_k P_0 g_k f_k} x_p x_{b,k}$ as interference and then cancel the interference $\sqrt{P_0 h_k} x_p$ since both h_k and x_p are known by the BR. On this basis, the transmitted information of the k-th BD $x_{b,k}$ can be decoded.

Based on (5), we can express the signal to interference plus noise ratio (SINR) at the BR for decoding x_p as

$$\gamma_{b,k}^{p} = \frac{P_0 h_k}{\alpha_k P_0 g_k f_k + \sigma^2}. \tag{6}$$

In order to ensure that the BR can decode x_p successfully, the following inequality should hold, given by

$$\gamma_{b,k}^{p} \geq \gamma_{th}, \tag{7}$$

where γ_{th} is the given threshold for decoding x_p, indicating the minimum SINR requirement for decoding x_p.

When $\gamma_{b,k}^{p} \geq \gamma_{\text{th}}$ holds, we can perform the SIC technology, and then the signal to noise ratio (SNR) at the BR for decoding $x_{b,k}$ is given by

$$\gamma_{b,k} = \frac{\alpha_k P_0 g_k f_k}{\epsilon P_0 h_k + \sigma^2}, \tag{8}$$

where ϵ with $0 \leq \epsilon \leq 1$ denotes the interference cancellation factor.

Correspondingly, the achievable throughput at the k-th BR in this subphase can be calculated as

$$\begin{aligned} C_{b,k} &= t_k W \log_2\left(1 + \xi \gamma_{b,k}\right) \\ &= t_k W \log_2\left(1 + \frac{\xi \alpha_k P_0 g_k f_k}{\epsilon P_0 h_k + \sigma^2}\right), \end{aligned} \tag{9}$$

where ξ denotes the performance gap reflecting the real modulation [5, 6].

For the k-th BD, its harvested energy is given by

$$E_k^b = t_k \phi_k \left[P_0 g_k \left(1 - \alpha_k\right)\right]. \tag{10}$$

Then, in the end of the first phase, the total harvested energy at the k-th BD can be computed as

$$\begin{aligned} E_k &= \left(\beta - t_k\right)\phi_k \left[P_0 g_k\right] + t_k \phi_k \left[P_0 g_k \left(1 - \alpha_k\right)\right] \\ &= \left(\beta - t_k\right)\left(\frac{a_k P_0 g_k + d_k}{P_0 g_k + v_k} - \frac{d_k}{v_k}\right) \\ &\quad + t_k \left(\frac{a_k P_0 g_k \left(1 - \alpha_k\right) + d_k}{P_0 g_k \left(1 - \alpha_k\right) + v_k} - \frac{d_k}{v_k}\right). \end{aligned} \tag{11}$$

As for the primary transmission, it also suffers from the co-channel interference from the k-th BD, and the received signals at the PR in the subphase t_k are expressed as

$$y_{\text{PR}}^{k} = \sqrt{P_0 h_p} x_p + \sqrt{\alpha_k P_0 g_k f_{p,k}} x_p x_{b,k} + N_{\text{PR}}. \tag{12}$$

Since the signal power of the PT is higher than that of the k-th BD, the PR will decode x_p first by treating $\sqrt{\alpha_k P_0 g_k f_{p,k}} x_p x_{b,k}$ as interference. Accordingly, the SINR at the PR for decoding x_p is given by

$$\gamma_k^{p} = \frac{P_0 h_p}{\alpha_k P_0 g_k f_{p,k} + \sigma^2}. \tag{13}$$

Then, the achievable throughput at the PR in the subphase t_k is determined by

$$\begin{aligned} C_k^{p} &= t_k W \log_2\left(1 + \gamma_k^{p}\right) \\ &= t_k W \log_2\left(1 + \frac{P_0 h_p}{\alpha_k P_0 g_k f_{p,k} + \sigma^2}\right). \end{aligned} \tag{14}$$

When the PT is idle, each BD uses its harvested energy to transmit information. Let p_k denote the transmit power of the k-th BD in the subphase τ_k. Then, the achievable throughput of the k-th BD in the subphase τ_k is given by

$$C_{a,k} = \tau_k W \log_2\left(1 + \frac{p_k f_k}{\sigma^2}\right). \tag{15}$$

3. Throughput Fairness for Wireless Powered Cognitive Hybrid Active-Passive Communications

In this section, we study the throughput fairness among different BDs for wireless powered cognitive networks with hybrid active-passive communications by designing an optimal resource allocation scheme. In particular, we formulate a throughput fairness optimization problem by jointly optimizing the transmit power and time of the PT, the EH time, the BackCom time, and power reflection coefficients of BDs, and the transmit power and time of each BD, subject to multiple constraints, i.e., QoS, energy causality, transmit power, and power reflection coefficient constraints. Then, an efficient iterative algorithm is developed to solve it.

3.1. Problem Formulation

3.1.1. Optimization Objective. The optimization objective is to guarantee the throughput fairness among different BDs. Toward this end, a max-min approach is adopted [6, 15]. Thus, we determine the optimization objective as maximizing the minimum achievable throughput of each BD. For the k-th BD, its total achievable throughput in the whole transmission block can be computed as

$$\begin{aligned} C_{\text{tot}}^{k} &= C_{b,k} + C_{a,k} \\ &= t_k W \log_2\left(1 + \frac{\xi \alpha_k P_0 g_k f_k}{\epsilon P_0 h_k + \sigma^2}\right) + \tau_k W \log_2\left(1 + \frac{p_k f_k}{\sigma^2}\right). \end{aligned} \tag{16}$$

On this basis, the optimization objective is determined by $\min_k C_{\text{tot}}^{k}$.

3.1.2. QoS Constraint for the Primary Transmission. This constraint is to ensure that the achievable throughput of the PT is not less than its minimum required throughput. Based on (4) and (14), the achievable throughput of the PT can be computed as

$$C_p = C_e^p + \sum_{k=1}^{K} C_k^p = t_e W \log_2\left(1 + \frac{P_0 h_p}{\sigma^2}\right) + \sum_{k=1}^{K} t_k W \log_2\left(1 + \frac{P_0 h_p}{\alpha_k P_0 g_k f_{p,k} + \sigma^2}\right). \tag{17}$$

Let $C_{\min}$ denote the minimum required throughput of the primary transmission. Then, the QoS constraint for the primary transmission can be expressed as

$$C_p = t_e W \log_2\left(1 + \frac{P_0 h_p}{\sigma^2}\right) + \sum_{k=1}^{K} t_k W \log_2\left(1 + \frac{P_0 h_p}{\alpha_k P_0 g_k f_{p,k} + \sigma^2}\right) \geq C_{\min}. \tag{18}$$

3.1.3. Energy-Causality Constraint for the BD's Transmission. This constraint is to ensure that each BD only uses its harvested energy to support the energy consumption for the passive and active communications so that the energy early stored in its battery is saved. Following reference [6], a fixed power consumption model is considered for BackCom, where the power consumption for the k-th BD is fixed as a constant, denoted by $P_{b,k}$. Then, the total energy consumption for BackCom at the k-th BD is given by $P_{b,k}t_k$. Let $p_{a,k}$ denote the constant circuit power consumption for active communications at the k-th BD. Then, the total energy consumption for active communications at the k-th BD can be computed as $p_k\tau_k + p_{a,k}\tau_k$. On this basis, the energy-causality constraint for the k-th BD can be represented as

$$P_{b,k}t_k + p_k\tau_k + p_{a,k}\tau_k \leq E_k = (\beta - t_k)\phi_k[P_0 g_k] + t_k\phi_k[P_0 g_k(1-\alpha_k)], \quad \forall k. \tag{19}$$

3.1.4. The Minimum Required SINR Constraint for Decoding x_p. This constraint is to ensure that the BR can decode x_p successfully. Without this constraint, the BR may not decode $x_{b,k}$, leading to $C_{b,k} = 0$. Thus, this constraint is necessary for the considered network. Accordingly, the minimum required SINR constraint for decoding x_p is given by

$$\gamma_{b,k}^p \geq \gamma_{th}, \forall k, \Leftrightarrow \frac{P_0 h_k}{\alpha_k P_0 g_k f_k + \sigma^2} \geq \gamma_{th}, \quad \forall k. \tag{20}$$

3.1.5. Transmit Power Constraint for the PT. Let $P_{\max}$ express the maximum allowed transmit power for the PT. Then, the transmit power constraint for the PT is given by

$$0 \leq P_0 \leq P_{\max}. \tag{21}$$

3.1.6. Throughput Fairness Optimization Problem. Based on (16), (18), (19), (20), and (21), the throughput fairness optimization problem is formulated as

$$\begin{aligned} \mathbf{P}_1: \max_{P_0, t_e, \beta, \{t_k\}_{k=1}^K, \{\alpha_k\}_{k=1}^K \{p_k\}_{k=1}^K, \{\tau_k\}_{k=1}^K} & \min_k C_{\text{tot}}^k \\ \text{s.t. C1: } & (18), \\ \text{C2: } & (19), \\ \text{C3: } & (20), \\ \text{C4: } & (21), \\ \text{C5: } & t_e + \textstyle\sum_{k=1}^K t_k \leq \beta, \sum_{k=1}^K \tau_k \leq T - \beta, \\ & 0 \leq \beta \leq T, t_e, t_k, \tau_k \geq 0, \quad \forall k, \\ \text{C6: } & 0 \leq \alpha_k \leq 1, \quad \forall k, \end{aligned} \tag{22}$$

where C1 expresses the QoS constraint for the primary transmission, C2 denotes the energy-causality constraint for each BD, C3 is the minimum required SINR for decoding x_p, C4 constrains the maximum transmit power of the PT, C5 is the constraint for the EH time, the BackCom time, etc., and C6 is the constraint for the power reflection coefficient of each BD.

As for $\mathbf{P}_1$, it is highly nonconvex and hard to solve. The reasons are as follows: firstly, the min function is involved in the objective function, which makes the objective function more complex and difficult to handle. Secondly, both the co-channel interference and the remaining part due to imperfect SIC exist, bringing the difference of convex (DC) structures in both the objective function and constraint C1 and leading to highly nonconvex objective function and C1. Thirdly, the use of the nonlinear EH model makes C2 more complex, which brings a new challenge to solve $\mathbf{P}_1$. Finally, except the above difficulties, several coupled relationships among multiple optimization variables exist, i.e., P_t, t_k, and α_k, leading to the nonconvex objective function and constraints, e.g., C1, C2, etc.

3.2. Solution. This subsection is provided to solve $\mathbf{P}_1$ efficiently. Firstly, in order to remove the min function in the objective function and simplify the objective function further, we introduce an auxiliary variable λ into $\mathbf{P}_1$ by letting $\lambda = \min_k C_{\text{tot}}^k$. Then, we can rewrite $\mathbf{P}_1$ as

$$\begin{aligned} \mathbf{P}_2: \max_{P_0, t_e, \beta, \lambda, \{t_k\}_{k=1}^K, \{\alpha_k\}_{k=1}^K \{p_k\}_{k=1}^K, \{\tau_k\}_{k=1}^K} & \lambda \\ \text{s.t. C1} - \text{C6}, & \\ \text{C7: } & C_{b,k} + C_{a,k} \geq \lambda. \end{aligned} \tag{23}$$

As for $\mathbf{P}_2$, it is still nonconvex since the DC structures, the nonlinear EH model, and the coupled relationships still exist. In order to handle the DC structure in $C_{b,k}$, we provide the following proposition to determine the optimal transmit power of the PT P_0^*.

Proposition 1. *In the considered network, the minimum throughput of each BD is maximized when the PT transmit its information with its maximum allowed transmission power, e.g.,* $P_0^* = P_{\max}$.

Proof. Here, we prove Proposition 1 by means of contradiction. We assume that $\{P_0^*, t_e^*, \beta^*, \lambda^*, \{t_k^*\}_{k=1}^K, \{\alpha_k^*\}_{k=1}^K, \{p_k^*\}_{k=1}^K, \{\tau_k^*\}_{k=1}^K\}$ is the optimal solution to $\mathbf{P}_2$, where both $P_0^* < P_{\max}$ and $\lambda^* = \min_k t_k^* W \log_2(1 + ((\xi \alpha_k^* P_0^* g_k f_k)/(\epsilon P_0^* h_k + \sigma^2))) + \tau_k^* W \log_2(1 + (p_k^* f_k/\sigma^2))$ hold. Then, another solution can be constructed, given by $P_0^+ = P_{\max}, t_e^+ = t_e^*, \beta^+ = \beta^*, t_k^+ = t_k^*, \alpha_k^+ = \alpha_k^*, p_k^+ = p_k^*, \tau_k^+ = \tau_k^*$. Obviously, the constructed solution is a feasible solution which satisfies all the constraints of $\mathbf{P}_2$. Accordingly, we can compute λ^+ as $\min_k t_k^+ W \log_2(1 + ((\xi \alpha_k^+ P_0^+ g_k f_k)/(\epsilon P_0^+ h_k + \sigma^2))) + \tau_k^+ W \log_2(1 + ((p_k^+ f_k)/\sigma^2))$. Since $P_0^+ = P_{\max} > P_0^*$ holds, we can prove that $\lambda^+ > \lambda^*$ is satisfied. The reasons are as follows: let $F_k(P_0) = t_k W \log_2(1 + ((\xi \alpha_k P_0 g_k f_k)/(\epsilon P_0 h_k + \sigma^2)))$.

Then, the first-order derivative of $F_k(P_0)$ with respect to P_0 is given by

$$\frac{\partial F_k(P_0)}{\partial P_0} = \frac{t_k W \xi \alpha_k f_k g_k \sigma^2}{\left(\epsilon P_0 h_k + \sigma^2\right)\left(\epsilon P_0 h_k + \sigma^2 + \xi \alpha_k f_k g_k P_0\right) \ln 2}. \tag{24}$$

Since $(\partial F_k(P_0)/\partial P_0) > 0$ always holds, $F_k(P_0)$ will increase with the increasing of P_0. That is to say, $\lambda^+ > \lambda^*$ holds, which contradicts the assumption that $P_0^* < P_{\max}$. Therefore, $P_0^* = P_{\max}$ holds when the minimum throughput of each BD is maximized for the considered network. Hence, the proof is complete.

Based on Proposition 1, we substitute $P_0 = P_{\max}$ into $\mathbf{P}_2$ and reformulate $\mathbf{P}_2$ as

$$\begin{aligned}
\mathbf{P}_3: \max_{t_e, \beta, \lambda, \{t_k\}_{k=1}^K, \{\alpha_k\}_{k=1}^K, \{p_k\}_{k=1}^K, \{\tau_k\}_{k=1}^K} \quad & \lambda \\
\text{s.t. C1}': \ & \sum_{k=1}^K t_k W \log_2\left(1 + \frac{P_{\max} h_p}{\alpha_k P_{\max} g_k f_{p,k} + \sigma^2}\right) \\
& + t_e W \log_2\left(1 + \frac{P_{\max} h_p}{\sigma^2}\right) \geq C_{\min}, \\
\text{C2}': \ & P_{\text{b},k} t_k + p_k \tau_k + p_{a,k} \tau_k \leq (\beta - t_k) \phi_k [P_{\max} g_k] \\
& + t_k \phi_k [P_{\max} g_k (1 - \alpha_k)], \forall k, \\
\text{C3}': \ & 0 \leq \alpha_k \leq \min\left(\frac{P_{\max} h_k - \gamma_{\text{th}} \sigma^2}{P_{\max} f_k g_k \gamma_{th}}, 1\right), \forall k, \text{C5}, \\
\text{C7}': \ & t_k W \log_2\left(1 + \frac{\xi \alpha_k P_{\max} g_k f_k}{\epsilon P_{\max} h_k + \sigma^2}\right) \\
& + \tau_k W \log_2\left(1 + \frac{p_k f_k}{\sigma^2}\right) \geq \lambda, \forall k,
\end{aligned} \tag{25}$$

where C3′ is the combination of C3 and C6.

It can be observed from $\mathbf{P}_3$ that $\mathbf{P}_3$ is still a nonconvex problem. To deal with the coupled relationships among different variables, i.e., α_k and t_k, p_k and τ_k, the following auxiliary variables, $x_k = \alpha_k t_k, y_k = p_k \tau_k, \forall k$, are introduced in $\mathbf{P}_3$ to replace variables $\alpha_k, p_k, \forall k$. Then, $\mathbf{P}_3$ is reformulated as

$$\begin{aligned}
\mathbf{P}_4: \max_{t_e, \beta, \lambda, \{t_k\}_{k=1}^K, \{x_k\}_{k=1}^K, \{y_k\}_{k=1}^K, \{\tau_k\}_{k=1}^K} \quad & \lambda \\
\text{s.t. C1}'': \ & \sum_{k=1}^K t_k W \log_2\left(1 + \frac{P_{\max} h_p t_k}{x_k P_{\max} g_k f_{p,k} + t_k \sigma^2}\right) \\
& + t_e W \log_2\left(1 + \frac{P_{\max} h_p}{\sigma^2}\right) \geq C_{\min}, \\
\text{C2}'': \ & P_{\text{b},k} t_k + y_k + p_{a,k} \tau_k \leq (\beta - t_k) \phi_k [P_{\max} g_k] \\
& + t_k \phi_k \left[\frac{P_{\max} g_k (t_k - x_k)}{t_k}\right], \forall k, \\
\text{C3}'': \ & 0 \leq x_k \leq t_k \times \min\left(\frac{P_{\max} h_k - \gamma_{\text{th}} \sigma^2}{P_{\max} f_k g_k \gamma_{th}}, 1\right), \forall k, \text{C5}, \\
\text{C7}'': \ & t_k W \log_2\left(1 + \frac{\xi x_k P_{\max} g_k f_k}{t_k\left(\epsilon P_{\max} h_k + \sigma^2\right)}\right) \\
& + \tau_k W \log_2\left(1 + \frac{y_k f_k}{\tau_k \sigma^2}\right) \geq \lambda, \forall k,
\end{aligned} \tag{26}$$

where $\alpha_k = (x_k/t_k), p_k = (y_k/\tau_k), \forall k$. □

Proposition 2. *As for* $\mathbf{P}_4$*, the objective function and all the constraints except* C1″ *are convex.*

Proof. It can be observed that the objective function and the constraint C3″ are linear, which are also convex. For the constraint C7″, using the fact that the perspective function can preserve convexity, we can conclude that the convexities of functions $t_k W \log_2(1 + ((\xi x_k P_{\max} g_k f_k)/(t_k(\epsilon P_{\max} h_k + \sigma^2))))$ and $\tau_k W \log_2(1 + (y_k f_k/\tau_k \sigma^2))$ are the same as those of $W \log_2(1 + ((\xi x_k P_{\max} g_k f_k)/(\epsilon P_{\max} h_k + \sigma^2)))$ and $W \log_2(1 + (y_k f_k/\sigma^2))$. Since both $W \log_2(1 + ((\xi x_k P_{\max} g_k f_k)/(\epsilon P_{\max} h_k + \sigma^2)))$ and $W \log_2(1 + (y_k f_k/\sigma^2))$ are concave functions, $t_k W \log_2(1 + ((\xi x_k P_{\max} g_k f_k)/(t_k(\epsilon P_{\max} h_k + \sigma^2))))$ and $\tau_k W \log_2(1 + (y_k f_k/\tau_k \sigma^2))$ are also concave. Thus, C7″ is a convex constraint.

For the constraint C2″, its convexity depends on the convexity of $t_k \phi_k[(P_{\max} g_k (t_k - x_k))/(t_k)]$. According to the perspective function, the convexity of $t_k \phi_k[(P_{\max} g_k (t_k - x_k))/(t_k)]$ is the same as that of $\phi_k[P_{\max} g_k (1 - x_k)]$. As pointed out in [6], $\phi_k[P_{\max} g_k (1 - x_k)]$ can be proved to be concave by using the properties of practical EH circuits. Thus, the constraint C2″ is convex.

Based on the above analysis, Proposition 2 is achieved, and the proof is complete.

We note that the nonconvexity of C1″ is due to the existence of the DC structure, i.e., $\sum_{k=1}^K t_k W \log_2(1 + ((P_{\max} h_p t_k)/(x_k P_{\max} g_k f_{p,k} + t_k \sigma^2)))$. To address this problem, we use the SCA method to deal with the nonconvexity of C1″. Specifically, we first replace $\sum_{k=1}^K t_k W \log_2(1 + (P_{\max} h_p t_k)/(x_k P_{\max} g_k f_{p,k} + t_k \sigma^2))$ with its first-order Taylor expression so that C1″ can be turned into

a linear constraint, which is always a convex constraint. Then, we can use the existing convex tools to solve the convex subproblem by changing $\mathbf{P}_4$ with the first-order Taylor expression. Finally, an efficient iterative algorithm is proposed based on the SCA method to solve $\mathbf{P}_4$, where the above subproblem is solved in each iteration.

Let $G_k[\alpha_k]$ denote $t_k W \log_2(1 + ((P_{\max} h_p)/(\alpha_k P_{\max} g_k f_{p,k} + \sigma^2)))$. Then, $\sum_{k=1}^{K} t_k W \log_2(1 + ((P_{\max} h_p t_k)/(x_k P_{\max} g_k f_{p,k} + t_k \sigma^2)))$ can be denoted by $\sum_{k=1}^{K} G_k[\alpha_k]$. By taking the first-order derivative of $G_k[\alpha_k]$ with respect to α_k, we have

$$\frac{\partial G_k[\alpha_k]}{\partial \alpha_k} = \frac{-P_{\max^2} h_p g_k f_{p,k} W t_k}{\left(\alpha_k P_{\max} g_k f_{p,k} + \sigma^2 + P_{\max} h_p\right)\left(\alpha_k P_{\max} g_k f_{p,k} + \sigma^2\right)\mathrm{In}2}. \tag{27}$$

Accordingly, the first-order Taylor expression of $G_k[\alpha_k]$ on a given value α_k^0 can approximate $G_k[\alpha_k]$ as

$$\begin{aligned} G_k[\alpha_k] &\approx \frac{\partial G_k\left(\alpha_k^0\right)}{\partial \alpha_k^0}\left(\alpha_k - \alpha_k^0\right) + G_k\left[\alpha_k^0\right] \\ &= \frac{-P_{\max^2} h_p g_k f_{p,k} W t_k\left(\alpha_k - \alpha_k^0\right)}{\left(\alpha_k^0 P_{\max} g_k f_{p,k} + \sigma^2 + P_{\max} h_p\right)\left(\alpha_k^0 P_{\max} g_k f_{p,k} + \sigma^2\right)\mathrm{In}2} + G_k\left[\alpha_k^0\right] \\ &= \frac{-P_{\max^2} h_p g_k f_{p,k} W\left(x_k - \alpha_k^0 t_k\right)}{\left(\alpha_k^0 P_{\max} g_k f_{p,k} + \sigma^2 + P_{\max} h_p\right)\left(\alpha_k^0 P_{\max} g_k f_{p,k} + \sigma^2\right)\mathrm{In}2} + G_k\left[\alpha_k^0\right], \end{aligned} \tag{28}$$

where α_k^0 will be updated in each iteration based on α_k obtained in the previous iteration.

Based on (28), the following subproblem can be obtained from $\mathbf{P}_4$, given by

$$\begin{aligned} \mathbf{P}_5: \max_{t_e,\beta,\lambda,\{t_k\}_{k=1}^K,\{x_k\}_{k=1}^K,\{y_k\}_{k=1}^K,\{\tau_k\}_{k=1}^K} \quad & \lambda \\ \text{s.t. } \mathrm{C1}''': \ & \sum_{k=1}^{K} \frac{-P_{\max^2} h_p g_k f_{p,k} W\left(x_k - \alpha_k^0 t_k\right)}{\left(\alpha_k^0 P_{\max} g_k f_{p,k} + \sigma^2 + P_{\max} h_p\right)\left(\alpha_k^0 P_{\max} g_k f_{p,k} + \sigma^2\right)} \\ & \times \frac{1}{\mathrm{In}2} + G_k\left[\alpha_k^0\right] + t_e W \log_2\left(1 + \frac{P_{\max} h_p}{\sigma^2}\right) \geq C_{\min}, \\ & \mathrm{C2}'', \mathrm{C3}'', C7''. \end{aligned} \tag{29}$$

□

Proposition 3. *$\mathbf{P}_5$ is convex and can be efficiently solved by using the existing convex tools.*

Proof. It can be observed that $\mathrm{C1}'''$ is a linear constraint, which is obviously convex. Combining with Proposition 2, $\mathbf{P}_5$ is convex, which can be efficiently solved by using the existing convex optimization tools. □

3.3. Iterative Algorithm. In this subsection, we propose an efficient iterative algorithm to solve $\mathbf{P}_4$, as shown in Algorithm 1. In particular, the subproblem $\mathbf{P}_5$ is optimally solved under given $\alpha_k^0, \forall k$ in each iteration, and the values of $\alpha_k^0, \forall k$ are updated based on the optimal power reflection coefficients $\alpha_k^*, \forall k$ obtained in the previous iteration. We note that for the first iteration, the values of $\alpha_k^0, \forall k$ are predefined. The optimal solution to $\mathbf{P}_4$ is achieved when the algorithm converges, namely, the stop condition $|\alpha_k^* - \alpha_k^0| \leq \varepsilon$ with the maximum tolerance ε is satisfied.

We provide the analysis of the computational complexity of Algorithm 1 as follows: we assume that the interior point method is applied to solve $\mathbf{P}_5$ with given $\alpha_k^0, \forall k$. Let m_1 and N_u denote the number of the inequality constraints of $\mathbf{P}_5$ and the number of iterations for Algorithm 1, respectively. Then, the computational complexity of Algorithm 1 can be calculated as $N_u O(\sqrt{m_1} \log(m_1))$ [16].

Algorithm 1: An efficient iterative algorithm for solving $\mathbf{P}_4$.
(1) Set the maximum tolerance ε and the maximum number of iterations $I_{\max}$;
(2) Set the iteration index $i = 1$ and the initial given values $\alpha_k^0, \forall k$;
(3) **repeat**
(4) Solve $\mathbf{P}_5$ with given $\alpha_k^0, \forall k$ via CVX, to obtain its optimal solution, denoted by $\left\{t_e^*, \beta^*, \lambda^*, \{t_k^*\}_{k=1}^K, \{x_k^*\}_{k=1}^K, \{\tau_k^*\}_{k=1}^K, \{y_k^*\}_{k=1}^K\right\}$;
(5) Compute α_k^* as $(x_k^*/t_k^*), \forall k$ and p_k^* as $(y_k^*/\tau_k^*), \forall k$;
(6) **if** $|\alpha_k^* - \alpha_k^0| \le \varepsilon$ **then**
(7) Set Flag = 1;
(8) **else**
(9) Set Flag = 0 and $i = i + 1$;
(10) Update α_k^0 as $\alpha_k^0 = \alpha_k^*, \forall k$;
(11) **end if**
(12) **until** Flag = 1 or $i = I_{\max}$.
(13) Output the optimal solution to $\mathbf{P}_4$ as $\left\{t_e^*, \beta^*, \lambda^*, \{t_k^*\}_{k=1}^K, \{\alpha_k^*\}_{k=1}^K, \{\tau_k^*\}_{k=1}^K, \{p_k^*\}_{k=1}^K\right\}$.

Algorithm 1

4. Numerical Results

In this section, we use computer simulations to verify the superiority of our proposed resource allocation scheme and the effectiveness of the proposed algorithm. Unless otherwise specified, the basic simulation parameters are set, as shown in Table 1 [11, 13]. According to [14], we set $a_k = 2.463$, $d_k = 1.635$, and $v_k = 0.826, \forall k$ to characterize the used nonlinear EH model clearly. Besides, all the channels are set as follows: Following the standard channel fading model, the channel gain is the product of the small-scale fading and the large-scale fading. Let g_k', f_k', h_k', $f_{p,k}'$,and h_p' denote the small-scale fadings of the PT-the k-th BD link, the k-th BD-its receiver link, the PT-the k-th BR link, the k-th BD-the PR link, and the PT-PR link, respectively. We denote $D_{1,k}$, $D_{2,k}$, $D_{3,k}$, $D_{4,k}$, and D_p as the distances of the PT-the k-th BD link, the k-th BD-its receiver link, the PT-the k-th BR link, the k-th BD-the PR link, and the PT-PR link, respectively. Then, we have $g_k = g_k' D_{1,k}^{-\zeta}$, $f_k = f_k' D_{2,k}^{-\zeta}$, $h_k = h_k' D_{3,k}^{-\zeta}$, $f_{p,k} = f_{p,k}' D_{4,k}^{-\zeta}$, and $h_p = h_p' D_p^{-\zeta}$, where ζ denotes the path loss exponent. Here, we set $\zeta = 3$, $D_{1,1} = 11$ m, $D_{1,2} = 12$ m, $D_{1,3} = 15$ m, $D_{1,4} = 14$ m, $D_{2,1} = 40$ m, $D_{2,2} = 35$ m, $D_{2,3} = 32$ m, $D_{2,4} = 35$ m, $D_{3,1} = 35$ m, $D_{3,2} = 32$ m, $D_{3,3} = 40$ m, $D_{3,4} = 35$ m, $D_{4,1} = 50$ m, $D_{4,2} = 55$ m, $D_{4,3} = 53$ m, $D_{4,4} = 52$ m, and $D_p = 30$ m.

In order to illustrate the superiority of the proposed scheme, we compare the performance under the proposed scheme with that of the following benchmark schemes: backscatter-assisted cognitive networks and wireless powered cognitive networks. For the backscatter-assisted cognitive networks, each BD only backscatters its information to its receiver, while for the wireless powered cognitive networks, each BD first harvests energy from the PT's signals when PT is busy and then uses its harvested energy to transmit its information to its receiver. We note that both backscatter-assisted cognitive networks and wireless powered cognitive networks can be regarded as special cases for the considered network and can be obtained after a few changes on the considered network. For example, let $p_k = 0$ and $\tau_k = 0$ and the backscatter-assisted cognitive networks can be achieved. That is to say, the proposed algorithm can also be used to obtain the optimal schemes under backscatter-assisted cognitive networks and wireless powered cognitive networks.

Figure 3 shows the convergence of Algorithm 1, where different settings of $P_{\max}$ are considered. Here, we set $P_{\max}$ as 0.5 W, 0.8 W, and 1 W, respectively. It can be seen that Algorithm 1 can always converge to a certain value after only a few iterations, i.e., 2 iterations, which illustrates the convergence of Algorithm 1 and shows that Algorithm 1 is computationally efficient. Besides, we can also see that with a larger $P_{\max}$, the minimum throughput achieved by BDs also increases. This is because the optimal transmit power of the PT is determined by $P_{\max}$ and a larger P_0 brings a higher throughput.

Figure 4 shows the minimum throughput achieved by BDs versus the maximum allowed transmit power of the PT $P_{\max}$, where $P_{\max}$ is varied from 0.5 W to 2.5 W. In order to demonstrate the superiority of the proposed scheme, we compare the performance under the proposed scheme with that of backscatter-assisted cognitive networks and wireless powered cognitive networks. It can be seen that the minimum throughput achieved by BDs increases with the increasing of $P_{\max}$ since a larger $P_{\max}$ means a higher P_0, which brings a higher throughput for each BD. Besides, comparing with backscatter-assisted cognitive networks and wireless powered cognitive networks, we can also find that the proposed scheme outperforms the other schemes as the proposed scheme has more flexibility to use resources efficiently, which also illustrates the advantages of the proposed scheme.

Figure 5 shows the minimum throughput among BDs versus the minimum required throughput for the PT, $C_{\min}$, where $C_{\min}$ is ranged from 1 Mbyte to 1.5 Mbytes. From this figure, it can be seen that the minimum throughput achieved by BDs decreases with the increasing of $C_{\min}$, as a larger $C_{\min}$ means a higher QoS requirement for the PT's transmission, and more resources will be allocated to the PT, leading to a reduction to the throughput achieved by each BD. By comparisons, we can see that the proposed scheme can achieve higher throughput than the other schemes, verifying the advantages of the proposed scheme.

Table 1: Basic simulation settings.

Parameters	Notation	Value
The entire time block	T	1 s
The system bandwidth	W	100 kHz
The constant circuit power consumption for BackCom at the k-th BD	$P_{b,k}$	10 μW
The constant circuit power consumption for active communications at the k-th BD	$p_{a,k}$	1 mW
The maximum transmit power at the PT	$P_{\max}$	1 W
The performance gap reflecting the real modulation for BackCom	ξ	−15 dB
The noise power	σ^2	−60 dBm
The number of BDs	K	4
The minimum required throughput of the PT	$C_{\min}$	1000 kbytes
The threshold required for decoding x_p	γ_{th}	20

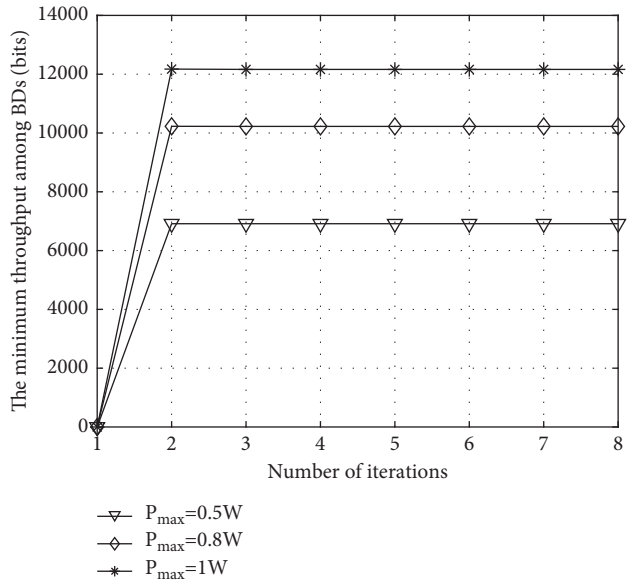


Figure 3: The convergence of Algorithm 1 under different settings of $P_{\max}$.

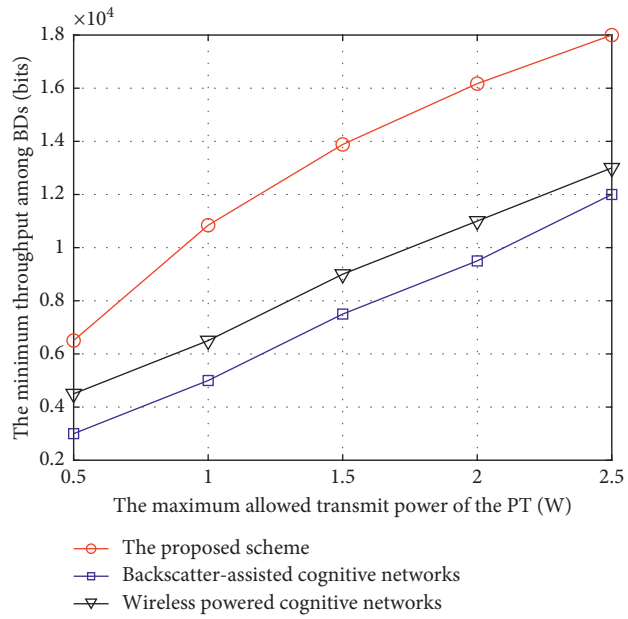


Figure 4: The minimum throughput among BDs versus the maximum allowed transmit power of the PT.

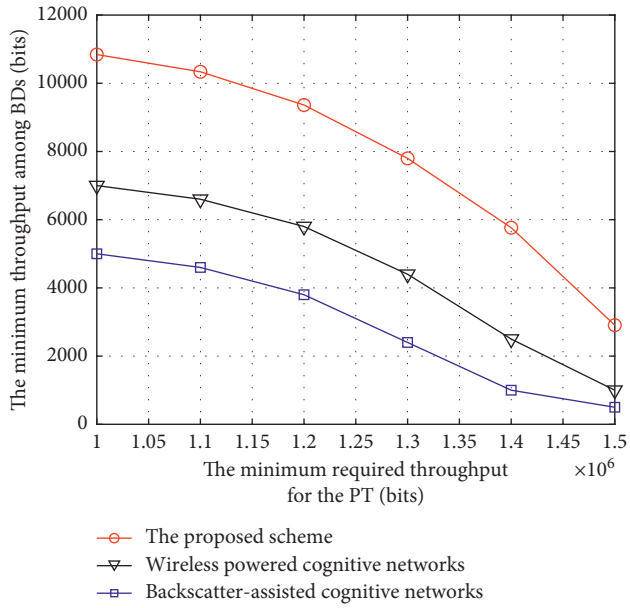


Figure 5: The minimum throughput among BDs versus the minimum required throughput for the PT.

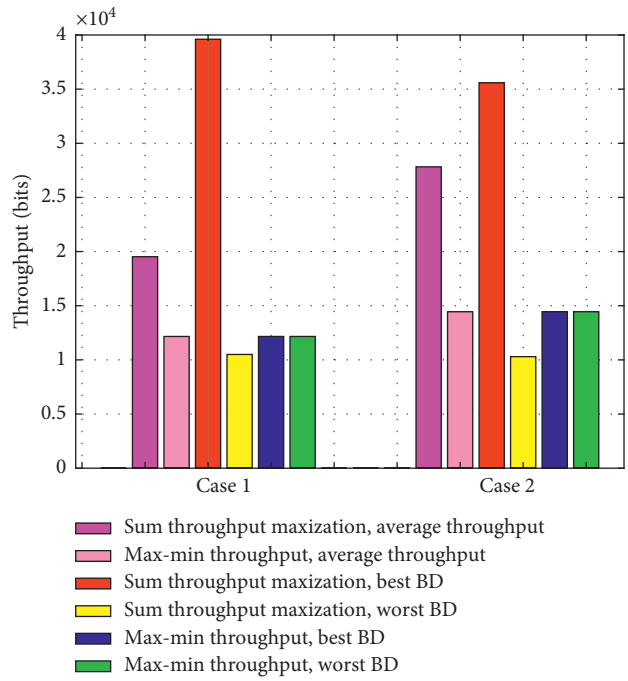


Figure 6: Fairness comparison.

Figure 6 compares the BD fairness achieved by the proposed scheme, denoted by max-min throughput, and the sum throughput maximization scheme, denoted by sum throughput maximization, under two cases, where the settings of the channels are different. It can be seen that there exists a tradeoff between the sum throughput maximization and the max-min throughput. Specifically, the proposed scheme can greatly improve the fairness among BDs while the average throughput among all BDs is reduced. This is because, for the proposed scheme, more resources will be allocated to the BD with a worst channel condition for good throughput fairness, while for the sum throughput

maximization, more resources will be allocated to the BD with a better channel condition for achieving the maximum throughput of all BDs.

5. Conclusions

In this paper, we have studied the throughput fairness for the wireless powered cognitive hybrid active-passive communication network while considering a nonlinear EH model. In particular, we have formulated an optimization problem to maximize the minimum throughput that achieved all BDs by jointly optimizing the transmit power and time of the PT, the BDs' time sharing among EH, BackCom and active communications, and the power reflection coefficient and transmit power of each BD subject to the QoS, energy causality, transmit power constraints, etc. In order to solve this problem, the optimal transmit power of the PT was firstly achieved by means of contradiction, and then an efficient iterative algorithm was developed to obtain the optimal solutions. Simulation results verified the quick convergence of the proposed algorithm and demonstrated the superiority of the proposed scheme in terms of throughput fairness.

Data Availability

The simulation data used to support the findings of this study are included within the article. The MATLAB code used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the scholarship from China Scholarship Council (no. 201708230301), the Science Foundation of Heilongjiang Province for the Excellent Youth (no. YQ2019F014), the Science Talent Support Program of Heilongjiang Bayi Agricultural University (no. ZRCQC201807), and the Scientific Research Foundation for Doctor of Heilongjiang Bayi Agricultural University (no. XDB2015-28).

References

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: a survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[2] J. M. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, 2009.

[3] Y. Xu, H. Sun, and Y. Ye, "Distributed resource allocation for swipt-based cognitive ad-hoc networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 4, pp. 1320–1332, 2021.

[4] A. Costanzo, D. Masotti, G. Paolini, and D. Schreurs, "Evolution of SWIPT for the iot world: near- and far-field solutions for simultaneous wireless information and power transfer," *IEEE Microwave Magazine*, vol. 22, no. 12, pp. 48–59, 2021.

[5] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, p. 1, 2021, https://ieeexplore.ieee.org/document/9562293.

[6] Y. Ye, L. Shi, X. Chu, and G. Lu, "Throughput fairness guarantee in wireless powered backscatter communications with HTT," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 449–453, 2021.

[7] T. D. Tran and L. B. Le, "Hybrid backscatter and underlay transmissions in RF-powered cognitive radio networks," in *Proceedings of the 2019 26th International Conference on Telecommunications (ICT)*, pp. 11–15, Hanoi, Vietnam, April 2019.

[8] X. Lu, D. Niyato, H. Jiang, D. I. Kim, Y. Xiao, and Z. Han, "Ambient backscatter assisted wireless powered communications," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 170–177, 2018.

[9] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Ambient backscatter: a new approach to improve network performance for RF-powered cognitive radio networks," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3659–3674, 2017.

[10] D. T. Hoang, D. Niyato, P. Wang, and D. I. Kim, "Optimal time sharing in rf-powered backscatter cognitive radio networks," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, France, May 2017.

[11] B. Lyu, H. Guo, Z. Yang, and G. Gui, "Throughput maximization for hybrid backscatter assisted cognitive wireless powered radio networks," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2015–2024, 2018.

[12] R. Kishore, S. Gurugopinath, P. C. Sofotasios, S. Muhaidat, and N. Al-Dhahir, "Opportunistic ambient backscatter communication in rf-powered cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 413–426, 2019.

[13] L. Shi, R. Q. Hu, J. Gunther, Y. Ye, and H. Zhang, "Energy efficiency for RF-powered backscatter networks using HTT protocol," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13932–13936, 2020.

[14] Y. Chen, N. Zhao, and M.-S. Alouini, "Wireless energy harvesting using signals from multiple fading channels," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 5027–5039, 2017.

[15] L. Tassiulas and S. Sarkar, "Maxmin fair scheduling in wireless ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 163–173, 2005.

[16] S. Boyd, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7061573, 9 pages
https://doi.org/10.1155/2022/7061573

Research Article

Multiple Prime Expansion Channel Hopping for Blind Rendezvous in a Wireless Sensor Network

Zhou Zhixin[1], **Yanjun Deng**[1], **Zhang Xiaohong**[1], **Zhang Xianfei**[1], **Hu Liqin**[2] and **Zhao Zhidong**[3]

[1]*School of Electronic Information, Hangzhou Dianzi University, Zhejiang, Hangzhou 310035, China*
[2]*Zhejiang College of Construction, Hangzhou, China*
[3]*School of Cyberspace, Hangzhou Dianzi University, Zhejiang, Hangzhou 310035, China*

Correspondence should be addressed to Zhao Zhidong; zhaozd@hdu.edu.cn

Received 24 November 2021; Accepted 15 January 2022; Published 9 February 2022

Academic Editor: Liqin Shi

A channel rendezvous is a significant aspect of communication. In this context, blind rendezvous is the process of selecting a common available channel and establishing a communication link for wireless devices in a wireless sensor network. The rendezvous of asymmetric and heterogeneous wireless devices is a challenge. Thus, to improve speeds and stability of rendezvous, we analyze time slot overlap and channel determinism in the rendezvous algorithm and propose a rendezvous algorithm named Multiple Prime Expansion (MPE). In a final simulation study, we compare the performance of the MPE with other existing algorithms in an asymmetric and heterogeneous scenario. Results show that MPE has excellent performance for the ATTR and MTTR.

1. Introduction

In recent years, numerous cities in China have built an Internet of Things projects to improve the mobility and carrying capacity of data in public fields. These projects can boost the speed of data collection and optimize city management and services [1]. Rail transportation services use data to adjust station populations and smartphone applications to improve the efficiency of targeted services based on user access data [2]. At the same time, cities protect the supply of water, food, and energy through billing data [3–6].

Currently, more and more data collection systems are based on Wireless Sensor Networks (WSNs) [7]. Thus, compared to wired transmission systems, WSN enhances flexibility but requires more resilient energy and timeliness cost. Energy consumption and data timeliness have important implications for WSNs.

Routing protocols are widely used in wireless surveillance as tools that can organize wireless nodes to collect information in an orderly manner [8]. A number of researchers have improved routing protocols to reduce wireless monitoring energy consumption. These are divided into several groups, such as node deployment, clustering techniques, and transmission methods [9, 10]. This method can be applied for both applications and engineering [11].

The connection speed of sensors is an important variable in data timeliness research. The rapid increase in sensor integration [12] and data demand greatly enhanced the demand for the number of wireless devices in WSN and hence wireless spectrum. This has led to a dramatic increase in demand for wireless spectra and so this has become a scarce resource. However, spectrum scarcity and low spectrum utilization persist in rendezvous for wireless devices has proven challenging in resource-scarce settings [13]. Thus, to achieve efficient data transmission among sensors, communicating parties must adopt an efficient and robust mechanism to complete data interaction. Blind channel rendezvous, without a common control channel (CCC), has garnered a growing interest [14]. Channel hopping is a typical technique used in most blind channel rendezvous. Wireless devices access the channels in different time slots according to a predefined frequency hopping sequence, and

rendezvous is successfully achieved when the two users access the same common available channel in a certain time slot.

Time to rendezvous (TTR) is also a key factor. However, early in development, some devices initially make random sequences to blind rendezvous, which would make rendezvous time unpredictable. Therefore, during the operation of the wireless device, we needed to identify a precondition that contained enough frequency of rendezvous for data transmission. Wireless devices needed to exchange data multiple times during operation, which made minimizing TTR a direction in research. For researchers, any channels between users can directly be used as a sequence without any preprocessing. In other words, the effectiveness of the channel sequence directly affects the stability of rendezvous. Details of the technical specifications and requirements for the algorithm are provided as follows.

(1) Degree of rendezvous (DR): This can be defined by dividing the number of channels that have completed rendezvous by the total number of channels. Thus, when DR = 1, this is referred to as complete rendezvous. This means that all channels in the sequence can be rendezvous, and high DR can avoid channel blocking and improve channel utilization.

(2) Frequency of rendezvous (FR): The number of rendezvous times divided by the frequency hopping slot is FR. FR describes the strength of rendezvous across the whole frequency hopping cycle.

(3) Average TTR (ATTR): ATTR refers to the average time between two consecutive rendezvous in the frequency hopping cycle, which is the average value of the rendezvous time. ATTR describes the speed of the rendezvous.

(4) Maximum TTR (MTTR): The longest time period from start to rendezvous is MTTR. Thus, if the sequence cannot guarantee rendezvous, MTTR is infinite. In other words, as the length of the hopping sequence increases, the MTTR is increasing obviously.

Our current approach meets these four above requirements and suggests interesting directions for research and applications. According to starting time, the hopping sequences can be classified as synchronous sequences and asynchronous sequences. Lin et al. proposed a multiradio channel-hopping scheme (CHS) that preserves network connectivity and synchronous sequences [15]. Another algorithm, ASCH (asymmetric synchronous channel hopping), divides the whole sets of channels into several levels to meet asynchronous sequences [16]. However, in most cases, the wireless devices have local clocks that make it difficult to achieve synchronization. Some asynchronous algorithms have been proposed, such as in the work of Liu, who considered the impact of network factors (channel availability and multiuser contention) when designing the frequency hopping sequence [17]. Wang proposed MAAPS based on rendezvous-success rate and variance [18].

Hopping sequences can be divided into homogeneous sequences and heterogeneous sequences based on different available channels [19, 20]. A homogeneous model includes symmetric/asymmetric channel sets such as GOS [21] and CRESQ [22].

We, therefore, investigate the two requirements in WSN for the rendezvous algorithm. On the one hand, there are various sensors with different monitoring data types and frequencies in WSN [23], which means that the available channels are heterogeneous. On the other, as data fusion is an essential requirement for WSN, hopping sequences are required to meet the asymmetric condition. Therefore, the nonpreprocessing rendezvous algorithm used in WSN needs to satisfy symmetric and asymmetric requirements on a heterogeneous basis.

The above analysis suggests that the investigation of the rendezvous algorithm is essential. Only a little information is currently available about the asymmetric and heterogeneous rendezvous algorithm of WSN. In this work, we propose a MPE (Multiple Prime Expansion) algorithm to match asymmetric and heterogeneous requirements. The MPE examined the feedback relationship between the length of a sequence and TTR to strengthen the discrimination subsequences. The rendezvous algorithm switches to another operation mode when a subsequence reaches certain conditions.

2. Problem Statement

The time slot communication system was measured using unit time, which includes time slots, channels, and frequency hopping rules. The slot in slot communication system corresponds to a minimum time interval to select a communication link in the communication network. Another study focused on channel frequency hopping sequence, which can be regarded as a specific access sequence for a rendezvous algorithm. In each process, wireless devices access the designated channel by channel frequency hopping sequence to establish contact with other devices on each channel. In this study, we introduce several difficulties relating to channel hopping technology in WSN.

2.1. Limitations of WSN Conditions. Wireless sensors were analyzed with channel information connected to other wireless devices via antennae and matched with local channels. There are, however, some difficulties in using the rendezvous algorithm in WSN compared with other fields. These existing rendezvous algorithms are not yet efficient enough to be used in WSN due to special constraints. The main constraints on the frequency hopping sequences are shown to be as follows.

(1) The rendezvous rules cannot be easily changed during frequency hopping, where WSNs are usually incorporated into a system together with servers, mobile applications, and databases [24].

(2) According to its different functions, different sensor matching schemes must be designed. In detail, WSNs contain various types of sensors, like power,

telecommunications, water supply, and natural gas sensors. It causes a heterogeneous relationship between wireless devices. Thus, symmetric and asymmetric requirements need to be considered in rendezvous algorithm design.

(3) Distinct working frequencies develop at various times under different types of sensors. Thus, symmetric and asymmetric requirements need to be considered in the rendezvous algorithm.

2.2. Problem Formulation. We assume that there are two wireless devices to rendezvous. In our sequence, each device is equipped with one antenna, which only tries to rendezvous with one device in a time slot. Thus, let $S_1(i)$ and $S_2(i)$ denote the sequences of two wireless devices in the rendezvous algorithm and let $j \in \{1, 2, \ldots, j, \ldots, S\}$ denote the rendezvous time from slot 1 to slot S, where S denotes the total number of slots. When the rendezvous occurs at slot j.

This means that the rendezvous problem in WSN can be formulated as follows:

$$\min_j \left(\sum_{i=1}^{j} \operatorname{sign}\left(\left| S_1(i) \right| - \left| S_2(i) \right| \right) \right). \quad (1)$$

3. Materials and Methods

We assessed the challenges above using the improved algorithm MPE as detailed below.

As WSN technology use has increased, more and more research has focused on the timeliness of data exchange. The use of the rendezvous algorithm is key to this task. Two issues need to be considered, however, when designing a rendezvous algorithm: (A) rendezvous within a limited time and (B) a reduction in rendezvous time.

The MPE rendezvous algorithm enables the inevitability of the rendezvous and reduces rendezvous time which enables rapid data exchange for the WSN. The algorithm consists of five phases: (a) given global channels; (b) any two wireless devices have the possibility of rendezvous (common channel); (c) achieve inevitability of rendezvous; (d) create subsequences; (e) subsequence is expanded into a frequency hopping sequence.

3.1. Network Model. We have made some assumptions based on the above constraints caused by WSN.

(1) We assume multiple nonoverlapping channels and each wireless device can sense a part of the channel.

(2) We allow each wireless device to be equipped with only one cognitive radio transceiver (antenna), which can sense the channel status and switch to different channels without auxiliary means.

(3) All wireless devices are anonymous in the network (without ID).

(4) Each wireless device has a local clock.

(5) The sensing device includes a network mapping function. Different wireless terminals use a channel of a unified set of channel indexes. However, the common mapping function is outside the scope of this paper.

(6) The length of each slot is the same.

(7) We allow M wireless devices. Thus, if we let $M_u (M_u \subseteq M, M_u \neq \varnothing)$ denote the available channel set of wireless devices u. For any two wireless devices u and v, their locally available channel sets C_u and C_v may have different channels. In order to ensure the rendezvous, we assume that any two nodes have the possibility of rendezvous. That is, there is at least one publicly available channel $(\forall u, v \in M, C_u \cap C_v \neq \varnothing)$.

3.2. Slot Symmetry. Channel hopping sequences are usually periodic. The smallest repeating unit of the channel hopping sequence is called the subsequence. As the rule for the construction of the channel hopping sequence is fixed, the structure in each subsequence has been restricted. The intersection of any two sequences shows periodicity with the repetition of the subsequence. Therefore, the rendezvous algorithm is the generation and connection method of the subsequence with rendezvous ability.

According to the start time of channel hopping sequences, the channel hopping sequences were divided into synchronous sequences and asynchronous sequences, as shown in Figure 1.

An asynchronous channel hopping sequence presents two different asynchronous situations, as shown in Figures 2(a) and 2(b).

The asynchronous time slot communication system is shown in Figure 2. The time slots in Figure 2(a) are P, and the two sequences are separated by $T2$ time slots. The time slots in Figure 2(b) are Q, and the gap of the start time of the two sequences in the figure is $(T2 + k)$ time slots.

It is clear that in Figure 2(b), $T2$ is a nonnegative integer, $k \in [0, 1]$, the time slots of these two sequences are not aligned. As there is no time slot alignment information, the wireless devices decide the moment when the wireless devices start to rendezvous. In other words, the rendezvous of the asynchronous time slot system is more complicated. Therefore, the asynchronous nonaligned time slot communication system can be transformed into an aligned time slot communication system.

As the communication system of the asynchronous time slot satisfies $Q = 2P$, an asymmetric asynchronous time slot can be considered a symmetric slot communication system.

We assume that the minimum time interval required for any two wireless devices to establish a communication link is Q time slots. In an asynchronous communication system, the overlap range of rendezvous time slots can be divided into the following two situations.

Situation 1: the range of k is $[0, Q/2)$, as shown in Figure 3:

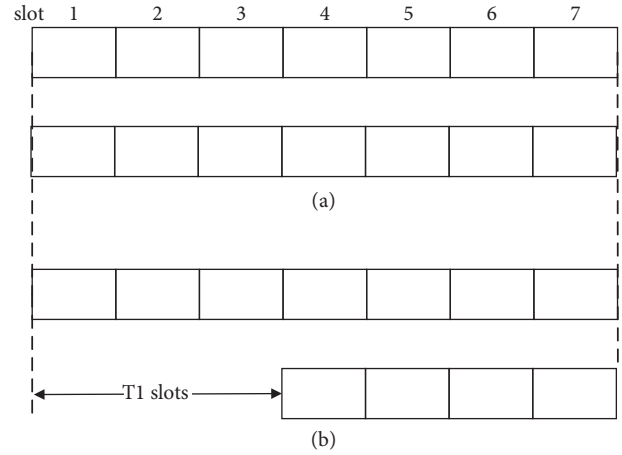


Figure 1: Synchronous/asynchronous channel hopping sequences.

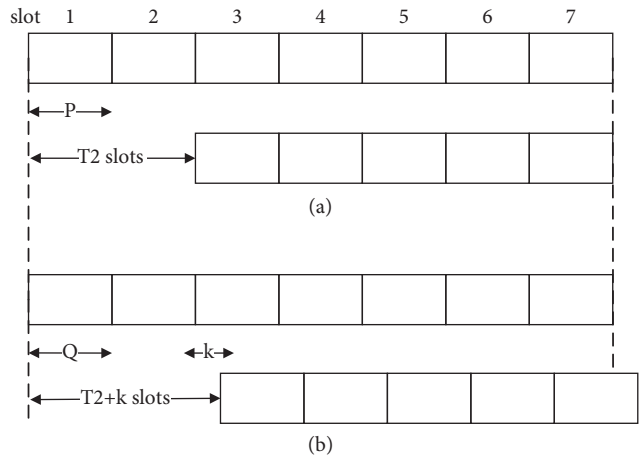


Figure 2: Asynchronous aligned/nonaligned time slots.

$$Q - k = 2P - k \geq P. \tag{2}$$

The length of the time slot shown in red in Figure 3 can satisfy the asynchronous channel hopping sequences to complete rendezvous.

Situation 2: The range of k is $[0, Q/2)$, as shown in Figure 4:

$$Q - (2P - k) = k \geq P. \tag{3}$$

The length of the time slot shown in red in Figure 4 also can allow asynchronous channel hopping sequences to complete rendezvous.

Section 3.2 verifies that the asymmetric time slot communication sequence can be regarded as the symmetric channel hopping sequence for rendezvous. In other words, when time slots in the network all have the same length, the time required to establish a communication link is at least two time slots.

3.3. Algorithm Processes. As mentioned in Section 2, the channel frequency hopping sequence is limited by the heterogeneity of wireless devices. Otherwise, it will waste a lot of time that uses the traditional rendezvous algorithm. As there are some differences between the available channels of wireless devices and the total number of channels. It is, therefore, crucial to design a suitable rendezvous algorithm for asymmetric channel environments. The algorithm model is shown in Figure 5.

The MPE algorithm needs to initially rendezvous in a limited time. Since the number of available channels for each wireless device is different, the channel frequency hopping sequence of the MPE algorithm is unequal in different devices. We assume that D and E (D and E are both periodic sequences), respectively, represent the channel sequences of two wireless devices. D is composed of d cycle of subsequence, and E is composed of the base sequence e cycle. x and y are coprime numbers, and both x and y are not less than the number of available channels, and c is the total number of channels. The verification of the inevitability of rendezvous in MPE is as follows.

$D = \{d_1, d_2, d_3, \ldots, d_y\}$ means the sequence containing y repeating subsequences $d = \{c_1, c_2, c_3, \ldots, c_x\}$. $E = \{e_1, e_2, e_3, \ldots, e_y\}$ means the sequence containing y repeating subsequences $e = \{c_1, c_2, c_3, \ldots, c_y\}$.

Assume that any element c in d (d is a subsequence of D) corresponds to c_t ($c_t \in E$). Then in the mth and nth subsequences, c corresponds to $c(t + mx)\%y$ and $c(t + nx)\%y$ in E, respectively. Thus, if

$$c(t + mx)\%y = c(t + nx)\%y, \tag{4}$$

then $[(m - n)x]\%y = 0$ can be derived. That is, $(m - n)x/y = k$, k is a natural number.

As x and y are prime numbers and $x \neq yx \neq y$, their relationship is elucidated by formula (5), as follows:

$$m - n = qy\,(q = 1, 2, 3, \ldots) \tag{5}$$

This can be established by calculating $(m - n)x/y = k$ only when $k = x$. Therefore, in the following y-1 cycles, $\forall c \in d$ will correspond to y different elements in E. If $\exists c \in D, c \in E$, the sequences D and E will inevitably meet within y cycles, and, therefore, q is a natural number.

Thus, if we assume that subsequences of D and E have common channels and the lengths are different prime numbers, there are rendezvous between D and E. When x and y are different prime numbers, there is $\forall c_i \in d$ corresponding to all elements in e. This means that sequences D and E will achieve rendezvous during a finite time.

Subsequent to research into rendezvous inevitability, more recently, the focus has shifted to reducing rendezvous time between two wireless devices. Since the rendezvous time depends on the channel size (number of channels), the aspect of reducing the rendezvous time based on global channels was the focus of the present work.

Assuming that the number of global channels is M. There are two wireless devices waiting to rendezvous in the network. Each wireless device only contained one antenna, and each antenna has only accessed a channel in a time slot. The available channels of the two wireless devices all belong to global channels. Let A and B be the number of available channels of the two wireless devices, respectively. In order to

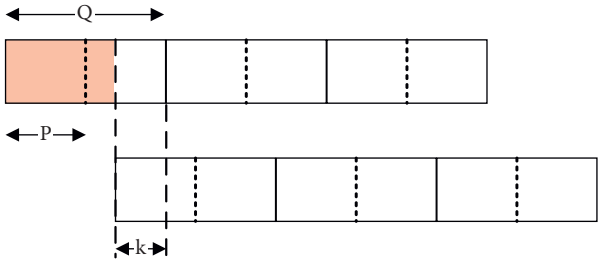


Figure 3: Asynchronous nonaligned-situation 1.

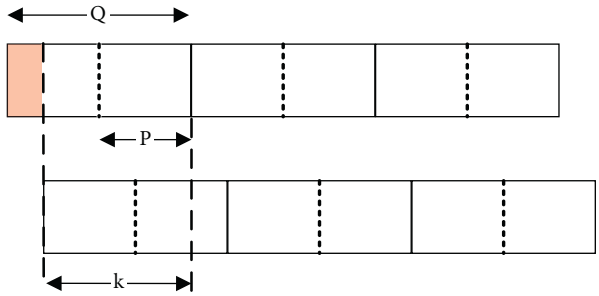


Figure 4: Asynchronous nonaligned-situation 2.

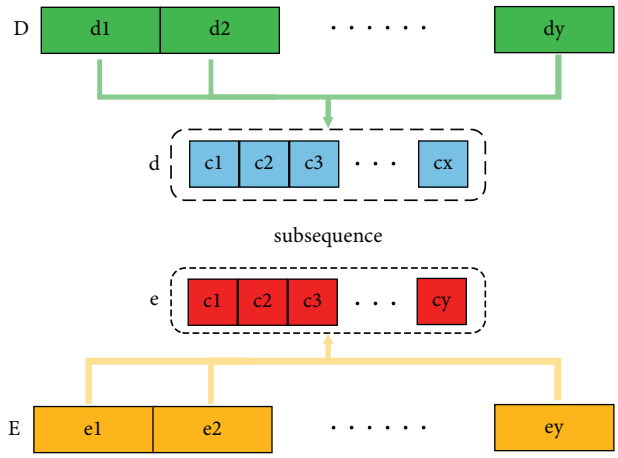


Figure 5: Common channel diagram.

achieve rendezvous, it is necessary to have at least one common channel between two wireless devices, as shown in Figure 6.

According to Section 3.3, when the length of the subsequences is different prime, the two sequences can achieve rendezvous in a limited time. Therefore, A and B need to be expanded into LA and LB with prime lengths. That is to say, both LA and LB are composed of two parts, the "original part" (A or B) and the "fill part" (FA or FB). The goal of the present work was to increase the probability of the common channel in the "fill part"(FA or FB). Different fill part conditions are caused by different lengths of original parts. Different methods to design "fill part" can be caused by different lengths of "original parts." The "fill part" selects at least one channel number from the original channel (A or B). Then when the "original part" is expanded into a subsequence with a prime number, the probability that a single subsequence contains a common channel can be increased. Two cases are considered in the following research to design the "fill part."

In the first case, when $A \neq B$. We let p denotes the probability that the "original part" contains common channels, and m denote the length of the "fill part". The "fill part" is generated from the "original part." The "fill part" contains the probability $p_{A\neq B}$ of the common channels, and the calculation process is shown in the following formula:

$$p_{A\neq B} = \frac{1}{2^m} \sum_{i=0}^{m} pC_m^i (1-p) C_m^{m-i}. \tag{6}$$

In the second case, when $A = B$, formula (7) has suggested there is more overlap part between A and B than in the previous case. In other words, the second case tends to have more common channels. We can fill in the "fill part" by random nonrepeated channels from the "original part." The "fill part" contains the probability $pA = B$ of the common channels, and the calculation process is shown in the following formula:

$$p_{A=B} = \frac{mp}{L_A}. \tag{7}$$

The designing process of MPE algorithm is presented in Table 1.

4. Performance Evaluation

4.1. Simulation Environment. In this section, we used simulation experiments to evaluate the performance of the proposed algorithm and verify the above assumptions. The parameters are shown in Table 2. MATLAB was used to simulate the rendezvous process. The time for a successful rendezvous is influenced, as mentioned earlier, by the number of global channels and the distribution of common channels. Therefore, the experiment mainly aimed to simulate both symmetric and asymmetric rendezvous algorithms under heterogeneous conditions. The entire simulation process was repeated 2000 times, and MTTR and ATTR over all the simulations were recorded.

4.2. Experimental Results and Analysis. This section simulates the MPE algorithm in the symmetrical and asymmetrical channel scenarios. The performance evaluation indicators are the longest convergence time and the average convergence time. When in a symmetrical/asymmetric situation, this article uses the MPE algorithm with JS (Jump-Stay) [25], ACH (asynchronous channel hopping rendezvous algorithm) [26], RW (receiver wait for rendezvous algorithm) [27], AHW (alternate hop-and-wait channel rendezvous algorithm), and SSB (short sequence-based rendezvous algorithm) [28] for comparison.

Figure 7 is the comparison result of MTTR in a symmetrical scenario. An upward trend of MTTR could be observed in all algorithms with the global channels increasing. Since the MTTR of CRSEQ is much higher than that of other algorithms, only a part of the data is shown in

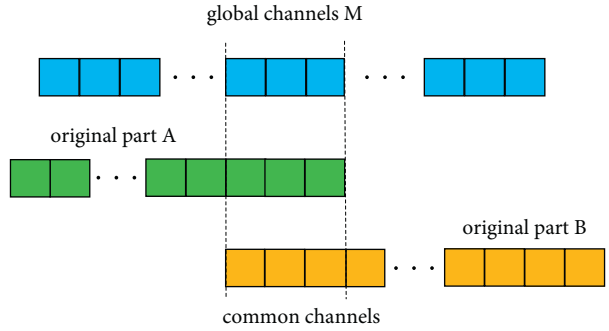


Figure 6: Common channel diagram.

Table 1: Multiple prime expansion algorithm.

Multiple prime expansion algorithm	
1	Input: M, A, B M: the total number of channels A and B: two wireless devices waiting to rendezvous $A \subset M, B \subset M, A \cap B \neq \varnothing$ $p_{A \neq B} = (1/2^m) \sum_{i=0}^{m} pC_m^i (1-p) C_m^{m-i}$ or $p_{A=B} = (mp/L_A)$; p: probability of A or B contains common channels m: the length of the "fill part"
2	$L_A(L_B)$ = the smallest prime number not smaller than $A(B)$
3	if A and B are prime number
4	$L_A = A$, $L_B = B$
5	else
6	if A is prime number, B is composite number
7	$L_A = A$, $L_B = [B, F_B]$
8	else
9	if A is composite number, B is prime number
10	$L_A = [A, F_A]$, $L_B = B$;
11	else
12	$L_A = [A, F_A]$, $L_B = [B, F_B]$;
13	**end**

Table 2: Parameter setting of simulating experiments.

Simulating parameter setting	Value
Number of channels	5~50
Experiment time	2000
Number of wireless devices	2
Analysis type	heterogeneous/symmetrical

Figure 7. As for MPE, although the length of the sequence of the MPE algorithm is slightly larger than other algorithms, the common channel ratio is increased in "fill part" in MPE, and thus rendezvous can be completed faster for MPE. In symmetrical instances, when there are fewer global channels, MPE's MTTR is close to RW and ACH. One likely reason for this experimental result is that the "fill part" of MPE does not effectively improve the probability of common channels. The MPE algorithm is better than other algorithms in MTTR when there are many common channels between wireless devices.

Figure 8 is a comparison result of ATTR in the symmetrical scenario. In the experiment, the ATTR of each algorithm has more obvious fluctuations compared with MTTR. Among them, after the number of global channels of CRSEQ exceeds 20, ATTR increases rapidly, which substantially limited rendezvous multiple times in a limited time. As shown in Figure 8, the common channels show an upward trend as global channel scale expansion, which makes the MPE algorithm show better rendezvous stability. Compared with SSB, the ATTR of the MPE algorithm is much smaller than the SSB algorithm, with an average decrease of 63.9%. In the entire simulation process, the MPE algorithm also achieves the shortest ATTR. Compared with the ACH and RW algorithms, respectively, the average decrease is 71.55% and 87.43%. This shows that the MPE algorithm has a strong continuous capability to rendezvous.

Figure 9 is the comparison result of MTTR in the asymmetric scenario. Compared with the symmetric result, there are two obvious characteristics. (1) The MTTR under asymmetric conditions shows a nonlinear upward trend. (2) In the symmetric scenario, there appears to be no obvious gap of MTTR between most rendezvous algorithms. However, in the asymmetric scenario, the rendezvous certainty of each algorithm will be affected by the symmetry scenario. The difficulty will be significantly increased with the fast loss of symmetry.

From Figure 9, we can see that the MTTR of JS and RW is much higher than the other algorithms (MTTR >10000) with an acceleration trend. The MTTR of MPE is very close

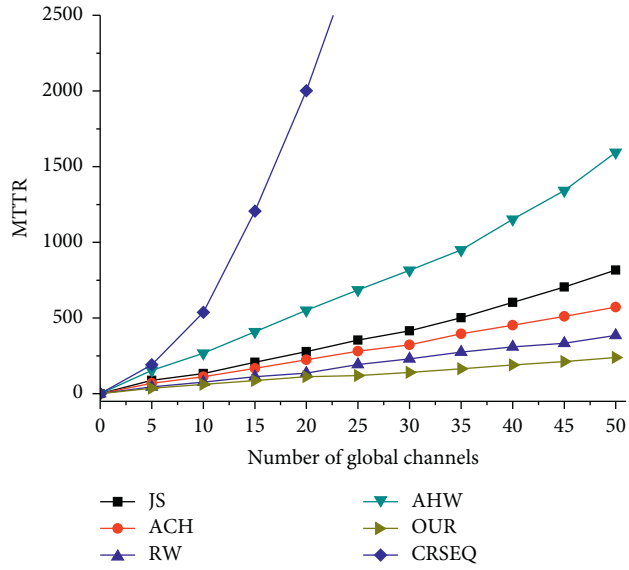


Figure 7: MTTR in the symmetrical scenario.

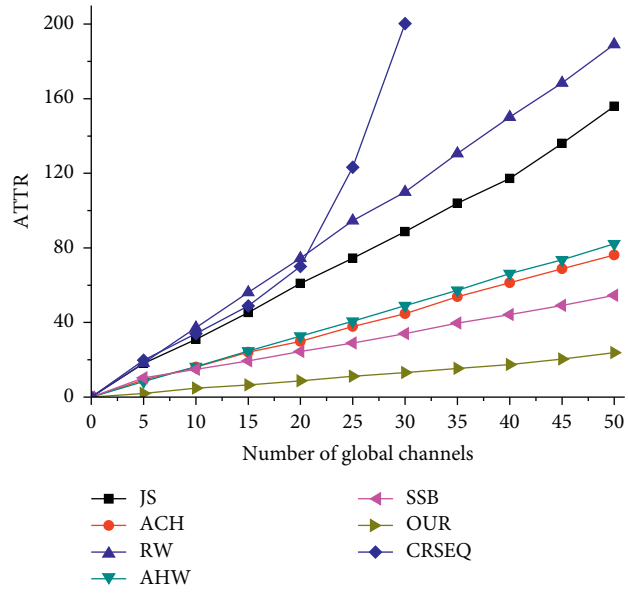


Figure 8: ATTR in a symmetrical scenario.

to AHW when global channels have a smaller scale. However, after the number of channels is more than 30, the MTTR of the MPE algorithm keeps steady. In contrast, all other algorithms showed varying degrees of rapid rise. For example, MTTR ($M = 50$) compared with MTTR ($M = 25$), AHW increased by 236.88%, and the average growth rate of SSB (every 5 channels) is as high as 59.3%. The MTTR of the MPE algorithm never exceeded 1500. This shows that the method adopted by the MPE algorithm has a significant effect in suppressing MTTR.

Figure 10 shows the comparison of ATTR of algorithms in the asymmetric scenario. The ATTR of MPE is the shortest in the experiment process. The comparison results show that ACH, AHW, and SSB with few communication loads (fewer global channels) all have better continuous rendezvous capabilities. It is worth noting that the increase in the number of global channels did increase the differentiation of continuous rendezvous capabilities of all the algorithms. The ATTR of the four algorithms (MPE, ACH, AHW, and SSB) is much smaller than that of the JS algorithm and the RW algorithm, and it is 83.3% lower than that of the RW algorithm on average. In the comparison of four algorithms (MPE, ACH, AHW, and SSB), MPE also achieves the shortest ATTR, which is up to 53.71% and 28.88% lower than the ACH and SSB algorithms. This indicates that the MPE algorithm can also maintain the stability of the rendezvous under asymmetric conditions.

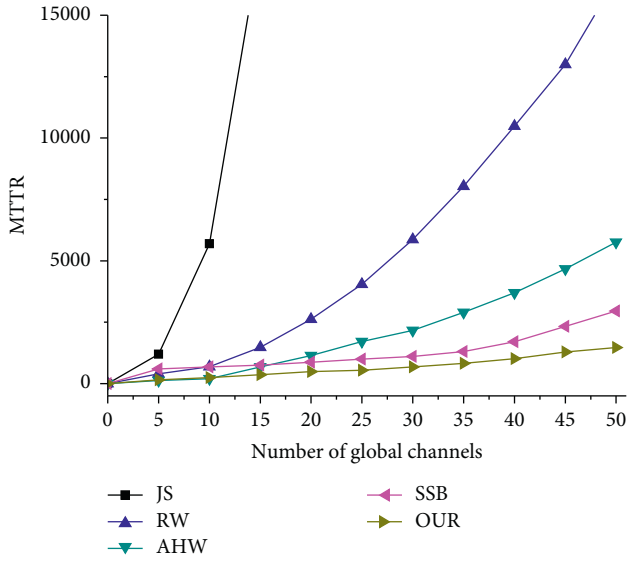


Figure 9: MTTR in an asymmetrical scenario.

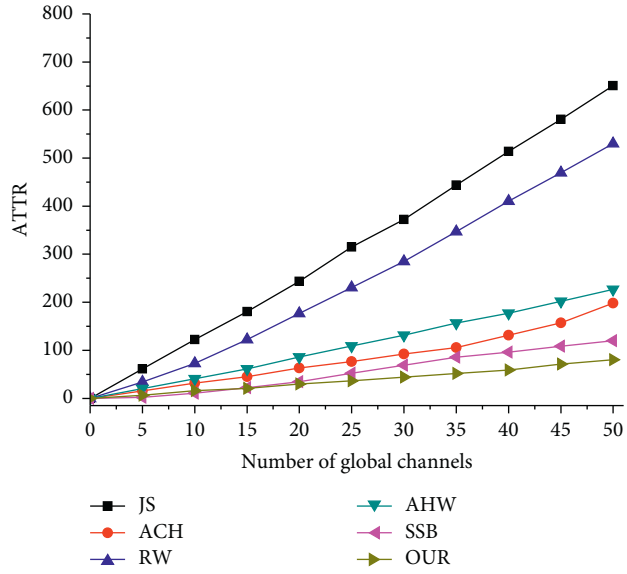


Figure 10: ATTR in an asymmetrical scenario.

5. Conclusions

In this work, asymmetric and heterogeneous scenarios are the two most determinative factors in the speed of rendezvous. We propose the MPE algorithm to address the above problems. Therefore, MPE can efficiently meet the rendezvous requirements of wireless devices in symmetric and asymmetric scenarios under the heterogeneous channel, resulting in that wireless devices can achieve rendezvous faster and save channel search energy. To evaluate the applicability of MPE, some suitable analytical methods, namely, MTTR and ATTR are also proposed. The results show that the performance of MPE is better than classical blind rendezvous algorithms, namely ACH and JS.

MPE discusses the rendezvous algorithm in heterogeneous and asymmetric cases only for the two wireless devices. It is not yet able to perform fast rendezvous in multi wireless devices. Therefore, proposing a blind rendezvous method for multiwireless devices in data interaction would be an important future direction for research.

Data Availability

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

Ethical Approval

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Conflicts of Interest

The authors declare that they have no conflicts of interest in this work.

Acknowledgments

This work was financially supported by the Scientific Research Project of National Natural Science Foundation of China (no. U1709212) and Zhejiang Province Public Welfare Project of China (no. LGF18F030005).

References

[1] K. Zhang, J. Ni, K. Yang, X. Liang, J. Ren, and X. S. Shen, "Security and privacy in smart city applications: challenges and solutions," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 122–129, 2017.

[2] C. Dai, X. Liu, J. Lai, P. Li, and H.-C. Chao, "Human behavior deep recognition architecture for smart city applications in the 5G environment," *IEEE Network*, vol. 33, no. 5, pp. 206–211, 2019.

[3] H. Zheng, S. Wang, and X. Ping, "Study of spinyhead croaker (collichthys lucidus) fat content forecasting model based on electronic nose and non-linear data resolution model," *Food Analytical Methods*, vol. 12, no. 1, 2019.

[4] C. Shao, H. Zheng, Z. Zhou et al., "Ridgetail white prawn (exopalaemon carinicauda) K value predicting method by using electronic nose combined with non-linear data analysis model," *Food Analytical Methods*, vol. 11, no. 11, pp. 3121–3129, 2018.

[5] J. Peng, L. Zheng, and J. Li, "Sucrose quantitative and qualitative analysis from tastant mixtures based on Cu foam electrode and stochastic resonance," *Food Chemistry*, vol. 197, p. 1168, 2015.

[6] H. Yang, Y. Ye, and X. Chu, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, vol. 99, 2021.

[7] Z. Zhou, C. Shao, and H. Zheng, "Simulating study on RHCRP protocol in utility tunnel WSN," *Wireless Networks*, vol. 26, no. 99, 2020.

[8] C. Alcaraz, J. Lopez, R. Roman, and H.-H. Chen, "Selecting key management schemes for WSN applications," *Computers & Security*, vol. 31, no. 8, pp. 956–966, 2012.

[9] S. Kumar, A. Kumar, and R. K. Vishwkarma, "A survey on routing protocol for wireless sensor network," *International*

Journal of Advanced Research in Computer Engineering & Technology, vol. 2, no. 2, 2013.

[10] O. Olayinka and A. Attahiru, "A survey on an energy-efficient and energy-balanced routing protocol for wireless sensor networks," *Sensors*, vol. 17, no. 5, p. 1084, 2017.

[11] F. Losilla, A. J. Garcia-Sanchez, and F. Garcia-Sanchez, "A comprehensive approach to WSN-based its applications: a survey," *Sensors*, vol. 11, no. 11, 2012.

[12] W. Fang, J. Wu, and Y. Bai, "Quantitative risk assessment of a natural gas pipeline in an underground utility tunnel," *Process Safety Progress*, vol. 38, 2019.

[13] O. G. Zabaleta, J. P. Barrangú, and C. M. Arizmendi, "Quantum game application to spectrum scarcity problems," *Physica A: Statistical Mechanics and Its Applications*, vol. 466, pp. 455–461, 2017.

[14] H. Zhang, N. Xu, and F. Xu, "Graph cut based clustering for cognitive radio ad hoc networks without common control channels," *Wireless Networks*, vol. 24, 2016.

[15] T. Y. Lin, K. R. Wu, and G. C. Yin, "Channel-hopping scheme and channel-diverse routing in static multi-radio multi-hop wireless networks," *IEEE Transactions on Computers*, vol. 64, no. 1, pp. 71–86, 2014.

[16] S. Mohapatra and P. K. Sahoo, "ASCH: A novel asymmetric synchronous channel hopping algorithm for Cognitive Radio Networks," in *Proceedings of the IEEE International Conference on Communications*, IEEE, Kuala Lumpur, Malaysia, May 2016.

[17] Q. Liu, X. Wang, B. Han, X. Wang, and X. Zhou, "Access delay of cognitive radio networks based on asynchronous channel-hopping rendezvous and CSMA/CA MAC," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 3, pp. 1105–1119, 2015.

[18] Y. T. Wang, G. C. Yang, and S. H. Huang, "Multi-MTTR asynchronous-asymmetric channel-hopping sequences for scalable cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 99, p. 1, 2018.

[19] T.-H. Lin, G.-C. Yang, and W. C. Kwong, "A homogeneous multi-radio rendezvous algorithm for cognitive radio networks," *IEEE Communications Letters*, vol. 23, no. 4, pp. 736–739, 2019.

[20] C. Chang, C. Chen, and D. S. Lee, "Efficient encoding of user IDs for nearly optimal expected time-to-rendezvous in heterogeneous cognitive radio networks," *IEEE/ACM Transactions on Networking*, vol. 23, 2017.

[21] N. C. Theis, R. W. Thomas, and L. A. Dasilva, "Rendezvous for cognitive radios," *IEEE Transactions on Mobile Computing*, vol. 10, no. 2, pp. 216–227, 2010.

[22] J. Shin, D. Yang, and C. Kim, "a channel rendezvous scheme for cognitive radio networks," *IEEE Communications Letters*, vol. 14, no. 10, pp. 954–956, 2010.

[23] Y. Deng, Z. Zhou, Z. Zhao et al., "Simulation study on ASCMP protocol in utility tunnel WSN," *IEEE Access*, vol. 7, pp. 168141–168150, 2019.

[24] H. Qian, Z. Zong, C. Wu, J. Li, and L. Gan, "Numerical study on the behavior of utility tunnel subjected to ground surface explosion," *Thin-Walled Structures*, vol. 161, Article ID 107422, 2021.

[25] H. Liu, Z. Lin, X. Chu, and Y.-W. Leung, "Jump-stay rendezvous algorithm for cognitive radio networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 10, pp. 1867–1881, 2012.

[26] K. Bian, "Maximizing rendezvous diversity in rendezvous protocols for decentralized cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1294–1307, 2012.

[27] H. Liu, Z. Lin, and X. Chu, "Ring-walk based channel-hopping algorithms with guaranteed rendezvous for cognitive radio networks," in *Proceedings of the 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing*, pp. 755–760, IEEE, Washington, DC, USA, December 2010.

[28] V. A. Reguera, E. O. Guerra, and R. D. Souza, "Short channel hopping sequence approach to rendezvous for cognitive networks," *IEEE Communications Letters*, vol. 18, no. 2, pp. 289–292, 2013.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7768588, 11 pages
https://doi.org/10.1155/2022/7768588

Research Article

Sum-Throughput Maximization in Backscatter Communication-Based Cognitive Networks

Qian Li

School of Computer Science and Technology, Zhoukou Normal University, Zhoukou, China

Correspondence should be addressed to Qian Li; liqian@zknu.edu.cn

Received 8 December 2021; Revised 2 January 2022; Accepted 11 January 2022; Published 4 February 2022

Academic Editor: Yinghui Ye

In this paper, we consider a backscatter communication (BackCom) based cognitive network that consists of one primary transmitter, one primary receiver, multiple secondary transmitters (STs), and one secondary receiver (SR). Each ST operates in the BackCom or energy harvesting model. Our goal is to jointly optimize the energy harvesting and backscatter time, the transmit power of the primary transmitter, and the power reflection coefficient of each ST to maximize the sum throughput of all the STs under a nonlinear energy harvesting model while satisfying multiple constraints, i.e., the energy causality of each ST, the Quality of Service of the primary transmitter, etc. The formulated problem is nonconvex due to the coupled variables and hard to solve. In order to address this problem, we decouple partial coupled variables by using the properties of the objective function and constructing auxiliary variables, and the remaining coupled variables are decoupled via successive convex approximation (SCA). On this basis, a SCA based iterative algorithm is developed to solve the formulated problem. Simulation results are provided to support our work.

1. Introduction

Internet of Things (IoT) is expected to deploy massive smart sensor nodes in the future communications to seamlessly connect the physical environment and the cyberspace for providing intelligent services [1]. However, such an approach requires huge spectrum resources, and this motivates us to consider high spectrum-efficient communication paradigms for IoT. In this context, cognitive radio has been proposed, whose key idea is to allow sensor nodes sharing spectrum with primary uses without causing any harmful factors to the primary transmission [2]. In cognitive radio, the smart sensor node (also called secondary user) transmits its own information by active radios that need the power-consuming components and consume a lot of energy, greatly shortening their lifespan and leading to an energy-constrained problem for smart sensor nodes.

In addition to the cognitive radio that improves the spectrum efficiency, backscatter communication (BackCom) is another key technology, and its main purpose is to overcome the energy-constrained problem [3–5]. BackCom allows a smart sensor node modulating its information on the incident signals and backscattering the modulated signals to its associated receiver by changing the power reflection coefficient so that the power-consuming components can be avoided, while harvesting energy from the incident signal for realizing energy self-sustainability [3–5]. Despite these superiorities, the communication performance of BackCom is limited as it uses the ambient signals as the incident signals, and the ambient signals introduce serious cochannel interference to the BackCom receiver [4]. Accordingly, researchers proposed to use the controllable signals as the incident signals, but such an approach requires an extra cost to deploy RF sources. Recall that cognitive radio is able to provide controllable signals for the BackCom transmitter (also referred to as the secondary transmitter (ST) in this paper). Recent works have integrated BackCom into cognitive radio, yielding a spectrum- and energy-efficiency paradigm, called BackCom based cognitive networks.

In [6], the authors formulated a problem to maximize the throughput of a BackCom link in a BackCom based cognitive network with a single ST by jointly optimizing the transmit

power of the primary user and the power reflection coefficient of the ST. In [7], the authors proposed to maximize the energy efficiency of the secondary link by jointly optimizing the transmit power of the primary transmitter (PT), the power reflection coefficient of the ST, and the time for energy harvesting and BackCom. Extending the single ST scenario [6, 7] into multiple STs [8], the authors maximized the sum rate of STs by jointly optimizing the PT's transmit power and the power reflection coefficient of each ST. The authors of [9] considered a full-duplex-enabled BackCom based cognitive network and proposed a joint time, transmit power, and power reflection coefficient scheduling to maximize the throughput of the BackCom system. In addition to the above works, the harvest-then-transmit (HTT) protocol has also been integrated into BackCom based cognitive networks, and various resource allocation schemes have been studied [10–12], where the main focus is to balance the time for energy harvesting, HTT, and backscattering.

After carefully examining the existing resource allocation schemes, we note that all of them were based on a linear energy harvesting model. As pointed out by existing works, the linear energy harvesting model does not match the behavior of a practical energy harvester. More specifically, the harvest power is a nonlinear function with respect to the input power. The previous works [13–15] (in which BackCom based cognitive network has not been considered) have proved that the mismatch of energy harvesting models will lead to performance degradation. Accordingly, it is necessary to design resource allocation for BackCom based cognitive networks with nonlinear energy harvesting model. Motivated by this, in this paper, we consider a BackCom based cognitive network with multiple STs and aim to maximize the total throughput of STs by jointly optimizing the energy harvesting time, BackCom time, power reflection coefficient, and PT's transmit power under a nonlinear energy harvesting model. Meanwhile, the energy-causality constraint of each ST and the Quality of Service (QoS) of both the primary link and the secondary links are considered. The main contributions of this paper are summarized as below.

We formulate an optimization problem to maximize the sum throughput of STs, and propose an efficient iterative algorithm to solve the problem. Since the formulated problem is nonconvex, the main challenge is how to transform the original problem into a convex one. Towards this end, we firstly determine the optimal transmit power of the PT by using the properties of the objective function, then introduce some auxiliary variables to decouple coupled variables, and lastly employ the successive convex approximation (SCA) to transform a nonconvex constraint into a linear one. We also provide computer simulation results to verify the proposed iterative algorithm and show the advantages of the proposed scheme over the baseline schemes.

2. System Model and Working Flow

As shown in Figure 1, we consider a BackCom based cognitive network, which consists of one PT, one PR, K energy-constrained STs, and one SR. In this network, the PT

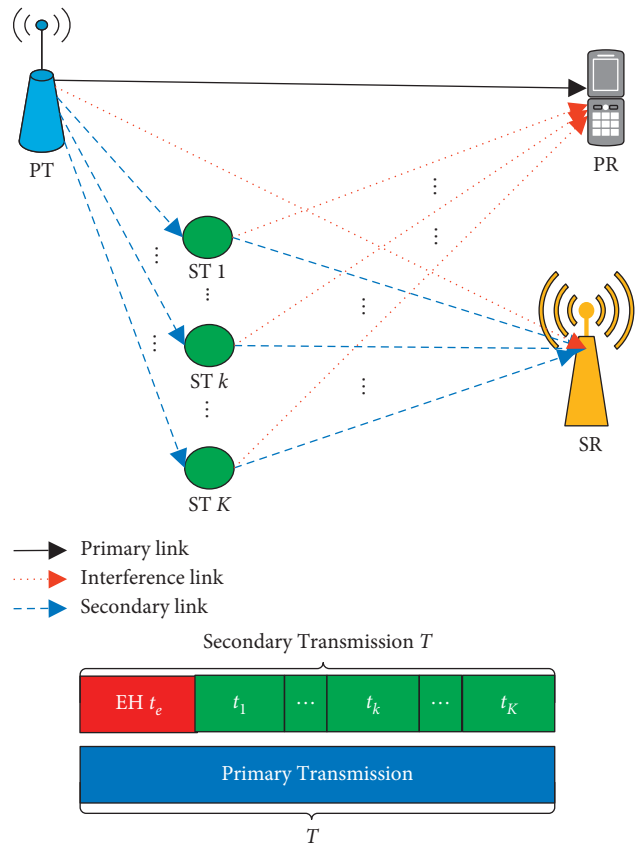


Figure 1: BackCom based cognitive network and its frame structure.

transmits its own information to the PR, while the RF signals transmitted by the PT can also be exploited by the K STs for energy harvesting (EH) and information transmission. Assume that all the devices are equipped with a single antenna and work in the half-duplex mode. Suppose that each ST is equipped with both the EH circuit and the BackCom circuit so that it can harvest energy from the received RF signals and backscatter the received signals for information transmission. In order to prolong the operation time of each energy-constrained ST, we assume that each ST only uses its harvested energy instead of the energy early stored in its battery to cover the energy consumption during information backscattering. All channels including the PT-PR link and the ST-SR links are assumed to follow quasi-static fading. In the beginning of each transmission block, the channel state information (CSI) of all links can be perfectly achieved by the PT via the existing advanced channel estimation methods, so that the PT can determine the optimal resource allocation scheme based on the obtained CSI and feed the optimal scheme back to all the STs.

Let T denote the duration of the whole transmission block. For the PT-PR link, the PT may transmit its signals in the whole transmission block. For the ST-SR links, the whole transmission can be divided into two phases, which are the EH phase and the BackCom phase, respectively. In the EH phase, the PT broadcasts its signals to the PR while all the STs will perform EH. In the BackCom phase, the PT keeps

broadcasting while all the STs take turns to perform BackCom in order to avoid the cochannel interference among STs.

2.1. EH Phase. Let t_e denote the duration of the EH phase. Denote P_t as the transmit power of the PT. Then, the received signal at the k-th ($k \in \mathcal{K} = \{1, 2, \ldots, K\}$) ST is given by

$$y_{ST}^k = \sqrt{P_t h_k} x_p + N_{ST}, \tag{1}$$

where h_k denotes the channel gain of the PT–k-th ST link, x_p with $\mathbb{E}[|x_p|^2] = 1$ is the information transmitted by the PT to the PR, and N_{ST} is the thermal noise at the k-th ST. Since the backscatter communication circuit consists only of passive components and takes few signal processing operations, the thermal noise is usually very small and can be ignored, i.e., $N_{ST \approx 0}$ [7].

For EH, we consider a more practical nonlinear EH model since the linear EH model cannot characterize the nonlinearity of the practical EH circuit [16]. The reasons of considering the above nonlinear EH model instead of the one considered in [17, 18] are as follows. Firstly, the nonlinear EH model proposed in [16] is accurate enough for characterizing the nonlinearity of practical EH circuits. Secondly, the use of the nonlinear EH model proposed in [16] can simplify the difficulty and reduce the complexity of solving the formulated optimization problem. Accordingly, we compute the harvested energy of the k-th as

$$E_e^k = t_e \left(\frac{a_k P_t h_k + d_k}{P_t h_k + v_k} - \frac{d_k}{v_k} \right), \tag{2}$$

where a_k, d_k, and v_k are the parameters of the nonlinear EH model.

For the PT-PR link, the received signal at the PR is expressed as

$$y_{PR}^e = \sqrt{P_t f_p} x_p + N_{PR}, \tag{3}$$

where f_p denotes the channel gain of the PT-PR link and N_{PR} is the additive white Gaussian noise (AWGN) at the PR with mean zero and variance σ^2. Accordingly, the achievable throughput at the PR in this phase can be computed as

$$R_e^p = t_e B \log_2 \left(1 + \frac{P_t f_p}{\sigma^2} \right), \tag{4}$$

where B is the system bandwidth.

2.2. BackCom Phase. The whole BackCom phase can be divided into K subphases. In each subphase, each ST performs BackCom to transmitted information. Let t_k denote the duration of the k-th subphase. Denote α_k with $0 \le \alpha_k \le 1$ as the power reflection coefficient of the k-th ST based on which the received RF signal at the k-th ST can be split into two parts: one part is used for BackCom and the other part is flowed into the EH circuit. Then, in the subphase t_k, the received signal at the SR is given by

$$y_{SR}^k = \sqrt{\alpha_k P_t h_k g_k} x_p x_{s,k} + \sqrt{P_t f_s} x_p + N_{SR}, \tag{5}$$

where g_k denotes the channel gain between the k-th ST and the SR, f_s is the channel gain of the PT-SR link, $x_{s,k}$ with $\mathbb{E}[|x_{s,k}|^2] = 1$ is the transmitted information of the k-th ST, and N_{SR} is the AWGN at the SR with mean zero and variance σ^2.

From (5), it can be observed that the cochannel interference from the PT-SR link always exists, which degrades the throughput achieved by the k-th ST via BackCom. In order to decode the transmitted information of the k-th ST correctly, the SR performs the successive interference cancellation (SIC) technology. Specifically, the SR first decodes the PT's transmitted information x_p by treating $\sqrt{\alpha_k P_t h_k g_k} x_p x_{s,k}$ as the cochannel interference and then uses the SIC technology to cancel the interference from the PT as well as decoding the transmitted information of the k-th ST $x_{s,k}$. Therefore, the signal to interference plus noise ratio (SINR) at the SR for decoding x_p is given by

$$\gamma_{s,k}^p = \frac{P_t f_s}{\alpha_k P_t h_k g_k + \sigma^2}. \tag{6}$$

After using SIC technology, the signal to noise ratio (SNR) at the SR for decoding $x_{s,k}$ is expressed as

$$\gamma_{s,k} = \frac{\alpha_k P_t h_k g_k}{\eta P_t f_s + \sigma^2}, \tag{7}$$

where η with $0 \le \eta \le 1$ is the interference cancellation factor.

Accordingly, the achievable throughput of the k-th ST-SR link can be computed as

$$R_{s,k} = t_k B \log_2 \left(1 + \xi \gamma_{s,k} \right), \tag{8}$$

where ξ expresses the performance gap reflecting the real modulation [19–21]. In this subphase, the harvested energy of the k-th ST is determined by

$$E_k^b = t_k \left(\frac{a_k (1 - \alpha_k) P_t h_k + d_k}{(1 - \alpha_k) P_t h_k + v_k} - \frac{d_k}{v_k} \right). \tag{9}$$

At the end of the BackCom phase, the total harvested energy of the k-th ST is given by

$$E_{tot}^k = E_k^b + \left(t_e + \sum_{i=1}^{K} t_i - t_k \right) \left(\frac{a_k P_t h_k + d_k}{P_t h_k + v_k} - \frac{d_k}{v_k} \right). \tag{10}$$

For the PT-PR link, it also suffers from the cochannel interference from the k-th ST, and the received signal at the PR is given by

$$y_{PR}^k = \sqrt{P_t f_p} x_p + \sqrt{\alpha_k P_t h_k f_k} x_p x_{s,k} + N_{PR}, \tag{11}$$

where f_k denotes the channel gain from the k-th ST to the PR. Then, the SINR at the PR for decoding x_p is computed as

$$\gamma_k^p = \frac{P_t f_p}{\alpha_k P_t h_k f_k + \sigma^2}. \tag{12}$$

Correspondingly, the achievable throughput of the PT-PR link in this subphase is given by $R_k^p = t_k B \log_2(1+\gamma_k^p)$.

3. Throughput Maximization for BackCom-Based Cognitive Networks

In this section, we design an optimal resource allocation scheme to maximize the total throughput of all the STs for the BackCom based cognitive network. In particular, we formulate a throughput maximization problem by jointly optimizing the transmit power of the PT, BackCom time, and power reflection coefficients of STs, as well as the EH time, subject to QoS, energy causality, latency, transmit power, and power reflection coefficient constraints, and then use the existing convex tools to solve it.

3.1. Problem Formulation. Before formulating the throughput maximization problem, we should determine the optimization objective and constraints.

3.1.1. Optimization Objective. The optimization objective is to maximize the total throughput of the STs which can be computed as

$$R_{\text{tot}}^s = \sum_{k=1}^{K} R_{s,k} = \sum_{k=1}^{K} t_k B \log_2\left(1+\frac{\xi\alpha_k P_{\text{t}} h_k g_k}{\eta P_{\text{t}} f_s + \sigma^2}\right). \tag{13}$$

3.1.2. QoS Constraints. There are two QoS constraints to constrain the throughput of the PT and each ST, respectively. For the QoS constraint of each ST, we should guarantee that the achievable throughput of each ST is not less than its minimum required throughput. Let $C_{\min,k}$ denote the minimum required throughput of the k-th ST. Then, the QoS constraint of the k-th ST can be expressed as

$$\begin{aligned} &R_{s,k} \geq C_{\min,k}, \quad \forall k, \\ \Longleftrightarrow &t_k B \log_2\left(1+\frac{\xi\alpha_k P_{\text{t}} h_k g_k}{\eta P_{\text{t}} f_s + \sigma^2}\right) \geq C_{\min,k}, \quad \forall k. \end{aligned} \tag{14}$$

For the QoS constraint of the PT, the total achievable throughput of the PT should not be less than the PT's minimum required throughput, denoted by $C_{\min}$. Therefore, the QoS constraint of the PT can be expressed as

$$\begin{aligned} R_e^p + \sum_{k=1}^{K} R_k^p = {} & t_e B \log_2\left(1+\frac{P_t f_p}{\sigma^2}\right) \\ & + \sum_{k=1}^{K} t_k B \log_2\left(1+\frac{P_t f_p}{\alpha_k P_t h_k f_k + \sigma^2}\right) \geq C_{\min}. \end{aligned} \tag{15}$$

3.1.3. Energy-Causality Constraint. The energy-causality constraint states that the energy consumption of each ST should not be larger than its harvested energy during the whole transmission block. Note that a rechargeable battery is equipped in each ST and that each ST may first use the energy stored in its battery to support BackCom and then use the harvested energy to power the battery. The energy-causality constraint ensures that the energy early stored in the battery is not reduced. Here, we consider fixed power consumption for BackCom by following [19–21]. Let $P_{\text{b},k}$ denote the power consumption for the k-th ST when performing BackCom. Therefore, the energy-causality constraint for the k-th ST is given by

$$\begin{aligned} P_{b,k} t_k \leq E_{tot}^k = {} & t_k f_k\left((1-\alpha_k) P_t h_k\right) \\ & + \left(t_e + \sum_{i=1}^{K} t_i - t_k\right) f_k\left(P_t h_k\right), \quad \forall k, \end{aligned} \tag{16}$$

where $f_k(x) = (a_k x + d_k/x + v_k) - (d_k/v_k)$. Please note that the power consumption for the EH circuit has been included in the EH model and thus has not been considered in (16).

3.1.4. Transmit Power Constraint. Let $P_{\max}$ denote the maximum allowed transmit power of the PT. Then, the PT's transmit power constraint can be expressed as

$$0 \leq P_{\text{t}} \leq P_{\max}. \tag{17}$$

Based on (13), (14), (15), (16), and (17), the throughput maximization problem can be formulated as

$$\begin{aligned} \text{P}_1: \ & \max_{P_t, t_e, \{t_k\}_{k=1}^K, \{\alpha_k\}_{k=1}^K} R_{tot}^s, \\ & s.t.\ C1:\ (14),\ (15), \\ & C2:\ (16), \\ & C3:\ (17), \\ & C4:\ \left(\gamma_{s,k}^p \geq \gamma_{th}, \forall k\right), \\ & C5:\ t_e + \textstyle\sum_{k=1}^{K} t_k \leq T, t_e, t_k \geq 0, \forall k, \\ & C6:\ 0 \leq \alpha_k \leq, \ \forall k, \end{aligned} \tag{18}$$

where C1 denotes the QoS constraints for each ST and the PT, C2 is the energy-causality constraint for each ST, C3 constrains the maximum transmit power of the PT, C4 ensures that each ST can decode x_p successfully and γ_{th} is the threshold required for decoding x_p, C5 is the latency constraint, and C6 is the constraint for the power reflection coefficient of each ST.

It can be observed that $\mathbf{P}_1$ is a highly nonconvex optimization problem and is difficult to solve due to the following reasons. Firstly, there exist several coupled relationships between multiple optimization variables, i.e., P_{t}, t_k, α_k, etc., leading to the nonconvex objective function and constraints, i.e., C1, C2, etc. Secondly, the cochannel interference causes the difference of convex (DC) structures in the objective function and C1, bringing new challenges to solving $\mathbf{P}_1$. Thirdly, the consideration of the nonlinear EH model is another difficulty for solving $\mathbf{P}_1$ since the nonlinear EH model makes C1 more complex. Therefore, it is hard to solve $\mathbf{P}_1$.

3.2. Solution. In order to remove DC structures existing in the objective function and C1 and simplify $\mathbf{P}_1$, the following lemma is introduced to determine the optimal transmit power of the PT.

Lemma 1. *The maximum throughput of all the STs for the considered network is achieved when the PT transmits its signals with its maximum transmit power, i.e., $P_t^* = P_{\max}$.*

Proof. Please see Appendix A.

By substituting $P_t = P_{\max}$ into $\mathbf{P}_1$, $\mathbf{P}_1$ can be reformulated as

$$\begin{aligned}
\mathrm{P}_2\colon & \max_{t_e,\{t_k\}_{k=1}^K,\{\alpha_k\}_{k=1}^K} \sum_{k=1}^K t_k B \log_2\left(1+\frac{\zeta\alpha k P_{\max} h_k g_k}{\eta P_{\max} f_s+\sigma^2}\right),\\
s.t.\ & C1'\colon t_k B\log_2\left(1+\frac{\zeta\alpha k P_{\max} h_k g_k}{\eta P_{\max} f_s+\sigma^2}\right)\geq C_{\min,k},\forall k,\\
& \sum_{k=1}^K t_k B \log_2\left(1+\frac{P_{\max} f_p}{\alpha_k P_{\max} h_k f_k+\sigma^2}\right)+t_e B\log_2\left(1+\frac{P_{\max} f_p}{\sigma^2}\right)\geq C_{\min}, \qquad (19)\\
& C2'\colon P_{b,k}t_k\leq\left(t_e+\sum_{i=1}^K t_i-t_k\right)f_k\left(P_{\max}h_k\right)+t_k f_k\left((1-\alpha_k)P_{\max}h_k\right),\forall k,\\
& C5,C7\colon 0\leq\alpha_k\leq\min\left(\frac{P_{\max}f_s-\gamma_{th}\sigma^2}{P_{\max}h_k g_k\gamma_{th}},1\right),\forall k,
\end{aligned}$$

where C7 is the combination of C4 and C6.

$\mathbf{P}_2$ is still nonconvex since the coupled relationships between different variables, e.g., α_k and t_k, still exist in the objective function and several constraints. To address this issue, we introduce the following auxiliary variables: $z_k = \alpha_k t_k, \forall k$, to replace the variables $\alpha_k, \forall k$, and rewrite $\mathbf{P}_2$ as

$$\begin{aligned}
\mathrm{P}_2\colon & \max_{t_e,\{t_k\}_{k=1}^K,\{z_k\}_{k=1}^K} \sum_{k=1}^K t_k B \log_2\left(1+\frac{\zeta z_k P_{\max} h_k g_k}{t_k\left(\eta P_{\max} f_s+\sigma^2\right)}\right),\\
s.t.\ & C1''\colon t_k B\log_2\left(1+\frac{\zeta z_k P_{\max} h_k g_k}{t_k\left(\eta P_{\max} f_s+\sigma^2\right)}\right)\geq C_{\min,k},\forall k,\\
& \sum_{k=1}^K t_k B \log_2\left(1+\frac{P_{\max} f_p}{z_k P_{\max} h_k f_k+\sigma^2}\right)+t_e B\log_2\left(1+\frac{P_{\max} f_p}{\sigma^2}\right)\geq C_{\min}, \qquad (20)\\
& C2''\colon P_{b,k}t_k\leq t_k f_k\left(\frac{(t_k-z_k)P_{\max}h_k}{t_k}\right)\left(t_e+\sum_{i=1}^K t_i-t_k\right)f_k\left(P_{\max}h_k\right),\forall k,\\
& C5,C7'\colon 0\leq z_k\leq t_k\times\min\left(\frac{P_{\max}f_s-\gamma_{th}\sigma^2}{P_{\max}h_k g_k\gamma_{th}},1\right),
\end{aligned}$$

where $\alpha_k = (z_k/t_k), \forall k$.

□

Proposition 1. *In $\mathbf{P}_3$, the objective function and all the constraints except $C1''$ are convex.*

Proof. Please see Appendix B.

In order to handle the nonconvex constraint C1′′ and solve $\mathbf{P}_3$, the SCA method is used, where the first-order Taylor expression is used to approximate the function $t_k B \log_2(1 + (t_k P_{\max} f_p / z_k P_{\max} h_k f_k + \sigma^2 t_k))$ in C1′′ and turn this nonconvex function into a linear function. Specifically, let $F_k(\alpha_k) = t_k B \log_2(1 + (P_{\max} f_p / \alpha_k P_{\max} h_k f_k + \sigma^2))$. By taking the first-order derivative of $F_k(\alpha_k)$ with respect to α_k, we have

$$\frac{\partial F_k(\alpha_k)}{\partial \alpha_k} = \frac{-P_{\max^2} f_p h_k f_k B t_k}{\left(\alpha_k P_{\max} h_k f_k + \sigma^2 + P_{\max} f_p\right)\left(\alpha_k P_{\max} h_k f_k + \sigma^2\right)\mathrm{In}2}. \tag{21}$$

Using the first-order Taylor expression, $F_k(\alpha_k)$ can be approximated as

$$\begin{aligned} F_k(\alpha_k) &\approx \frac{\partial F_k\left(\alpha_k^0\right)}{\partial \alpha_k^0}\left(\alpha_k - \alpha_k^0\right) + F_k\left(\alpha_k^0\right) = \frac{-P_{\max^2} f_p h_k f_k B t_k \left(\alpha_k - \alpha_k^0\right)}{\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2 + P_{\max} f_p\right)\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2\right)\mathrm{In}2} + F_k\left(\alpha_k^0\right), \\ &= \frac{-P_{\max^2} f_p h_k f_k B\left(z_k - \alpha_k^0 t_k\right)}{\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2 + P_{\max} f_p\right)\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2\right)\mathrm{In}2} + F_k\left(\alpha_k^0\right), \end{aligned} \tag{22}$$

where α_k^0 is the given value for α_k and can be updated iteration by iteration.

By substituting (22) into C1′′, the QoS constraint for the PT's transmission can be rewritten as

$$\sum_{k=1}^{K}\left(\frac{-P_{\max^2} f_p h_k f_k B\left(z_k - \alpha_k^0 t_k\right)}{\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2 + P_{\max} f_p\right)\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2\right)\mathrm{In}2} + F_k\left(\alpha_k^0\right)\right) + t_e B \log_2\left(1 + \frac{P_{\max} f_p}{\sigma^2}\right) \geq C_{\min}. \tag{23}$$

Accordingly, $\mathbf{P}_3$ can be transformed into the following subproblem, given by

$$\begin{aligned} \mathrm{P}_4: &\max_{t_e, \{t_k\}_{k=1}^K, \{z_k\}_{k=1}^K} \sum_{k=1}^{K} t_k B \log_2\left(1 + \frac{\zeta z_k P_{\max} h_k g_k}{t_k\left(\eta P_{\max} f_s + \sigma^2\right)}\right), \\ s.t.\ &C1''': t_k B \log_2\left(1 + \frac{\zeta z_k P_{\max} h_k g_k}{t_k\left(\eta P_{\max} f_s + \sigma^2\right)}\right) \geq C_{\min,k}, \forall k, \\ &\sum_{k=1}^{K} t_k B \log_2\left(1 + \frac{-P_{\max}^2 f_p h_k f_k B\left(z_k - \alpha_k^0 t_k\right)}{\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2 + P_{\max} f_p\right)\left(\alpha_k^0 P_{\max} h_k f_k + \sigma^2\right) In2} + F_k\left(\alpha_k^0\right)\right) + t_e B \log_2\left(1 + \frac{P_{\max} f_p}{\sigma^2}\right) \geq C_{\min}, \\ &C2''', C5, C7'. \end{aligned} \tag{24}$$

Proposition 2. $\mathbf{P}_4$ *is proved to be convex, which can be efficiently solved by using the existing convex optimization tools.*

Proof. After using the first-order Taylor expression, the nonconvex QoS constraint for the PT's transmission in C1′′′ can be turned into a linear constraint. Combining with Proposition 1, $\mathbf{P}_4$ can be proved to be convex and can be efficiently solved by using the existing convex optimization tools. □

□

(1) Set the maximum tolerance ε and the maximum number of iterations $I_{\max}$;
(2) Set the iteration index $i = 1$ and the initial given values $\alpha_k^0, \forall k$;
(3) Based on Lemma 1, the optimal transmit power of the PT P_t^* is set as $P_{\max}$;
(4) repeat
(5) Solve the optimization problem $\mathbf{P}_4$ with given $\alpha_k^0, \forall k$, to obtain the optimal solutions, denoted by $t_e^*, \{t_k^*\}_{k=1}^K, \{z_k^*\}_{k=1}^K$;
(6) Compute α_k^* as $(z_k^*/t_k^*), \forall k$;
(7) Compute the value of R_{tot}^s based on (13);
(8) if $|\alpha_k^* - \alpha_k^0| \le \varepsilon$ then
(9) Set Flag = 1;
(10) else
(11) Set Flag = 0 and $i = i + 1$;
(12) Update α_k^0 as $\alpha_k^0 = \alpha_k^*, \forall k$;
(13) end if
(14) until Flag = 1 or $i = I_{\max}$.
(15) Output P_t^*, $t_e^*, \{t_k^*\}_{k=1}^K, \{\alpha_k^*\}_{k=1}^K$ and R_{tot}^s.

Algorithm 1: SCA based iterative algorithm.

3.3. Design of a SCA Based Iterative Algorithm. In this subsection, we propose a SCA based iterative algorithm to solve $\mathbf{P}_3$ efficiently. The detailed process of the proposed algorithm is shown in Algorithm 1.

As shown in Algorithm 1, in each iteration, we should use the existing convex tools, i.e., CVX, to optimally solve the subproblem $\mathbf{P}_4$ with given $\alpha_k^0, \forall k$. Then the optimal solutions to $\mathbf{P}_4$ are obtained, denoted by $t_e^*, \{t_k^*\}_{k=1}^K, \{\alpha_k^*\}_{k=1}^K$, where α_k^* is computed as $(z_k^*/t_k^*), \forall k$. If the stop condition, namely, $|\alpha_k^* - \alpha_k^0| \le \varepsilon$ with the maximum tolerance ε, is satisfied, then the solution to $\mathbf{P}_3$ is $t_e^*, \{t_k^*\}_{k=1}^K, \{\alpha_k^*\}_{k=1}^K$. Otherwise, we should update the value of α_k^0 as α_k^* and repeat the above steps until the stop condition is satisfied.

4. Numerical Results

In this section, both the effectiveness and the superiority of the proposed algorithm are verified via computer simulations. The key simulation parameters, unless otherwise specified, are provided in Table 1. Following [16], the parameters of the considered nonlinear EH model at the k-th ST are set as $a_k = 2.463$, $d_k = 1.635$, and $\nu_k = 0.826, \forall k$. For the settings of all channels, we consider a standard channel fading model. Specifically, the channel gain of the PT-PR link is modeled by $f_p = f_p' D_p^{-\beta}$, where f_p' denotes the small-scale fading of the PT-PR link, D_p is the distance from the PT to the PR, and β denotes the path loss exponent. The channel gain of the PT–k-th ST link is given by $g_k = g_k' D_{1,k}^{-\beta}, \forall k$, where g_k' and $D_{1,k}$ are the small-scale fading and the distance from the PT to the k-th ST, respectively. The channel gain of the k-th ST–SR link is modeled by $h_k = h_k' D_{2,k}^{-\beta}, \forall k$, where h_k' and $D_{2,k}$ are the small-scale fading and the distance from the k-th ST to the SR, respectively. The channel gain of the k-th ST–PR link is $f_k = f_k' D_{3,k}^{-\beta}, \forall k$ with the small-scale fading f_k' and the distance $D_{3,k}$. The channel gain of the PT-SR link is modeled by $f_s = f_s' D_s^{-\beta}, \forall k$, with the small-scale fading f_s' and the distance D_s. In the simulations, we set $\beta = 2.7$, $D_p = 30$ m, $D_s = 30$ m, $D_{1,1} = 12$ m, $D_{1,2} = 10$ m, $D_{1,3} = 15$ m, $D_{1,4} = 13$ m, $D_{2,1} = 20$ m, $D_{2,2} = 15$ m, $D_{2,3} = 20$ m, $D_{2,4} = 15$ m, $D_{3,1} = 25$ m, $D_{3,2} = 30$ m, $D_{3,3} = 30$ m, and $D_{3,4} = 30$ m.

Figure 2 demonstrates the convergence of Algorithm 1, where different settings of $C_{s,\min}$ are considered and $C_{\min^1} = C_{\min^2} = C_{\min^3} = C_{\min^4} = C_{s,\min}$. We set $C_{s,\min}$ as 10 bits, 20 bits, and 30 bits. It can be observed that the proposed algorithm in Algorithm 1 can always converge to a certain value after only a few iterations, i.e., 3 iterations. This indicates that the proposed algorithm is convergent and computationally efficient. Besides, it can also be seen that a larger $C_{s,\min}$ brings a lower total throughput of all STs. This is because a larger $C_{s,\min}$ means a higher QoS requirement for the ST's transmission, and more resources will be allocated to the STs with worse channels, leading to a reduction in the total throughput of all STs.

Figure 3 plots the total throughput of all STs versus the minimum required throughput for each ST $C_{s,\min}$, and $C_{s,\min}$ is varied from 5 bits to 30 bits. In order to illustrate the advantages of the proposed scheme, we compare the performance achieved by the proposed scheme with the fixed scheme, where the power reflection coefficient of each ST is fixed as 0.5, 0.8, and 0.9, respectively. As shown in this figure, the total throughput of all STs decreases with the increase of $C_{s,\min}$, since a larger $C_{s,\min}$ brings a higher QoS requirement for the ST's transmission, and more resources will be allocated to the STs with worse channels, leading to a reduction in the total throughput of all STs. By comparisons, we can see that the total throughput of all STs under the proposed scheme is higher than that under the fixed scheme, as the proposed scheme provides more flexibilities to utilize resources efficiently for maximizing the total throughput of all STs. This also demonstrates the superiority of the proposed scheme.

Figure 4 shows total throughput of all STs versus the minimum SINR threshold required for decoding x_p, γ_{th}, where γ_{th} ranges from 20 to 100. The power reflection coefficient of each SN under the fixed scheme is set as 0.5, 0.8, and 0.9. It can be observed from this figure that the total throughput of all STs decreases when γ_{th} increases. This is because a higher γ_{th} brings a higher requirement for

Table 1: Key simulation settings.

Parameter	Notation	Value
The entire time block	T	1 second
The communication bandwidth	B	100 kHz
The constant circuit power consumption for BackCom at the k-th ST	$P_{b,k}$	10 μ W
The maximum transmit power at the PT	P_{max}	1 W
The performance gap reflecting the real modulation for BackCom	ξ	−15 dB
The noise power	σ^2	−60 dBm
The number of STs	K	4
The minimum required throughput of the k-th ST	$C_{min,k}$	10 bits
The minimum required throughput of the PT	C_{min}	100 bits
The threshold required for decoding x_p	γ_{th}	20

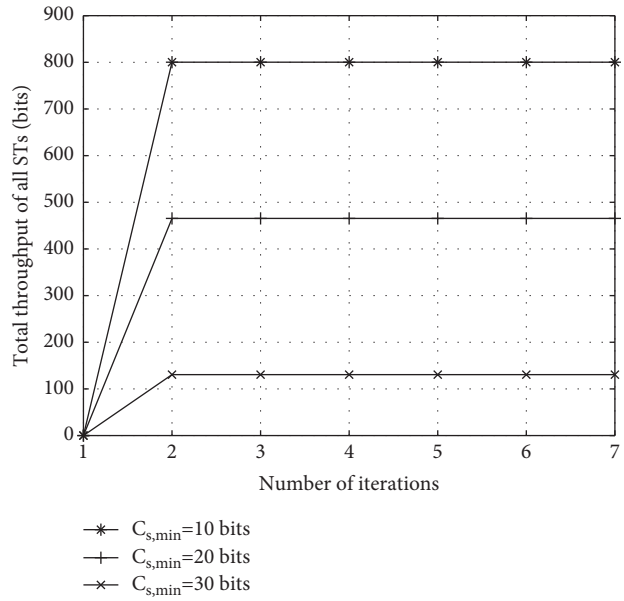


Figure 2: The convergence of Algorithm 1 under different settings of $C_{s,\min}$, where $C_{\min^1} = C_{\min^2} = C_{\min^3} = C_{\min^4} = C_{s,\min}$.

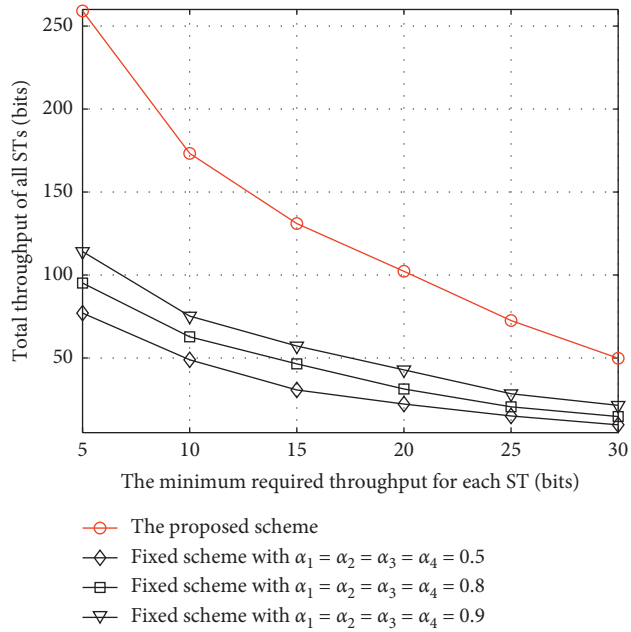


Figure 3: Total throughput of all STs versus the minimum required throughput for each ST.

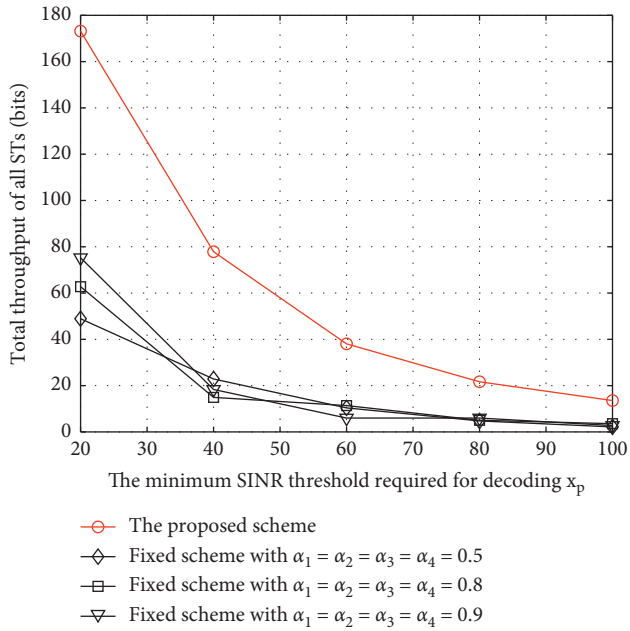


Figure 4: Total throughput of all STs versus the minimum SINR threshold required for decoding x_p.

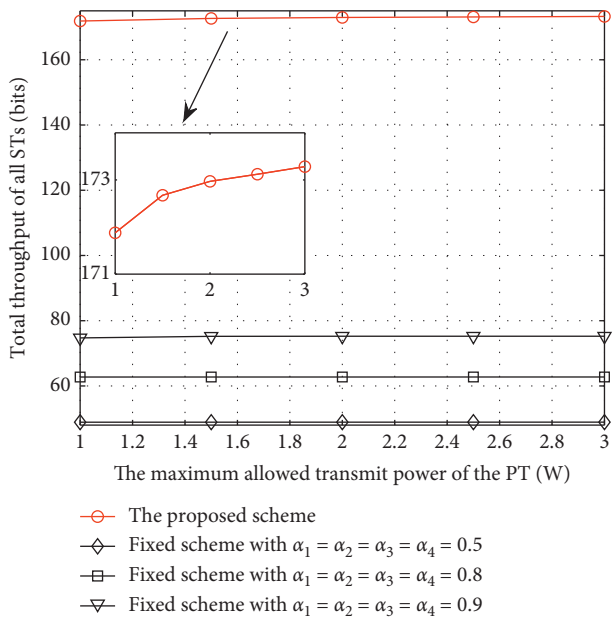


Figure 5: Total throughput of all STs versus the maximum allowed transmit power at the PT.

decoding x_p, leading to a reduction in the total throughput of all STs. By comparisons, we can observe that the proposed scheme outperforms the other schemes in terms of the total throughput of all STs, which illustrates the advantages of the proposed scheme.

Figure 5 shows the total throughput of all STs versus the maximum allowed transmit power at the PT $P_{\max}$ under different schemes. Here, $P_{\max}$ varies from 1 W to 3 W. It can be observed that the total throughput of all STs under all the schemes increase with the increase of $P_{\max}$. Based on Lemma 1, the optimal transmit power of the PT is determined by $P_{\max}$, and a higher transmit power of the PT allows STs to harvest more energy for supporting the energy consumption of BackCom and to backscatter signals with a higher power, resulting in an improvement for the total throughput of all STs. Besides, we also observe that the total throughput of all STs under the proposed scheme is the highest among these schemes, which also demonstrates the superiority of the proposed scheme in terms of total throughput of all STs.

5. Conclusions

In this paper, we have studied the throughput maximization for the BackCom based cognitive network while considering a nonlinear EH model. Specifically, we have formulated an optimization problem to maximize the total throughput of all STs by jointly optimizing the EH time, the transmit power of the PT, the BackCom time, and the power reflection coefficient of each ST under the QoS, energy causality, latency, transmit power, and power reflection coefficient constraints. In order to solve the nonconvex problem, we first determined the optimal transmit power of the PT by using the properties of the objective function and then proposed a SCA based iterative algorithm to obtain the proposed scheme. The simulation results verified the quick convergence of the proposed algorithm and showed that the proposed scheme outperforms the other schemes in terms of total throughput of all STs.

Appendix

A. Proof of Lemma 1

Let $C_k(P_t) = t_k B \log_2(1 + (\xi \alpha_k P_t h_k g_k / \eta P_t f_s + \sigma^2))$ and $R^s_{\text{tot}} = \sum_{k=1}^{K} C_k(P_t)$. By taking the first-order derivative of R^s_{tot} with respect to P_t, we have

$$\frac{\partial R^s_{tot}}{\partial P_t} = \sum\nolimits_{k=1}^{K} \frac{\partial C_k(P_t)}{\partial P_t} = \sum\nolimits_{k=1}^{K} \frac{t_k B \xi \alpha_k h_k g_k \sigma^2}{\left(\eta P_t f_s + \sigma^2\right)\left(\eta P_t f_s + \sigma^2 + \xi \alpha_k h_k g_k P_t\right) In2}. \tag{A.1}$$

Since $(\partial R^s_{\text{tot}} / \partial P_t) > 0$ always holds, R^s_{tot} is a monotone increasing function with respect to P_t. That is to say, a larger P_t brings a larger R^s_{tot}. In order to maximize R^s_{tot}, the optimal transmit power of the PT, denoted by P^*_t, should equal the maximum within its feasible region.

By observing the constraints C1, C2, C3, and C4, we can find that the lower bound of P_t is determined by C1, C2, and C4 while the upper bound of P_t is always $P_{\max}$. The reasons are as follows. As for C1, similar to R^s_{tot}, we can prove that the functions $C_k(P_t)$ and $\sum_{k=1}^{K} t_k B \log_2(1 + (P_t f_p / \alpha_k P_t h_k f_k + \sigma^2)) + t_e B \log_2(1 + (P_t f_p / \sigma^2))$ are monotone increasing functions with respect to P_t. The proof process is omitted here for brevity. Therefore, C1 determines a lower bound of P_t.

As for C2, since the harvested power of the EH circuit increases with the increase of the input power and then converges to the maximum value when the input power is large enough, $f_k(x)$ is a monotone increasing function with respect to x. That is, the right side of C2 is also a monotone increasing function with respect to P_t, and P_t determines another lower bound of P_t.

As for C4, based on (6), we can transform C4 as $P_t \geq (\gamma_{\text{th}} \sigma^2 / f_s - \gamma_{\text{th}} \alpha_k h_k g_k)$, $\forall k$, which is also a lower bound of P_t.

Therefore, the upper bound of P_t is only determined by $P_{\max}$, and the optimal transmit power of the PT is given by $P^*_t = P_{\max}$. Then, the proof of Lemma 1 is complete.

B. Proof of Proposition 1

After carefully analyzing $\mathbf{P}_3$, it is not hard to conclude that both constraints C5 and C7′ are linear constraints. Thus, $\mathbf{P}_3$ is convex if and only if the objective function is a concave function and constraints C1′′ and C2′′ are convex.

For the objective function, using the fact that the perspective function can preserve convexity, we can find that the convexity of the function $t_k B \log_2(1 + (\xi z_k P_{\max} h_k g_k / t_k (\eta P_{\max} f_s + \sigma^2)))$ is the same as that of the function $\log_2(1 + (\xi z_k P_{\max} h_k g_k / \eta P_{\max} f_s + \sigma^2))$. Since $\log_2(1 + (\xi z_k P_{\max} h_k g_k / \eta P_{\max} f_s + \sigma^2))$ is a concave function with respect to z_k, $t_k B \log_2(1 + (\xi z_k P_{\max} h_k g_k / t_k (\eta P_{\max} f_s + \sigma^2)))$ is a concave function jointly with respect to z_k and t_k. Thus, the objective function is a concave function.

For the QoS constraint for the ST's transmission in C1′′, since $t_k B \log_2(1 + (\xi z_k P_{\max} h_k g_k / t_k (\eta P_{\max} f_s + \sigma^2)))$ is a concave function jointly with respect to z_k and t_k, the QoS constraint for the ST's transmission is convex. For the QoS constraint for the PT's transmission in C1′′, $t_k B \log_2(1 + (t_k P_{\max} f_p / z_k P_{\max} h_k f_k + \sigma^2 t_k))$ is neither convex nor concave, leading to the nonconvex QoS constraint for the PT's transmission and the nonconvex constraint C1′′.

For the constraint C2′′, its convexity depends on the convexity of $t_k f_k((t_k - z_k) P_{\max} h_k / t_k)$. Based on the perspective function, the convexity of $t_k f_k((t_k - z_k) P_{\max} h_k / t_k)$ is the same as that of $f_k((1 - z_k) P_{\max} h_k)$. By taking the first-order derivative of $f_k((1 - z_k) P_{\max} h_k)$ with respect to $(1 - z_k) P_{\max} h_k$, we have

$$\frac{\partial f_k((1 - z_k) P_{\max} h_k)}{\partial (1 - z_k) P_{\max} h_k} = \frac{a_k \nu_k - d_k}{((1 - z_k) P_{\max} h_k + \nu_k)^2}. \tag{B.1}$$

Since the harvested power of the EH circuit increases with the input power, $(\partial f_k((1 - z_k) P_{\max} h_k) / \partial (1 - z_k) P_{\max} h_k)$ should be always larger than or equal to 0. Therefore, $a_k \nu_k - d_k \geq 0$ always holds. Then, we take the second-order derivative of $f_k((1 - z_k) P_{\max} h_k)$ with respect to z_k; we have

$$\frac{\partial^2 f_k((1 - z_k) P_{\max} h_k)}{\partial z_k^2} = \frac{-2 (P_{\max} h_k)^2 (a_k \nu_k - d_k)}{((1 - z_k) P_{\max} h_k + \nu_k)^3}. \tag{B.2}$$

According to $a_k \nu_k - d_k \geq 0$, we have $\partial^2 f_k((1 - z_k) P_{\max} h_k) / \partial z_k^2 \leq 0$. Therefore, $f_k((1 - z_k) P_{\max} h_k)$ is a concave function with respect to z_k. Correspondingly, $t_k f_k((t_k - z_k) P_{\max} h_k / t_k))$ is also a concave function jointly with respect to z_k and t_k, and the constraint C2′′ is convex.

Based on the above analysis, Proposition 1 is obtained, and the proof is complete.

Data Availability

The simulation data used to support the findings of this study are included within the article. The MATLAB code used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the University Key Scientific Research Project of Henan Province (no. 22A520052).

References

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: a survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[2] J. M. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, 2009.

[3] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: a contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2889–2922, 2018.

[4] Y. Ye, L. Shi, X. Chu, and G. Lu, "On the outage performance of ambient backscatter communications," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7265–7278, 2020.

[5] L. Shi, R. Q. Hu, Y. Ye, and H. Zhang, "Modeling and performance analysis for ambient backscattering underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 6563–6577, 2020.

[6] X. Kang, Y.-C. Liang, and J. Yang, "Riding on the primary: a new spectrum sharing paradigm for wireless-powered iot devices," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, France, May 2017.

[7] Y. Ye, L. Shi, R. Qingyang Hu, and G. Lu, "Energy-efficient resource allocation for wirelessly powered backscatter communications," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1418–1422, 2019.

[8] J. Wang, H.-T. Ye, X. Kang, S. Sun, and Y.-C. Liang, "Cognitive backscatter noma networks with multi-slot energy causality," *IEEE Communications Letters*, vol. 24, no. 12, pp. 2854–2858, 2020.

[9] S. Xiao, H. Guo, and Y.-C. Liang, "Resource allocation for full-duplex-enabled cognitive backscatter networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3222–3235, 2019.

[10] D. T. Hoang, D. Niyato, P. Wang, and D. I. Kim, "Optimal time sharing in rf-powered backscatter cognitive radio networks," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, France, May 2017.

[11] R. Kishore, S. Gurugopinath, P. C. Sofotasios, S. Muhaidat, and N. Al-Dhahir, "Opportunistic ambient backscatter communication in RF-powered cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 413–426, 2019.

[12] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Ambient backscatter: a new approach to improve network performance for RF-powered cognitive radio networks," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3659–3674, 2017.

[13] T. Wang, G. Lu, Y. Ye, and Y. Ren, "Dynamic power splitting strategy for SWIPT based two-way multiplicative AF relay networks with nonlinear energy harvesting model," *Wireless Communications and Mobile Computing*, vol. 2018, no. 1, pp. 1–9, 2018.

[14] E. Boshkovska, D. W. K. Ng, N. Zlatanov, A. Koelpin, and R. Schober, "Robust resource allocation for mimo wireless powered communication networks based on a non-linear eh model," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 1984–1999, 2017.

[15] H. Yang, Y. Ye, X. Chu, and M. Dong, "Resource and power allocation in SWIPT-enabled device-to-device communications based on a nonlinear energy harvesting model," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 10813–10825, 2020.

[16] Y. Chen, N. Zhao, and M.-S. Alouini, "Wireless energy harvesting using signals from multiple fading channels," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 5027–5039, 2017.

[17] Y. Liu, Y. Ye, H. Ding, F. Gao, and H. Yang, "Outage performance analysis for SWIPT-based incremental cooperative NOMA networks with non-linear harvester," *IEEE Communications Letters*, vol. 24, no. 2, pp. 287–291, 2020.

[18] L. Shi, Y. Ye, R. Q. Hu, and H. Zhang, "Energy efficiency maximization for SWIPT enabled two-way DF relaying," *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 755–759, 2019.

[19] Y. Ye, L. Shi, X. Chu, and G. Lu, "Throughput fairness guarantee in wireless powered backscatter communications with HTT," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 449–453, 2021.

[20] S. H. Kim and D. I. Kim, "Hybrid backscatter communication for wireless-powered heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6557–6570, 2017.

[21] H. Yang, Y. Ye, X. Chu, and S. Sun, "Energy efficiency maximization for UAV-enabled hybrid backscatter-harvest-then-transmit communications," *IEEE Transactions on Wireless Communications*, vol. 1, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1400114, 10 pages
https://doi.org/10.1155/2022/1400114

Research Article

Influence of Sublevel Unloading Excavation with Deep Consideration of the Superposition Effect on Deformation of an Existing Tunnel under an Intelligent Geotechnical Concept

Xiangling Tao,[1,2,3] **Pinzhi Luan**,[2] **Jinrong Ma**,[2] **and Weihua Song**[4]

[1]*College of Innovation and Entrepreneurship, Jiangsu Vocational Institute of Architectural Technology, Xuzhou 221116, China*
[2]*State Key Laboratory for Geomechanics and Deep Underground Engineering, China University of Mining and Technology, Xuzhou 221116, China*
[3]*Anhui Huizhou Geology Security Institute Co. Ltd., Hefei 230000, China*
[4]*Xuzhou New Town State-Owned Assets Management Co. Ltd., China*

Correspondence should be addressed to Xiangling Tao; taoxl@cumt.edu.cn

Received 29 November 2021; Revised 17 December 2021; Accepted 22 December 2021; Published 31 January 2022

Academic Editor: Liqin Shi

The construction of a deep and large foundation will inevitably impose stress and deformation on the existing shield tunnel structure. Based on the two-stage analysis and the Pasternak foundation beam model, the analytical solution for the existing tunnel deformation is derived by taking the grouting pressure and water buoyancy into account in the total tunnel deformation. Using FLAC 3D software, based on the soil-structure interaction model, the variation law of the tunnel uplift value under partition excavation of the foundation pit is studied and the variation law of the stress field and displacement field of the tunnel under the MJS method is analyzed and compared. The results of this paper are worthy of reference for similar projects.

1. Introduction

With the rapid construction of urban underground space development, it is often encountered that the new foundation pit or tunnel project is adjacent to the existing tunnel construction. Therefore, the deformation analysis of the built operation tunnel by underground space unloading development has become a key problem to be solved in projects and scholars are gradually keen on analyzing the influence of excavation on existing tunnels.

The unloading of foundation pit excavation will cause the change of the displacement field and stress field of the underlying shield tunnel. When the vertical displacement or lateral convergence deformation of the shield is large, it will seriously affect the safe use and operation of the existing tunnel [1–5]. At present, many scholars have used theoretical calculation [6–8], numerical simulation [9–12], and measured analysis [13, 14] to conduct in-depth analysis on the deformation of the underlying tunnel caused by foundation pit excavation under different geological conditions or construction schemes. In addition, many factors are also considered in the research on the mechanism of stress and displacement of a tunnel caused by loading or unloading. Zhang et al. [15] considered the role of the foundation pit supporting structure, established the additional confining pressure variation model of adjacent shield tunnels considering the influence of longitudinal deformation, and obtained the variation law of the influence of foundation pit excavation on the lateral force of shield tunnels.

Jiang [16] introduced the rheological deformation of soil in the viscoelastic-plastic theoretical model and analyzed the deformation law of existing tunnels under different relative distances, excavation areas, and construction procedures in the process of foundation pit excavation. Zhang et al. [17] considered the dewatering effect of the foundation pit and obtained the theoretical calculation formula of additional stress and vertical deformation of the underlying existing tunnel caused by upper excavation and dewatering. Hefny

and Chua [18] used PLAXIS software combined with specific working conditions to study the influence of the new tunnel close crossing construction on the existing tunnel. Tao et al. [19] took the segment floating of Q3 clay subway tunnel excavation in Xuzhou as the research object and obtained the analytical solution of segment floating caused by shield tunnel excavation. Liang et al. [20] considered the bending and shear effects of the shield tunnel in solving the tunnel deformation caused by foundation pit excavation and compared it with the simulation results and measured values to verify the correctness of the theoretical solution. Tan et al. [21] analyzed the influence of partition excavation of foundation pit on subway tunnel around foundation pit from the perspective of engineering measured data under sensitive geological conditions of hard clay. Sun et al. [22] studied the deformation law of tunnel under different excavation methods of foundation pit. Zhang et al. [23] simplified the existing shield tunnel as a Timoshenko beam placed on the Pasternak foundation, and the longitudinal deformation formula of the existing shield tunnel induced by ground surcharge was derived. In summary, the existing research considers many factors when solving the stress and deformation law of the tunnel caused by the excavation of the foundation pit but few scholars consider the deformation caused by the excavation of the shield tunnel into the total deformation of the tunnel. Especially when the construction time of the foundation pit and the tunnel is close, the superposition construction effect of tunnel excavation and foundation pit excavation will cause large deformation of the tunnel. Therefore, based on the above research results, this paper analyzes the disturbance effect of the partition unloading of the foundation pit on the existing tunnel. Based on two-stage analysis method and Pasternak foundation beam model, a theoretical model of segment floating is established by spatializing the time axis. Solving the deformation law of the tunnel under superposition provides a more accurate calculation method of tunnel deformation for such projects. In addition, based on the overlapping passage project of Xuzhou Metro Line 2, this paper uses FLAC3D to carry out simulation and makes a more in-depth analysis of the internal force change and deformation mechanism of the tunnel under partition unloading, in order to provide a basis for the rational planning and construction of similar foundation pit projects.

2. Analytical Solution of Tunnel Vertical Deformation Caused by Excavation Unloading

2.1. Calculation Method of Additional Stress of the Tunnel Structure Caused by Excavation. Since the excavation time of the foundation pit is after the completion of tunnel construction, the interaction between the two on tunnel deformation is weak, so the tunnel deformation during tunnel construction and during foundation pit excavation is calculated separately in this paper. The excavation of the foundation pit in the existing subway tunnel belongs to the internal unloading effect of the soil. When calculating the tunnel deformation, the Mindlin elastic solution [24] can be used to calculate the vertical load acting on the tunnel structure first.

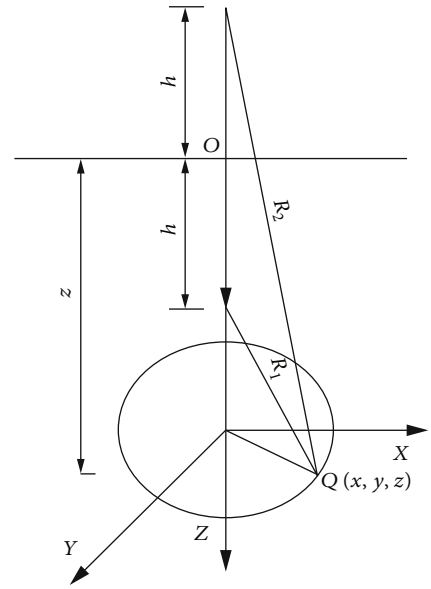


Figure 1: Mindlin solution diagram.

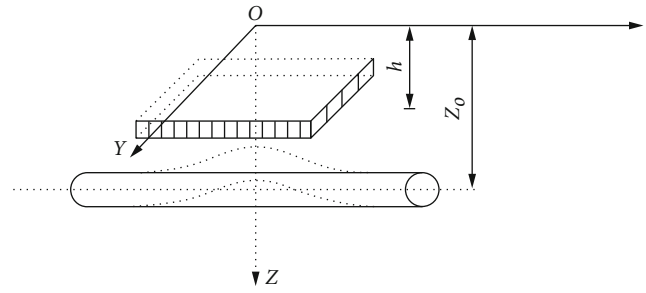


Figure 2: Mindlin diagram of the foundation pit above the tunnel.

Figure 1 shows the single-tunnel diagram of Mindlin solution. On this basis, we expanded it, combined with the actual working situation on the site, and conducted theoretical derivation again. Figure 2 shows the Mindlin solution diagram of the foundation pit above the tunnel.

When a point (ξ, η) acts on a concentrated force $pd\xi d\eta$ at depth h, the vertical additional stress at any point Q (x, y, z) in a semi-infinite elastic space soil is

$$\begin{aligned}\sigma_z = \frac{p}{8\pi(1-\nu)}\Bigg\{&(1-2\nu)(z_0-h)\iint_D \frac{d\xi d\eta}{R_1^3} \\ &+ 3(z_0-h)^3\iint_D \frac{d\xi d\eta}{R_1^5} - (1-2_\nu)(z_0-h)\iint_D \frac{d\xi d\eta}{R_2^3} \\ &+ \left[3(3-4_\nu)z_0(z_0+h)^2 - 3h(z_0+h)(5z_0-h)\right] \\ &\times \iint_D \frac{d\xi d\eta}{R_2^5} + 30hz_0(z_0+h)^3\iint_D \frac{d\xi d\eta}{R_2^7}\Bigg\},\end{aligned} \tag{1}$$

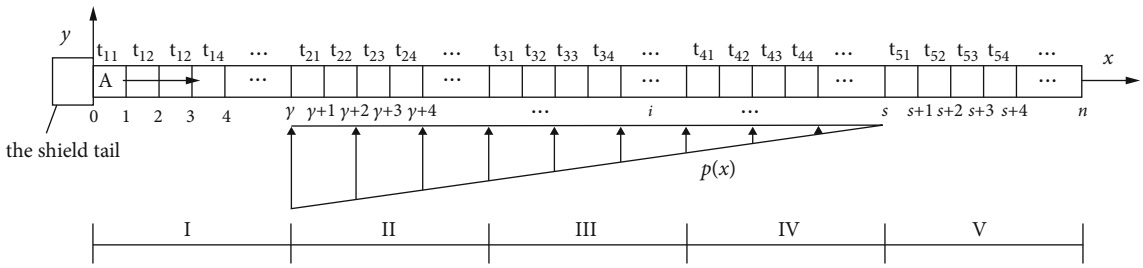


Figure 3: Calculation model of shield floating considering grouting pressure and water buoyancy.

with

$$R_1 = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z_0 - h)^2},$$
$$R_2 = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z_0 + h)^2}, \tag{2}$$

where ν is the Poisson's ratio, h is the excavation depth of the foundation pit, and z is the distance between the tunnel center line and the surface.

2.2. Solution of Vertical Deformation of the Tunnel Structure Caused by Foundation Pit Excavation. When calculating the vertical deformation at any position of the tunnel caused by the excavation of the foundation pit, considering the consistency with the calculation method of the floating amount caused by the tunnel construction, this paper uses the Pasternak foundation model to calculate the deformation of the existing tunnel under the additional load and the Pasternak foundation model can simulate the shear force between the soil springs and the continuity of the deformation; thus, the calculation results are more accurate. The stress of the tunnel is substituted into the model, and after a series of simplifications, the vertical deformation of the tunnel caused by the excavation of the foundation pit at any point can be obtained.

$$\omega(x) = e^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x) + e^{-\alpha x}(C_3 \cos \beta x + C_4 \sin \beta x). \tag{3}$$

3. Analytical Solution of Tunnel Vertical Deformation considering Floating Factor

Figure 3 shows the upward deformation caused by grouting pressure, and water buoyancy often occurs during shield tunnel construction.

In the paper, the segment uplift is divided into five stages: (I) the ungrouted stage, (II) the fast uplift stage, (III) the slow uplift stage, (IV) the equilibrium stage, and (V) grout solidification.

According to the Pasternak foundation beam model, the differential equation of segment deflection caused by tunnel construction is

$$EI\frac{d^4y}{dx^4} - GB\frac{d^2y}{dx^2} + KBy = P(x), \tag{4}$$

where EI is the bending stiffness of the foundation beam (kN·m^2), y is the floating capacity (m), K is the foundation bed coefficient (kN/m^3), x is the longitudinal axis (m), G is the shear stiffness of foundation soil (kN/m), B is the width of the tube ring (m), P is buoyancy (kN), t is the time axis (m) after spatialization.

General solution of homogeneous equation of differential equation of torsion curve can be obtained according to formula (4)

$$y = e^{\alpha x}[C_1 \cos(\beta t) + C_2 \sin(\beta t)] + e^{-\alpha x}[C_3 \cos(\beta t) + C_4 \sin(\beta t)], \tag{5}$$

where C_1, C_2, C_3, and C_4 in the expression are constants to be determined and their values are determined by the boundary conditions. Since in equation (4), it is a special solution of equation (4). So the general solution of formula (4) is:

$$y = e^{\alpha t}[C_1 \cos(\beta t) + C_2 \sin(\beta t)] + e^{-\alpha x}[C_3 \cos(\beta t) + C_4 \sin(\beta t)] + \frac{at+b}{KB}. \tag{6}$$

The influence of the shield tail on point 0 is regarded as hinged support. Since the slurry has been solidified as a fixed end at point 3, the vertical displacement and bending moment of node 0 are equal to 0, the vertical displacement and rotation angle of node 3 are equal to 0, and the boundary conditions are

$$y_{10}|_{t=0} = 0,$$
$$y'_{10}|_{t=0} = 0,$$
$$y_{33}|_{t=3} = 0,$$
$$y'_{33}|_{t=3} = 0. \tag{7}$$

As shown in Figure 2, the connection points 1 and 2 of segment assembly (I), segment floating (II, III, and IV), and slurry (V) fully solidified; the deflection, rotation angle, bending moment, and shear force of point 1 of segment I, point 1 of segment II, point 2 of segment IV, and point 2

of segment V are the same, so the coordination equation between foundation beams is

$$\begin{aligned} y_{ij} &= y_{(i+1)j}, \\ \frac{dy_{ij}}{dt} &= \frac{dy_{(i+1)j}}{dt}, \\ E_i I_i \frac{d^2 y_{ij}}{dt^2} &= E_{i+1} I_{i+1} \frac{d^2 y_{(i+1)j}}{dt^2}, \\ E_i I_i \frac{d^3 y_{ij}}{dt^3} &= E_{i+1} I_{i+1} \frac{d^3 y_{(i+1)j}}{dt^3}, \end{aligned} \tag{8}$$

where y_{ij} is the deformation of the beam and j node in the i section (i is the beam element number, j is the beam node number, $i = j = 1, 2$). $E_i I_i$ is the equivalent stiffness of each segment. For formula (5), we obtain the 1st, 2nd, and 3rd derivatives, respectively, and then replace them with the boundary conditions and the coordination equation to obtain 12 square matrices:

$$[K][C] = [B]. \tag{9}$$

In the formula,

$$[K] = \begin{bmatrix} [K_{10}] & & & \\ [K_{11}] & -[K_{21}] & & \\ & -[K_{22}] & [K_{32}] & \\ & & & [K_{33}] \end{bmatrix},$$

$$[C] = [C_{11}\ C_{12}\ C_{13}\ C_{14}\ C_{21}\ C_{22}\ C_{23}\ C_{24}\ C_{31}\ C_{32}\ C_{33}\ C_{34}]^T,$$

$$[B] = \left[0 \quad 0 \quad \frac{at+b}{K_2 B_2} \quad \frac{a}{K_2 B_2} \quad 0 \quad 0 \quad \frac{at+b}{K_2 B_2} \quad \frac{a}{K_2 B_2} \quad 0 \quad 0 \quad 0 \quad 0\right]^T,$$

$$[K_{10}] = \begin{bmatrix} \gamma_{10} & \varepsilon_{10} & \lambda_{10} & \eta_{10} \\ \gamma'_{10} & \varepsilon'_{10} & \lambda'_{10} & \eta'_{10} \end{bmatrix},$$

$$[K_{11}] = \begin{bmatrix} \gamma_{11} & \varepsilon_{11} & \lambda_{11} & \eta_{11} \\ \gamma'_{11} & \varepsilon'_{11} & \lambda'_{11} & \eta'_{11} \\ \gamma''_{11} & \varepsilon''_{11} & \lambda''_{11} & \eta''_{11} \\ \gamma'''_{11} & \varepsilon'''_{11} & \lambda'''_{11} & \eta'''_{11} \end{bmatrix},$$

$$[K_{21}] = \begin{bmatrix} \gamma_{21} & \varepsilon_{21} & \lambda_{21} & \eta_{21} \\ \gamma'_{21} & \varepsilon'_{21} & \lambda'_{21} & \eta'_{21} \\ \gamma''_{21} & \varepsilon''_{21} & \lambda''_{21} & \eta''_{21} \\ \gamma'''_{21} & \varepsilon'''_{21} & \lambda'''_{21} & \eta'''_{21} \end{bmatrix},$$

$$[K_{22}] = \begin{bmatrix} \gamma_{22} & \varepsilon_{22} & \lambda_{22} & \eta_{22} \\ \gamma'_{22} & \varepsilon'_{22} & \lambda'_{22} & \eta'_{22} \\ \gamma''_{22} & \varepsilon''_{22} & \lambda''_{22} & \eta''_{22} \\ \gamma'''_{22} & \varepsilon'''_{22} & \lambda'''_{22} & \eta'''_{22} \end{bmatrix},$$

$$[K_{32}] = \begin{bmatrix} \gamma_{32} & \varepsilon_{32} & \lambda_{32} & \eta_{32} \\ \gamma'_{32} & \varepsilon'_{32} & \lambda'_{32} & \eta'_{32} \\ \gamma''_{32} & \varepsilon''_{32} & \lambda''_{32} & \eta''_{32} \\ \gamma'''_{32} & \varepsilon'''_{32} & \lambda'''_{32} & \eta'''_{32} \end{bmatrix},$$

$$[K_{33}] = \begin{bmatrix} \gamma_{33} & \varepsilon_{33} & \lambda_{33} & \eta_{33} \\ \gamma'_{33} & \varepsilon'_{33} & \lambda'_{33} & \eta'_{33} \end{bmatrix}. \tag{10}$$

The undetermined constant C_{ij} ($i = 1,2,3$) can be obtained by solving formula (9). $j = 1,2,3,4$ and the undetermined constants are substituted back to formula (5), that is, the theoretical solution of the static uplift of each beam section at the stress stage at that time. By multiplying the theoretical solution of the static buoyancy of each beam segment by the T-hour compression reduction coefficient of the soil $[H]$, the theoretical solution of the segment buoyancy can be obtained:

$$L = \int y_i \cdot H_i dt + C_l, \tag{11}$$

where C_1 is the floating volume of the segment caused by other construction factors (including the grouting method and slurry properties) and it is a fixed constant under certain working conditions. The total uplift of the tunnel can be obtained by superposition of the vertical deformation of the tunnel obtained by the two methods.

4. Project Overview and Numerical Model Establishment

The foundation pit of the passage project crossing Xuzhou Metro Line 2 is constructed by the open cut and sequential method. The excavation area of the foundation pit is 1600 m^2, and the excavation depth is 10.8 m. The bottom of the foundation pit is only 2.4 m away from the roof of the tunnel. The soil inside and outside the foundation pit is reinforced by the MJS method, and the foundation pit is supported by the combination of the bored pile with Φ800 mm @1000 mm and the internal support. The diameter of the existing shield tunnel is 6.2 m, the thickness of the segment is 0.3 m, and the net distance between the left and right tunnels is about 8 m, as shown in Figure 4.

Considering the influence on the shield tunnel, the safety level of the foundation pit is first grade. Due to the large range of foundation pit excavation, in order to reduce the influence of the overall one-time excavation on the existing tunnel structure, the foundation pit is divided into six partitions for block excavation considering the actual working conditions, as shown in Figure 5.

The actual working condition of the site foundation pit is a parallelogram, but in order to facilitate the theoretical derivation and the convenience of the subsequent FLAC 3D modeling, the shape of the foundation pit is regularized to become a rectangle.

According to the geological survey report, the strata involved can be mainly divided from top to bottom into

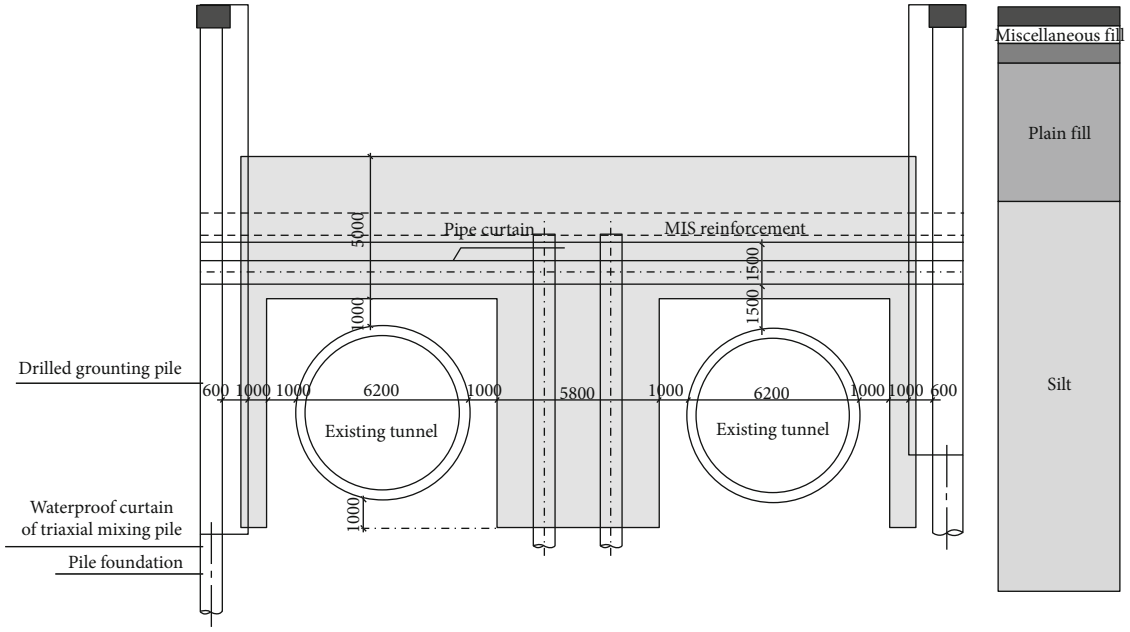


Figure 4: Vertical stratigraphic labeling on the left side of the longitudinal section of the foundation pit crossing the subway node.

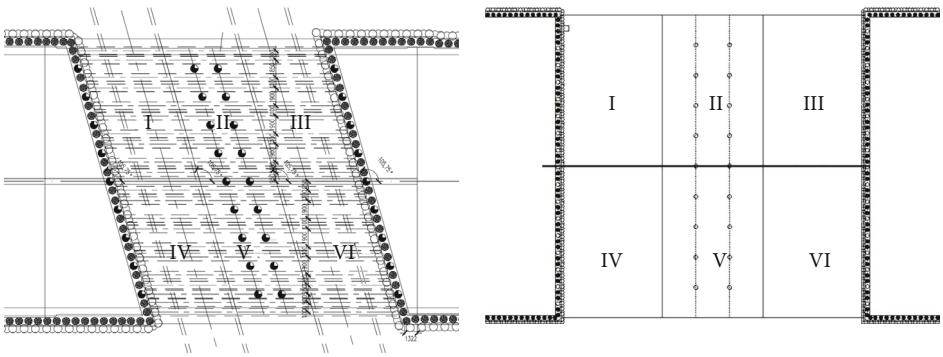


Figure 5: Construction process diagram of the foundation pit.

miscellaneous fill, plain fill, clay, silt, and moderately weathered limestone. The solid element is used for the simulation of the stratum, reinforced soil, shield segment, and pile body. The beam structural unit is used for the simulation of the support in the foundation pit, and the liner structural unit is used for the simulation of the diaphragm wall. The parameters of the constitutive model are shown in Table 1.

In order to ensure the safety of the existing subway tunnel structure, the soil around the tunnel is usually reinforced by the MJS method in the project. As presented in Figure 6, the thickness of the solid added above the tunnel is 4.5 m and the thickness of the solid added on both sides of the tunnel is 2–6 m. There is 1 m safe distance between the solid added and the tunnel. Two rows of uplift piles are set between the left and right tunnels to connect with the bottom plate of the foundation pit structure, and the length of bored piles on both sides of the tunnel is appropriately increased.

The simulation is divided into 5 analysis steps according to different construction stages: (1) applying gravity stress field, balancing ground stress, and clearing displacement; (2) excavating the tunnel soil, activating and assigning the elastic element parameters to the tunnel segment, and using the "apply nstress" command to force the surface to simulate the grouting pressure; (3) MJS reinforcement of the soil around the tunnel and construction of uplift piles; (4) foundation pit partition construction; and (5) connecting the construction structural floor with the uplift pile.

The optimal construction sequence of each partition needs to be determined after it is necessary to determine the optimal construction sequence of each zone. Zone II and zone V are suitable for the first construction because the excavation area is small and the disturbance effect on soil and surrounding existing structures is weak. The overall construction sequence of area I and area IV should be earlier than that of area III and area VI, because there are many

Table 1: Calculation parameters of the soil layer.

Soil type	Thickness (m)	Bulk modulus (MPa)	Shear modulus (MPa)	Density (kg/m^3)	Cohesion (KPa)	Friction angle (°)
Miscellaneous fill	2.2	5	3.2	1700	10	8
Plain fill	4.9	12.5	5.0	1860	21	10
Silt	13.8	16.0	8.6	1950	12	18
Clay	10.5	18.6	9	1960	75	22
Medium weathered limestone	8.6	840	490	2300	240	37

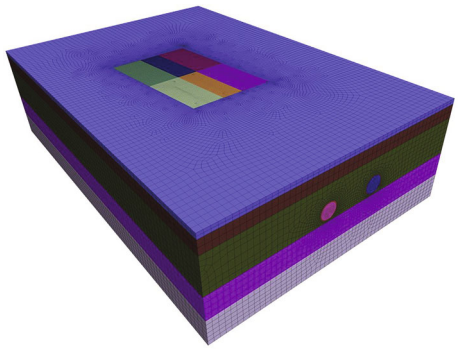


Figure 6: 3D numerical calculation model.

Table 2: Excavation sequence table and final maximum uplift value of the tunnel.

Scheme number	Excavation sequence	Tunnel maximum uplift value (mm)
1	II ⟶ V ⟶ I ⟶ VI ⟶ IV ⟶ III	12.86
2	II ⟶ I ⟶ III ⟶ V ⟶ IV ⟶ VI	12.13
3	V ⟶ II ⟶ IV ⟶ III ⟶ I ⟶ VI	13.65
4	V ⟶ IV ⟶ VI ⟶ II ⟶ I ⟶ III	13.25

buildings nearby and the surrounding environment is poor. Based on this, the different excavation sequences and the maximum uplift values of the final tunnel obtained by the above simulation analysis steps of each scheme are shown in Table 2.

It can be seen in the table that the influence of the change of the excavation sequence of the foundation pit on the deformation of the tunnel is different. The final maximum uplift value of the tunnel under scheme 2 is the smallest, and the excavation sequence of scheme 2 should be selected for foundation pit construction. The subsequent numerical simulation also selects the construction sequence under the scheme to study the deformation and stress law of the tunnel under different working conditions.

5. Calculation Results and Comparative Analysis

5.1. Tunnel Deformation Analysis without Reinforcement Measures. In order to verify and compare the vertical deformation results of the tunnel obtained by the theoretical method and to explore the weakening effect of the MJS method on the stress and deformation transfer of the soil around the tunnel, it is necessary to carry out the excavation simulation of the foundation pit without reinforcement measures. The vertical deformation of the tunnel under this condition is shown in Figure 7.

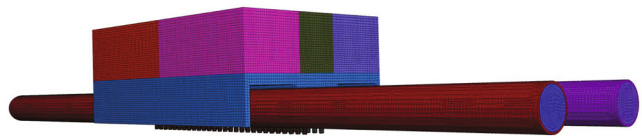


Figure 7: 3D grid division diagram of tunnel reinforcement.

The calculation shows that after the excavation of the foundation pit, the maximum uplift of the tunnel without reinforcement measures reaches 28.57 mm, which occurs in the position of the left line tunnel near the center line of the foundation pit, and the maximum uplift of the right line tunnel is 26.83 mm, which obviously exceeds the allowable deformation of the subway shield tunnel, which seriously affects the safe use and normal operation of the existing subway tunnel. Since the excavation of the upper area of the left line tunnel is earlier than that of the upper area of the right line tunnel, the uplift value of the left line tunnel is greater than that of the right line. As the position close to the center line of the foundation pit, the tunnel uplift is gradually increasing. The maximum uplift deformation of the tunnel under different construction stages is shown in Figure 8.

It is shown that the maximum uplift of the tunnel increases gradually with the construction steps. After the completion of the tunnel construction, the maximum uplift value is 7.69 mm, accounting for 26.9% of the total deformation, indicating that the grouting pressure has caused a certain amount of uplift of the segment in the tunnel shield construction. The increase of tunnel uplift caused by excavation construction in II, I, and III zones in the north of the foundation pit is 13.31 mm, which is about 1.8 times of that caused by excavation construction in V, IV, and VI zones in the south of the foundation pit. In the construction process of the north of the foundation pit, more attention should be paid to the control of tunnel deformation. In addition, in the process of foundation pit construction, due to the large excavation area and the early construction sequence, the excavation of area I has the greatest impact on the uplift value of tunnel segments. The deformation increase in this

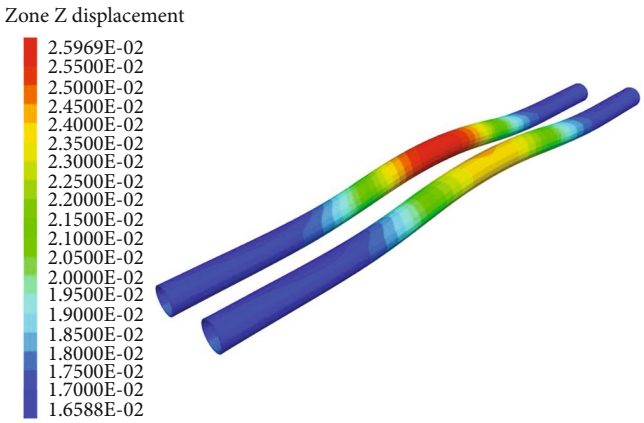


Figure 8: Cloud images of tunnel uplift deformation without reinforcement measures.

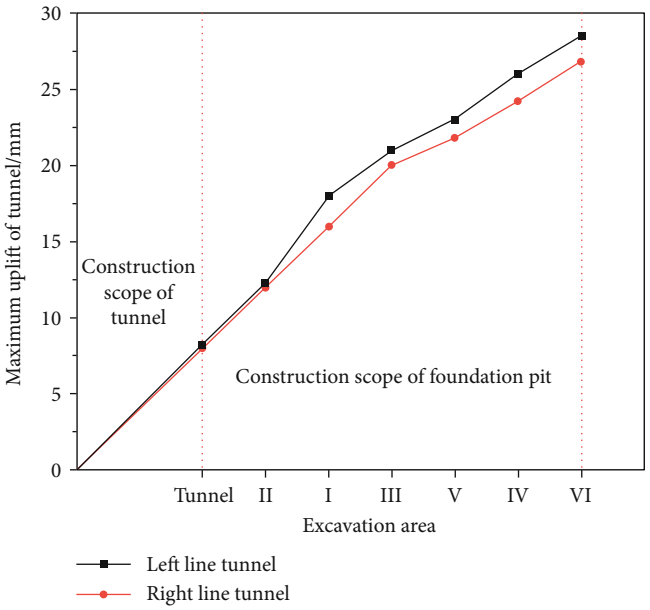


Figure 9: Maximum tunnel uplift deformation under different construction stages.

stage is 5.74 mm, accounting for about 20.1% of the tunnel uplift deformation in the whole construction process. Therefore, the reinforcement strength can be appropriately increased below this area.

5.2. Tunnel Deformation and Stress Analysis under Reinforcement Measures. In order to further study the control effect of the reinforcement method on tunnel deformation, the vertical deformation nephogram and horizontal deformation nephogram of the tunnel under the condition of reinforcement measures are extracted as shown in Figure 9.

It can be found that the uplift deformation of the tunnel is effectively limited after the reinforcement measures are taken and the maximum uplift is 12.13 mm, which is reduced by 57.84% compared with that before the reinforcement, which meets the limit displacement requirements of the subway tunnel. It shows that the MJS reinforcement has obvious control effect on the deformation of the tunnel structure. The maximum uplift position of the tunnel is transferred to the bottom of the tunnel. This is because after the reinforcement measures are taken, the uplift deformation of the tunnel is mainly caused by the grouting pressure during the shield construction of the tunnel and the excavation unloading of the tunnel itself, so that the overall floating amount below the tunnel is greater than that above. At the same time, as is known that the horizontal displacement of the tunnel is small, the maximum value is 2.35 mm, which also meets the tunnel deformation limit. The horizontal displacement of the inner part of the left and right line tunnel under the foundation pit is small due to the blocking effect of the solid and uplift pile, which verifies the effectiveness of the reinforcement scheme.

In order to further study the effect of the reinforcement scheme on the tunnel, the cloud picture of the vertical stress field around the tunnel structure at the section where the maximum deformation after foundation pit excavation is located is extracted as shown in Figure 10.

It can be seen from the analysis that in the area far from the excavation of the tunnel foundation pit, the stress distribution is still approximately linear with the soil depth and the vertical stress field of the soil around the excavation area is greatly disturbed. Under the action of the gravity of the reinforcement, the soil stress in the area below the tunnel presents a "groove type" distribution. Due to the self-weight stress of the soil transmitted by the upper structure and under the combined action of the grouting pressure, the vertical stress is slightly larger than that of the upper part and the maximum value appears at the arch waist of the tunnel.

5.3. Comparative Analysis of Calculation Results. The parameters such as grouting pressure, Poisson's ratio, and foundation pit excavation depth are substituted into formulas (3) and (11), and the cumulative buoyancy of the tunnel can be obtained by superposition of the two calculation results. In order to verify the accuracy and reliability of the theoretical calculation formula and the simulation parameters in this paper, the deformation law of the tunnel is further analyzed; therefore, the comparative curves of the tunnel uplift under the analytical calculation results, the measured results, and the numerical calculation results under the above two working conditions are drawn, as shown in Figure 11.

The theoretical values are also shown in Figure 12. It can be found that the tunnel under different calculation methods and working conditions shows a bending state of large deformation in the middle and small deformation at both ends. The curve of numerical simulation results under the condition of reinforcement measures can better fit the variation law of measured values, indicating that the value of simulation parameters and the establishment of the model are more in line with the actual working conditions. The maximum uplift deformation of the tunnel obtained by field measurement is 11.24 mm, which is 7.3% smaller than the numerical simulation results. This is because of the obvious

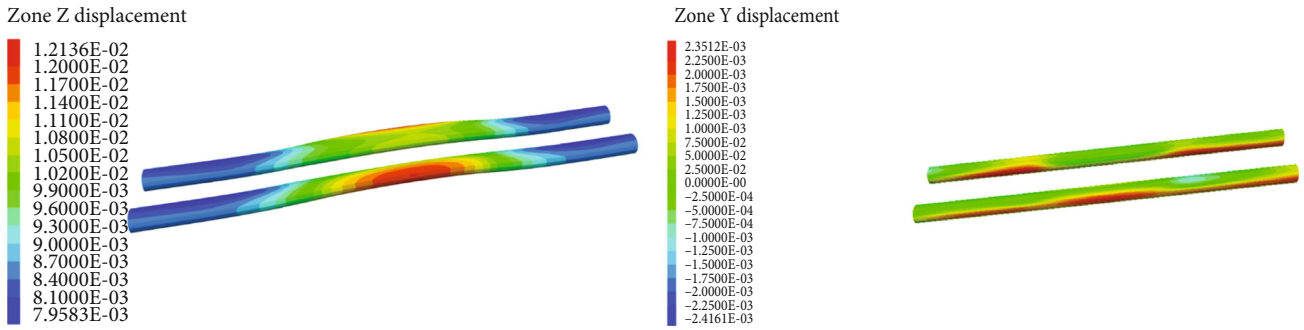


Figure 10: Tunnel deformation under reinforcement measures.

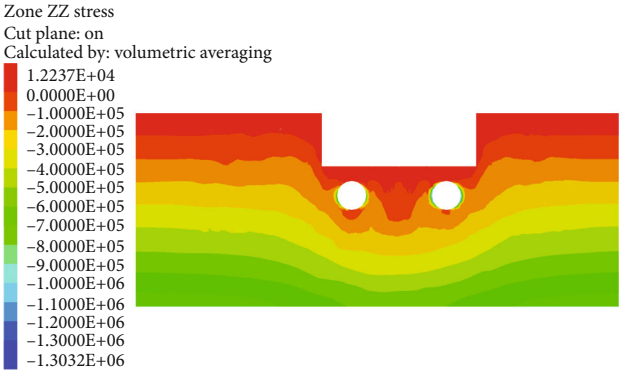


Figure 11: Vertical stress field around the tunnel structure after foundation pit excavation.

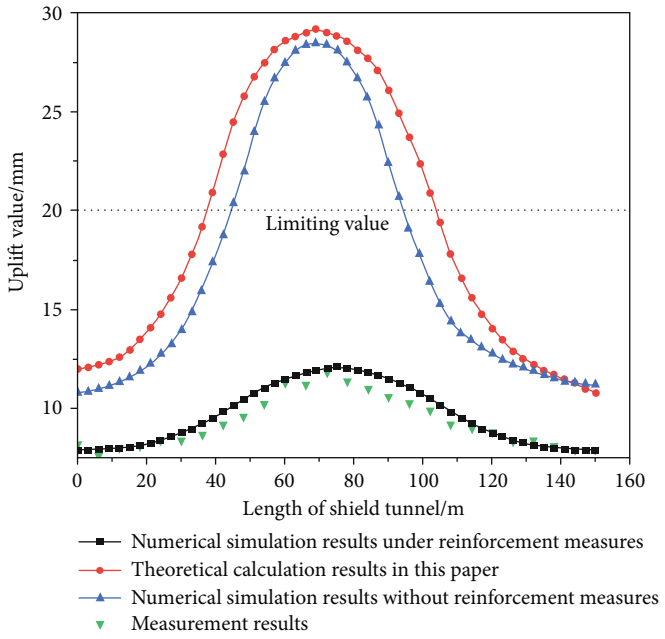


Figure 12: Comparison of tunnel uplift values obtained by different methods.

space-time effect and the looseness of the soil in the actual construction process and the influence depth of foundation pit excavation is smaller than that of the simulation. Therefore, the measured uplift is smaller than the simulation value after the reinforcement measures are taken but the variation law and value of the two values are close, indicating that the parameter value in the simulation is reasonable. From an overall perspective, the variation law of the tunnel uplift value obtained by theoretical calculation and numerical simulation under the condition without reinforcement measures is basically consistent, indicating that the theoretical calculation can better predict the deformation law of the tunnel.

Because the superposition effect of the tunnel and foundation pit construction is ignored in the theoretical calculation and the stiffness of the tunnel segment is larger in the numerical simulation, the theoretical value is slightly larger than the simulation value. The maximum uplift deformation of the tunnel obtained by the theory is 29.52 mm, and the value obtained by the simulation is 3.2% smaller than the theoretical value, indicating that the theoretical calculation results are more conducive to safety.

6. Conclusion

A theoretical calculation method is proposed to obtain the analytical solution of the floating amount of the existing shield tunnel under the superposition effect of the foundation pit and tunnel construction, which provides a more accurate prediction method for the tunnel deformation law under such projects. Through numerical simulation analysis, the deformation law of the tunnel under partition excavation and MJS reinforcement is studied and compared.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

The authors acknowledge the support received from the Natural Science Fund for Colleges and Universities in Jiangsu Province (no. 20KJA560003), China Postdoctoral Science Foundation (no. 2020M681769), and Guidance Project of Housing and Urban-Rural Development Department of Jiangsu Province (2019ZD079 and 2017ZD094). This work was also supported by the Qing Lan Project and Postdoctoral Workstation of Hefei, as well as the JiangSu Collaborative Innovation Center for Building Energy Saving and Construction Technology (Grant: SJXTBS1710) and Xuzhou Science and Technology Project (Grant: KC18133).

References

[1] Z. G. Zhang, M. S. Huang, and W. D. Wang, "Evaluation of deformation response for adjacent tunnels due to soil unloading in excavation engineering," *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research*, vol. 38, pp. 244–253, 2013.

[2] J. S. Sharma, A. M. Hefny, J. Zhao, and C. W. Chan, "Effect of large excavation on deformation of adjacent MRT tunnels," *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research*, vol. 16, no. 2, pp. 93–98, 2001.

[3] X. Wen and C. R. Pang, "Influence of foundation pit excavation on existing shield tunnel and its protection range," *Applied Mechanics & Materials*, vol. 580-583, pp. 1258–1263, 2014.

[4] M. Devriendt, L. Doughty, P. Morrison, and A. Pillai, "Displacement of tunnels from a basement excavation in London," *Proceedings of the Institution of Civil Engineers*, vol. 163, no. 3, pp. 131–145, 2010.

[5] C. Jing, G. W. Qian, and H. X. Zuo, "Analysis for influence of excavation on adjacent existing shield tunnel," in *Proceedings of the 2016 4th International Conference on Mechanical Materials and Manufacturing Engineering*, Wuhan, People's Republic of China, 2016.

[6] J. T. Qiu, J. Jiang, X. J. Zhou, Y. F. Zhang, and Y. D. Pan, "Analytical solution for evaluating deformation response of existing metro tunnel due to excavation of adjacent foundation pit," *Journal of Central South University*, vol. 28, no. 6, pp. 1888–1900, 2021.

[7] D. M. Zhang, X. C. Xie, Z. L. Li, and J. Zhang, "Simplified analysis method for predicting the influence of deep excavation on existing tunnels," *Computers and Geotechnics*, vol. 121, no. 2, article 103477, 2020.

[8] J. W. Liu, C. H. Shi, and M. F. Lei, "Analytical method for influence analysis of foundation pit excavation on underlying metro tunnel," *Journal of Central South University: Science and Technology*, vol. 50, no. 9, pp. 2215–2225, 2019.

[9] S. Q. Yan, C. H. Qiu, and J. F. Xu, "Numerical simulation of a deep excavation near a shield tunnel," *Tehnički vjesnik*, vol. 25, 2018.

[10] S. H. Ye and Z. H. Zhao, "Deformation analysis and safety assessment of existing metro tunnels affected by excavation of a foundation pit," *Underground Space*, vol. 6, no. 4, pp. 421–431, 2021.

[11] Z. T. Yu, H. Y. Wang, and W. J. Wang, "Experimental and numerical investigation on the effects of foundation pit excavation on adjacent tunnels in soft soil," *Mathematical Problems in Engineering*, vol. 2021, 11 pages, 2021.

[12] S. Y. Fan, Z. P. Song, and T. Xu, "Tunnel deformation and stress response under the bilateral foundation pit construction: a case study," *Engineering*, vol. 21, no. 3, p. 77, 2021.

[13] J. S. Ding, Y. Q. Xian, and T. J. Liu, "Study on the influence of the shield tunnel deformation due to excavation of foundation pit with field data," *Advanced Materials Research*, vol. 16, no. 3, pp. 446–449, 2012.

[14] Y. Gui, Z. Zhao, and X. Qin, "Study on deformation law of deep foundation pit with the top-down method and its influence on adjacent subway tunnel," *Advances in Civil Engineering*, vol. 2020, no. 8, p. 15, 2020.

[15] X. H. Zhang, G. Wei, and C. W. Jiang, "The study for longitudinal deformation of adjacent shield tunnel due to foundation pit excavation with consideration of the retaining structure deformation," *Deformation*, vol. 12, no. 12, p. 2103, 2020.

[16] Y. Jiang, "Influence on metro tunnel structure caused by foundation pit excavation considering coupling effect," *Fresenius Environmental Bulletin*, vol. 17, no. 4, 2019.

[17] X. M. Zhang, X. F. Ou, J. S. Yang, and J. Y. Fu, "Deformation response of an existing tunnel to upper excavation of foundation pit and associated dewatering," *International Journal of Geomechanics*, vol. 17, no. 4, 2017.

[18] A. M. Hefny, H. C. Chua, and J. Zhao, "Parametric studies on the interaction between existing and new bored tunnels," *Tunnelling and Underground Space Technology*, vol. 19, no. 4-5, p. 471, 2004.

[19] X.-L. Tao, Y.-H. Su, Q.-Y. Zhu, and W.-L. Wang, "Pasternak model-based tunnel segment uplift model of Subway shield

tunnel during construction," *Advances in Civil Engineering*, vol. 2021, no. 5, p. 10, 2021.

[20] R. Liang, T. Xia, M. Huang, and C. Lin, "Simplified analytical method for evaluating the effects of adjacent excavation on shield tunnel considering the shearing effect," *Computers and Geotechnics*, vol. 81, no. 3, pp. 167–187, 2017.

[21] Y. Tan, X. Li, Z. Kang, J. Liu, and Y. Zhu, "Zoned excavation of an oversized pit close to an existing metro line in stiff clay: case study," *Journal of Performance of Constructed Facilities*, vol. 29, no. 6, 2014.

[22] H. Sun, Y. Chen, J. Zhang, and T. Kuang, "Analytical investigation of tunnel deformation caused by circular foundation pit excavation," *Computers and Geotechnics*, vol. 106, no. 5, pp. 193–198, 2019.

[23] K. C. Shu and N. Xiong, "Research on strong agile response task scheduling optimization enhancement with optimal resource usage in green cloud computing," *Future Generation Computer Systems*, vol. 124, pp. 12–20, 2021.

[24] W. Shu, K. Cai, and N. N. Xiong, "A ShortTerm traffic flow prediction model based on an improved gate recurrent unit neural network," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9830151, 1 page
https://doi.org/10.1155/2023/9830151

Retraction

Retracted: Research on Multimodal Image Fusion Target Detection Algorithm Based on Generative Adversarial Network

Wireless Communications and Mobile Computing

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Wu, X. Wu, Y. Zhu et al., "Research on Multimodal Image Fusion Target Detection Algorithm Based on Generative Adversarial Network," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1740909, 10 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1740909, 10 pages
https://doi.org/10.1155/2022/1740909

Research Article

Research on Multimodal Image Fusion Target Detection Algorithm Based on Generative Adversarial Network

Zhaoli Wu,[1,2,3,4] **Xuehan Wu**,[2,3] **Yuancai Zhu**,[2,3,4] **Jingxuan Zhai**,[2,3,4] **Haibo Yang**,[2] **Zhiwei Yang**,[2] **Chao Wang**,[2] **and Jilong Sun**[2]

[1]*China University of Mining and Technology, School of Computer Science and Technology, Xuzhou 221116, China*
[2]*Jiangsu Vocational Institute of Architectural Technology, School of Information and Electronics Engineering, Xuzhou 221116, China*
[3]*Xuzhou Intelligent Machine and Visual Application Technology Engineering Research Center, Xuzhou 221116, China*
[4]*Xuzhou Big Data Analysis and Data Security Engineering Research Center, Xuzhou 221116, China*

Correspondence should be addressed to Zhaoli Wu; lb20170009@cumt.edu.cn

Received 7 December 2021; Revised 21 December 2021; Accepted 27 December 2021; Published 24 January 2022

Academic Editor: Liqin Shi

In this paper, we propose a target detection algorithm based on adversarial discriminative domain adaptation for infrared and visible image fusion using unsupervised learning methods to reduce the differences between multimodal image information. Firstly, this paper improves the fusion model based on generative adversarial network and uses the fusion algorithm based on the dual discriminator generative adversarial network to generate high-quality IR-visible fused images and then blends the IR and visible images into a ternary dataset and combines the triple angular loss function to do migration learning. Finally, the fused images are used as the input images of faster RCNN object detection algorithm for detection, and a new nonmaximum suppression algorithm is used to improve the faster RCNN target detection algorithm, which further improves the target detection accuracy. Experiments prove that the method can achieve mutual complementation of multimodal feature information and make up for the lack of information in single-modal scenes, and the algorithm achieves good detection results for information from both modalities (infrared and visible light).

1. Introduction

With the rapid development of deep learning, the task of target detection in computer vision tasks has made great progress. However, the task of target detection is very difficult to apply in some real-world scenarios. In the military, security, and other fields, traditional visible light images have very obvious limitations. In recent years, scholars have found that the introduction of multimodal data can significantly improve the accuracy of detection algorithms. Multimodality refers to image pairs formed by applying different imaging principles to the same scene. With the successful application of deep convolutional neural networks in target detection tasks, scholars have produced many excellent results in multimodal research. The author uses a convolutional neural network to fuse two modal information and discusses the impact of different fusion stages on the target detection results [1]. The author believes that only fusing two modal information for target detection is imperfect, and it is necessary to retain the unique information of the two modalities [2]. Therefore, the author adds two modules to the network based on the idea of probability, and one module is used. To output the degree of dependence of the current image on the respective features and fusion features of the two modalities, the second module uses the output of the module 1 as the weight, and the respective output results of the two modalities and the output results of the fusion feature are weighted to obtain the discrimination probability. Konig et al. use the faster RCNN target detection algorithm, but use the fused feature layer and the two-modal feature layer in the training process. Literature [3] adjusts the fusion weight of each modal under

different lighting conditions by designing a light perception network to simulate day and night illumination, but the detection accuracy is very dependent on the light perception network.

In response to the above problems, this paper starts from the perspective of the adversarial discriminant domain [4], uses an unsupervised learning method to reduce the modal difference between bimodal images, and proposes a modal information fusion detection algorithm based on a generative adversarial network. In the improved generative confrontation network, the generator is designed with local detail features and global semantic features to extract source image details and semantic information, and perceptual loss is added to the discriminator to keep the data distribution of the fused image consistent with the source image and improve fusion image accuracy. The fused features enter the interest pooling network for rough classification, and the generated candidate frame is mapped to the feature map, and finally, the target classification and positioning are completed through the fully connected layer.

2. Algorithm Structure

In traditional infrared and visible light image fusion methods, a hybrid model is usually established to combine the advantages (salience) of multiple parties. Although the image fusion performance is improved, the fusion rules need to be manually designed. Generative adversarial networks (GAN) have inherent advantages in the field of image generation and can fit and approximate the real data distribution without supervision. The use of generators and discriminators for confrontation makes the fusion image retain richer information, and the end-to-end network structure no longer needs to manually design fusion rules.

2.1. Information Fusion Network Framework. The generative confrontation network was proposed by Goodfellow in 2014 [5] and is widely used in the field of deep learning. Generative adversarial network is a two-person zero-sum game idea, which can effectively estimate the distribution of data characteristics and generate new samples. The generative confrontation network includes a generative model (G) and an identification model (D). The generative model has the ability to fit the distribution of image data, and the discrimination model can estimate the probability that the input sample is real data. The purpose of the generator is to generate sample data. The sample data distribution is P_z. The training process of generating a confrontation network is to make the data distribution P_z of the generated data infinitely close to the real data distribution P_r. The specific formula is as follows.

$$\min_G \max_D V_{\mathrm{GAN}}(D,G) = \mathrm{E}_{x\sim p_r}[\log D(x)] + \mathrm{E}_{z\sim p_z}[\log(1 - D(G(z)))]. \tag{1}$$

It can be seen from the above formula that PG cannot show that if the discriminator is trained too well or too poorly, the generator will not get effective gradient descent, and the two cannot be updated synchronously, which will cause the GAN training to collapse. To solve this problem, the solution is to make the discriminator meet the Lipschitz [6] continuity condition (Lipschitz continuity):

$$\left|f(x_1) - f(x_2)\right| \le K\left|x_1 - x_2\right|. \tag{2}$$

For f, the smallest constant K is called the Lipsch constant of f. Limit the gradient of the discriminator to a certain range, so that the discriminator can gradually update the gradient in a small range.

This paper establishes a dual discriminator GAN for multimodal image fusion, and the overall framework is shown in Figure 1.

The generator in the figure above represents the generator of the fusion image, the input channel is connected with the visible light image and the infrared image, and the fusion image is input to the discriminator in a single-channel manner. The dual discriminators discriminator I and discriminator V are used to distinguish between fusion image and infrared image and fusion image and visible light image, respectively. After the continuous confrontation and iterative update between the generator and the discriminator, the trained generator is obtained. Single channel represents the single channel that contains the source image and the fusion image when the input of each discriminator is input. If the input contains both the fusion image and the corresponding source image as the dual channel of conditional information, the task of the discriminator will be simplified to whether the input image is same. This is too simple for the discriminatory network, and it is impossible to establish an effective confrontational relationship between the generator and the discriminator.

2.1.1. Generator Network Structure. The generator contains a total of six convolution modules, each of which contains a convolution layer and an activation function and uses the same $3 * 3$ convolution kernel. The number of convolution kernels for the first 5 convolution modules is 32. This can ensure that the network structure fully extracts image features. The generator structure diagram is shown in Figure 2.

2.1.2. Discriminator Network Structure. The discriminator contains a total of 6 convolution modules. Each convolution module contains a convolution layer and an activation function. The convolution size is set to 3, and the number of convolution kernels is set to 64, 128, and 256, respectively. Finally, it contains two fully connected layers. The discriminator structure diagram is shown in Figure 3.

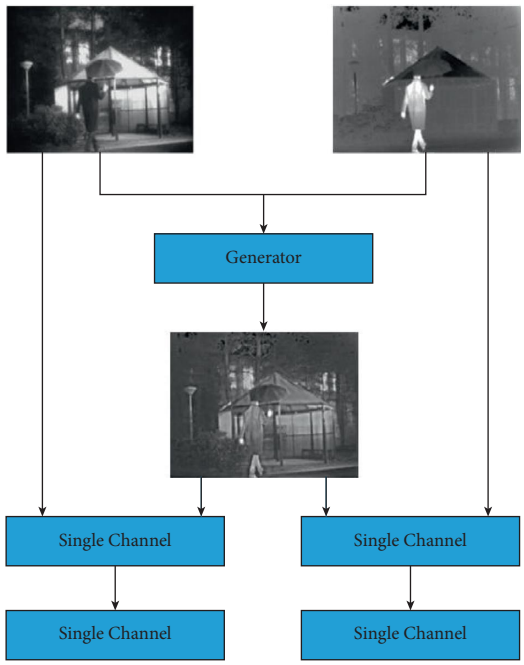


Figure 1: Image fusion based on GAN.

2.1.3. Loss Function

(1) Generator Loss Function. The generator loss function is defined as follows:

$$L = L_{\text{advers}}(G) + \lambda_1 L_{\text{content}}. \tag{3}$$

L is the total loss of the generator, $L_{\text{advers}}(G)$ represents the confrontation loss, L_{content} represents the content loss, and 1 is the coefficient. In order to make the generated image, save the infrared and visible light information as much as possible, the content loss of the generator is defined as L_{content}.

$$L_{\text{content}} = \frac{1}{HW}\left(\mu\left\|I_f - I_r\right\|_F^2 + \gamma\left\|\text{LBP}\left(I_f\right) - \text{LBP}\left(I_v\right)\right\|_F^2\right). \tag{4}$$

Here, μ and γ are the coefficients, H and W are the length and width of the input image, and I_f, I_r, and I_v are the fusion image, infrared image, and visible light image. The first item in the bracket is for the fusion image to save more information from the infrared image, and the second LBP function is defined as shown in formula (5), the purpose is to make the fusion image save more texture information from the visible light image.

$$\text{LBP}\left(x_c, y_c\right) = \sum_{p=0}^{p-1} 2^p s\left(i_p - i_c\right). \tag{5}$$

Here, (x_c, y_c) is the central pixel, and its pixel intensity value is i_c, i_p. $L_{\text{advers}}(G)$ stands for confrontation loss. The confrontation loss consists of two parts: the confrontation loss between the generator and the discriminator 1 and the confrontation loss between the generator and the discriminator 2, and the definition is shown as follows:

$$L_{\text{advers}}(G) = -\sum_{i=1}^{N} \text{E}_{z\sim p_g}\left[D_i(z)\right], \tag{6}$$

where z represents the generated data, p_g represents the distribution of the generated data, and N represents the number of discriminators (N takes 2).

(2) Discriminator Loss Function. Although the fusion image generated by the generator can save infrared and visible light information to a certain extent, it still needs to use the generated image and the source image to save more detailed information through the discriminator. The discriminator loss function is shown as follows:

$$L_{D_{ir}} = -\text{E}_{x\sim p_{ir}}\left[D_{ir}(x)\right] + \text{E}_{z\sim p_g}\left[D_{\text{vis}}(z)\right] + \lambda_3 \text{E}_{\tilde{x}}\left[\left(\left\|\nabla_{\tilde{x}} D_{ir}(\tilde{x})\right\|_2 - 1\right)\right]. \tag{7}$$

Here, $L_{D_{ir}}$ represents the loss of the visible light image and the generated fusion image as the input of the discriminator, p_{vis} and p_g represent the visible light image distribution and the distribution of the generated image, and λ_3 is the hyperparameter.

2.2. Improved Target Detection Algorithm. The target detection task based on the deep convolutional neural network has made great progress with the rapid development of deep learning, and the detection accuracy has been significantly improved compared with traditional detection methods. Many scholars have designed many detection networks. In general, the detection network is roughly divided into two-stage target detection and single-stage target detection. The two-stage target detection network has a candidate frame extraction step. Compared with the single-stage, the accuracy is higher, but the network prediction speed is slower. From R-CNN [7] to faster R-CNN [8], network detection accuracy is getting higher and higher, and the detection speed is getting faster and faster. Faster R-CNN is a classic structure in a two-stage target detection network. The network structure diagram is shown in Figure 4.

The faster R-CNN target detection algorithm is to define convolution feature extraction, candidate frame selection, candidate frame classification, and bounding box regression in a network, which can be regarded as faster R-CNN is the RPN (region proposal network) network and fast R—the combination of CNN network and the convolutional layer of RPN [9] is shared with fast R-CNN. The specific method is shown in Figure 5.

Many scholars have successively proposed improvement strategies for deep convolutional neural networks, some articles have improved the loss function, and some have proposed new improvement ideas such as deformable convolution and expanded convolution. This article focuses on the improvement ideas of the nonmaximum supression (NMS) [10] algorithm. The function of the NMS algorithm is to remove redundant detection results, and only keep a bounding box as the output of the detection result, which has very important significance for target detection. The original detection results of the network often produce multiple bounding boxes near the same target. At this time, it is necessary to sort according to the probability value

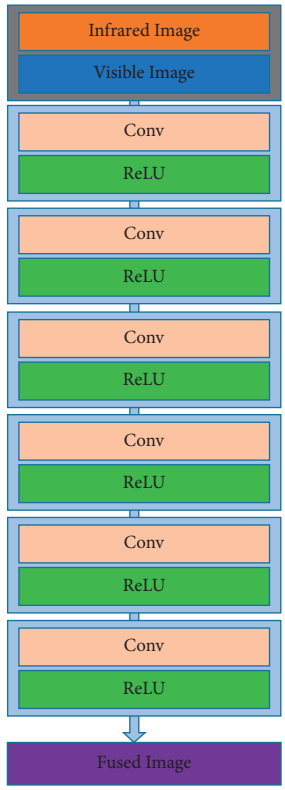


Figure 2: Generator network structure.

of the bounding box classification and select the bounding box with the highest score as the final detection result of the target at that location. If the remaining bounding boxes are such that if the IoU value of the selected bounding box is greater than the set threshold, it will be eliminated directly. The NMS algorithm is shown as follows:

$$S_i = \begin{cases} s_i, & \text{IoU} < N_t, \\ 0, & \text{IoU} \geq N_t. \end{cases} \tag{8}$$

The disadvantage of the NMS algorithm is that if the two targets on the image are relatively close, the IoU value of the bounding box between the target and the target is very large, which is very easy to cause a target to be undetected. Therefore, in view of the above shortcomings, this paper adopts the soft-NMS algorithm. The purpose is to select a bounding box with the highest current score and then update the score according to the IoU [11] value between the surrounding bounding box and the boundary with the highest score. A bounding box with a larger IoU value has a lower update score; a bounding box with a not too large IoU value will not have a too low score after the update, so that problems caused by NMS can be avoided to a certain extent. Soft-NMS is shown as follows:

$$S_i = \begin{cases} S_i, & \text{IoU} < N_t, \\ S_i e^{(\text{IoU}^2/\delta)}, & \text{IoU} < N_t. \end{cases} \tag{9}$$

2.3. Multimodal Information Fusion Detection. The algorithm in this paper regards the entire image fusion process as a process of confrontation between the generator and the discriminator. The training of the network model is not exactly the same as the test. Only the trained generator is needed during the test, and no discriminator is required to participate. In the confrontation generation network of the double discriminator, the first discriminator is mainly used to discriminate infrared images and generate images, and the other discriminator discriminates visible light images and generates images. The purpose is to enable the generated images to save infrared image temperature information and visible light gradient information, to avoid problems such as insufficient storage of single discriminator information and rely on the confrontation generation network to map visible light image information and infrared image information to the same feature space. At this time, the target detection task is similar to the visible light target detection. The feature extraction network and the classification network are completed. The detection framework is shown in Figure 6.

3. Experimental Results and Analysis

The experimental environment is configured as Ubuntu16.04 operating system, Pytorch deep learning framework; the hardware environment is NVIDIA GTX 1080ti graphics card ×2, Intel Core i7 processor. The experimental part uses FLIR [12] infrared data set for algorithm verification. The data set has two domains: infrared domain and visible light domain. The infrared image contains 7153 images, and the visible light image contains 6936 images. The detection categories are divided into three categories: people, cars, and bicycles.

3.1. Fusion Experiment. The fusion method in this paper is analyzed and compared with other fusion algorithm methods.

3.1.1. Qualitative Evaluation. The experimental results show that the fusion algorithm used in this paper has richer background detail information, such as the sky information in the two images, which obviously saves more texture information. In addition, compared with the single discriminator FusionGAN [13] algorithm, it can obviously get a better fusion effect, and it can reflect the prominent target and detailed features of the source image better, which is helpful for the next target detection. The fusion effect of different algorithms is shown in Figure 7.

3.1.2. Quantitative Evaluation. It mainly uses quantitative evaluation index fusion methods such as information entropy (EN), standard deviation (SD), mutual information (MI), and peak signal-to-noise ratio (PSNR) [14].

It can be seen from Figure 8 that the fusion algorithm in this paper has achieved obvious advantages in the three indicators of MI, EN, and SD, especially in the SD indicator, which shows great superiority. This can reflect to a certain extent that the fusion algorithm in this paper not

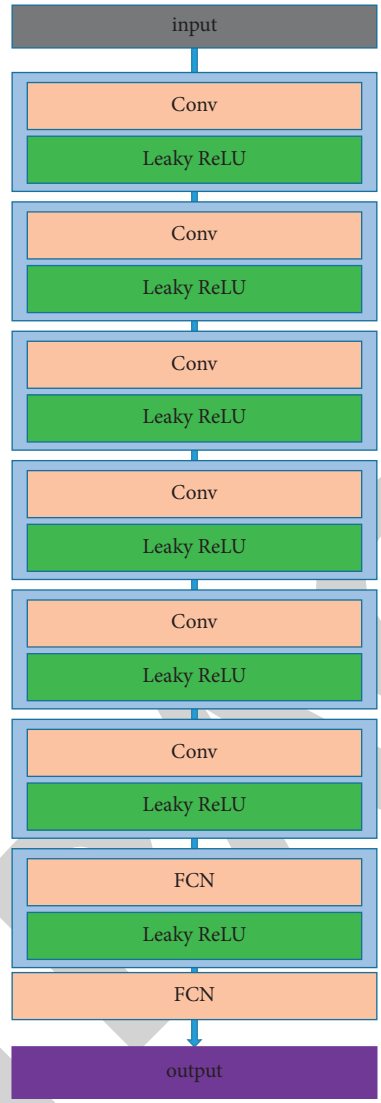


Figure 3: Discriminator network structure.

only has better visual effects but also has obvious advantages in quantitative evaluation.

3.2. Target Detection Model

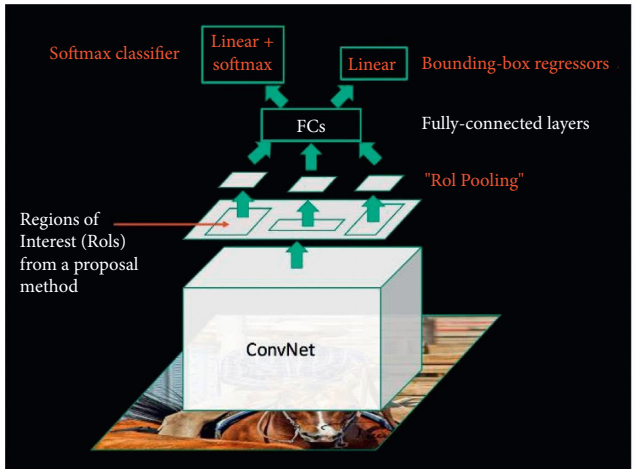
3.2.1. Target Detection Model Training. First, use the abovementioned trained generator to fuse visible light and infrared images to obtain a fusion image containing multimodal information and then use the image data set to train the fusion image target detection model.

After 50,000 iterations of training, the training loss of the visible light target detection model is about 0.5, and its loss transformation curve is shown in Figure 9.

The change curve of the intersection ratio between the predicted bounding box and the actual bounding box is shown in Figure 10. The abscissa represents the number of iterations, and the ordinate represents the intersection ratio of the predicted bounding box and the actual bounding box. As the number of iterations increases, the intersection ratio becomes the overall. The upward trend is finally close to 1, which means that the predicted bounding box in the visible light scene is very close to the actual bounding box, which meets the training requirements.

3.2.2. Target Detection Experiment. The target detection model generally uses mAP (mean average precision) [15, 16] index for evaluation, which is the average value of the average detection accuracy (average precision, AP) of multiple types of objects. The test sets under the visible light and infrared scenes were tested, respectively, and the results are shown in Tables 1–3.

It can be seen from the above table that the method in this paper has a high accuracy rate in the overall structure, can effectively fuse the bimodal information, and realize the accurate description of the scene information. The model detection effect is shown in Figure 11.



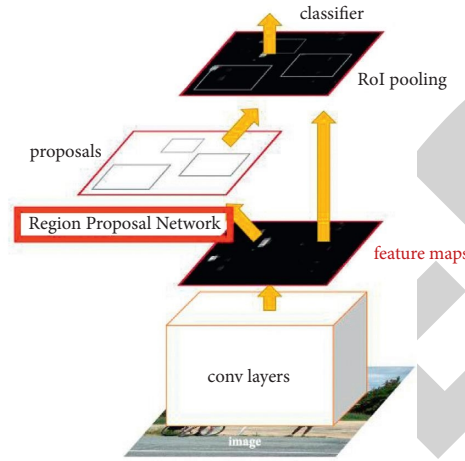


Figure 4: Faster RCNN network structure.

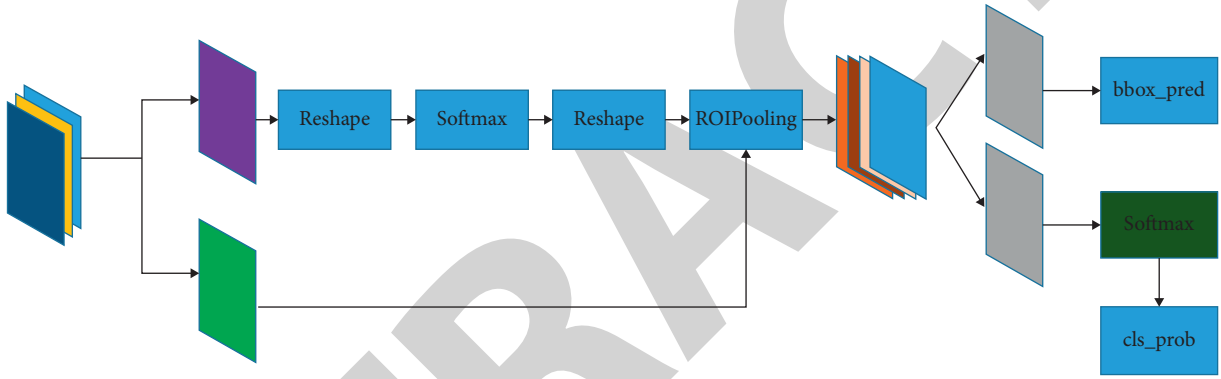


Figure 5: RPN network structure.

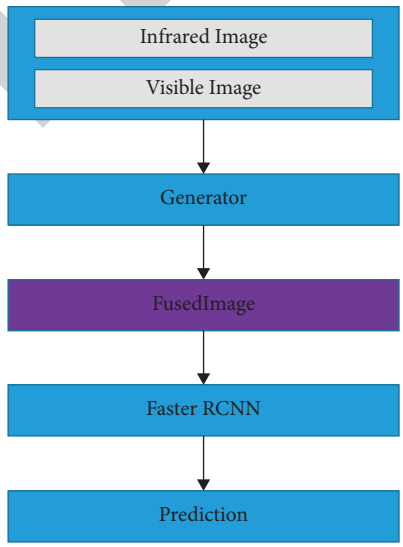


Figure 6: Multimodal information fusion detection algorithm.

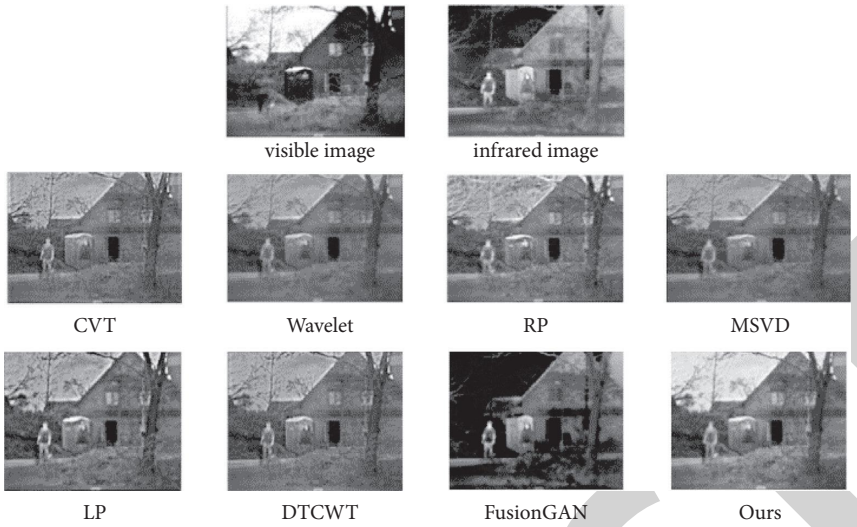


Figure 7: Comparison of the effect of different fusion algorithms.

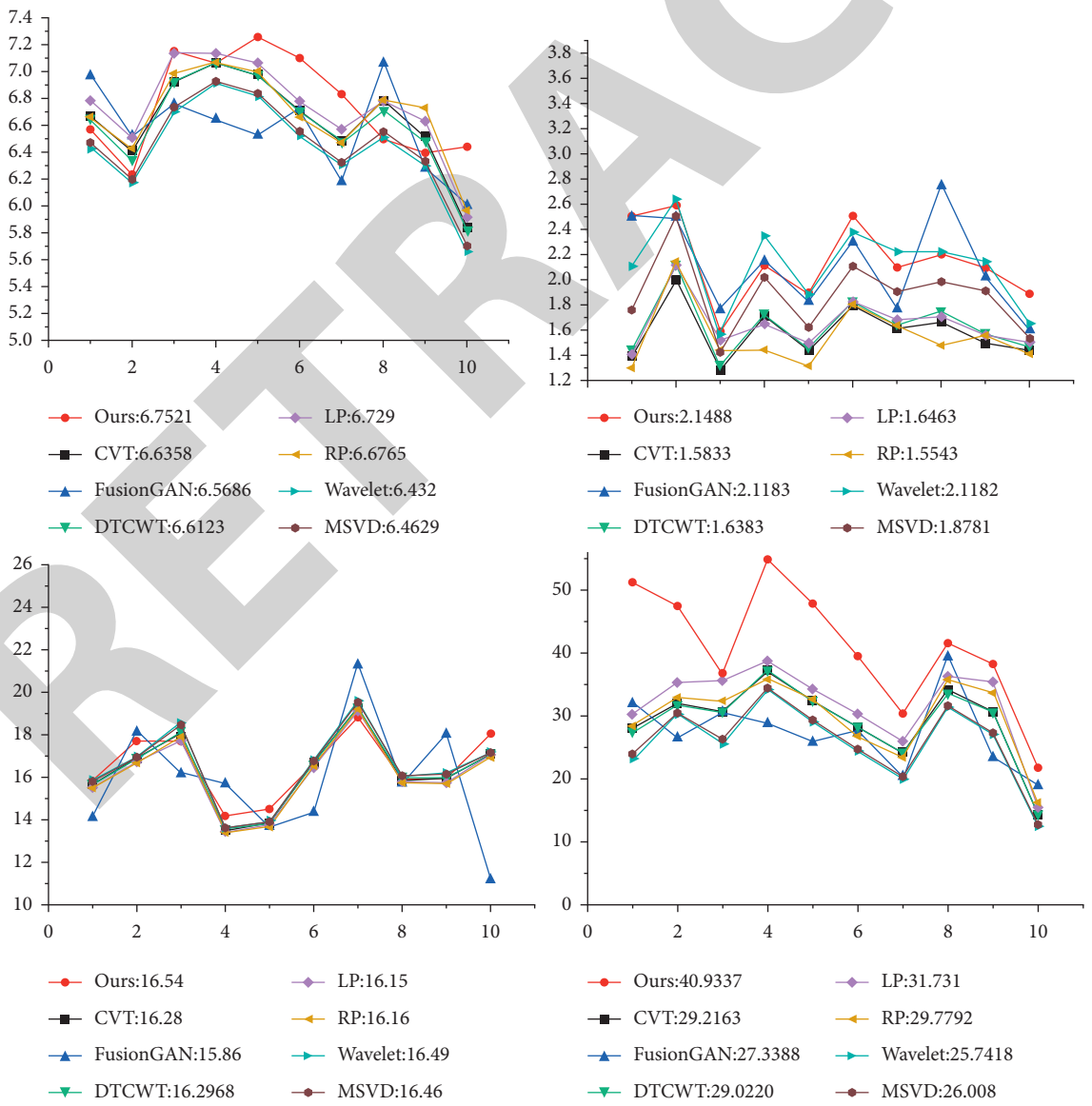


Figure 8: Quantitative evaluation.

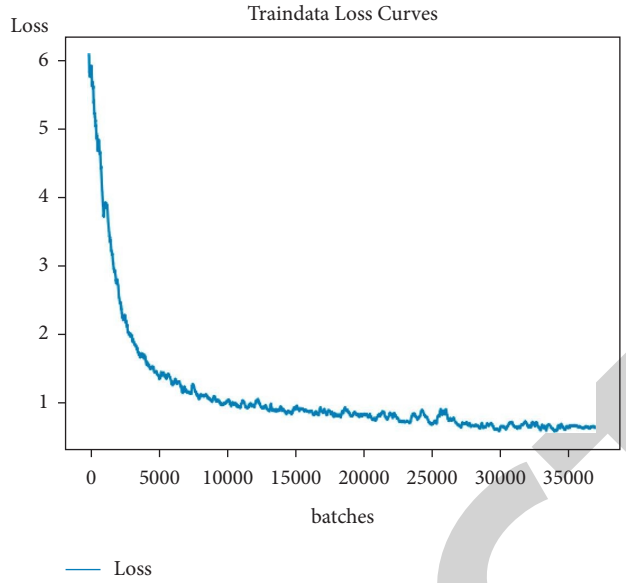


Figure 9: Training loss curve of detection model.

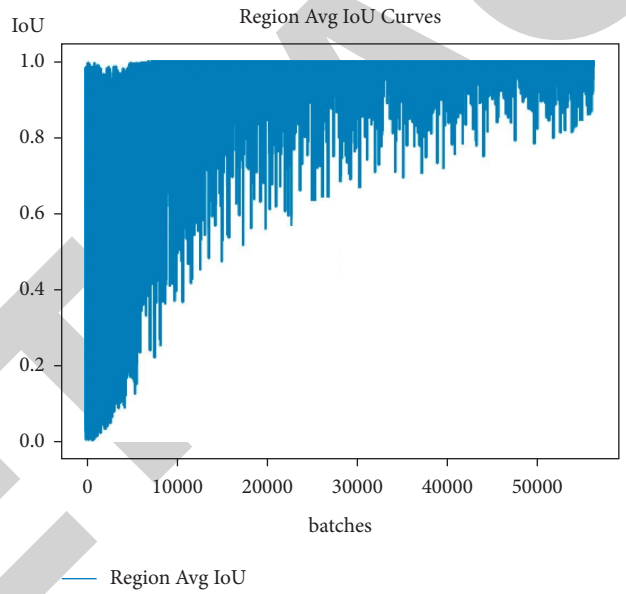


Figure 10: Curve of change of intersection ratio of detection model.

Table 1: Visible.

Model	AP (%)
Faster RCNN (+NMS)	86.28
Faster RCNN (+soft-NMS)	89.35
Ours	93.82

Table 2: Infrared.

Model	AP (%)
Faster RCNN (+NMS)	84.39
Faster RCNN (+soft-NMS)	87.16
Ours	93.36

Table 3: Fused.

Model	AP (%)
Faster RCNN (+NMS)	74.39
Faster RCNN (+soft-NMS)	79.16
Ours	95.36

Figure 11: Fusion model detection effect diagram.

4. Conclusion

Multimodal information fusion has a wide range of application scenarios. This paper designs a confrontation generation network that can realize end-to-end training to fuse multimodal information to improve the complementarity and low redundancy among multimodal information features and improve the accuracy of target detection and classification based on fusion features. Multimodal information fusion provides richer target information than single-modality, but it also greatly increases the amount of calculation, which makes it difficult to achieve real-time detection effects in application scenarios with limited computing resources.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the Development and Application of Identification Control System in Epidemic Prevention and Control Area (Grant nos. JYJFZX20-01) and National Educational Technology Research Project of Central Audio Visual Education Center (Grant no. 186130061). Xuzhou city will promote the special Key Research and Development Plan for Scientific and Technological Innovation (industrial key technology research and development) project "R&D and Application of Water Resources Cloud Control Platform at River Basin Level (Grant no. KC21108), special policy guidance plan for scientific and technological innovation (industry-university-research cooperation) Big Data-Based Multi-Objective Coordinated and Balanced Allocation of Large-Region Water Resources (Grant no. KC21335), school level mixed teaching team of computer network technology specialty group by the Academic Affairs Office of Jiangsu Construction Vocational and Technical College (Grant no. jw2021-8), and the research and practice of demonstration vocational education group—Taking Xuzhou Huaihai Service Outsourcing Vocational Education Group as an example (Grant no. ES2021-2).

References

[1] Z. Wang, "A new approach for segmentation and quantification of cells or nanoparticles," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 962–971, 2016.

[2] Z. Wang, J. Xiong, Y. Yang, and H. Li, "A flexible and robust threshold selection method," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 9, pp. 2220–2232, Sept. 2018.

[3] T. Mantecón, C. R. del Blanco, F. Jaureguizar, and N. García, "Hand gesture recognition using infrared imagery provided by leap motion controller," in *Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems, ACIVS 2016*, pp. 47–57, Lecce, Italy, October 2016.

[4] E. Persoon and K.-S. Fu, "Shape discrimination using Fourier descriptors," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 3, pp. 170–179, 1977.

[5] C. Li, H. Cheng, S. Hu, X. Liu, J. Tang, and L. Lin, "Learning collaborative sparse representation for grayscale-thermal

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9898456, 1 page
https://doi.org/10.1155/2023/9898456

Retraction

Retracted: A Lightweight Face Verification Based on Adaptive Cascade Network and Triplet Loss Function

Wireless Communications and Mobile Computing

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Lin, C. Ye, W. Liu et al., "A Lightweight Face Verification Based on Adaptive Cascade Network and Triplet Loss Function," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3017149, 10 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3017149, 10 pages
https://doi.org/10.1155/2022/3017149

Research Article

A Lightweight Face Verification Based on Adaptive Cascade Network and Triplet Loss Function

Jianhong Lin,[1,2] Chaoyang Ye,[3] Weinan Liu,[4] Siqi Ren,[5] Ye Wang,[5] Wenrui Ma,[5] Bin Xu,[5] and Yifan Ding[2]

[1]*College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China*
[2]*Zhejiang Ponshine Information Technology Co., Ltd., Hangzhou 311100, China*
[3]*National (Hangzhou) New-Type Internet Exchange, Hangzhou 310009, China*
[4]*Business & Tourism Institute, Hangzhou Vocational & Technical College, Hangzhou 310018, China*
[5]*School of Computer and Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, China*

Correspondence should be addressed to Siqi Ren; rensiqi@zjgsu.edu.cn

Received 8 December 2021; Revised 24 December 2021; Accepted 28 December 2021; Published 20 January 2022

Academic Editor: Liqin Shi

In the past few years, with the continuous breakthrough of technology in various fields, artificial intelligence has been considered as a revolutionary technology. One of the most important and useful applications of artificial intelligence is face detection. The outbreak of COVID-19 has promoted the development of the noncontact identity authentication system. Face detection is also one of the key techniques in this kind of authentication system. However, the current real-time face detection is computationally expensive which hinders the application of face recognition. To address this issue, we propose a face verification framework based on adaptive cascade network and triplet loss. The framework is simple in network architecture and has light-weighted parameters. The training network is made of three stages with an adaptive cascade network and utilizes a novel image pyramid based on scales with different sizes. We train the face verification model and complete the verification within 0.15 second for processing one image which shows the computation efficiency of our proposed framework. In addition, the experimental results also show the competitive accuracy of our proposed framework which is around 98.6%. Using dynamic semihard triplet strategy for training, our network achieves a classification accuracy of 99.2% on the dataset of Labeled Faces in the Wild.

1. Introduction

Artificial intelligence is one of the hottest topics in computer science which studies theories and methods used to make machine as intelligent as human beings. With the continuous breakthrough of technology in various fields, artificial intelligence has been considered as a new revolutionary technology to make progress in scientific, technological, and industrial revolution. Driven by a large amount of data, better artificial intelligence algorithms, and more powerful computing equipment, a variety of artificial intelligence applications have been used in both industries and people's daily life. Artificial intelligence industry refers to an industry that provides intelligent products and technical services to the society based on artificial intelligence technology. It has derived various new intelligent applications through enabling manufacturing, agriculture, medical, and other industries, such as subtopic detection [1], intelligent agriculture [2], and smart grid [3]. These applications have become a new engine to promote high-quality economic and social development.

With the continuous development of mobile phone applications, digital image processing and recognition have attracted more and more attention. Before the development of artificial intelligence technology, image recognition was mainly based on statistical decision-making and template matching. These traditional image recognition methods have their own limitations, and in the process of image recognition, it is necessary to manually preprocess the image and extract relevant features. If the image to be recognized has

large deformity or strong noise interference, the traditional recognition methods cannot get the expected results, resulting in poor accuracy of image recognition. With the development of artificial intelligence technology, various deep learning methods have raised image recognition technology to a new level, which has greatly improved both accuracy and real-time efficiency. The deep learning method represented by convolutional neural network, relying on its self-learning ability and computer computing ability, has achieved very good image recognition results in various application fields.

Face recognition is an important research field of digital image processing and recognition. Because of its wide application in business and security fields, face recognition becomes more and more important. For example, in the context of national fitness, people have more opportunities to enter the Asian Games venues for exercise. If efficient face recognition machines are configured in these venues, users can quickly get in and out of the sports venues. Face is the biometric information of users, which can be bound with the account to better understand the data of users' subsequent fitness, so as to further guide users to better carry out national fitness. Another useful scenario for face recognition is mobile payment. Through the face payment functionality of Alipay (the payment application of Alibaba group), users can complete payment conveniently and quickly, all of which benefit from efficient and safe face recognition algorithm. Another technique popularizes the application of face recognition is Internet of Things (IoT). IoT is one of the important application areas of 5G and future wireless communication systems. Backscatter communication can realize the low power consumption information transmission of IoT. Face recognition is an important application of IoT. The lightweight face recognition framework can be widely used in low-power devices to enrich the application scenarios of IoT. The lightweight face recognition framework combined with backscatter communication can in turn better popularize IoT. Face recognition detection and verification usually go through within two stages. One is to detect face, including face detection and face alignment, which has important practical value and significance, and has made a lot of research results [4]. The second stage is face classification. There are still many problems and challenges in this stage of research. For example, (1) face has strong variability. In different environments, the skin color of face will change due to the influence of environment. The first challenge requires face detection approach to be applicable in different scenarios; (2) the variability of face position is due to the fact that face can exist in any position in picture space or appear in a picture of any size. The second challenge requires face detection approach to check out as many faces as possible in real application. With the continuous progress of deep learning, the research heat of face recognition algorithm is rising again. Compressed convolutional neural network can complete real-time high-quality face detection on mobile platform [5]. Cascade convolutional neural network (CNN), which belongs to deep convolutional neural network (DCNN), can detect face more quickly by relying on lightweight module [6].

The main contributions of this paper can be described as following:

(1) A triplet loss function and a neural network are presented, and base on them, a lightweight face detection approach is constructed. An adaptive scale selection mechanism in first stage of the proposed face detection approach is proposed to avoid prohibitive computation which makes the approach efficient

(2) Our proposed approach achieves competitive accuracy on the dataset of Labeled Faces in the Wild (LFW) while keeps real time performance

(3) Our solution has the advantage of being lightweight and can be widely used in IoT scenarios. By incorporating the encrypted face authentication information to improve the identity authentication protocols and achieve the goal of personalized privacy protection, our approach can be applied into security field

The organization of this paper is as follows. Section 2 presents the related work. Section 3 introduces the building blocks including relevant parameters and cascade CNN. Section 4 introduces the network structure, including model training, training definition, and triple and training method selection. Section 5 shows the experimental results, including the experimental results of self-built database and LFW training set. Section 6 summarizes the paper.

2. Related Work

Before the blooming of deep learning, the performance of traditional face recognition task is advanced by handcrafted features or adjusted parameter, such as the famous local binary pattern (LBP) [7] and SIFT [8]. Ahonen et al. presented LBP texture feature-based face representation approach, in which the face features are extracted according to the LBP feature distributions and then concatenated into one single vector. Lowe paper proposed an approach for image feature extraction. The approach transforms image data into scale-invariant coordinates which can be used to extract local features. Therefore, this approach is be distinguished by Scale Invariant Feature Transform (SIFT). However, as these types of traditional methods usually take advantages of shallow network, the accuracy is relatively low. For example, the LBP can only obtain 95.17% in terms of accuracy on LFW.

With the development of CNN [9] and ImageNet, the research on face detection is on the rise again. Currently, face detection algorithms usually are based on cascade structure. For example, Mathias et al. addressed the face detection issue and presented an approach which takes advantage of deformable part model (DPM) and enjoy good performance [10]. However, as the approach requires annotations on the training data set, it suffers from large overhead of computation. Sun et al. proposed a new face detection approach which combines faster region-based CNN (RCNN) framework and a variety number of strategies, such as feature

concatenation [11]. The approach achieves remarkable performance on Face Detection Data Set (FDDB) benchmark and becomes one of the state of the art approaches in the aspect of receiver operating characteristic curves. Shi et al. pointed out that with a progressive calibration network (PCN), it is easily to distinguish face frames from nonface frames, and based on this novel PCN, they presented a rotation-invariant based face detection approach [12]. Liu et al. proposed an object detection method which discretizes the output space of bounding boxes in the image and rearranges them into a set of default boxes. The approach requires only a single deep neural network (DNN) which is faster and more accurate than the famous You Only Look Once approach [13].

In the past few years, the face detection approaches have improved. The previous methods are like the one in [14] while the new approaches take DCNN into consideration. The cascade CNN is among the most used and researched neural network in the area of face detection. To address both the effective and accurate issues in real-world face detection applications which usually have large visual variations and require discriminative detection, Li et al. proposed a cascade CNN-based face detection approach which achieves remarkable detection capability and also enjoy good performance [15]. The approach can detect the background regions at the first fast stage with low resolution and rejects these detected parts. Then, in the second stage, the approach checks high resolution part in the image to select the possible candidates for face detection. Dong and Wu focused on the Gaussian distribution and presented a face alignment approach which is based on Adaptive Cascade Deep Convolutional Neural Networks (ACDCNN) [16]. According to the Gaussian distribution among the image blocks, the approach can dynamically select the most relevant training blocks, taking advantage of an adaptively cascade CNN structure, with which, the approach enjoys high performance in accuracy, low complexity in model structure, and high robustness. To address the task handling with extensive facial landmark localization, traditional convolutional network becomes insufficient. Therefore, Zhou et al. proposed a novel approach with four-level cascade CNN [17]. Each level in the cascade CNN can predict position and rotation angles of specific image blocks and generate a coarse-to-fine detection way. Besides, this approach has the ability to process video streams immediately. To estimate the apparent age, Chen et al. proposed an approach combining a coarse-to-fine strategy and an error correction module [18]. The approach is also based on DCNN, and the used DCNN has the ability to classify the age of a detected face and can obtain a fine-grained age which further will be corrected with the error check module. The approach is relatively complex, but the performance is very good. The classic CNN-based face detection method simply stacks different types of filter layers where shallower filters can effectively check out simple non-ace samples, while deeper filters can distinguish face blocks from nonface blocks which are difficult to detect. Zhou et al. proposed a data routing mechanism that allows different layers to pass different types of samples and introduced a dual-stream context CNN architecture, which adaptively uses body part information to enhance face detection [19]. Based on them, the authors proposed an inside cascaded structure-based face detection approach where there are different classifier layers in the same CNN. The approach achieves good results in the challenging FDDB and WIDER FACE benchmark tests. Aiming at simultaneously handling four types of task, i.e., face detection, landmarks localization, pose estimation and gender recognition, Ranjan et al. presented a DCNN-based approach, i.e., HyperFace [20]. In addition, two variants based on HyperFace were proposed. The former is HyperFace-ResNet which uses the idea of residual network [21] and enjoys high performance. The latter is called Fast-HyperFace with the import of a high speed face detector. Both the two methods achieve competitive scores in the four tasks.

Usually, face detection approach needs to operate a large number of images and requires high computation devices, such as GPU cluster [22]. Guo et al. proposed an elaborately designed CNN-based face detection approach, which operates on the complete feature maps and is fast in the detection speed [23]. The authors conducted some experiments which illustrates that the approach works well on popular datasets. To better detect faces in images with nonface inputs and low-quality faces, Yu et al. proposed a novel face detection approach based on uncertainty prediction and the L2-norm of features which can reliably detect face elements from out of distribution samples and enjoys good performance [24]. The detection of small face based on DCNN usually suffers from low performance, and Ke et al. proposed a regional cascade multiscale detection approach to solve this issue [25]. The approach is made of one global face detector and some local face detectors. The product generated by the former detector on the original training set will be delivered into the latter local detectors; with this mechanism, the approach enjoys high performance. As cascade face detectors fail to achieve high accuracy and the performance of anchor-based face detectors highly depends on pretrained dataset, Yu and Tao proposed a face detection framework with efficient anchor cascade [26]. The framework takes advantage of contextual information and enjoys both efficiency and accuracy on face detection task. The experimental result shows it work better than the popular MTCNN [27].

In recent face verification algorithms, Hermans et al. compared the effects of triplet with its variant on the results [28]. Florian et al. presented a face detection system which is based on a compact Euclidean space to map face information and compute the face similarity [29]. The system which is called FaceNet utilizes triplets of face patches based on the method in [30] and achieves state-of-the-art face detection performance on the LFW dataset. Deng et al. proposed an Additive Angular Margin Loss (ArcFace) to obtain highly discriminative features, with geometric interpretation, for face recognition [31]. Lu et al. proposed the Deep Coupled ResNet (DCR) model, the backbone network was used to extract robust features that are resolution invariant, and the coupled mapping (CM) loss function was proposed to optimize the model parameters of the two branches,

respectively, on high- and low-resolution pictures [32]. Xi et al. proposed an alternating training regimen to achieve less biased classifiers and more discriminative feature representation [33]. Yu et al. used the binarization image denoising method to vanish complexity of locating feature pts, which can accurately extract facial features [34].

The application of face verification also attracts the attention of researchers. Lightweight face verification approaches be widely used in IoT scenarios, such as intelligent transportation [35] [36], especially in traffic flow prediction [37], Android applications [38], and AI-supported IoT systems [39]. To improve the identity authentication protocols and achieve the goal of personalized privacy protection, face verification approaches can be applied into security field, such as in browsers [40], social platforms [41], and cloud computing [42] by incorporating the encrypted face authentication information.

Liu et al. proposed a novel approach which takes advantage of a modified cascade CNN [43]. The proposed approach is made of three stages when training face dataset. Aiming at achieving fast face detection and higher accuracy, we introduce a triplet loss function in Section 3.1 and novel network architecture in Section 4 which constructs a new face detection approach. Using dynamic semihard triplet strategy for training, our network achieves a classification accuracy of 99.2% on the LFW dataset.

3. Building Blocks

In this section, the building blocks are presented, which consists of two parts. The former is related parameters including intersection over union, nonmaximum suppression, classifier, loss function, and triplet loss. The latter presents a three stages cascade CNN and an adaptive scale selection mechanism.

3.1. Relevant Parameters

3.1.1. Intersection over Union. Intersection over Union (IoU) is a concept used in target detection which calculates the overlap rate between the "predicted border" and "true border," i.e., the ratio of their intersection to union. Equation (1) shows the definition of IoU.

$$\mathrm{IoU} = (A \cap B)/(A \cup B). \tag{1}$$

3.1.2. Nonmaximum Suppression. The essence of nonmaximum suppression (NMS) is to search for local maximums and suppress nonmaximum elements, and IoU is used to compute NMS. When performing face detection, a window sliding method is generally adopted to generate a lot of candidate frames on the face image, and then these candidate frames are feature extracted and sent to the classifier, usually a cascade CNN. Generally, a score will be calculated on each detected face block or box. As there will be scores on many boxes, all these scores will be sorted, and one box with the highest score will be selected according to the degree of overlap, i.e., IoU, and between other frames and the current frame. In addition, the target box will not be selected if the degree of IoU is greater than a certain threshold. And except for the boxes exceeding the threshold, the high-scoring frame is selected as the detected face.

3.1.3. Classifier and Loss Function. Classifier is a function or model which conducts some mapping operations and put one item into one category. Classifier which can be applied to data prediction application is a general term to define method with classifying functionalities, such as decision trees, logistic regression, and neural network. Loss function is used to evaluate the difference between the predicted value of the classifier and the true value.

In this paper, we define two loss functions. The first one is used by the classifier, i.e., our CNN network, while the second one is used to detect face frames.

(a) The first loss function uses confidence map and bounding regression map to conduct the training job, and crossentropy is used as the loss function which is defined as Eq. (2)

$$H(y) = -\sum_{i=1}^{n} y_i' * \log(y_i), \tag{2}$$

where y_i is the predicted label which is calculated by the neural network, y_i' represents the true value of one image which is labeled in the dataset, and i is the number of elements in the dataset.

The reverse derivation of $H(y)$ is used to obtain the partial differentiation of the weights of different neural network layers.

(b) The second loss function is to address the regression issue in the task of frame detection, and we use Euclidean loss function for border regression which calculates the distance between the predicted value $y_n{}^{\wedge}$ and the label value y_n. The Euclidean loss function, $y_n{}^{\wedge}$, and y_n are defined in Eqs. (3)–(5), respectively

$$L = \frac{1}{2n}\sum_{n=1}^{N}\left\|\hat{y}_n - y_n\right\|_2^2, \tag{3}$$

$$\hat{y} = \left(x_1^{\mathrm{det}}, y_1^{\mathrm{det}}, w^{\mathrm{det}}, h^{\mathrm{det}}\right), \tag{4}$$

$$y = \left(x_1^{gt}, y_1^{gt}, w^{gt}, h^{gt}\right). \tag{5}$$

The elements in the tuple of $(x_1^{\mathrm{det}}, y_1^{\mathrm{det}}, w^{\mathrm{det}}, h^{\mathrm{det}})$ in Eq. (4) represent the x and y coordinates and height and width of the predicted face detection box while the elements in the tuple of $(x_1^{gt}, y_1^{gt}, w^{gt}, h^{gt})$ in Eq. (5) represent the x and y coordinates and height and width of the correct box in the face image.

3.1.4. Triplet Loss. In our proposed approach, the output of the cascade neural network will be the input of the triplet

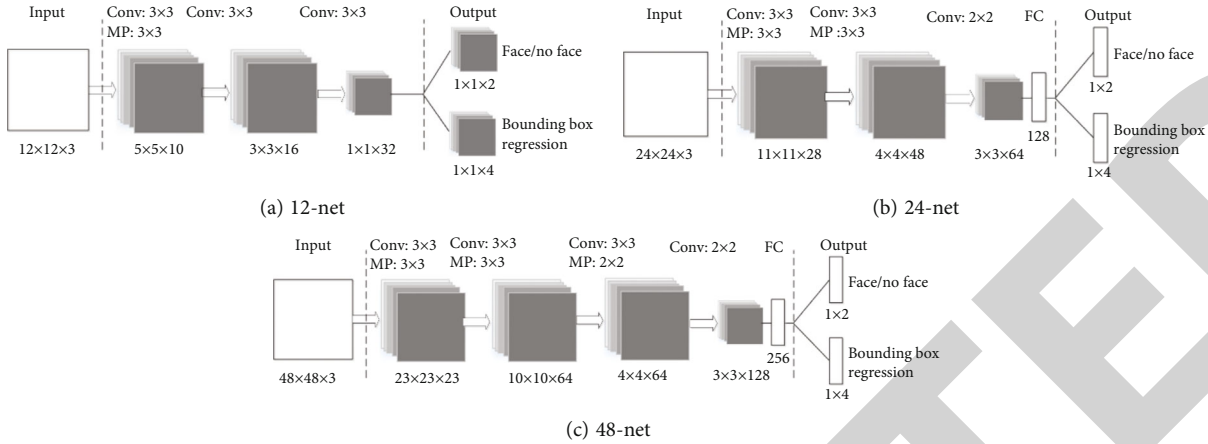


Figure 1: Neural networks.

loss function, which is an embedding mapping function represented as $f(x) \in R^d$. The triplet loss maps an image x into a d-dimensional Euclidean space. In addition, L2-normalization is used to make sure its coordinates locate on a unit hypersphere. Furthermore, to ensure that an image x_i^a of a specific person is more similar to his other image x_i^p than that of any image x_i^n of other people, the following loss function is defined as Eq. (6)

$$\text{Loss} = \sum_i^N \left[\left\| f(x_i^a) - f(x_i^p) \right\|_2^2 - \left\| f(x_i^a) - f(x_i^n) \right\|_2^2 + \alpha \right], \quad (6)$$

where $\|f(x_i^a) - f(x_i^p)\|_2^2$ is the distance between image x_i^a and image x_i^p in a d-dimensional Euclidean space, denoted by $d(a, p)$, and similarly, $\|f(x_i^a) - f(x_i^n)\|_2^2$ is the distance between image x_i^a and image x_i^n, denoted by $d(a, n)$. Additionally, the superscript a means anchor, p means positive, and n means negative.

3.2. Cascade CNN. The cascade CNN consists of two components, i.e., neural network and adaptive scale selection mechanism. Three types of neural networks are used in our proposed approach. In addition, a selection mechanism is used to decide type of neural networks to apply.

3.2.1. Neural Networks. Figure 1 shows the three types of neural network. Figure 1(a) shows structure of 12-net, Figure 1(b) shows structure of 24-net, and Figure 1(c) shows structure of 48-net. Each neural network includes one input with three parameters, i.e., width, height and channel, hidden layers, and one output. The size of network is determined by the input image size. The hidden layers are generated through two types filters, i.e., convolutional filter (Conv) and max-pooling filter (MP), each of which contains different sizes. Note that FC is full connection layer.

The first stage of the cascade CNN is a 12-net. The output is a feature map with size $1 \times 1 \times 32$ in 12-net which will further be calculated into two tensors. One is a confidence map with size $1 \times 1 \times 2$ which shows whether there exists a face or not in the input image. And the other is whether there exists a face or not in the input image. And the other is the bounding box regression with size $1 \times 1 \times 4$ which shows how the window should be adjusted in size and orientation to get a candidate frame if the input image contains a face. An adaptive scale selection mechanism is used in this stage to obtain all the candidate frames, which will be further input into 24-net to get more specified classification results and more accurate bounding boxes.

The second stage of the cascade CNN is a 24-net. The output is a feature map with size $3 \times 3 \times 64$ in 24-net which will further be calculated into two arrays. One is a confidence map with size 1×2 which shows whether there exists a face or not in the input image. And the other is whether there exists a face or not in the input image. And the other is the bounding box regression with size 1×4 which will be used to restrain the margin of a bounding box for the generated candidate frames. In this stage, if one candidate frame has a confidence score greater than 0.9 and NMS less than 0.7 with other candidate frames, then it will be kept in the candidate frame list which is defined as L_s and will be used in the finally stage.

The third stage of the cascade CNN is a 48-net. The output is a feature map with size $3 \times 3 \times 128$ in 48-net which will further be calculated into two arrays. One is a 1×2 confidence array and a 1×4D bounding info array. In this stage, if one candidate frame has a confidence score greater than 0.95 and NMS less than 0.7 with other candidate frames, then it will be regarded as the final outputs.

3.2.2. Adaptive Scale Selection Mechanism. To detect all the possible faces from a given image P with the pix size $(H \times W)$, usually, image pyramid is used which is made of different scales of the same image P. However, if too many scales are used, the computation overhead becomes insufferable. To solve this issue, in this paper, we propose an adaptive scale selection mechanism. Assume that there exists a scale set defined as $S = [S_1, S_2..S_n]$. With the scale S_i, the original image P can be transformed to another image P_i with resolution $(H_i \times W_i)$, where $H_i = H \times S_i$, $W_i = W \times S_i$. The new image P_i then becomes the input of the cascade CNN

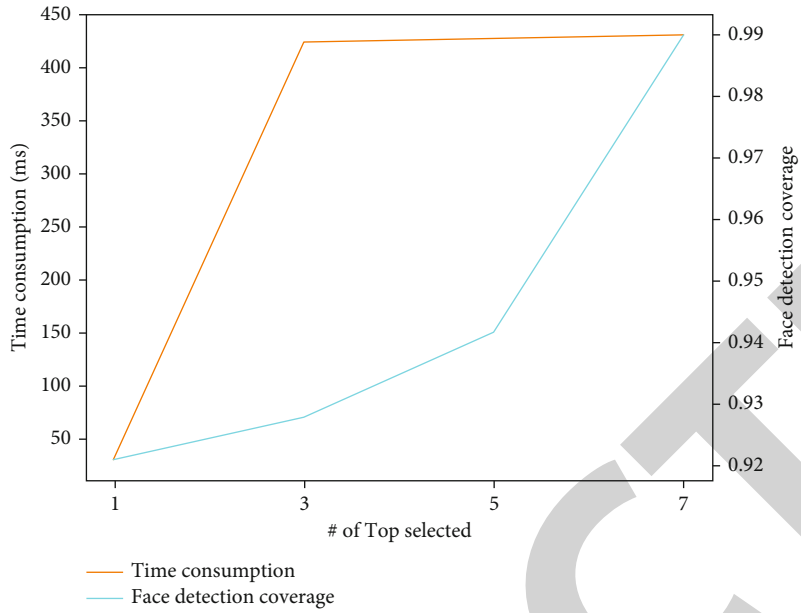


Figure 2: Relation between time consumption and coverage of found face along with the increase of top number.

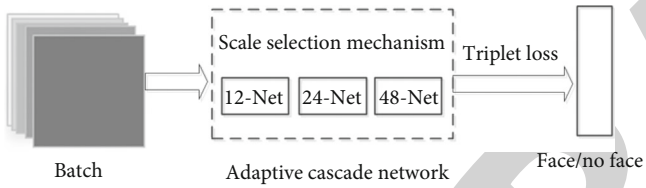


Figure 3: Network architecture.

which is described in Section 3.2. After the usage of the cascade CNN on the whole set S, we can obtain a sorted list from high value to low value. Then, we need to make a trade-off between detection speed and detection coverage. Therefore, a most appropriate number N which will be used to select top candidate frames should be decided. We conduct the experiment on the dataset and obtain the following Figure 2.

4. Network Architecture

In this section, we present the core network architecture which can be seen from Figure 3. The network includes a batch input layer and a face detection network based on adaptive cascade network. The adaptive cascade network is made of two parts. One is the three types of networks, i.e., 12-net, 24-net, and 48-net. The other is the scale selection mechanism. Both of them have been presented in Section 3.2. The last softmax layer of face detection network is replaced by a 1024-size fully connected layer (denoted as the triplet loss layer). Then, through L2 normalization, and get the embedding vectors, the triplet loss is calculated based on this feature representation.

4.1. Model Training. The introduction of triplet loss is to allow the network to learn an embedding. The network is trained using the squared L2 distances for the purpose to obtain face similarity. The face verification is completed by comparing whether the Euclidean distance of the image vector to be verified is less than a certain threshold or comparing with the known face vector in the library.

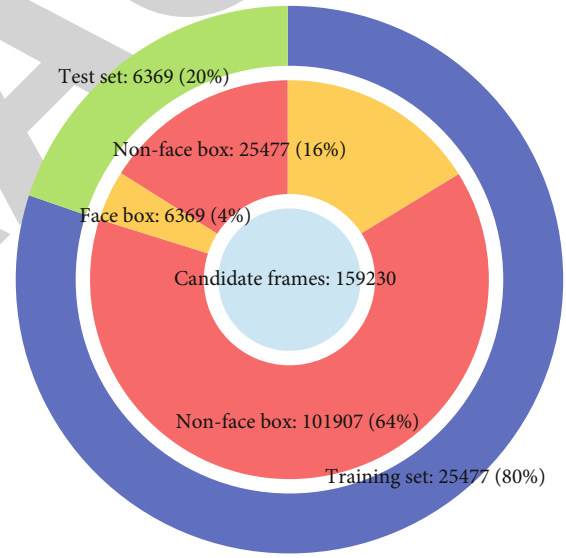


Figure 4: The construction of the self-built database.

Table 1: Approach comparison in terms of speed and accuracy.

Index	Value
CPU	Core (TM) i5-7200U 2.50GHz
Memory	8 GB
Graphics card	NVIDIA GeForce 920MX

4.2. Training Definition. Assume that a set of images input during training is in the form of$<a, p, n>$, where a and p

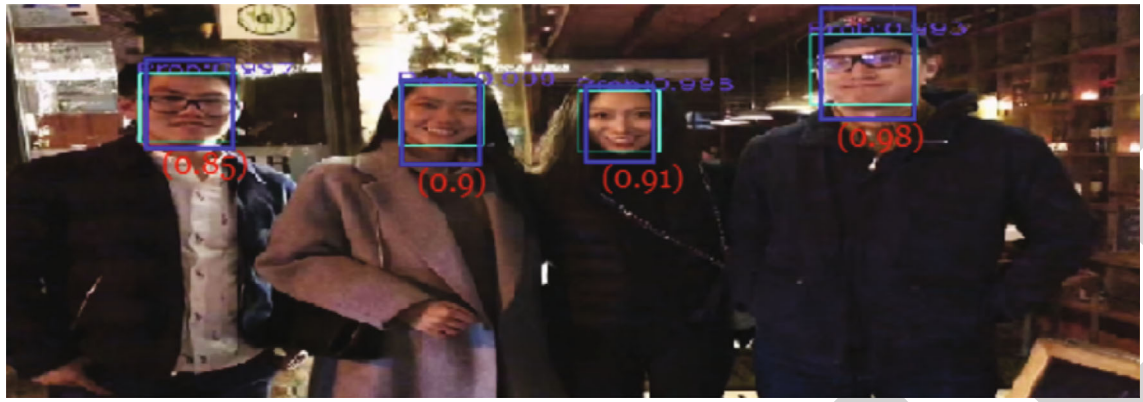


Figure 5: The detection result with and without the triplet loss function.

Table 2: Approach comparison in terms of time consumption and accuracy.

Model	Time consumption (milliseconds)	Accuracy (percentage)
Cascade network	578	93.4%
SIFT	365	95.2%
Our approach with top 1 candidate selected	31	94.5%
Our approach with top 3 candidates selected	75	96.9%
Our approach with top 5 candidates selected	152	98.6%

correspond to the same id, n corresponds to different id. The goal is to train the triplet loss layer parameters as Eq. (7).

$$d(a,p) + \alpha < d(a,n). \tag{7}$$

During the training process, the learnable parameters of all layers except triplet loss are in a frozen state, which is only used to complete feature conversion.

4.3. Triple and Training Method Selection. Since there is no target during the training, we find that there is a great influence of triplet selection of the model convergence and experimental results. Therefore, we introduce the definitions of easy triplet, hard triplet, and semihard triplet. Easy triplet: $L = 0$ is $d(a,p) + \alpha < d(a,n)$, which means that the distance between anchor and positive is less than the distance between anchor and negative. Hard triplet: $d(a,n) < d(a,p)$ means that the distance between anchor and positive is great than the distance between anchor and negative. Semihard triplet: $d(a,p) < d(a,n) < d(a,p) + \alpha$ means that the distance between anchor and positive is closely to the distance between anchor and negative.

The training method is divided into online and offline methods. The goal is to make the loss in formula (5) continue to decrease in the iterative process. The offline training method is to select all the triples in the training set and use the loss to gradient back propagation, but the distance between some anchors and negatives is very large, the calculation efficiency of using the full amount is low, and the embedding parameters cannot converge because the gradients generated by the anchors and negatives are too large; so, use online learning dynamically selects triples to solve the problems of low computational efficiency and nonconvergence of parameters.

5. Experiments

In this section, we conduct some experiments on two different face datasets. The images in the first dataset are collected from Internet while the second face dataset is the famous LFW dataset.

5.1. Performance on Self-Built Dataset. In this paper, we construct a face dataset, the images in which are all collected from Internet. We collect 12880 images totally, and based on them, 159230 candidate frames are generated. In these 159230 candidate frames, there are 31846 frames with face frames and 127384 nonface frames. The ratio of face boxes to nonface boxes is 1 : 4. In addition, in these 31846 candidate frames, the ratio of training set to test set is 8 : 2. We can see the detail construction of the self-built database from Figure 4.

The experiment is conducted in the following platform as Table 1:

We conduct the training on the self-built image dataset, the training result shows that classification accuracy of the cascade CNN achieves 99.7%, and the regression r-square is as high as 0.94. Figure 5 is the experimental result on one image of self-built image dataset with and without the triplet loss function. The green frames are without the triplet loss function, and the blue frames are with triplet loss function. We can see that the blue frames have more face contents.

Table 3: Comparison of different triplets and α.

Types	$\alpha = 0.5$	$\alpha = 1.0$
Hard triplet	97.1%	97.3%
Semihard triplet	97.5%	97.8%
Dynamic semihard triplet	99.1%	99.2%

We compare three approaches, the last of which has three variations, on the self-built dataset. We conduct this experiment 50 times and use the average as the final results.

Table 2 shows the experiments result. We can see from the result that our approach has less time consumption with all the three parameters, and the one with top 1 candidate selected works best, which only need 31 milliseconds. In addition, our approach with top 3 and 5 candidates selected work better than the common cascade network and SIFT. And the approach with top 5 candidates selected achieves a competitive score, 98.6%.

5.2. Performance on LFW. LFW dataset consists of 13233 images and 5749 individuals totally, and each image has the same resolution 250×250. Using the LFW dataset to train the embedding layer, we use different strategies of choosing triplets and different values of α to get the experimental results as Table 3. The first method is to select all hard triples to train the parameters in embedding. The second method is to select all semihard triples to train the parameters in embedding. The third method is to calculate all possible anchor-positive combinations at first, then use minibatch as the unit to calculate the distance of $d(a, p)$ in each minibatch, calculate the distance between all anchors and negative as $d(a, n)$, store them in a list, arrange the values of $d(a, n)$ in the list and select the smallest value, and calculate the distance between $d(a, p)$ and $d(a, n)$, if and only if $d(a, n) < d(a, p) + \alpha$; the a, p, n will be counted as a set of training data. When the value is 1.0, the accuracy rate for the hard triplet is 97.3%, the accuracy rate for the semihard triplet is 97.8%, and the accuracy rate for the dynamic semihard triplet is 99.2%.

6. Conclusions

This paper proposes a framework based on adaptive cascade CNN network and triplet loss for face detection and verification with fast speed and high accuracy. The framework firstly calculates the input through an image pyramid at a low resolution and adaptively selects the candidate frames. Secondly, those selected candidate frames are processed by more accurate detection network with high resolution. Finally, triple loss is calculated to conduct precise identification. The framework is very robust against complex backgrounds. We train the face verification model and complete the verification within 0.15 second for processing one image which shows the computation efficiency of our proposed framework. In addition, the experimental results also show that the competitive accuracy of our proposed framework which is around 98.6%. Using dynamic semihard triplet strategy for training, our network achieves a classification accuracy of 99.2% on the Labeled Faces in the Wild dataset.

Our future work will consider applying face verification to access control in smart grids and spatial crowdsourcing [44]. In addition, we will consider incorporating the encrypted face authentication information to improve the identity authentication protocols and achieve the goal of personalized privacy protection in face verification applications. At last, the combination of face verification and backscatter communication in IoT is another future research direction; we will consider to popularize the applications of IoT.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62072403, 61906167, in part by the 2019 Industrial Internet Innovation Development Project-Industrial Internet Network Security Public Service Platform Project (TC190H3WN), in part by the Key Research and Development Program of Zhejiang Province under Grant 2020C01076, in part by the Natural Science Foundation of Zhejiang Province under Grant LTY21F020001 and LY21F020011, and in part by the Research Project of Zhejiang Federation of Social Sciences (2022B19).

References

[1] L. Dong, M. N. Satpute, W. Wu, and D. -Z. Du, "Two-phase multidocument summarization through content attention-based subtopic detection," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 6, pp. 1379–1392, 2021.

[2] J. Yuan, W. Liu, J. Wang, J. Shi, and L. Miao, "An efficient framework for data aggregation in smart agriculture," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 10, article e6160, 2021.

[3] S. Zhao, F. Li, H. Li et al., "Smart and practical privacy-preserving data aggregation for fog-based smart grids," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 521–536, 2021.

[4] B. Yma, A. Lw, A. Zl, and C. Fl, "A novel face presentation attack detection scheme based on multi-regional convolutional neural networks," *Pattern Recognition Letters*, vol. 131, pp. 261–267, 2020.

[5] Y. Cai, Y. Lin, L. Xia, X. Chen, and H. Yang, "Long live time: improving lifetime and security for nvm-based training-in-memory systems," *IEEE Transactions on Computer-Aided*

Design of Integrated Circuits and Systems, vol. 39, no. 12, pp. 4707–4720, 2020.

[6] M. Mathias, R. Benenson, M. Pedersoli, and L. Van Gool, "Face detection without bells and Whistles," in *European Conference on Computer Vision*, pp. 720–735, Zurich, Switzerland, 2014.

[7] D. Shi and H. Tang, "Face recognition algorithm based on self-adaptive blocking local binary pattern," *Multimedia Tools and Applications*, vol. 80, no. 16, pp. 23899–23921, 2021.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[9] S. Cebollada, L. Payá, X. Jiang, and O. Reinoso, "Development and use of a convolutional neural network for hierarchical appearance-based localization," *Artificial Intelligence Review*, pp. 1–28, 2021.

[10] W. Liu, D. Anguelov, D. Erhan et al., "SSD: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Amsterdam, The Netherlands, 2016.

[11] X. Sun, P. Wu, and S. Hoi, "Face detection using deep learning: an improved faster rcnn approach," *Neurocomputing*, vol. 299, pp. 42–50, 2018.

[12] X. Shi, S. Shan, M. Kan, S. Wu, and X. Chen, "Real-Time rotation-invariant face detection with progressive calibration networks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City,USA, 2018.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016.

[14] J. Li, T. Wang, and Y. Zhang, "Face detection using surf cascade," in *IEEE International Conference on Computer Vision Workshops*, pp. 2183–2190, Barcelona, 2011.

[15] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, Massachusetts, 2015.

[16] Y. Dong and Y. Wu, "Adaptive Cascade deep convolutional neural networks for face alignment," *Computer Standards & Interfaces*, vol. 42, pp. 105–112, 2015.

[17] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin, "Extensive facial landmark localization with coarse-to-fine convolutional network cascade," in *Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops*, Sydney, Australia, 2013.

[18] J. C. Chen, A. Kumar, R. Ranjan, V. M. Patel, A. Alavi, and R. Chellappa, "A cascaded convolutional neural network for age estimation of unconstrained faces," in *IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, Niagara Falls, NY, USA, 2016.

[19] K. Zhang, Z. Zhang, H. Wang, Z. Li, Y. Qiao, and W. Liu, "Detecting faces using inside cascaded contextual CNN," in *IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017.

[20] R. Ranjan, V. M. Patel, and R. Chellappa, "HyperFace: a deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 41, no. 1, pp. 121–135, 2019.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *European Conference on Computer Vision*, Amsterdam, The Netherlands, 2016.

[22] J. Xu, J. Wang, Q. Qi, H. Sun, and D. Yang, "Effective scheduler for distributed dnn training based on mapreduce and gpu cluster," *Journal of Grid Computing*, vol. 19, no. 1, 2021.

[23] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, "A fast face detection method via convolutional neural network," *Neurocomputing*, vol. 395, pp. 128–137, 2020.

[24] C. Yu, X. Zhu, Z. Lei, and S. Z. Li, "Out-of-distribution detection for reliable face recognition," *IEEE Signal Processing Letters*, vol. 27, pp. 710–714, 2020.

[25] X. Ke, J. Li, and W. Guo, "Dense small face detection based on regional cascade multi-scale method," *IET Image Processing*, vol. 13, no. 14, pp. 2796–2804, 2019.

[26] B. Yu and D. Tao, "Anchor cascade for efficient face detection," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2490–2501, 2019.

[27] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[28] A. Hermans, L. Beyer, and B. Leibe, "In defense of the triplet loss for person re-identification," 2017, https://arxiv.org/abs/1703.07737.

[29] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: a unified embedding for face recognition and clustering," in *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, Massachusetts, 2015.

[30] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proceedings of ICML*, New York, NY, USA, 2009.

[31] J. Deng, J. Guo, and S. Zafeiriou, "ArcFace: additive angular margin loss for deep face recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4690–4699, Long Beach, CA, 2019.

[32] Z. Lu, X. Jiang, and C. C. Kot, "Deep coupled ResNet for low-resolution face recognition," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 526–530, 2018.

[33] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, "Feature transfer learning for face recognition with under-represented data," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, 2019.

[34] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, "Improvement of face recognition algorithm based on neural network," in *2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Changsha, China, 2018.

[35] G. Xu, X. Li, L. Jiao et al., "BAGKD: A batch authentication and group key distribution protocol for VANETs," *IEEE Communications Magazine*, vol. 58, no. 7, pp. 35–41, 2020.

[36] G. Xu, W. Zhou, A. K. Sangaiah et al., "A security-enhanced certificateless aggregate signature authentication protocol for InVANETs," *IEEE Network*, vol. 34, no. 2, pp. 22–29, 2020.

[37] W. Shu, K. Cai, and N. N. Xiong, "A short-term traffic flow prediction model based on an improved gate recurrent unit neural network," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[38] G. Xu, W. Wang, L. Jiao et al., "SoProtector: safeguard privacy for native SO files in evolving mobile IoT applications," *IEEE Internet of Things Journal*, vol. 7, no. 4, 2020.

[39] G. Xu, Y. Zhao, Y. Jiao et al., "TT-SVD: an efficient sparse decision making model with two-way trust recommendation in the

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1103022, 11 pages
https://doi.org/10.1155/2022/1103022

Research Article

Fusion Deep Learning and Machine Learning for Heterogeneous Military Entity Recognition

Hui Li, Lin Yu, Jie Zhang, and Ming Lyu

School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China

Correspondence should be addressed to Jie Zhang; zhangjie_njust@njust.edu.cn

Received 21 November 2021; Revised 13 December 2021; Accepted 17 December 2021; Published 17 January 2022

Academic Editor: Liqin Shi

With respect to the fuzzy boundaries of military heterogeneous entities, this paper improves the entity annotation mechanism for entity with fuzzy boundaries based on related research works. This paper applies a BERT-BiLSTM-CRF model fusing deep learning and machine learning to recognize military entities, and thus, we can construct a smart military knowledge base with these entities. Furthermore, we can explore many military AI applications with the knowledge base and military Internet of Things (MIoT). To verify the performance of the model, we design multiple types of experiments. Experimental results show that the recognition performance of the model keeps improving with the increasing size of the corpus in the multidata source scenario, with the *F*-score increasing from 73.56% to 84.53%. Experimental results of cross-corpus cross-validation show that the more types of entities covered in the training corpus and the richer the representation type, the stronger the generalization ability of the trained model, in which the recall rate of the model trained with the novel random type corpus reaches 74.33% and the *F*-score reaches 76.98%. The results of the multimodel comparison experiments show that the BERT-BiLSTM-CRF model applied in this paper performs well for the recognition of military entities. The longitudinal comparison experimental results show that the *F*-score of the BERT-BiLSTM-CRF model is 18.72%, 11.24%, 9.24%, and 5.07% higher than the four models CRF, LSTM-CRF, BiLSTM-CR, and BERT-CRF, respectively. The cross-sectional comparison experimental results show that the *F*-score of the BERT-BiLSTM-CRF model improved by 6.63%, 7.95%, 3.72%, and 1.81% compared to the Lattice-LSTM-CRF, CNN-BiLSTM-CRF, BERT-BiGRU-CRF, and BERT-IDCNN-CRF models, respectively.

1. Introduction

The US military attaches great importance to the unified management of battlefield resources. In recent years, the US Department of Defense Advanced Research Projects Agency (DAPRA) has issued a number of basic research topics on the unified management of battlefield resources. The topic guides focus on the importance of building the expert knowledge base and knowledge graph for the battlefield resources [1]. They point out the research route and direction of the unified management of battlefield resources for the US military in the future. Relevant guidelines also point out that it is necessary to focus on expanding the coverage of military entity data and improving the quantity and quality of the knowledge base of military entities by combining data from different data sources like the existing data sets of the US Army, military canonical books, reliable professional websites, and military blogs. In terms of specific applications, the US military has achieved many practical applications; the most famous case is that Palantir applies the intelligence analysis technology and knowledge graphs to assist the US Intelligence Agency in capturing Osama bin Laden and uncovering the Ponzi scheme. Currently, Palantir is working with DAPRA to conduct in-depth research on the application of knowledge graphs and intelligence analysis technology to assist in intelligence gathering, processing, and military resource control. In the area of unified resource management, the PLA Academy of Military Sciences, in conjunction with the Department of Equipment Development, has carried out a number of studies on the management of military resources. In recent years, they have published "Distributed Resource Management Methodology," "Research on End Resource Management Technology," and "Research on Resource Management Technology Based on Edge

Computing," and other related topics. With these recognized military entities, many military apps about military management could be constructed, such as military resource planning, military resource scheduling, and military resource monitoring, according to the topics issued by the DAPRA and the PLA Academy of Military Sciences. Furthermore, the military resource posture could be drawn in real time with the development of the MIoT technology.

Inspired by the topic guidelines of the US military and the PLA Academy of Military Sciences, we perform battlefield resource entities extraction works from multisource heterogeneous data. The entities, we mainly concern, are computing and storage, battlefield perception, communication networks, weapon platform, and logistics support, as well as integrated mission environment information, combatant, combat agency, combat time, combat location, and event information involved in combat resources. These military entities can be virtualized, and then, various virtual military resource services and military IoT applications can be built in cloud services [2]. With these cloud services, military staffs can carry out studies on the virtual scheduling of military resources, thus improving the utilization of military resources.

Military entity recognition technology is based on general entity recognition technology, which has experienced the development process from dictionary, rule, and machine learning to deep learning [3]. At present, the recognition model based on pretraining language model and deep learning algorithm is the main stream of entity recognition in military and general fields with the increasing computing power of small computers, the maturity of deep learning technology, and the development of pretraining language model [4]. However, it is a challenge work for the military entity recognition with the distinct domain characteristics of military field. Firstly, it is lack of solid and reliable corpus for military entity recognition [5]. Since the data stored in military information systems can only be accessed by the staff with special rights, and fewer AI researchers in military field, since that the quantity and quality of the data sets of military fields cannot be compared with open domain. Secondly, there are few types of corpuses for military entity recognition. Military entities have distinct domain characteristics that multiple terms, multiple abbreviations, multiple types of expressions, multiple nested expressions, and multiple fresh words [3]. As the few types of corpus and data sources, a few characteristics of military entities could be covered with these corpuses, and thus, the military entity recognition on cross data source could not be conducted. Last but not least, it is lack of unified criteria for entity division. A nested entity could be divided into different granularity with different worker, so we need a unified criteria to guide the practice for military entity recognition.

To solve the problems discussed above, three types, abbreviated, scientific or English name, and novel and casual, of military corpuses are constructed, and then, we can explore different research work on military entity recognition with them, such as crossing data source entity recognition. We improve the entity labeling mechanism for entity with fuzzy boundary based on the military entity recognition works conducted by other researchers. With the constructed corpuses, three kinds of experiments have been carried out, the experiments of applying the BERT-BiLSTM-CRF model to military entity recognition for different size corpus sets, the experiments of applying the BERT-BiLSTM-CRF model to military entity recognition across corpus sets, and the experiments of comparing the performance of multimodel military entity recognition. The experimental results show that, the model used in this paper outperforms the listed models, such as CRF, LSTM-CRF, Lattice-LSTM-CRF, and other models, for military entity recognition. Furthermore, the generalization ability of the model is verified with these experiments; the experimental results provide a reliable reference for researchers to use BERT-BiLSTM-CRF model for military entity recognition.

2. Related Works

Military entity recognition has attracted much attention, and many research works have been conducted. Researchers mainly applied machine learning combined with dictionaries or rules for military recognition at earlier times.

Jiang et al. [6] have proposed a model combined CRF with rules to extract military entities from combat instruments for automated generation of combat orders, which combines external lexical features, grammatical rules of military expressions, and the rule learning capability of CRF. The model achieves an F-value of 75.48% on the corpus constructed by the authors with 300 combat instruments. Feng et al. [7] have proposed a semisupervised military entity recognition method based on CRF for identifying entity military information such as military ranks, military equipment, military facilities, and military institutions. The proposed method takes use of the basic features of military texts to construct a grammatical feature set of military texts and fuses them with the CRF model. Experimental had been carried out on the corpus consisting of combat documents, duty documents, military documents, military online news, military blogs, and military reviews with this model, and the results show that the highest F-value of entity extraction was 90.9%. Shan et al. [8] have proposed a CRF-based military entity recognition method under a small granularity strategy for the military named entities with complex internal nested relationships and inconspicuous grammatical distinctions. This model applies a small granularity strategy, combines it with a CRF model to identify small nondivisible military entities, and finally, the small granular entities are integrated to obtain complete military entities. Experimental carried out on a manually constructed annotated corpus of combat instruments, and the model achieves an F-value of 78% on the corpus.

With the continuous development of deep learning techniques and the increasing computing power of minicomputers, applications of deep learning in the field of military entity recognition are emerging. In [9], a recognition method based on deep neural network models has been studied; the method applies the word vectors and word states as features for weapon name recognition and achieves an F-value of 91.02% on a corpus set constructed with

military website data. Liu et al. [10] apply a BiLSTM-CRF model to identify weaponry equipment names, and this achieves an F-value of 93.88% on the corpus constructed from military documents. Wang et al. [11] propose a character+BiLSTM+CRF model to extract military entity from military corpus, which aims to solve the problem that the complexity of artificial construction features and the inaccuracy of military text segmentation in the traditional methods of military named entity recognition; the experimental results show that the model proposed in this paper outperforms the traditional methods. Other applications for the military entity recognition have been studied, which are based on deep learning, such as recognition models based on a combination of a self-attentive mechanism and BiLSTM-CRF model [12], military entity recognition models based on multineural network collaboration [13], and military entity recognition models based on transfer representation learning [14]. Where the multineural network-based model applies word vectors obtained from BERT pretraining as input and combines with BiLSTM-CRF model for military entity recognition, which achieves an F-value of 84.07% on the corpus constructed based on military websites and military blogs. However, the authors did not consider the sentence contextual features in the training process, which easily led to contextual coreference that could not be effectively processed. Single data source was used, and the scenario of multiple corpus intersection was not considered in the work, which could not verify the generalization ability of the mode. Furthermore, the entity type coverage was relatively small, and the effect of corpus set size on the experimental results and the effect of heterogeneous data on the experimental results were not illustrated in the experimental session. At present, entity recognition based on pretrained models and attention mechanism in the field of generic entity recognition is the mainstream [15–20], which gives important insight into the direction of entity recognition technology development in the military field.

Recognition models based on lexicons, rules, and machine learning are traditional methods, which rely on powerful feature engineering and cost a lot for large-scale applications. Entity recognition models based on pretraining and deep learning do not need to rely on the support of basic feature engineering, and they are the main research direction for future military entity recognition. Some research results for named entity recognition have been achieved with these methods, but there are also many problems, such as no standard corpus to measure the merits of the models, no experimental tests across scenes and corpus, and no comprehensive consideration of all information in the context.

3. Construction of Multisource Heterogeneous Military Corpus

3.1. Analysis of Data Sources and Data Characteristics. Multisource heterogeneous data has the characteristics of wide fields, large span, and rich information. We can take advantage of this and then construct a relatively complete set of military entities for the research work of military entity recognition. There are two kinds of military text data; one is the nonpublic data, such as combat documents, military documents, military documents, reconnaissance intelligence, military teaching plans, simulation training task scenarios, and simulation logs; the other one is the open-source data, such as military blogs, military news, well-known military websites (such as Phoenix military, Jane's Defense Weekly, and Hanhe defense), and arms dealer websites. These data can be divided into three categories, according to the specific forms of military entity data: (1) abbreviation type, it is commonly used in military combat documents, such as double 35, which represents double barrel 35 mm artillery; (2) scientific or English name type, many military entities recorded as the forms of scientific or English name in military books, such as 7.62 mm sniper rifle, F-35, and Su-30; (3) novel and casual type, the expression of military entities in network terms is relatively casual, and with many fresh and cool words are used, such as the network terms about j-20, including "Wei Long," "J-00," and "door-to-door." In addition, the entities in these corpuses have the characteristics of fuzzy boundary and multinesting.

3.2. Data Preprocessing and Corpus Construction. In this work, we mainly select four kinds of nonpublic data and three kinds of open-source data as the main sources of the experimental corpus. Nonpublic data contains representative military documents, reconnaissance intelligence, simulation training mission scenario, and military books; open-source data contains military blogs, military reviews, and well-known military websites. The data in the selected data sources covers the three types of data discussed above, which can support the experimental needs for cross data scenarios. Most of the nonpublic data is text type, which needs to be electronic, and then, the original text is segmented according to the punctuation marks of ",", "。", "!", ";", and "?" and serialized into data in CSV or TSV format.

We apply crawler to access web data from military blogs, military reviews, and well-known military websites, then extract the text information by text density, and finally, the data is serialized by the proposed method above. The serialized corpus is the original unmarked corpus. We label the original corpus in a word level way [13]. In this way, three kinds of military entity recognition corpus sets, abbreviated, scientific or English name, and novel and casual, three military Entity Recognition Corpus sets, are obtained, covering 12 categories of entity objects including personnel name, military place name, time, military event, military institution, military facility, combat environment, computing and storage, battlefield perception, communication network, weapon platform, and logistics support. The size information of these corpus is shown in Table 1, and the entity information of them is shown in Table 2.

3.3. Data Labeling. There are various forms of military entities in multisource heterogeneous data; in [13], five rules have been proposed to solve the problem for recognizing military entities with fuzzy boundaries, but only partial cases can be solved for the problem with these rules. With this

Table 1: 3 types of corpus size information.

Corpus type	Number of sentences	Number of words
Abbreviation type	278149	12655807
Scientific or English name type	340786	14926444
Novel and casual type	686774	47044046
Total	1305709	74626298

problem, we propose some rules considering abbreviations and standardized expressions for the other cases.

Rule 1: numbers are connected with weapons or equipment (or place names, military institution names), and they can be labeled as weapon or equipment entities (place name entities, military institution entities), such as 1130 antiaircraft gun, 591 highlands, and the 38th army

Rule 2: numbers and length units are connected with the weapon entity, and they can be marked as the weapon entity, such as a 7.62 mm sniper gun

Rule 3: adjectives are connected with weapon entities, and we label them as weapon entities, such as long endurance UAV

Rule 4: the abbreviations of English or Chinese characters of weapon entities are connected with numbers, and they can be marked as weapon entities, such as J-20 and J16

Rule 5: personal names are connected with military institutions, and they can be marked as military institutional entities, such as the "Yang Gensi company"

With these rules proposed in this work and in [13], we can solve the problem that determining the entity boundary in heterogeneous corpus.

Named entities in the military field are characterized by multiple types, multiple professional terms, and less ambiguity [10]. The BIO labeling method is suitable for labeling entities with these characteristics, which is a concise and efficient labeling mechanism. The specific labeling scheme is shown in Table 3.

4. Military Entity Recognition Model

In this work, we apply the BERT-BiLSTM-CRF model to recognize battlefield resource entity recognition from military text. This model uses the word vectors obtained by BERT pretraining as input information and integrates bidirectional LSTM (Long Short-term Memory) and CRF to identify entities from the input information. The model is divided into BERT pretraining language model layer, BiLSTM layer, and the layer, and its structure is illustrated in Figure 1.

Let D denote the military corpus sets, thus $D = \{d_1, d_2, \cdots, d_n\}$, where d_i denotes the ith corpus set, $d_i = \{s_{i1}, s_{i2}, \cdots, s_{im}\}$, s_{ij} denotes the sentence j in corpus i, $s_{ij} = \{x_{ij1}, x_{ij2}, \cdots, x_{ijn}\}$, x_{ijk} denotes the kth word in sentence s_{ij}. At the beginning, the input unit transforms the x_{ijk} to C_k, and then, C_k is transformed to E_k by the transformer encoders; next, E_k goes through the forward LSTM units, $F = \{F_1, F_2, \cdots, F_n\}$, and backward LSMT units, $B = \{B_1, B_2, \cdots, B_n\}$; then, we get a feature matrix P_k; at the end, the CRF layer captures the dependencies between adjacent labels according to feature vectors and outputs corresponding labels y_k.

4.1. BERT Pretraining Model. Word vectors applied to military entity recognition can be trained from military corpus with the pretraining model of BERT. This model uses bidirectional transformer network structure to learn semantic feature information of military text context. In particular, two kinds of unsupervised pretraining tasks are developed, which inspired by the idea of cloze filling, to learn more context information from the text during the training process. One is Masked Language Model (MLM), and the other one is Next Sentence Predict (NSP), which are introduced to overcome the unidirectional problem that meets by most word vector generation models. As a result, more fully information can be extracted from the military corpus by the BERT pretraining language model which applies the organic combination of the transformer encoder structure and the unsupervised training task.

4.1.1. Input Unit. The input unit of the BERT model consists of Word Embedding, Segment Encoding, and Positional Encoding of the input sequences. Word features, sentence features, and position features of each word in the input sentence $s_{ij} = \{x_{ij1}, x_{ij2}, \cdots, x_{ijn}\}$ should be calculated before they are transformed into the BERT layer. Where the word features are denoted as $(e^w_{ij1}, e^w_{ij2}, \cdots, e^w_{ijn})$, sentence features are denoted as $(e^s_{ij1}, e^s_{ij2}, \cdots, e^s_{ijn})$, and position features are denoted as $(e^p_{ij1}, e^p_{ij2}, \cdots, e^p_{ijn})$. The word feature of the word x_{ijk} is provided by the corresponding word vectors in the vocabulary trained by Google. e^s_{ijk} denotes segment number of the input sequences; e^s_{ijk} takes either 0 or 1. We apply absolute position mode for the position feature; that is, $e^p_{ijk} = k$, where k denotes the position feature of the word at the kth input position.

The word feature, sentence feature, and position feature of the corresponding position in the input sequence are added together to obtain $C_k = e^w_{ijk} + e^s_{ijk} + e^p_{ijk}$, where $C_k \in C$, and $C = (C_1, C_2, \cdots, C_n)$. The input feature composition and calculation method are illustrated in Figure 2.

4.1.2. Transformer Encoder Unit. The BERT model consists of a bidirectional transformer encoder network (encoder structure is shown in Figure 3); the structure is illustrated in Figure 4. It takes sequence of $C = (C_1, C_2, \cdots, C_n)$ as input, and then, the multilayer transformer encoding unit pretrains the sequence into dynamic word vectors.

The encoder unit includes self-attention network unit, feedforward neural network unit, and basic normalized network unit. The self-attention network unit is used to learn the features of input sequences. The application of the self-attention mechanism helps the BERT to get rid of the problem of long-term dependence on recurrent neural nets, and thus, it can be used to perform parallel computing [21].

Table 2: Information on the 3 types of corpus entities.

Entity type	Abbreviation	Scientific or English name	Novel and casual	Total
Personnel name (P)	892	1527	2096	4515
Military place name (L)	8327	4172	20384	32883
Time (T)	1477	3091	4773	9341
Military event (E)	4529	3784	5846	14159
Military establishment (G)	5297	9533	12734	27564
Military installation (F)	2175	1732	2209	6116
Operational environment (S)	1639	873	1194	3706
Computational storage (C)	792	1399	2087	4278
Battlefield perception (R)	3327	1894	2256	7477
Communication network (N)	1079	970	1768	3817
Weapon platform (W)	3096	6718	13895	23709
Logistic service (H)	882	1349	2297	4528
Total	33512	37042	71539	142093

Table 3: BIO labeling system for military entity.

Entity category	Entity start	Inside entity
Personnel name (P)	B-P	I-P
Military place name (L)	B-L	I-L
Time (T)	B-T	I-T
Military event (E)	B-E	I-E
Military establishment (G)	B-G	I-G
Military installation (F)	B-F	I-F
Operational environment (S)	B-S	I-S
Computational storage (C)	B-C	I-C
Battlefield perception (R)	B-R	I-R
Communication network (N)	B-N	I-N
Weapon platform (W)	B-W	I-W
Logistic service (H)	B-H	I-H

The principle, the self-attention unit used for learning sentence features, is that it enables each word in a sequence to perform attention operations on each other to capture the input features. The formula for attention calculation is given in equation (1).

$$\text{Attention}(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (1)$$

In which, Q is the query vector, K is the key vector, V is the value vector, and d_k is the input vector dimension.

It is known that the ability is limited that single attention unit learns input features, so the multihead attention mechanism is applied by the transformer encoder unit. Its working principle is to perform different linear mappings of Q, K, and V and then calculate their attention values, respectively, and then fuse the h attention information obtained. Equations (2) and (3) are the calculation formulas of multiple attention.

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \cdots, \text{head}_h)W^O, \quad (2)$$

$$\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right). \quad (3)$$

Word order of input sequences is not considered for the self-attention unit, and the positional coding unit helps to solve this problem. At time t, the sum of Word Embedding, Positional Encoding, and Segment Encoding is the actual input for the BERT model, which has been discussed in the input unit. We can benefit from the input that the relative position and segment encoding added to ensure that the actual input word vector is different when the same word vector appears in different positions in different sequences.

4.1.3. Unsupervised Training Tasks. Inspired by clotting, the BERT model applies two unsupervised training tasks in the pretraining stage; one is the Masked Language Model, and the other one is the Next Sentence Predict.

In the Masked Language Model task, the model will randomly "remove" 15% of the words in the input sequences and then make the model actively learn the contextual semantic relations of the input sequences from different directions. Through iterative training, the probability of reasoning to get the correct answer is as large as possible, so as to achieve the purpose of learning the text semantics. In the task of Next Sentence Predict, the model randomly selects sentence pairs from the training text, in which the positive and negative samples account for 50%, respectively. Then, the training task is carried out on the training set of sentence pairs, and the BERT model is allowed to judge their correlation, so as to learn the relations between two sentences.

4.2. BiLSTM Layer. The BERT pretraining model provides dynamic word vectors for the whole recognition system, but it is slightly insufficient in learning sentence features.

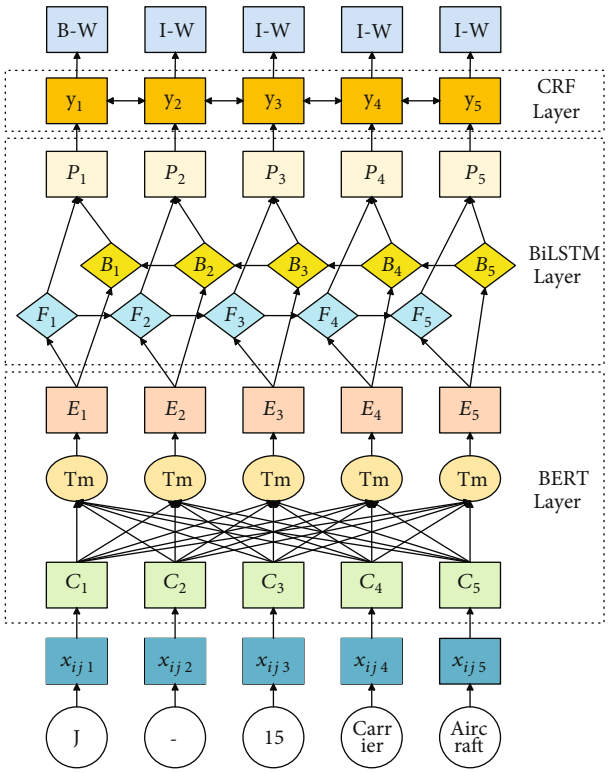


Figure 1: Military entity recognition model based on BERT-BiLSTM-CRF model.

Therefore, the BiLSTM layer, illustrated in Figure 2, is introduced to model sentences and learn the features of different input sequences. The input vector $E = (E_1, E_2, \cdots, E_n)$ which is the output of the BERT model goes through the forward LSTM layer $F = \{F_1, F_2, \cdots, F_n\}$, and the backward LSTM layer; then, it is transformed to a matrix $P \in R^{n \times m}$, where P_{kz} denotes the probability of the output label z that corresponds to the input word feature x_{ijk}.

The LSTM model, illustrated in Figure 5, adds "forgetting gate f_t," "input gate i_t," "output gate o_t," and "cell state C_t" to the structure of the RNN (recurrent neural networks) model, so it is known as improved RNN model. With the added units, the problems that long distance dependence and gradient disappearance of the recurrent neural network could be solved. Meanwhile, the units can adjust the memory function of the network help, which helps to maintain and update the state of the whole network.

The calculation formulas corresponding to the LSTM network units are shown as follows.

$$\begin{aligned} f_t &= \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\ \tilde{C}_t &= \tan h(W_c \cdot [h_{t-1}, x_t] + b_c), \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t, \\ h_t &= o_t * \tan h(C_t). \end{aligned} \tag{4}$$

In which, σ is sigmoid function, x_t is input vector, h_t is output vector, w is parameter matrix, and b is offset parameter.

4.3. CRF Layer. Dependency between tags is a common sense in the input sequence; here, we take the BIO labeling system to illustrate this sense. The starting label of each word in the input sequence is "B-" or "O," and it is usually that "I-X" follows "B-X," and "I-X" is used as the ending label of the word. However, "I-" cannot be used as the starting label. For example, a legal annotation sequence is "B-L I-L I-L," which together represents a location information. Illegal labels such as "B-G I-L" may appear, if the labeling process is not controlled. Unfortunately, the BiLSTM layer focuses on the context information and sentence features of the input sequence and cannot learn these annotation rules.

The CRF layer takes the output characteristic matrix of BiLSTM layer as input and outputs the global optimal label sequence, that is $y = (y_1, y_2, \cdots, y_n)$, the most possible sequence annotation. The CRF layer transforms the dependency information between tags into constraints when predicting tags, so as to ensure the accuracy of prediction. Such as the "I-L" label cannot appear after the "B-W" label, the dependency constraint relationship between the labels will be automatically learned by the CRF layer in the data training stage. The label constraint relationship is represented by the transfer matrix A, where A_{ij} represents the dependence intensity between the ith label and the jth label. The higher the score, the greater the intensity and vice versa. In the actual prediction process, the start state and end state will be added into the input sequence. Therefore, the actual matrix is $R^{(k+2)\times(k+2)}$. In a tag sequence with a length equals to the length of the input sequence, the model scores the tag y of the input sequence x, and this is calculated as follows:

$$\text{score}(x, y) = \sum_{i=1}^{N} P_{i,y_i} + \sum_{i=1}^{N} A_{y_{i-1}, y_i}. \tag{5}$$

After the score is calculated, and then the normalized probability is calculated with softmax unit:

$$P(y \mid x) = \frac{\exp\left(\text{score}(x, y)\right)}{\sum_{y'} \exp(\text{score}(x, y'))}. \tag{6}$$

Let Y denotes the set of all labels, where $\forall y' \in Y$, the denominator of equation (6) denotes that all possible transfer scores are obtained. Take logarithms on both sides of equation (6) to obtain that the log likelihood with respect to the input sequence x.

$$\log\left(P(y|x)\right) = \text{score}(x, y) - \log\left(\sum_{y'} \exp(\text{score}(x, y'))\right). \tag{7}$$

In the training process, the set of maximum probability labels in the sequence is selected by obtaining the value

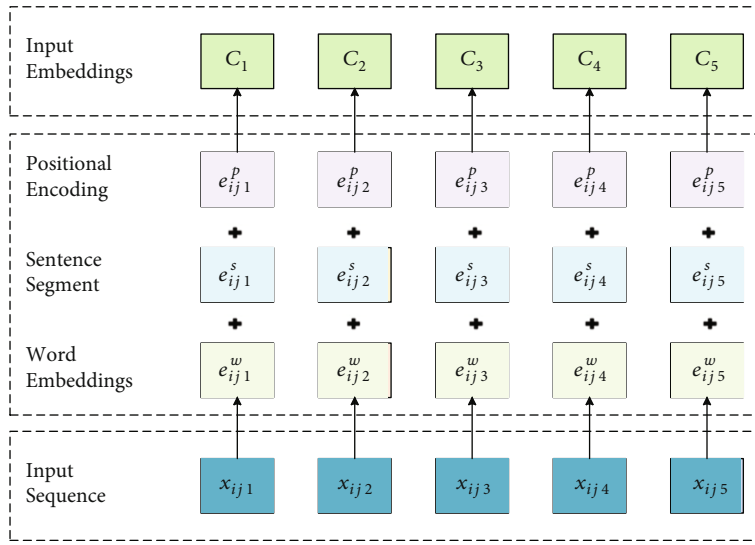


Figure 2: Input unit structure.

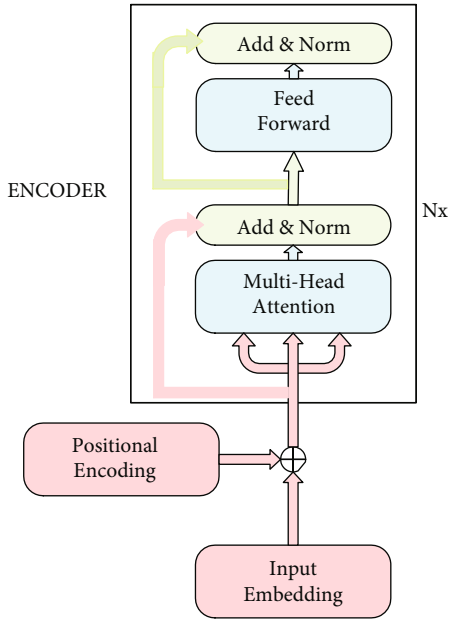


Figure 3: Transformer encoder unit.

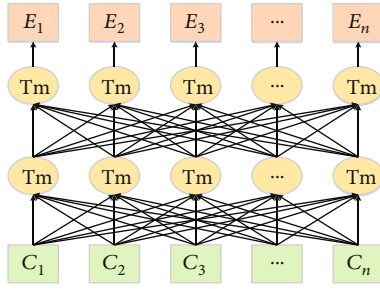


Figure 4: BERT pretraining language model.

of the maximum likelihood function, that is, the annotation sequence for the input sequence x predicted by CRF layer:

$$y^* = \underset{y \in Y}{\operatorname{argmax}}(\operatorname{score}(x, y)). \tag{8}$$

5. Experiments

We have conducted extensive literature research, and then, we find that few authors have considered the heterogeneous characteristics of military texts; thus, they have ignored the performance of these characteristics on the precision of deep learning model. Therefore, we have carried out relevant studies and conduct extensive experiments on related aspects. Experiments and discussions are shown as follows.

5.1. The Evaluation Metrics. In this work, precision (P), recall (R), and F-score (F) are selected as metrics for evaluating the performance of military entity extraction. The metrics are defined as follows:

$$\begin{aligned} P &= \frac{T_P}{T_P + F_P} \times 100\%, \\ R &= \frac{T_P}{T_P + F_n} \times 100\%, \\ F &= \frac{2 * P * R}{P + R} \times 100\%. \end{aligned} \tag{9}$$

In which, T_p is the number of positive entities identified by models, F_p is the number of negative entities identified by models, and F_n is the number of effective entities not detected by models.

5.2. Experimental Parameters. We carry out experimental with the BERT-Base model provided by an open-source project of Google. The information of hyperparameters in the

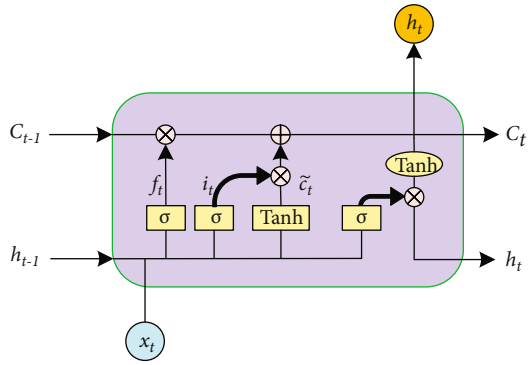


Figure 5: LSTM network structure.

Table 4: Model hyperparameters.

Parameters	Values
Batch size	64
Clip	0.5
Dropout rate	0.5
Learning rate	10^{-5}
Maximum sequence length	512
BiLSTM hidden layer dimension	256
Optimizer algorithm	SGD

training process is listed in Table 4. The deep learning framework and basic runtime environment are Pytorch V1.6.0 and Python V3.6.2. Experimental hardware configuration: 256G memory, 4 Nvidia GeForce RTX 3070 GPU.

5.3. Experimental Design and Experimental Result Analysis

5.3.1. Experiments on Corpus of Different Sizes. The three types of corpuses are divided into three data subsets of similar size by stratified sampling method, and experiments are carried out on them.

Experiment one (EP-ONE), one subset is extracted from each kind of corpus set to form a new corpus set, resulting in a total of 27 groups of experimental corpus sets. For each new corpus set, it is divided into five groups of similar size by the stratified sampling method, and then, using the cross-validation mechanism to carry out experimental, that four groups are selected as the training group, and the remaining one is used as the test group. With the cross-validation mechanism, five experiments are carried out on each new corpus set, and a total of 135 experiments are conducted. Finally, the mean of all experiments is taken as the test result.

Experiment two (EP-TWO), two subsets are extracted each time to form a new corpus set, resulting in 27 groups of experimental corpus sets. The following process is consistent with the EP-ONE.

Experiment three (EP-THREE), all the data is selected, and the corpus is divided into five groups of similar size by stratified sampling method. Then, the cross-validation mechanism is used as it is used in EP-ONE. A total of five experiments are carried out, and finally, the mean of five experiments is taken as the test result. The experimental results are listed in Table 5.

The experimental results show that the performance of military entity recognition with BERT-BiLSTM-CRF model is constantly improving with the increase of the training corpus, where the *F*-value of the EP-TWO is 6.6% higher than that of the EP-ONE, the *F*-value of the EP-THREE is 10.97% higher than that of the EP-ONE, and the *F*-value of the EP-THREE is 4.37% higher than that of the EP-TWO. To explain this situation, we analyze the data features of the corpus constructed in this work. The analysis result shows that the military entities distributed in the corpus have characteristics of multiple types and variety of representations. These characters make the distribution of the entities sparse, and the smaller the corpus set is the data sparsity is more obvious. Therefore, the evaluation metric values are low with a small training corpus set, such as the result of EP-ONE, of which the *F*-value for military entity recognition is only 73.56%. However, with the data increase of the corpus set, the sparse problem is gradually solved. Such as the EP-THREE, the *F*-value of military entity recognition reaches 83.53%. Therefore, in order to ensure the performance of military entity recognition, it is necessary to build a relatively large corpus for training the BERT-BiLSTM-CRF model.

5.3.2. Cross-Validation Experiments. Cross-validation experiments are conducted on the three types of corpus; the experiments are divided into two parts: Selecting any one of them as the training set and the other two as the testing sets and selecting any two of them as the training set and the other as the test set.

The experiments are conducted according to the above two strategies, and for the convenience of representing the results, the corpus types of the experiments are represented by numbers, where "1" represents the abbreviated corpus, "2" represents the scientific or English name corpus, and "3" represents the novel and casual corpus, and the experimental results are listed in Table 6.

The experimental results show that it is not robust that the generalization ability of the BERT-BiLSTM-CRF model proposed in this work, when it is used on different corpus sets, since the large difference in entity representation between different corpus sets. Such as when the model is trained on the abbreviated corpus and tested on the other two types of corpuses, the recall rate is only 49.78%, and the *F*-value is 56.85%. However, when it was trained on the novel casual type corpus and tested on both the abbreviated and scientific or English name type corpus, it gets better performance; the recall rate and the *F*-value are 74.33% and 76.98%, an increase of 24.55% and 20.13%, respectively, over the first group. We sample entities from different corpus and make a comparison on the distribution. Then, we find that the novel casual corpus contains a larger span of entity types and entity representation types, so when the model trained on it performs well. When cross-corpus training is conducted, the generalization ability of the model is significantly improved, with *F*-values reaching 74.36%, 82.79%, and 87.39%, respectively. The purpose of this experiment is to

Table 5: Experimental results of the BERT-BiLSTM-CRF model on corpus sets with different size.

	P (%)	*R* (%)	*F* (%)
EP-ONE	75.03	72.14	73.56
EP-TWO	81.29	79.06	80.16
EP-THREE	**85.73**	**83.37**	**84.53**

Table 6: Results of the cross-validation experiment on three corpus sets.

	P (%)	*R* (%)	*F* (%)
1-(2,3)	66.25	49.78	56.85
2-(1,3)	70.18	58.36	63.73
3-(1,2)	79.83	74.33	76.98
(1,2)-3	77.94	71.09	74.36
(1,3)-2	84.37	81.26	82.79
(2,3)-1	**89.13**	**85.72**	**87.39**

illustrate that when it is necessary to extract entities across scenes and across corpus sets, one needs to pay attention to the entity features, and the distribution of entities in different scenes corpus sets and choose the training corpus set reasonably. The experimental results also show that when hardware conditions are limited, users could choose a corpus with more entity types as the training set to train the model and would achieve better performance.

5.3.3. Comparison Experiments. The comparison models are divided into horizontal and vertical two groups, where vertical group contains models, such as CRF, LSTM, and BiLSTM-CRF, and horizontal group contains models Lattice-LSTM-CRF, BERT-BiGRU-CRF, and BERT-IDCNN-CRF. Experiments are carried out on the whole corpus, and the experimental process is the same as the EP-THREE mentioned above. The models, for comparison, have achieved state of the art in the development of named entity recognition; thus, the experimental results would be more convincing when compared to these models.

In experimental process, the CRF model is trained and tested with the open source CRF++ (v0.58) tool. The LSTM-CRF model and BiLSTM-CRF model use word vectors trained with word2vec as input, and the vector dimension is 300 [22]. The experimental super parameters are consistent with the literature [13]. The Lattice-LSTM-CRF model adopts the super parameters in literature [23]. The input of CNN-BiLSTM-CRF model adopts word2vec word vectors as the input, the vector dimension is 300, and other super parameters are consistent with literature [24]. The experimental results are listed in Tables 7 and 8.

The experimental results show that the BERT-BiLSTM-CRF model applied in this work outperforms the other models listed above; it gets better performance for the recognition of military entities on the corpus. In the longitudinal comparison experiments, we compare the metric values of the BERT-BiLSTM-CRF model with any other model in the longitudinal group. Where compared to the CRF model, the recall rate and *F*-value are increased by 23.76% and 18.72%, respectively; compared to the LSTM-CRF model, the recall rate and *F*-value are increased by 11.65% and 11.24%, respectively; compared to the BiLSMT-CRF model, the recall rate and *F*-value are increased by 8.94% and 9.24%, respectively; and compared to the BERT-CRF model, the recall rate and *F*-value are increased by 5.05% and 5.07%, respectively. Word2vec is a static word vector, and its application in dynamic word sense transformation scenarios is poor for entity recognition, whereas dynamic word vectors can perform well in this scene. Therefore, the LSTM-CRF and BiLSTM-CRF entity recognition models are not as effective as the BERT-BiLSTM-CRF model with dynamic pretrained word vectors for entity recognition on a corpus with "five-plus" types. Although the BERT-CRF model introduces a pretraining mechanism, it is not as effective as the BERT-BiLSTM-CRF model with a bidirectional long and short-term memory network in learning the contextual features of sentences. In the transverse comparison experiment, the experimental process is the same as longitudinal experiment. And we analyze their performance as follows: compared to the Latch-LSTM-CRF model, the recall rate and *F*-value are increased by 7.08% and 6.63%, respectively; compared to the CNN-BILSTM-CRF model, the recall rate and *F*-value are increased by 7.81% and 7.95%, respectively; compared to the CNN-BILSTM-CRF model, the recall rate and *F*-value are increased by 2.9% and 3.72%, respectively; and compared to the CNN-BILSTM-CRF model, the recall rate and *F*-value are increased by 1.75% and 1.81%, respectively. Where the Lattice-LSTM-CRF model uses the improved grid LSTM element to integrate word and word order information, which can avoid word segmentation errors not transmitted in the network. The CNN-BILSTM-CRF model uses CNN to extract character-level features of words, which improves the vector representation ability of words and reduces the influence of segmentation errors. Although these two models have been improved in character and word feature extraction, they cannot fundamentally

Table 7: Experimental results for longitudinal comparison.

Model	*P* (%)	*R* (%)	*F* (%)
CRF	73.45	59.61	65.81
LSTM-CRF	74.93	71.72	73.29
BiLSTM-CRF	76.16	74.43	75.29
BERT-CRF	80.64	78.32	79.46
BERT-BiLSTM-CRF	**85.73**	**83.37**	**84.53**

Table 8: Experimental results for transverse comparison.

Model	*P* (%)	*R* (%)	*F* (%)
Lattice-LSTM-CRF	79.57	76.29	77.90
CNN-BiLSTM-CRF	77.63	75.56	76.58
BERT-BiGRU-CRF	81.16	80.47	80.81
BERT-IDCNN-CRF	83.86	81.62	82.72
BERT-BiLSTM-CRF	**85.73**	**83.37**	**84.53**

overcome the weakness of static word vector, and the performance improvement and application scenarios will be limited to some extent. The BERT model, which skillfully combines multiattentional mechanism and unsupervised subtask training task, can integrate characters, words, sentences, and word order to learn contextual information and then complete the comprehension task. The experimental results show that there is little difference between the models with pretraining structure, but BiLSTM performs well than IDCNN and BiGRU in the modeling of serialized text features, so the recognition performance of BERT-BilSTM-CRF model is better in the entity recognition task of military language data.

6. Conclusion

In this work, we construct a set of corpuses and propose five rules for identifying the fuzzy boundary of military entities, which are applied to solve problems for military entity recognition, such as the lack of corpus, the single type of corpus, and the disunity of entity boundary division. These corpus have been divided into three types, abbreviation type, scientific or English name type, and novel and casual type. With these corpuses, we have conducted three types of experiments: (1) military entity recognition experiments using BERT-BiLSTM-CRF model on different sizes of corpus sets, (2) military entity recognition experiments using BERT-BiLSTM-CRF model across corpus sets, and (3) comparison experiments of multimodel military entity recognition. The experimental results illustrate the effect of different size data sets on the accuracy of the entity recognition model and the effect of data distribution on the accuracy of the recognition model and also validate the effectiveness of the BERT-BiLSTM-CRF model for military entity recognition.

At present, due to the limitation of data access, the amount of nonopen data available is limited, and many military entities are not yet covered. In the future, more new data samples will be generated by learning from existing sample data and combining the descriptions of some data patterns by domain experts using adversarial neural networks, which will be used to enrich the existing corpus. In addition, methods such as migration learning, knowledge distillation, and unsupervised learning are considered to reduce the reliance on corpus size and accuracy and to build lightweight military entity recognition models. In terms of practical applications, military knowledge graph plays a critical role in promoting the development of military intelligence, and military intelligence applications based on military knowledge graph will become more popular. In the future, intelligent resource management and scheduling technology will be widely used in the field of unmanned combat and in the field of military chess deduction.

Data Availability

The simulation experiment data of military texts used to support the findings of this study are restricted by the School Security Office in order to protect Military Intelligence Information. Data is available from Hui Li, mcclane@njust.edu.cn, for researchers who meet the criteria for access to confidential data.

Conflicts of Interest

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

Acknowledgments

The work by Ming Lyu is funded by the Natural Science Foundation of Jiangsu Province (Grant No. BK20180467).

References

[1] S.-H. Jo, H.-J. Kim, S.-Y. Jin, and W.-S. Lee, "A study on building knowledge base for intelligent battlefield awareness service," *Journal of the Korea Society of Computer and Information*, vol. 25, no. 4, pp. 11–17, 2020.

[2] W. Shu, K. Cai, and N. Xiong, "Research on strong agile response task scheduling optimization enhancement with optimal resource usage in green cloud computing," *Future Generation Computer Systems*, vol. 124, pp. 12–20, 2021.

[3] J. Li and P. Wang, "Military text named entity recognition based on deep learning," in *International Conference on Applications and Techniques in Cyber Security and Intelligence*, pp. 808–816, Shanghai, China, 2018.

[4] Y. Lu, R. Yang, X. Jiang, C. Yin, and X. Song, "A military named entity recognition method based on pre-training language model and BiLSTM-CRF," *Journal of Physics: Conference Series*, vol. 1693, no. 1, pp. 12161–12165, 2020.

[5] B. Zhou, H. Zhang, R. Zhang, Y. Feng, and Y. Xu, "Construction of military corpus for entity annotation," *Computer Science*, vol. 46, no. 6A, pp. 540–546, 2019.

[6] W. Jiang, J. Gu, and L. Cong, "Research on CRF and rules based military named entity recognition," *Command Control & Simulation*, vol. 33, no. 4, pp. 13–15, 2011.

[7] Y. Feng, H. Zhang, and W. Hao, "Named entity recognition for military texts," *Computer Science*, vol. 42, no. 7, 2015.

[8] H. Shan, H. Zhang, and Z. Wu, "A military named entity recognition method based on CRFs with small granularity strategy," *Journal of Armored Force Engineering Institute*, vol. 31, no. 1, pp. 84–89, 2017.

[9] F. You, J. Zhang, D. Qiu, and M. Yu, "Weapon named entity recognition based on deep neural network," *Computer System Application*, vol. 27, no. 1, pp. 239–243, 2018.

[10] C. Liu, Y. Yu, X. Li, and P. Wang, "Named entity recognition in equipment support field using tri-training algorithm and text information extraction technology," *IEEE Access*, vol. 9, pp. 126728–126734, 2021.

[11] X. Wang, R. Yang, and W. Zhu, "Military named entity recognition method based on deep learning," *Journal of Armored Force Engineering Institute*, vol. 32, no. 4, pp. 94–98, 2018.

[12] X. Zhang, X. Cao, and M. Zhang, "Military named entity recognition based on self-attention mechanism," *Command Control & Simulation*, vol. 41, no. 6, pp. 29–33, 2019.

[13] X. Yin, H. Zhao, J. Zhao, W. Yao, and Z. Huang, "Multi-neural network collaboration for Chinese military named entity recognition," *Journal of Tsinghua University (Science and Technology)*, vol. 60, no. 8, pp. 648–655, 2020.

[14] W. Liu, B. Zhang, W. Chen, C. Zhang, Y. Chen, and R. Pan, "Military named entity recognition based on transfer representation learning," *Command Information System and Technology*, vol. 11, no. 2, pp. 64–69, 2020.

[15] Z. Zhang, S. Wu, D. Jiang, and G. Chen, "BERT-JAM: maximizing the utilization of BERT for neural machine translation," *Neurocomputing*, vol. 460, pp. 84–94, 2021.

[16] L. Li and Y. Jiang, "Integrating language model and reading control gate in BLSTM-CRF for biomedical named entity recognition," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 3, pp. 841–846, 2020.

[17] I. Ashrafi, M. Mohammad, A. S. Mauree et al., "Banner: a cost-sensitive contextualized model for Bangla named entity recognition," *IEEE Access*, vol. 8, pp. 58206–58226, 2020.

[18] J. Qiu, Y. Zhou, Q. Wang, T. Ruan, and J. Gao, "Chinese clinical named entity recognition using residual dilated convolutional neural network with conditional random field," *IEEE Transactions on Nanobioscience*, vol. 18, no. 3, pp. 306–315, 2019.

[19] T. Zhou, W. Wu, L. Peng et al., "Evaluation of urban bus service reliability on variable time horizons using a hybrid deep learning method," *Reliability Engineering and System Safety*, vol. 217, article 108090, 2022.

[20] L. Dong, M. N. Satpute, W. Wu, and D.-Z. Du, "Two-phase multidocument summarization through content attention-based subtopic detection," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 6, pp. 1379–1392, 2021.

[21] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 2017, pp. 5999–6009, 2017.

[22] Z. Yin and Y. Y. Shen, "On the dimensionality of word embedding," *Advances In Neural Information Processing Systems 31 (NIPS 2018)*, vol. 31, pp. 1–12, 2018.

[23] Y. Zhang and J. Yang, "Chinese NER using lattice LSTM," in *Proceedings of the 56th annual meeting of the association for computational linguistics (ACL), VOL 1*, pp. 1554–1564, Melbourne, Australia, 2018.

[24] D. Gao, L. Peng, and Y. Bai, "HAZOP text named entity recognition using CNN-BilSTM-CRF model," in *2020 Chinese Automation Congress (CAC)*, pp. 6159–6164, Shanghai, Peoples R China, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6883591, 10 pages
https://doi.org/10.1155/2021/6883591

Research Article

Research on RFID Anticollision Algorithms in Industrial Internet of Things

Haizhong Qian

Information Engineering College, Jiangsu Maritime Institute, Nanjing 211199, China

Correspondence should be addressed to Haizhong Qian; hzhqian@jmi.edu.cn

Received 1 October 2021; Accepted 30 November 2021; Published 14 December 2021

Academic Editor: Yinghui Ye

As a perception enabling technology of the Internet of Things, RFID can quickly identify target objects. The tag-to-tag collision problem seriously affects the identification performance of the RFID system, which causes the reader to be unable to accurately identify any tag within the specific time. The mainstream anticollision algorithms are limited by the performance bottleneck under the standard framework. In this paper, we analyze the features and merits of three kinds of algorithms in detail and propose a new algorithm architecture for RFID anticollision. Through the extensive experimental results comparison, we prove that the new architecture is effective to improve the performance of DFSA algorithms. Finally, we summarize the future research trends in the RFID anticollision algorithms.

1. Introduction

Radio Frequency Identification (RFID) [1–3] technology is an emerging backscatter communication technology [4], which uses radio frequency signals to carry out information exchange in the wireless channel and realizes the noncontact identification between objects. A typical RFID system consists of readers, tags, and backend components. In many applications based on RFID technology, a mass of passive RFID tags is deployed in the RFID system. Due to the nature of passive tags, tags cannot communicate with each other, so all tags can only receive signals from readers. Since multiple tags share the same wireless channel to communicate with readers, when multiple tags send data to readers at the same time, the phenomenon of collision between more tags will occur, resulting in the failure of readers to identify tags. Obviously, the collision brings great challenges to information collection. Therefore, the research on RFID electronic tag information collection is very suitable for the background of the development of the times [5, 6] and has important reference value and practical significance. Therefore, when there are multiple tags at the same time in the covering area of the reader, the reader must use anticollision algorithm to improve the efficiency of tag recognition.

The current RFID multitag anticollision algorithms can be mainly divided into three categories, namely, probabilistic algorithms [7–10], deterministic algorithms [11–15], and hybrid algorithms [16–18]. Probabilistic algorithms are mainly derived from Aloha-based ideas. An Aloha-based algorithm takes the way of tag answering first and lets the tag randomly select a time period to respond to the query request of reader. If only one tag responded to the query in the current time period, the tag was successfully identified. In the Aloha-based algorithm, the most commonly used is the dynamic framed slotted Aloha (DFSA) algorithm. The DFSA algorithm formulates the concept of frame length (the frame length is equal to the number of time slots allowed by the tag), and then, the tag that receives the reader command will randomly select a time slot in a frame to respond and reply with its own ID. When a frame is over, the reader uses mathematical methods to estimate the number of remaining tags based on the slot status counted in the previous frame and starts the next frame identification until all tags are successfully recognized. The performance bound of the DFSA algorithm is 0.368 [19]. All DFSA algorithms are easy to implement and have low equipment cost but may have tag starvation. Deterministic algorithms are also called tree-walking algorithms. The core idea of this type

of algorithm is to use bit tracking technology [20, 21] to lock the specific location of ID information collisions, so that the reader can adjust the query prefix in time to successfully identify subsequent tags. As the query progresses, the tree splits and all the child nodes of the query tree are retrieved. In other words, as long as there are enough queries, readers can accurately query the ID of each tag. Thus, tree-walk algorithm has a 100% recognition rate and can avoid the problem of tag starvation. However, once the tag IDs are not evenly distributed, the tree-based algorithm will generate many collision slots and lead to a gradual deterioration in performance. In order to maximize the multitag recognition performance of the RFID system, many researchers have merged the advantages of different types of algorithms and proposed a series of hybrid anticollision algorithms. The representative algorithms are BSTSA [16], BTSA [17], and GBSA [18]. Compared with traditional anticollision algorithm, the above algorithms have been significantly improved in tag recognition performance, but at the same time, they also caused some problems, such as high hardware cost, compatibility with existing RFID standards, and other issues. It is noted that there are also two kinds of collisions in the RFID systems, namely, reader-to-reader collision. The reader can use some scheduling methods to avoid such collisions. And this kind of collision is not the focus of this paper.

This paper mainly analyzes the performance of the typical representatives of the existing several types of anticollision algorithms and summarizes and analyzes the performance bottlenecks and other shortcomings of the existing methods. Based on the analysis, we propose a new algorithm architecture in which the system throughput can break through the bottleneck of existing DFSA algorithms. Different from the traditional DFSA algorithms, the proposed new method can separately cope with each collision slot with the independent small-size frame. Through theoretical derivation and analysis, we proved that the proposed new method can make the anticollision algorithm break the performance limit under the constraints of the existing framework. Finally, the future research trends of anticollision algorithms are summarized.

2. Related Works

Taking into account factors such as RFID equipment cost and implementation complexity, currently, in RFID readers, multitag conflict avoidance technologies based on time division multiplexing are mainly used, which mainly include three types of methods based on Aloha, binary splitting, and query tree. Among the methods based on Aloha, currently, the most widely used is the dynamic framed slotted Aloha algorithm, which has been adopted by the UHF international standard EPC C1 Gen2. At present, most research focuses on two aspects. One is to maximize the utilization of time slots by dynamically adjusting the frame length of each round, and the other is to avoid waste of resources by terminating low-utilization frames early. The EACAEA [22] algorithm proposed by Chen considers the difference in time duration between different types of time slots to optimize the frame length to maximize system throughput. Chen's research shows that when the length of the collision slot is 5 times that of the idle slot, the optimal frame length is 1.89 times the number of tags to be identified. However, this method only adjusts the frame length once in one round of recognition, which causes its performance to show large fluctuations when the number of tags changes, and its stability is poor. Su et al. [4] designed a frame length adjustment strategy based on time and energy saving and, based on this strategy, proposed a multitag identification algorithm named TES-FAS suitable for the EPC C1 Gen2 standard. The TES-FAS algorithm combines subframe observation and adaptive frame adjustment mechanisms and can achieve better recognition performance under different reader parameter configurations. However, the frame length optimization of the reader in TES-FAS algorithm is based on a static RFID environment, which makes it not suitable for dynamic RFID environments. Taking into account the mobility of tags, Zhu et al. [23] designed a scheduling-based RFID anticollision algorithm that satisfies a high recognition rate under a certain tag moving speed. Although the method takes into account the movement characteristics of the tag, its design is based on an ideal communication channel. In actual RFID application scenarios, there are path loss, signal attenuation, shadowing effects, etc., resulting in unsatisfactory communication channels between the reader and the tag. For this reason, many researchers [24–26] have studied multitag identification methods suitable for nonideal channels. The binary splitting method is intuitively a random access algorithm. Different from the Aloha-based algorithm, the collided tag set will be separated by a fixed probability of 0.5. In binary splitting methods, the tag starvation will be significantly weakened.

Law first used the query tree [21] algorithm to solve the signal conflict problem in the RFID tag identification process. The tag ID is essentially a sequence of 0/1 binary numbers. The recognition process of the query tree algorithm is similar to a virtual binary tree, the height of the tree is the ID length of the tag, and each branch is marked with the "left 0 right 1" method. The reader first sends a query command with a prefix of 0, and those tags whose ID prefix is 0 will respond to the reader and return their own ID. When a collision occurs, the reader divides the collided tag set into two subsets based on tag ID collision bit. These subsets get smaller and smaller until each subset contains only one tag. This type of algorithm requires a stack in the reader to store the query prefix information. The reader will continuously update the query prefix according to the collision bit and push the query prefix onto the stack. The entire recognition process will not end until the stack is null. The disadvantage of the query tree algorithm is that when the tags are dense, the signal conflicts are obviously intensified, which affects the system's recognition performance. Pan and Wu [27] proposed the STT algorithm, which learns the distribution of tags online based on the previous query results, thereby dynamically selecting the query depth of the query tree. The shortcoming of the STT algorithm is that it assumes that the tag distribution of the subsequent query is consistent with the previously learned tag distribution.

Table 1: Summary of various types of anticollision algorithms.

Types	Representatives	Features	Development trends
Aloha-based	MAP [19]	High estimation accuracy and high complexity	Reduce the complexity
	ECD [10]	Pulse detection	Improve the feasibility
	FEIA [22]	Slot-by-slot adjustment	Enhance initial performance
Tree-based	CT [29]	Remove the empty queries	Further improve the efficiency
	DPPS [30]	Identify multiple tags in the same time slot	Optimize query mechanism
	CwT [12]	Window mechanism to reduce the transmitted bits	Reduce the slot number
Hybrid-based	BSTSA [17]	First proposed hybrid architecture	Performance improvement
	GBSA [18]	Performance limit of UHF RFID	Optimize the architecture

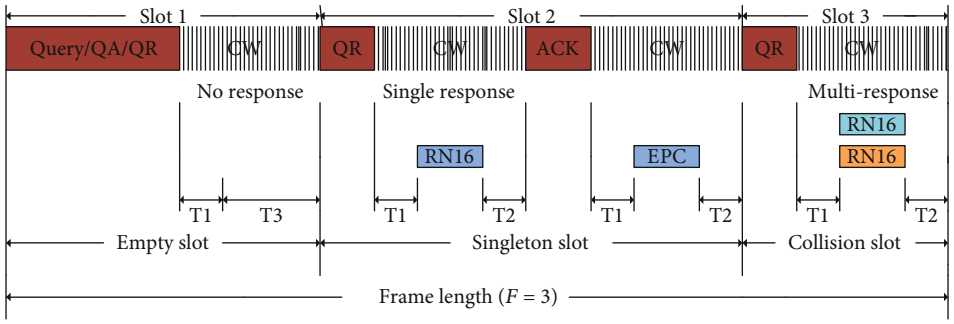


Figure 1: The timing link of EPC C1 Gen2 UHF RFID standard.

Shahzad and Liu [28] questioned this hypothesis and proposed the TH algorithm, which first uses the tag number estimation method to estimate the number of remaining tags and then estimates the ID distribution of unidentified tags and then jumps directly to the optimal layer for subsequent queries. In order to optimize the query prefix and improve the recognition efficiency of the query tree algorithm, subsequent researchers have presented a series of work on this basis, including the CT [29], CCMA [20], CwT [12], DPPS [30], and BQMT [13] algorithms. In Table 1, we summarize the features and development trends for existing representative anticollision algorithms.

3. Performance Analysis of Anticollision Algorithms

3.1. Analysis of DFSA Algorithms. The DFSA algorithm is a classical representative of Aloha-based algorithm, which strictly follows the timing link of EPC C1 Gen2 standard as illustrated in Figure 1. Based on the frame time slot algorithm, the DFSA algorithm estimates the number of tags in the current recognition stage through the tag estimation algorithm and adjusts the frame length at the beginning of each frame recognition stage according to the estimated number of tags to be recognized. Therefore, the key to the improvement of DFSA algorithm is (1) improve the accuracy of the estimation of the number of tags and adjust the frame length reasonably and (2) reduce the number of idle time slots and collision time slots. The workflow of the DFSA algorithm is illustrated in Figure 2.

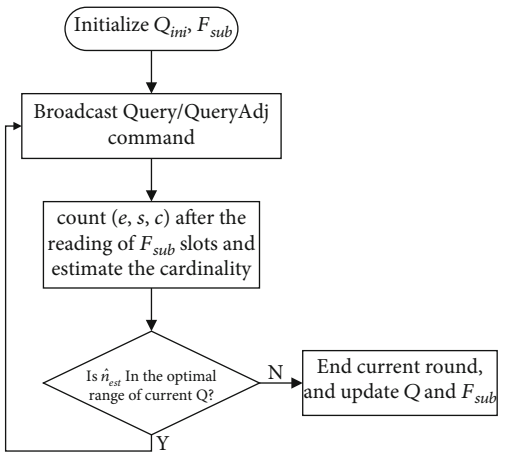


Figure 2: The flowchart of the DFSA algorithm.

We suppose the number of effective tags is$N$, and the length of frame size is$F$, since the probability of choosing the same time slot for any$r$tags can be expressed as

$$P_r = C_N^r \left(\frac{1}{F}\right)^r \left(1 - \frac{1}{F}\right)^{N-r}. \quad (1)$$

For a successful slot, there is only one tag that sends its IDs to the reader. So, the probability of successful slot is

$$P_s = C_N^1 \left(\frac{1}{F}\right)^1 \left(1 - \frac{1}{F}\right)^{N-1}. \quad (2)$$

When the current time slot has no tag send ID, it is defined as an idle time slot, and its probability can be expressed as

$$P_e = \left(1 - \frac{1}{F}\right)^N. \tag{3}$$

There are only three slot statuses during the recognition process: successful slot, idle slot, and collision slot. Thus, the probability of collision slot is

$$P_c = 1 - P_s - P_e = 1 - \frac{N + F - 1}{F}\left(1 - \frac{1}{F}\right)^{N-1}. \tag{4}$$

According to Equations (1), (2), and (3) above, we can get the expected values of successful slot, idle slot, and collision slot as follows:

$$\begin{cases} E_s = N(1 - 1/F)^{N-1}, \\ E_e = F(1 - 1/F)^N, \\ E_c = F - (N + F - 1)(1 - 1/F)^{N-1}. \end{cases} \tag{5}$$

We denote the system throughput as the ratio of successful slots to total slots F, denoted as U. Thus, U is expressed as

$$U = \frac{N}{F}\left(1 - \frac{1}{F}\right)^{N-1}. \tag{6}$$

In order to satisfy U maximum, we take the derivative of F on both sides of Equation (6) and make it equal to 0; we then have

$$\frac{dU}{dF} = \frac{N(N - F)(F - 1)^{N-2}}{F^{N+1}} = 0. \tag{7}$$

From Equation (7), we can know that in order to maximize U, there are two possibilities for the value of F, either $F = 1$ or $F = N$. Considering the actual situation, the number of tags in the coverage of the reader is usually much greater than 1. Therefore, we find that the F value that satisfies maximal U is N. That is, when the given frame length is equal to the number of tags to be identified, the anticollision algorithm can achieve the maximum system throughput.

Figure 3 reveals the relationship between U, P_e, and P_c under different tag numbers and different frame lengths. We can see from Figure 2 that each curve about U corresponds to a peak point, which is approximately equal to 0.368, which is obtained when $F = N$. Since the reader does not know the specific number of tags before identification, the purpose of the DFSA algorithm is to estimate the number of remaining tags so as to approach the performance to 0.368. Therefore, 0.368 has also become the performance bottleneck of the traditional DFSA algorithm. My research found that the reason for this bottleneck is that the current DFSA algorithm framework is aimed at all unread tags, and the reader sets a relatively large frame length to identify

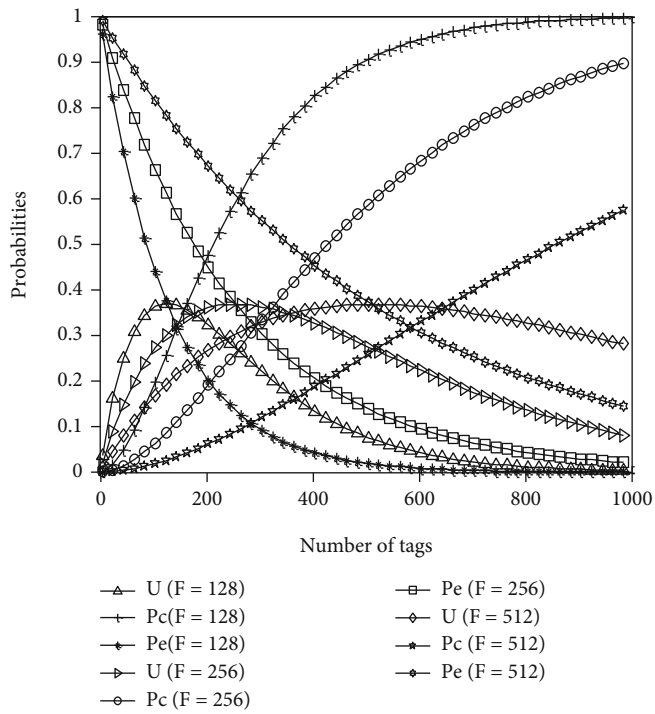


Figure 3: U, P_e, and P_c for different N and F values.

them. From Equation (6), we know that when F approaches infinity, the value of U approaches 0.368. When F is small, the value of U can be greater than 0.368. For example, when $F = N = 2$, the value of U is 0.5. This phenomenon stimulates the design of an independent frame recognition scheme, so that system throughput of the anticollision algorithm exceeds 0.368. Let us do a simple theoretical derivation to verify this point.

According to the definition of system throughput, we know that it can be redefined as N divided by the total number of time slots to identify N tags. Accordingly, Equation (6) can be rewritten as the following formula:

$$U = \frac{N}{F + A}, \tag{8}$$

where A represents the expected number of remaining time slots required for the reader to recognize N tags except for the initial frame F. Similarly, according to the calculation formula of system throughput, we can get

$$A = \frac{N_{\text{rest}}}{(1 - (1/N_{\text{rest}}))^{N_{\text{rest}}-1}}, \tag{9}$$

in which N_{rest} means the number of remaining tags. We know that before the end of the entire tag identification process, there are always collision slots, and the sum of the tags involved in all collision slots is the remaining tags. Thus, we have

$$N_{\text{rest}} = \sum_{i=1}^{m} r_i. \tag{10}$$

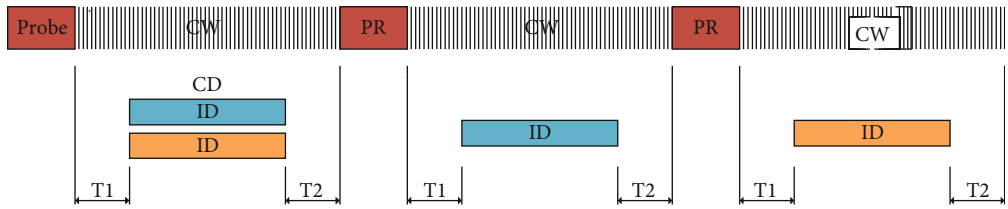


Figure 4: The timing link of ISO 18000-6B RFID standard.

Herein, m is the number of collision slots counted in the initial frame, and r_i represents the number of tags involved in i-th collision slot. We assume that the idea of the new anticollision algorithm (designed in this paper) is to independently allocate a small frame for each collision slot to identify it; then, the system throughput of the algorithm can be derived as follows:

$$U_{\text{new}} = \frac{N}{F+B}, \qquad B = \sum_{i=1}^{m} \frac{k_i}{(1-(1/k_i))^{k_i-1}}. \tag{11}$$

Assume that among all collision slots, the lth time slot involves the largest number of tags, and the number of tags included is k_l. So, we have

$$B < B^* = \frac{\sum_{i=1}^{m} k_i}{(1-(1/k_l))^{k_l-1}}. \tag{12}$$

Obviously, the function of $(1-(1/x))^{x-1}$ is a monotonically decreasing function. From this, we can deduce

$$U^*_{\text{new}} = \frac{N}{F+B^*} > U_{\text{DFSA}} = \frac{N}{F+A}. \tag{13}$$

The reason why the above formula holds is that $N_{\text{rest}} > k_l$. Therefore, we have

$$U_{\text{new}} > U^*_{\text{new}} > U_{\text{DFSA}}. \tag{14}$$

We prove through the above derivation that in the same multitag recognition scenario, the system throughput of the new algorithm can break through the performance bottleneck of the traditional DFSA algorithm. It is worth noting that the new algorithm here only refers to the concept and idea we put forward, and the detailed algorithm design is not the focus of this paper.

3.2. Analysis of Query Tree Algorithms. Unlike the DFSA algorithms, the query tree- (QT-) based algorithms do not have a common performance upper bound. Therefore, different optimization methods for query prefixes may result in large differences in final performance. The timing of QT-based algorithms usually follows the ISO-18000-6B standard, as shown in Figure 4. The workflow of QT-based algorithm is illustrated in Figure 5. Since there will be differences in the performance of each QT algorithm, we choose several representative QT algorithms for performance analysis here.

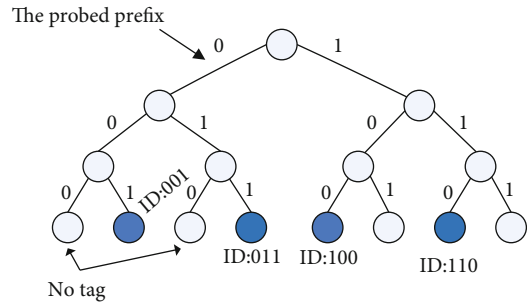


Figure 5: The workflow of the QT-based algorithm.

(1) Collision tree (CT) algorithm

The basic principle of the CT algorithm is that the reader sets the highest collision bit of the received string to 0 and 1, respectively; generates a new query prefix; and sends a new query command in the next time slot. The tag that receives the query command only needs to return the remaining ID information after matching the prefix. Compared with the traditional QT algorithm, the advantage of the CT algorithm is that it eliminates idle time slots in the query process and reduces the amount of data transmission in the time slot.

Assuming that there are 5 tags (A, B, C, D, and E) in the current reader's working domain, their IDs are "0010," "0101," "1101," and "1110." The reader uses the QT algorithm to recognize them, and the detailed recognition process is shown in Table 2.

We can observe in Table 2 that it takes 9 time slots for the reader to recognize these 5 tags using the QT algorithm. The recognition process of the QT anticollision algorithm is similar to traversing a binary tree, and each node on the tree will be detected. We can see that in the QT algorithm, each tag replies with a complete ID, which undoubtedly increases the communication complexity in the recognition process. In addition, when the number of tags to be identified is large, the QT algorithm will generate a large number of idle time slots, thereby further reducing performance. Since the CT algorithm eliminates idle nodes on the basis of the QT algorithm, the binary traversal tree corresponding to the CT algorithm only contains collision nodes (intermediate nodes) and leaf nodes, thereby improving the recognition efficiency.

Table 2: An recognition example by using QT algorithm.

Slot	Query prefix	Responding data from tags	Slot status
1	Empty string Ɛ	xxxx	Collided
2	0	0xxx	Collided
3	00	0010	Successfully identify A
4	01	0101	Successfully identify B
5	1	11xx	Collided
6	10		Idle
7	11	11xx	Collided
8	110	1101	Successfully identify C
9	111	1110	Successfully identify D

Assuming that there are n tags to be identified in the RFID system, the reader uses CT algorithm to identify them; the total number of time slots required can be expressed as

$$N_t = N_c + n, \tag{15}$$

where N_c represents the number of collision nodes in the traversal tree. We know that a collision node will produce 2 child nodes, so N_t can be further rewritten as

$$N_t = 2 \times N_c + 1. \tag{16}$$

Comparing Equations (15) and (16), we can have

$$N_c = n - 1. \tag{17}$$

Therefore, the total number of slots required by the CT algorithm to identify n tags can be expressed as

$$N_t = 2 \times n - 1. \tag{18}$$

According to the definition of system throughput, we can know that the maximum system throughput of CT algorithm is

$$U_{CT} = \lim_{n \longrightarrow \infty} \frac{n}{2n-1} = \lim_{n \longrightarrow \infty} \frac{1}{2 - (1/n)} = \frac{1}{2}. \tag{19}$$

Through formula (19), we know that the performance of the CT algorithm is relatively stable, and its system throughput is maintained at 50%. However, it is essentially a binary tree search. When the collision is obvious, the search process cannot be accelerated.

(2) CCMA algorithm

The CCMA algorithm is a multiary search algorithm, which introduces a custom query command and a collision string mapping mechanism. The mapping relationship of collision data is shown in Table 3. The main idea of the CCMA algorithm is that if the reader detects that the first and second collision bits are continuous, the reader will send a custom query command, namely, QueryP in the next time slot to make the involved tag return a 4-bit mapping string to replace the original ID prefix information. This mapping string can accurately reflect the collision information of the tag. The reader can accurately identify the first 2 bits of collision information of the tag through the received mapping string to determine the next query command. If the first and second collision bits are not consecutive, the reader will use the CT algorithm to identify the tag.

Table 3: The mapping table used in CCMA algorithm.

Collision information	Mapped string
00	0001
01	0010
10	0100
11	1000

The authors in [20] only give the simulation results of the CCMA algorithm. In order to better evaluate the performance of the QT algorithm, we do the following analysis of the CCMA algorithm. Similarly, we assume that there are n tags in the RFID system, and the reader uses the CCMA algorithm to recognize them. The total number of time slots required is

$$N_t = P_{cc} \times N_{cc} + P_{sc} \times N_{sc}, \tag{20}$$

where N_{cc} is the number of time slots required to identify n tags when all collisions are continuous collisions and N_{sc} is the number of time slots required to identify n tags when all collisions are noncontinuous. P_{cc} and P_{sc} are the probability of continuous collision and discontinuous collision, respectively.

Lemma 1. *Assuming that all collisions in the multitag recognition process are noncontinuous collisions, the number of time slots required by the CCMA algorithm to identify n tags is* $N_{ccma} = 2 \times n - 1$.

Proof. In the entire tag recognition process, when all the detected collisions are noncontinuous, the recognition process of the CCMA algorithm is similar to the CT algorithm, at this time, $N_{sc} = 2 \times N_c + 1 = 2 \times n - 1$. Thus, Lemma 1 is proved. □

Lemma 2. *In the entire tag recognition process, when all the detected collisions are continuous, the number of time slots used by the CCMA algorithm to recognize n tags is* $N_{cc} = (20 \times n - 11)/9$.

Proof. When all detected collisions are continuous, the recognition process of the CCMA algorithm is similar to a complete quaternary traversal tree. The collision node in the tree will produce 4 child nodes, so $N_4 = N_l + N_c = 4 \times N_c + 1$, where N_l and N_c denote the number of leaf nodes and collision nodes in the tree; we have

$$N_l = 3 \times N_c + 1. \tag{21}$$

In the quaternary traversal tree, the leaf nodes may be successful nodes or idle nodes, so the above formula can be rewritten as

$$n + N_i = 3 \times N_c + 1. \quad (22)$$

In the CCMA algorithm, the number of idle nodes that a continuous collision node may produce is between 0 and 2. The CCMA algorithm can eliminate idle time slots through the QueryP command. Next, we consider the following three scenarios, respectively.

Scenario A: when the number of idle nodes generated by a continuous collision node is 0, $N_i = 0$, and $N_c = (n-1)/3$ can be obtained from Equation (22). We substitute it into $N_4 = 4 \times N_c + 1$ to get $N_4 = (4 \times n - 1)/3$. According to the principle of the CCMA algorithm, the number of sending the custom command QueryP in the CCMA algorithm is the number of consecutive collisions (i.e., N_c), so $N_{cc} = N_4 - N_i + N_c = (5 \times n - 2)/3$.

Scenario B: when the number of idle nodes generated by a continuous collision node is 1, $N_i = N_c$, $N_c = (n-1)/2$ can be obtained from Equation (22), and substituting it into $N_4 = 4 \times N_c + 1$, we can find $N_4 = 2 \times n - 1$. Further, we can find $N_{cc} = 2 \times n - 1$.

Scenario C: when the number of idle nodes generated by a continuous collision node is 2, similarly, $N_{cc} = 3 \times n - 2$ can be obtained.

Since the above three scenarios appear with equal probability, we can get $N_{cc} = (20 \times n - 11)/9$; therefore, Lemma 2 is proved. □

Considering that in the process of tag recognition, the probability of continuous collision and discontinuous collision is 0.5, combining Lemmas 1 and 2, formula (20) can be rewritten as

$$N_t = \frac{1}{2} N_{sc} + \frac{1}{2} N_{cc} = \frac{19 \times n - 10}{9}. \quad (23)$$

Therefore, the maximum system throughput that the CCMA algorithm can achieve can be expressed as

$$U_{CCMA} = \lim_{n \longrightarrow \infty} \frac{9n}{19n - 10} = \frac{9}{19 - (10/n)} \approx 0.4736. \quad (24)$$

From Equation (24), we can know that the system throughput that the CCMA algorithm can maintain is slightly lower than that of the CT algorithm, but the amount of data transmission required is lower than that of the CT algorithm.

(3) DPPS algorithm

The performance analysis of the current mainstream QT anticollision algorithms shows that the recognition performance of the QT algorithm has a higher room for improvement. A typical idea is to optimize the response mechanism of the tag, and its representative is the DPPS algorithm. The basic principle of the DPPS algorithm is that the reader sends a query command to all tags within its coverage area, and when the tag receives this command, it will return its own complete ID. Once a collision is detected, the reader will update the query prefix based on the parsed data and send the PROBE_EQ command to probe the tag in the next time slot. There are 3 key parameters in the PROBE_EQ command. First is COM_Str, which represents the data part of the tag ID before the collision bit; second is Pre1, which represents prefix 1; and third is Pre2, which represents prefix 2. The value of Pre2 is equal to the value of Pre1 minus one. We can also make the following derivation to analyze the performance of the DPPS algorithm. In the DPPS algorithm, we use N_{slots} to represent the sum of the number of successful slots, complete collision slots, and identifiable collision slots. Let $E(N_{slots})$ be defined as the expected value of N_{slots}. According to the principle of the DPPS algorithm, $E(N_{slots})$ can be expressed as

$$E(N_{slots}) = 1 + N_c, \quad (25)$$

where N_c is the number of ID collisions during the recognition process.

Table 4: The experimental parameters used in this paper.

Parameter	Value
The number of tags	(100, 1000)
The length of IDs	96
Communication rate	40 kbps
T1	25 μs
T2	25 μs
T3	12.5 μs

Lemma 3. *Similarly, we assume that there are n tags to be identified in the system, and the reader uses the DPPS algorithm to identify them; then, the expected total number of time slots $E(N_{slots})$ is equal to n.*

Proof. The recognition process of the DPPS algorithm is similar to a variant of a binary traversal tree, and each node on the tree corresponds to a time slot. For the convenience of analysis, we treat both identifiable collisions and complete collisions as collision nodes. Obviously, a collision node will produce 2 child nodes, so the total number of nodes on the traversal tree is $2N_c + 1 = N_c + n$. It can be seen that $N_c = n - 1$. According to formula (25), we know that $E(N_{slots}) = 1 + n - 1 = n$. Therefore, Lemma 3 is proved. □

Above, we analyzed the performance of three typical QT algorithms. Through analysis, we know that the performance improvement of QT algorithms mainly depends on the query mechanism of the reader and the response mechanism of the tag. By optimizing these two mechanisms, the performance of tag recognition can be continuously improved. Through Lemma 3, we know that the DPPS algorithm only needs n time slots to identify n tags, and its system throughput reaches 100%. However, this does not mean that the performance of QT anticollision algorithms has reached the limit.

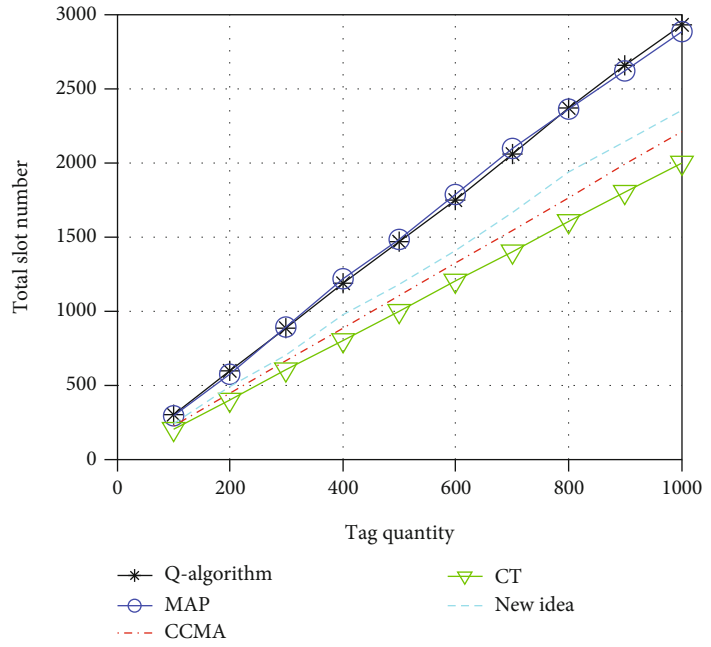


Figure 6: Experimental results: the total slot numbers by various algorithms.

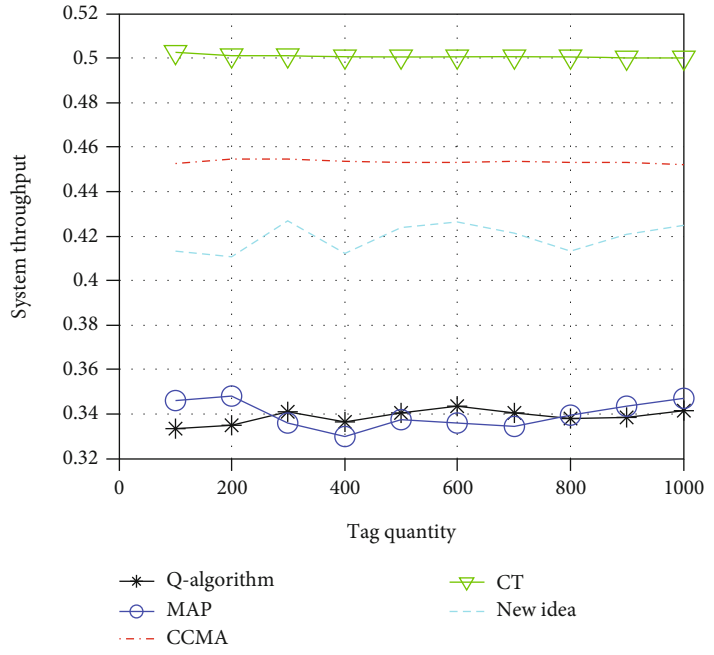


Figure 7: Experimental results: system throughput of various algorithms.

The reason is that in the DPPS algorithm, the length of a single time slot is longer than the length of the time slot in the traditional algorithm, so the actual improvement effect is not so high. Therefore, there is still much room for improvement in the performance of QT algorithms.

4. Experimental Results and Analysis

4.1. Experiment Setup. In the experiment section, we measure the performance of the RFID anticollision through some common metrics such as system throughput, total number of time slots, and recognition efficiency and compared with our analysis results to verify the effectiveness of our theoretical analysis. The simulations are performed on the desktop computer with Intel Core i5-4590 CPU and 8 GB RAM. All experiments are realized through MATLAB program. Since the performance analysis results of most DFSA algorithms are similar and close (below 0.368), we choose the more classic Q-algorithm and MAP algorithm as reference. In the QT-based algorithms, we choose the

CCMA and CT algorithm as the reference algorithms. The experimental parameters are listed in Table 4 [31–33]. In order to ensure the convergence of the results, we independently repeat the experiments 1000 times and then take the average value as the final result [34].

4.2. Result Analysis. Figure 6 compares the total number of time slots required by different algorithms to recognize the same quantity of tags. The comparative algorithms include Q-algorithm [9], MAP [19], CCMA [20], CT [29], and the new design concept proposed in this paper (named new idea in the simulations). It can be observed from the figure that when the QT algorithm recognizes the same number of tags, the number of time slots spent is significantly lower than that of the DFSA algorithm. Among the two QT algorithms, the CT algorithm spends less time slots than the CCMA algorithm, which is also consistent with the theoretical analysis results. Among three DFSA algorithms, the number of time slots consumed by the Q-algorithm and the MAP algorithm is very close, while the number of time slots consumed by the new idea is significantly lower than that by the other two DFSA algorithms.

Figure 7 depicts the system throughput that can be achieved under different algorithms. From the simulation results, we can further observe that the system throughput of QT algorithm is generally higher than that of DFSA algorithms. Through our previous theoretical analysis, we can see that the system throughput of most DFSA algorithms is lower than 0.368. This experimental result also further verifies our theoretical analysis. The throughput of the Q-algorithm and MAP algorithm is very close. Their average throughput is 0.3388 and 0.3398, respectively, which is 8.6% and 8.3% away from the theoretical maximum value. We can further observe that the average throughput that new idea algorithm can achieve is 0.4192 and 13.9% higher than the maximum throughput of the existing DFSA algorithms. The average throughput of the CT and CCMA algorithm is 0.5 and 0.453, respectively. The gap between the average throughput of the CCMA algorithm and the theoretical analysis is 4%. The reason is that the number of simulations is not set large enough.

4.3. Challenges and Future Trends. We have verified through theoretical analysis and experiments that many of the current anticollision algorithms can alleviate the problem of multitag recognition to a certain extent, but there are still some problems in the recognition efficiency, complexity, stability, etc., which cannot meet the large-scale RFID systems. In summary, there are following problems:

(1) The existing tag quantity estimation methods cannot be applied to low-cost RFID systems. Existing research on anticollision algorithms is excessively pursuing accuracy in the estimation of the tag quantity, which makes the complexity of the algorithm increase sharply. Taking into account the limitation of the computing power of the RFID platform, the high-complexity algorithm cannot be applied to the low-cost RFID systems

(2) Existing anticollision technologies are facing performance bottlenecks. Most anticollision algorithms are based on a channel with only one tag reply, so that the reader can decode it correctly. The collision information is discarded, and the tag needs to be retransmitted. This undoubtedly wastes a lot of useful information and restricts the further improvement of the recognition efficiency of the anticollision algorithm

5. Conclusions

This paper analyzes the current mainstream anticollision algorithms and conducts theoretical analysis and simulation on their performance. Based on the analysis of the reasons for the limited performance of the existing DFSA algorithm, a new DFSA algorithm architecture is proposed, and theoretical analysis and simulation results prove that it can break through the performance bottleneck of the current DFSA algorithm. Finally, we summarize the existing research on anticollision algorithms and present the current challenges and future research trends.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest to report regarding the present study.

References

[1] R. Jayadi, Y. C. Lai, and C. C. Lin, "Efficient time-oriented anticollision protocol for RFID tag identification," *Computer Communications*, vol. 112, pp. 141–153, 2017.

[2] J. Su, Z. Sheng, A. Liu, Z. Fu, and Y. Chen, "A time and energy saving-based frame adjustment strategy (TES-FAS) tag identification algorithm for UHF RFID systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 2974–2986, 2020.

[3] X. Liu, J. Zhang, S. Jiang et al., "Accurate localization of tagged objects using mobile RFID-augmented robots," *IEEE Transactions on Mobile Computing*, vol. 20, no. 4, pp. 1273–1284, 2021.

[4] Y. Ye, L. Shi, R. Qingyang Hu, and G. Lu, "Energy-efficient resource allocation for wirelessly powered backscatter communications," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1418–1422, 2019.

[5] L. Zhu and T. S. P. Yum, "A critical survey and analysis of RFID anti-collision mechanisms," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 214–221, 2011.

[6] D. Klair, Kwan-Wu Chin, and R. Raad, "A survey and tutorial of RFID anti-collision protocols," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 3, pp. 400–421, 2010.

[7] M. Benssalah, M. Djeddou, B. Dahou, K. Drouiche, and A. Maali, "A cooperative Bayesian and lower bound estimation in dynamic framed slotted ALOHA algorithm for RFID

systems," *International Journal of Communication Systems*, vol. 31, no. 13, pp. 1–13, 2018.

[8] A. Bekkali, S. Zou, A. Kadri, M. Crisp, and R. V. Penty, "Performance analysis of passive UHF RFID systems under cascaded fading channels and interference effects," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1421–1433, 2015.

[9] Z. He and H. Luo, "An efficient early frame breaking strategy for RFID tag identification in large-scale industrial Internet of Things," *Scientific Programming*, vol. 2021, 6 pages, 2021.

[10] J. Su, R. Xu, S. Yu, B. Wang, and J. Wang, "Idle slots skipped mechanism based tag identification algorithm with enhanced collision detection," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 5, pp. 2294–2309, 2020.

[11] C. N. Yang, L. J. Hu, and J. B. Lai, "Query tree algorithm for RFID tag with binary-coded decimal EPC," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1616–1619, 2012.

[12] H. Landaluce, A. Perallos, E. Onieva, L. Arjona, and L. Bengtsson, "An energy and identification time decreasing procedure for memoryless RFID tag anticollision protocols," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4234–4247, 2016.

[13] J. Su, Y. Chen, Z. Sheng, Z. Huang, and A. X. Liu, "From M-ary query to bit query: a new strategy for efficient large-scale RFID identification," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2381–2393, 2020.

[14] Y. C. Lai, L. Y. Hsiao, H. J. Chen, C. N. Lai, and J. W. Lin, "A novel query tree algorithm with bit tracking in RFID tag identification," *IEEE Transactions on Mobile Computing*, vol. 12, no. 10, pp. 2063–2075, 2012.

[15] J. Shin, B. Jeon, and D. Yang, "Multiple RFID tags identification with M-ary query tree scheme," *IEEE Communications Letters*, vol. 17, no. 3, pp. 604–607, 2013.

[16] T. F. la Porta, G. Maselli, and C. Petrioli, "Anticollision protocols for single-reader RFID systems: temporal analysis and optimization," *IEEE Transactions on Mobile Computing*, vol. 10, no. 2, pp. 267–279, 2011.

[17] H. Wu, Y. Zeng, J. Feng, and Y. Gu, "Binary tree slotted ALOHA for passive RFID tag anticollision," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 19–31, 2013.

[18] J. Su, Z. Sheng, A. X. Liu, and Y. Chen, "A partitioning approach to RFID identification," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2160–2173, 2020.

[19] W. T. Chen, "An accurate tag estimate method for improving the performance of an RFID anticollision algorithm based on dynamic frame length ALOHA," *IEEE Transactions on Automation Science and Engineering*, vol. 6, no. 1, pp. 9–15, 2009.

[20] J. Su, G. Wen, and J. Han, "An efficient RFID anti-collision algorithm for ISO 18000-6B algorithm," *Acta Electronica Sinica*, vol. 42, no. 12, pp. 2515–2519, 2014.

[21] K. W. Chiang, C. Hua, and T. S. P. Yum, "Prefix-randomized query-tree algorithm for RFID systems," in *Proceedings of IEEE International Conference on Communications*, pp. 1653–1657, Istanbul, Turkey, 2006.

[22] W. T. Chen, "Optimal frame length analysis and an efficient anti-collision algorithm with early adjustment of frame length for RFID systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3342–3348, 2016.

[23] W. Zhu, J. Cao, H. C. B. Chen, X. Liu, and V. Raychoudhury, "Mobile RFID with a high identification rate," *IEEE Transactions on Computers*, vol. 63, no. 7, pp. 1778–1792, 2014.

[24] Y. Chen, Q. Feng, X. Jia, and H. Chen, "Modeling and analyzing RFID generation-2 under unreliable channels," *Journal of Network and Computer Applications*, vol. 178, article 102937, 2021.

[25] L. Xie, B. Sheng, C. C. Tan, H. Han, and Q. Li, "Efficient tag identification in mobile RFID systems," in *29th IEEE Conference on Computer Communications (INFOCOM)*, pp. 1–9, San Diego, CA, USA, 2010.

[26] L. Yang, Y. Chen, X. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: real-time tracking of mobile RFID tags to high precision using COTS devices," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, pp. 237–248, Hawaii, US, 2014.

[27] L. Pan and H. Wu, "Smart trend-traversal: a low delay and energy tag arbitration algorithm for large RFID systems," in *28th IEEE Conference on Computer Communications (INFOCOM)*, pp. 2571–2575, Rio De Janeiro, Brazil, 2009.

[28] M. Shahzad and A. X. Liu, "Probabilistic optimal tree hopping for RFID identification," *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 796–809, 2015.

[29] X. Jia, Q. Feng, and L. Yu, "Stability analysis of an efficient anti-collision protocol for RFID tag identification," *IEEE Transactions on Communications*, vol. 60, no. 8, pp. 2285–2294, 2012.

[30] J. Su, Z. Sheng, G. Wen, and V. C. M. Leung, "A time efficient tag identification algorithm using dual prefix probe scheme (DPPS)," *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 386–389, 2016.

[31] J. Su, Z. Sheng, A. X. Liu, Y. Han, and Y. Chen, "Capture-aware identification of mobile RFID tags with unreliable channels," *IEEE Transactions on Mobile Computing*, pp. 1–14, 2020.

[32] J. Su, Z. Sheng, A. X. Liu, Z. Fu, and C. Huang, "An efficient missing tag identification approach in RFID collisions," *IEEE Transactions on Mobile Computing*, pp. 1–12, 2021.

[33] J. Su, Z. Sheng, L. Xie, G. Li, and A. X. Liu, "Fast splitting-based tag identification algorithm for anti-collision in UHF RFID system," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2527–2538, 2019.

[34] J. SU, X. ZHAO, D. HONG, Z. LUO, and H. CHEN, "Q-value fine-grained adjustment based RFID anti-collision algorithm," *IEICE Transactions on Communications*, vol. E99.B, no. 7, pp. 1593–1598, 2016.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 1508150, 6 pages
https://doi.org/10.1155/2021/1508150

Research Article

Research on News Text Classification Based on Deep Learning Convolutional Neural Network

Yunlong Zhu

Commission for Discipline Inspection, Zhengzhou Technical College, Zhengzhou 450121, China

Correspondence should be addressed to Yunlong Zhu; zhuyunlong123@zcmu.edu.cn

Received 2 September 2021; Revised 31 October 2021; Accepted 9 November 2021; Published 8 December 2021

Academic Editor: Yinghui Ye

Aiming at the problems of low classification accuracy and low efficiency of existing news text classification methods, a new method of news text classification based on deep learning convolutional neural network is proposed. Determine the weight of the news text data through the VSM (Viable System Model) vector space model, calculate the information gain of mutual information, and determine the characteristics of the news text data; on this basis, use the hash algorithm to encode the news text data to calculate any news. The spacing between the text data realizes the feature preprocessing of the news text data; this article analyzes the basic structure of the deep learning convolutional neural network, uses the convolutional layer in the convolutional neural network to determine the change value of the convolution kernel, trains the news text data, builds a news text classifier of deep learning convolutional neural network, and completes news text classification. The experimental results show that the deep learning convolutional neural network can improve the accuracy and speed of news text classification, which is feasible.

1. Introduction

With the rapid development of Internet technology, we are in the era of information explosion. While enjoying the convenience brought by rich online information, we are also facing the severe challenge of how to quickly and effectively extract data information from massive information. Therefore, people pay more and more attention to the research and analysis in the direction of information retrieval technology and data mining technology, which makes the research related to natural language processing develop rapidly. Among them, the main task of text classification technology, which plays a key role in massive text data processing, is to solve the problem of chaotic text information to a certain extent. The basic principle is to extract relevant text features from the original text content and finally judge the category label, which has attracted more and more attention from scholars. And then, there are many applications and research [1]. However, with the rise of big data, massive online Internet text information has a series of new features, such as more complex formats, more cumbersome types, faster update speed, and more difficult labeling. Especially after mobile phone users are popularized on a large scale, microblog, headline news, and other social life are enriched, and various short texts are also growing rapidly. All these changes will bring new challenges to text classification [2]. With the advent of the era of artificial intelligence, machine learning is one of the important disciplines in this field. In these decades of development, the research of machine learning-related algorithms has been quite mature, and a series of breakthroughs have been made in practical application. 2006 was the first year of deep learning. After its concept was proposed again, it quickly became the core research field of scholars all over the world. Up to now, as the core research object in the field of machine learning, deep learning has attracted extensive attention from contemporary Internet big data and artificial intelligence [3]. Since Google, Microsoft, IBM, Baidu, and other large Internet technology companies began to focus on the research and development of deep learning technology, it has made great breakthroughs in the fields of image, speech recognition, and natural language processing. Deep learning establishes a multilayer neural network structure by simulating the hierarchical structure of human brain, extracts the

distributed features of input data layer by layer from the bottom to the top, and finally establishes a good mapping function to describe the abstract relationship from the bottom signal to the high semantics [4]. Obviously, as an algorithm rising in the whole big data environment, deep learning will become one of the hot research directions in the future. Therefore, using artificial intelligence algorithm to classify news text has become a hot issue in this field.

Literature [5] proposed a short text classification method based on keyword similarity. Firstly, the word 2vec word vector model is obtained through a large number of corpus training. Then, the keywords of each type of text are obtained through textrank, and the de duplication operation is carried out in the keyword set as the feature set. For any feature, the similarity between each word in the short text and the feature is calculated through the word vector model. The maximum similarity is selected as the weight of the feature. Finally, k-nearest neighbor (KNN) and support vector machine (SVM) are selected as the classifier training algorithm. Based on the Chinese news title dataset, the classification effect is improved by about 6% on average compared with the traditional short text classification method, which verifies the effectiveness of the method. However, this method does not preprocess the text before text classification, resulting in more similar information, which affects the classification results. Reference [6] proposes an ACO-WNB classification algorithm based on improved information gain. Firstly, according to the word frequency distribution of feature words in the dataset, an adjustment factor is added to enhance/suppress the contribution/interference of feature words, select features with strong discrimination to form feature subsets, and improve the accuracy of Ig processing unbalanced datasets. Then, the ant colony optimization algorithm (ACO) is combined with the weighted naive Bayesian model, and ACO is used to iterate and globally optimize the weights to generate ACO-WNB classifier to improve the classification efficiency of text data. The improved algorithms are compared and analyzed with typical news datasets. The experiments show that Ig can effectively remove redundant high-frequency features and has better feature selection ability for unbalanced datasets. ACO-WNB classifier has higher accuracy and better classification efficiency for actual text data. The classification process of this method is complex and has some limitations. Literature [7] proposed CRNN text classification algorithm based on attention mechanism. Taking the pretrained word vector as the input, the convolutional neural network (CNN) is used to extract the features of the text vector; Bi gating loop unit (BI Gru) is used to capture the word order information in the text, extract the context dependency of the text, and identify the importance of different features combined with the attention mechanism. The highway network is used for feature optimization. The model is tested on three English corpora: 20 newsgroups, sst-1, and sst-2. The experimental results show that the model effectively improves the accuracy of classification tasks. However, this method has the problem of high noise in classification.

In view of the shortcomings of the above methods, this paper proposes a research on news text classification based on deep learning convolutional neural network. The weight of news text data is determined by VSM vector space model, and the information gain of mutual information is calculated to determine the characteristics of news text data. On this basis, the hash algorithm is used to encode the news text data, calculate the spacing between any news text data, and realize the feature preprocessing of news text data. This paper analyzes the basic structure of deep learning convolution neural network, determines the change value of convolution kernel with the help of convolution layer in convolution neural network, trains news text data, constructs news text classifier of deep learning convolution neural network, and completes news text classification. The technical route of this paper is as follows:

Step 1: determine the weight of news text data through VSM vector space model, calculate the information gain of mutual information, and determine the characteristics of news text data.

Step 2: the hash algorithm is used to encode the news text data, calculate the spacing between any news text data, and realize the feature preprocessing of news text data.

Step 3: analyze the basic structure of deep learning convolution neural network, determine the change value of convolution kernel with the help of convolution layer in convolution neural network, train news text data, construct news text classifier of deep learning convolution neural network, and complete news text classification.

Step 4: experimental analysis

Step 5: conclusion

2. Classification of News Text Based on Deep Learning Convolutional Neural Networks

2.1. News Text Data Feature Extraction. Due to the complex format of news text data directly acquired by the Internet, computers cannot directly understand the text content, so the representation of news text characteristics is to transform the unstructured or structured news text of the Internet into structured text intelligible by the computer. In this paper, we first characterize its feature [8] with the help of the VSM vector space model. The model is based on the number of feature words appear statistics, mainly including three steps: first, calculate the word frequency, then calculate the inverse document frequency, and finally calculate the TF-IDF. The model is based on the feature representation of the word vector, if the reference source is not found in error. In presenting the weight size of each feature item in each text set, each text can be measured by the reference source [9]. If a feature word corresponds to a word vector, all feature words in all text sets correspond to the corresponding dimension of the space. The weights are calculated here using the TF-IDF method based on word frequency, the number of single appearances, giving a large weight [10] to persuasive features in each document. The TF-IDF weights are calculated as follows:

$$w_{ij} = tf_{ij} \times idf_{ij} = tf_{ij} \times \log\left(\frac{N}{n_j}\right). \tag{1}$$

Among them, w_{ij} represents feature terms, tf_{ij} represents the number of news text appearances, idf_{ij} represents the frequency of news text occurrence, $\log(N/n_j)$ represents the derivative of text occurrence, and N represents the total number of news text.

After the feature weights of the calculated news text data, the entropy changes of the news text data. Information gain is an entropy-based method. First, calculate the change of information entropy when the feature item appears, that is, the information gain, and then select the information gain according to the size value to measure the importance of the information of the final classification; feature importance is proportional to the amount of information, that is, the more the feature carries information represents the greater the word feature importance [11]. The following formula is the calculation method of the information gain:

$$g(t) = -\sum_{i=1}^{m} P(C_i) lgP(C_i) + P(t) \sum_{i=1}^{m} (C_i), \tag{2}$$

where $P(C_i)$ represents the proportion or probability of C category documents in all sets of documents, $P(t)$ represents the probability of having feature items in a document, and m is the number of document categories.

Mutual information is a variable that reflects the correlation between two variables. It works as follows: mutual information size represents the degree of correlation between characteristics and category, and the larger the association, the closer the former. Mutual information metric method is as follows:

$$M(t, C) = \text{Log} \frac{P(t \pm c)}{p(t) \times p(c)}. \tag{3}$$

Among them, t represents the feature term and C represents the category.

According to the calculation of the mutual information of the above news text, the extraction of the data characteristics of the news text is completed, i.e.,

$$\varphi_j = \arg \min \sum_{i=1} \vartheta_I + \gamma |Y|^2. \tag{4}$$

In the formula, γ represents the proportional coefficient of news text feature data characteristics and the key value of news text features [16].

2.2. News Text Data Feature Preprocessing. To realize the effective classification of news text, this paper introduces the deep hash algorithm to effectively preprocess the above news text features. The deep hash algorithm can effectively handle the characteristics of news text. When the characteristics of news text differ, the algorithm of [12] by removing Hamming distance has the advantage of simple operation and high outlet efficiency [17].

The basic idea of hash function is that in the original news text feature data, the feature points are relatively close, and the impact of these two points occurs more frequently, when setting the set of news text features to be processed as follows:

$$H = \{h_i\}_{i=1}^{n}. \tag{5}$$

Set the distance metric function for two points in the news text feature to the following:

$$\vartheta_i = h_i \sum y_p. \tag{6}$$

According to the distance between the two points obtained above, since the hash algorithm has certain unidirectional, it is irreversible, i.e.,

$$\sigma \longrightarrow R(\sigma). \tag{7}$$

In the formula, $R(\sigma)$ represents the output after the hash and σ represents the original news text.

The deep hash algorithm is considered as a density function [13] when dealing with noise in features in news text, and each density function corresponds to features of multiple news text, i.e.,

$$\text{size}(R(\sigma)) < \text{size}(\sigma). \tag{8}$$

According to the features of the corresponding news text, [14] is converted to obtain the following:

$$E = \{R_I(\sigma)\} > 0. \tag{9}$$

Finally, the preprocessing of the news text feature data based on the deep hashing algorithm is implemented, i.e.,

$$V_I = \frac{1}{2} S_I Y(h). \tag{10}$$

In the formula, V_I represents the preprocessing results of news text feature data and $Y(h)$ represents the weight value of preprocessing [18].

2.3. Classification of News Text Based on Deep Learning Convolutional Neural Networks. Deep learning convolutional neural network is an AI learning algorithm based on the neural network topology and optimizes the convolutional layer in neural networks into a convolutional kernel and simultaneously completes the propagation [15] in different directions in signal processing. The technical core of the method is in the transformation of the convolutional wave, through which detailed information of the data can be obtained in any case. Its expansion structure is shown in Figure 1.

In the study of this paper, the change of convolutional kernels in the convolutional layer completes the classification of news text. The convolutional layer structure is shown in Figure 2.

In Figure 2, the $a_1, a_2 \cdots a_n$ represents the feature data that the structure needs to be entered, and n is the number of feature data; $g_1, g_1 \cdots, g_n$ represents the number of convolutional kernels, b and c represent different weight values

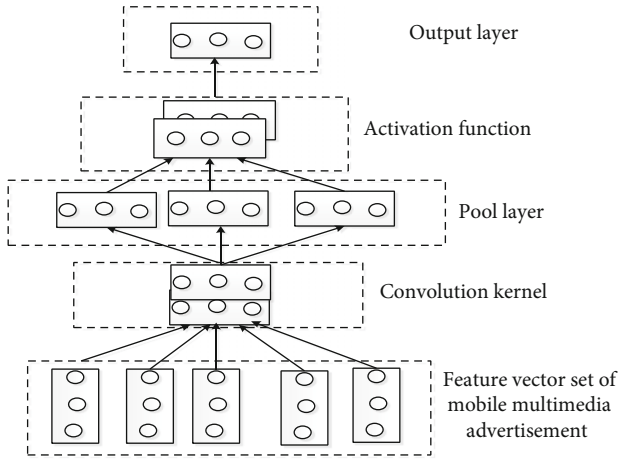


Figure 1: Topological diagram of the convolutional neural network.

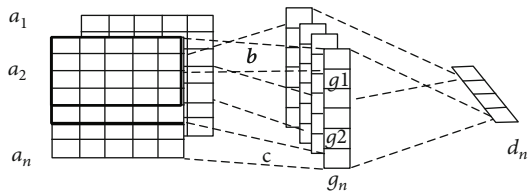


Figure 2: Convolutional layer structure.

of the structure, and $d_1, d_2 \cdots d_n$ represents the amount of the final output of this topology [19].

In a convolutional neural network, the $a_1, a_2 \cdots a_n$ news text feature data is input into it, and the processed data features are as follows:

$$Q = \left(\left(\frac{\sum_{a=1}^{Nn} ba_i - c_j}{a_i} \right) \right). \tag{11}$$

In the formula, c_j represents the translation factor of the convolutional kernel number, a_i represents the scaling factor, ba_i represents the weight value after the feature data input, and ε represents the excitation function [19].

Calculate the output values obtained from the above feature data to obtain the final training sample data, i.e.,

$$T_k = \sum_{i=1}^{n} \mu_{ik} Q. \tag{12}$$

In formula (12), the μ_{ik} represents the weight value of the convolutional kernel.

In this network, because it is prone to certain errors in the forward propagation, the forward propagation in the neural network is corrected to achieve a more accurate classification of news text. Therefore, the error correction is accomplished by using the gradient descent method.

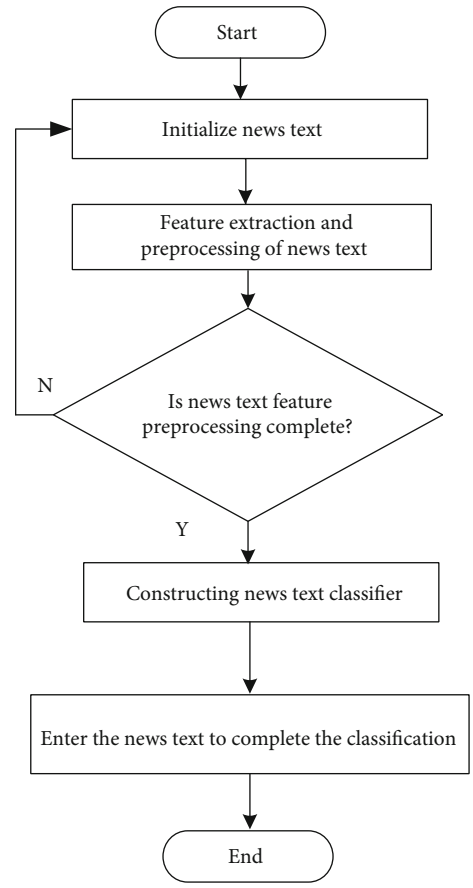


Figure 3: News text classification process based on deep learning convolutional neural networks.

Table 1: Experimental dataset.

Dataset	C	SN	N	$\lvert V \rvert$	Test
DBMC-1	3	4	7056	20852	0.25
DBMC-2	3	6	3578	19874	0.35
DBMC-3	3	8	2258	16874	0.35
DBMC-4	3	10	2451	21457	0.25

The error W of the forward propagation of the first extracted convolutional neural network is as follows:

$$W = \sum_{i=1}^{n} (x_i - y_i). \tag{13}$$

In formula (13), x_i is the ideal output of the K feature data and y_i is the actual output value of the K feature data.

In the error values obtained above, their weight values and scaling factors are corrected according to the gradient descent method, i.e.,

$$\begin{aligned} b &= b_j^i + \Delta b_j^{i+1}, \\ \Delta \mu_j^{i+1} &= -\tau \frac{\partial W}{\partial \mu_j^1}. \end{aligned} \tag{14}$$

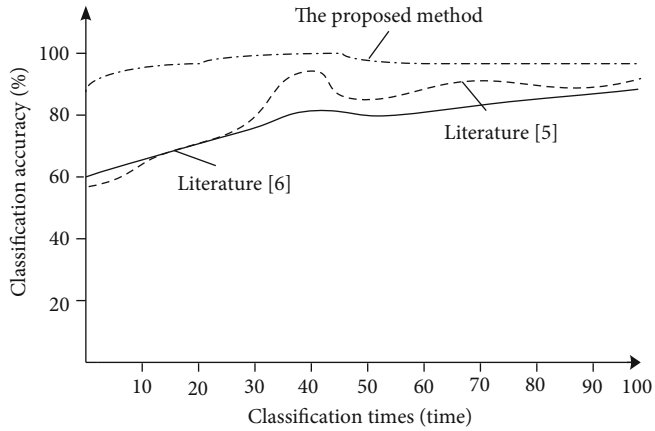


Figure 4: Comparison of news text classification accuracy for different methods.

In the formula, τ represents the learning efficiency values [20].

From the above analysis, a convolutional neural network classifier constructs the data trained on news text, i.e.,

$$F_i = \begin{cases} \dfrac{\sum_{i=1}^{n} u_i}{dg_i} \\ 1 \end{cases}. \tag{15}$$

The classification is completed according to the news text classifier obtained by formula (15), and the classification process of the method is shown in Figure 3.

3. Experimental Analysis

3.1. Design of Experimental Scheme. Experimental environment of this paper: the operating system is Ubuntu14.04. The processor is the Intel Core i5. The dataset used by the development tool for Pycharm community edition 3.4 experiments was collected from recent hot news content from Sina Weibo and manually annotated. The four datasets were DBMC-1, DBMC-2, DBMC-3, and DBMC-4. Among them, C represents the number of categories of comment content. Here is the positive and negativity of the comments' content. The SN represents the number of sentences per news text. The average length of each sentence is 7. The N represents the size of the dataset. The $|V|$ represents the size of the forming dictionary. Test represents the proportion of the test set to the dataset. Details are shown in Table 1.

The convolutional neural network structure used in this experiment mainly consists of 1 word embedding layer, 1 convolutional layer, and 1 pooling layer, the dimension of the word vector is set to 128, the size of the convolutional kernel window is set to 3×128,4×128,5×128, etc., and the number of convolutional nuclei is 128.

3.2. Design of Experimental Indicators. Verify the effectiveness and accuracy of the classification method. Among them, the classification accuracy is a percentage, the higher the value represents the higher the efficiency of classification, about low classification time consuming, the better the classification efficiency.

Table 2: Time-consuming analysis of sample news text classification (s).

Number of iterations/times	Methods of this paper	Document [5] methods	Document [6] methods
20	2.4	4.9	3.7
40	2.4	4.0	3.5
60	2.5	4.9	3.2
80	2.4	4.8	3.6
100	2.4	4.7	3.2

3.3. Analysis of Experimental Results. The proposed method, the literature [5] method, and the results are shown in Figure 4.

From the data in Figure 4, the proposed method, literature [5] method, and literature [6] method classified more than 90%, while the other two methods showed the rise, which is due to the spacing between any news text data; code any news text data, train the news text classifier of deep learning convolutional neural network, and complete the classification of news text.

To further verify the effectiveness of this method, the time-consuming classification of the proposed method, the literature [5] method, and the literature [6] method are compared, and the results are shown in Table 2.

Analyzing the data of experimental results in Table 2, we can see that with the number of iterations, the time-consuming method of sample news text classification is different with the proposed method, literature [5] method, and literature [6] method. When the number of iterations is 20, the proposed method has a time-consuming N. Text classification of sample news is about 2.4 s. The literature [5] method takes about 4.9 s, of sample news text classification. The literature [6] method takes about 3.7 s, for the classification of sample news text, When the number of iterations is 60, the proposed method takes about 2.5 s. The literature [5] method takes about 4.9 s, of sample news text classification. The literature [6] method takes about 3.6 s. As shown by the comparison of the experimental data, the classification of this paper is shorter and has faster speed. This is due to analyzing the basic structure of deep learning

convolutional neural network, determining the change value of its convolutional layers in the convolutional neural network, training news text data, building a news text classifier of deep learning convolutional neural networks, and completing the classification of news text.

4. Conclusion

In order to improve the classification accuracy of news text, a new classification method based on deep learning convolutional neural network is proposed. Determine the weight of the news text data through the VSM vector space model, calculate the information gain of mutual information, determine the characteristic news text data of the news text data, calculate the distance between any news text data, and analyze the basic structure of the deep learning convolutional neural network. Determine the change value of the convolution kernel, train the news text data, and build a news text classifier. The experimental results show that the deep learning convolutional neural network can improve the accuracy of news text classification and increase the classification speed.

Data Availability

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

Conflicts of Interest

The author declares that he has no conflict of interest.

References

[1] R. Asgarnezhad, M. Soltanaghaei, and S. A. Monadjemi, "An application of MOGW optimization for feature selection in text classification," *The Journal of Supercomputing*, vol. 16, no. 15, pp. 154–160, 2020.

[2] A. Esuli, A. Moreo, and F. Sebastiani, "Funnelling: a new ensemble method for heterogeneous transfer learning and its application to cross-lingual text classification," *ACM Transactions on Information Systems*, vol. 37, no. 3, pp. 1–30, 2019.

[3] M. Oleynik, A. Kugic, Z. Kasáč, and M. Kreuzthaler, "Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification," *Journal of the American Medical Informatics Association*, vol. 26, no. 11, pp. 1247–1254, 2019.

[4] M. A. Ibrahim, M. U. Ghani Khan, F. Mehmood, M. N. Asim, and W. Mahmood, "GHS-NET a generic hybridized shallow neural network for multi-label biomedical text classification," *Journal of Biomedical Informatics*, vol. 116, no. 1, article 103699, 2021.

[5] Z. Zhenhao, G. Yi, H. Meiqi, and W. Jixiang, "Research on short text classification based on keyword similarity," *Application Research of Computers*, vol. 37, no. 1, pp. 26–29, 2020.

[6] N.-j. Qiu, P. Gao, P. Wang, and Y. Tao, "Research on ACO-WNB classification algorithm based on improved information gain," *Computer Simulation*, vol. 36, no. 1, pp. 295–299, 2019.

[7] R. Chen, C.-g. Ren, Z.-y. Wang, Z.-j. Qu, and H.-p. Wang, "Attention based CRNN for text classification," *Computer Engineering and Design*, vol. 40, no. 11, pp. 3151–3157, 2019.

[8] Z. Chen and J. Ren, "Multi-label text classification with latent word-wise label information," *Applied Intelligence*, vol. 14, no. 9, pp. 1–14, 2020.

[9] A. Gcs and B. Ml, "Stacked DeBERT: all attention in incomplete data for text classification," *Neural Networks*, vol. 136, no. 1, pp. 87–96, 2021.

[10] J. Krishna, H. Purohit, and H. Rangwala, "Diversity-based generalization for neural unsupervised text classification under domain shift," *ECML-PKDD*, vol. 15, no. 24, pp. 1145–1152, 2020.

[11] B. Cowley, M. Filetti, and K. Lukander, "The psychophysiology primer: a guide to methods and a broad review with a focus on human–computer interaction," *Foundations and Trends in Human–Computer Interaction*, vol. 9, no. 3, pp. 151–308, 2016.

[12] I. Baldini, D. Wei, K. N. Ramamurthy, M. Yurochkin, and M. Singh, "Your fairness may vary: group fairness of pretrained language models in toxic text classification," *International Conference on Information Technology in Medicine & Education*, vol. 45, no. 1, pp. 15–21, 2021.

[13] J. Jang, Y. Kim, K. Choi, and S. Suh, "Sequential Targeting: an incremental learning approach for data imbalance in text classification," *International Conference on Information Technology in Medicine and Education*, vol. 45, no. 1, pp. 147–154, 2020.

[14] I. Chalkidis, M. Fergadiotis, S. Kotitsas, P. Malakasiotis, N. Aletras, and I. Androutsopoulos, "An empirical study on large-scale multi-label text classification including few and zero-shot labels," *International Conference on Information Technology in Medicine and Education*, vol. 45, no. 2, pp. 123–128, 2020.

[15] T. Igamberdiev and I. Habernal, "Privacy-preserving graph convolutional networks for text classification," *International Conference on Information Technology in Medicine and Education*, vol. 7, no. 1, pp. 868–871, 2021.

[16] Y. Xu, L. Gui, and T. Xie, "Intelligent recognition method of turning tool wear state based on information fusion technology and BP neural network," *Shock and Vibration*, vol. 2021, no. 8, Article ID 7610884, p. 10, 2021.

[17] C. Y. Peng, U. Raihany, S. W. Kuo, and Y. Z. Chen, "Sound detection monitoring tool in CNC milling sounds by K-means clustering algorithm," *Sensors*, vol. 21, no. 13, 2021.

[18] D. C. Mackintosh-Franklin, "An evaluation of formative feedback and its impact on undergraduate student nurse academic achievement," *Nurse Education in Practice*, vol. 50, no. 4, article 102930, 2021.

[19] Y. Chen, X. Wang, and X. Du, "Diagnostic evaluation model of English learning based on machine learning," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 2, pp. 2169–2179, 2021.

[20] X. Liu, "Feature recognition of English based on deep belief neural network and big data analysis," *Computational Intelligence and Neuroscience*, vol. 2021, no. 6, Article ID 5609885, p. 10, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9806161, 1 page
https://doi.org/10.1155/2023/9806161

Retraction

Retracted: Deep Learning and Collaborative Filtering-Based Methods for Students' Performance Prediction and Course Recommendation

Wireless Communications and Mobile Computing

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Liu, C. Yin, Y. Li, H. Sun, and H. Zhou, "Deep Learning and Collaborative Filtering-Based Methods for Students' Performance Prediction and Course Recommendation," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2157343, 13 pages, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 2157343, 13 pages
https://doi.org/10.1155/2021/2157343

Research Article

Deep Learning and Collaborative Filtering-Based Methods for Students' Performance Prediction and Course Recommendation

Jinyang Liu,[1] **Chuantao Yin**,[2] **Yuhang Li**,[2] **Honglu Sun**,[2] **and Hong Zhou**[1]

[1]*School of Economics and Management, Beihang University, Beijing 100191, China*
[2]*Sino-French Engineer School, Beihang University, Beijing 100191, China*

Correspondence should be addressed to Chuantao Yin; chuantao.yin@buaa.edu.cn

Received 9 September 2021; Revised 25 October 2021; Accepted 3 November 2021; Published 2 December 2021

Academic Editor: Yinghui Ye

At the beginning of a new semester, due to the limited understanding of the new courses, it is difficult for students to make predictive choices about the courses of the current semester. In order to help students solve this problem, this paper proposed a hybrid prediction model based on deep learning and collaborative filtering. The proposed model can automatically generate personalized suggestions about courses in the next semester to assist students in course selection. The two important tasks of this study are course recommendation and student ranking prediction. First, we use a user-based collaborative filtering model to give a list of recommended courses by calculating the similarity between users. Then, for the courses in the list, we use a hybrid prediction model to predict the student's performance in each course, that is, ranking prediction. Finally, we will give a list of courses that the student is good at or not good at according to the predicted ranking of the courses. Our method is evaluated on students' data from two departments of our university. Through experiments, we compared the hybrid prediction model with other nonhybrid models and confirmed the good effect of our model. By using our model, students can refer to the different recommendation lists given and choose courses that they may be interested in and good at. The proposed method can be widely applied in Internet of Things and industrial vocational learning systems.

1. Introduction

The Internet of Things (IoT) is a huge network formed by combining all kinds of information sensing devices and networks to realize the interconnection of people, machines, and things anytime and anywhere. The rapid development of IoT has brought massive data support to machine learning, and the combination of IoT and deep learning methods will be the general trend in the future.

So far, there has been a lot of research on this. Huang et al. used a deep learning algorithm instead of manually monitoring the wearing of a safety helmet onsite [1]. Jiang et al. obtained semantic information of the scene by using the improved Faster-RCNN model [2]. Liao et al. used the improved SSD to carry out occlusion gesture recognition, realizing the interaction between machine and nature [3]. Gao et al. distinguish human left and right hands through deep convolution and feature extraction and also realize hand positioning and detection [4, 5]. In addition to these, the combination of IoT and deep learning can also help improve the efficiency of education systems. Mobile devices can collect data of students, and deep learning methods can be used to predict and explain students' progress and achievements. Deep learning can also be used for personalized recommendation modules to recommend more relevant content to educators. In this paper, we have carried out a thorough and detailed study on the last point.

In higher education, students will have many courses, including the required courses arranged by the university or the department and the elective courses that students can choose based on their needs. Reasonable choices of elective courses and being well-prepared for the courses of the coming semester can help a student to learn more and have better results. When choosing the elective courses, the university or the department will provide many choices for students, but

before studying these courses, students' understanding of these courses is limited, so it is hard for them to decide by themselves which courses are suitable for them. Our method combines deep learning methods and traditional methods to predict students' performances and interests based on history data. This method will provide each student several lists of courses which include the list of elective courses which match students' interests and which the student might be good at, the list of elective courses which match the students' interests and which the student might not be good at, and the lists of required courses which the student might be good at and which the student might not be good at.

Student performance prediction is always a major concern in the education domain. Many machine learning algorithms have been applied to predict students' performance in previous studies, like support vector machine [6], decision tree [7], linear regression [8], and random forest [9]. With the development of deep learning, many deep learning algorithms have achieved better performances than traditional machine learning algorithms on many different domains. Two domains that are most influenced by deep learning methods are computer vision [10] and natural language processing [11]. Basing on a large amount of image data and their associated labels, convolutional neural networks [12] can be trained to extract meaningful feature vectors from images, and these feature vectors can be further used for many different tasks, like classification [13] and object detection [14]. Basing on a large amount of text data, by deep learning algorithms, we can train a word embedding model, which projects each word into a vector in the latent space. The distance between different vectors in this latent space measures the semantic similarity between words. In both of these two domains, by using deep learning algorithms, the original data is transformed into latent representation (original input is projected to a vector in the latent space), which can be further used for other missions. So, it is important to study the application of deep learning algorithms in the education domain to see whether we can obtain the latent vectors of students and the latent vectors of courses basing on history data, which can measure the semantic similarity between students and between courses, respectively.

Experiments with different train/test data split modes show that the Course2Student algorithm can improve the accuracy of students' ranking prediction. Another reason for which we propose this new algorithm is to improve the interpretability of the model. When using models like neural collaborative filtering [15], each student and each course are associated with a latent vector, but the physical meanings of values in the latent vector of student and the latent vector, of course, are hard to explain. On the contrary, by using the Course2Student algorithm, even though the exact meanings of the obtained latent vectors are still hard to know, we can know the contribution of each associated course when predicting the ranking of a student on a new course, which can make us have more understanding of how the prediction result is made. On the other hand, some previous studies about student performance prediction only concentrate on one or several related courses [16], and the inputs of models are usually fixed, so to apply these methods to our scenario, multiple models with different inputs must be trained [17]. Moreover, our Course2Student is more flexible, and predictions can always be made no matter which courses a student has learned. And we only need one course embedding model for the prediction of all courses.

This study also compares the Course2Student algorithm with the nonparametric algorithm: user-based collaborative filtering [18]. To compare these algorithms, we also use different train/test data split modes. Experiment results show that Course2Student is better than user-based collaborative filtering on these data. To further improve the accuracy and reliability of the prediction result, we use the hybrid prediction method by combining the prediction result of Course2Student and the prediction result of user-based collaborative filtering [19]. Experiment results show that the hybrid prediction method can achieve higher accuracy than the single prediction method by selecting the prediction results with high confidence.

Most of course recommendation algorithms in the previous studies are commonly used recommendation algorithms, which can also be used in other domains like movie recommendation and product recommendation, and which are mainly based on collaborative filtering methods [20]. For the ranking prediction problem, Liu et al. proposed an improved probabilistic latent semantic analysis model (PLSA) [21] and a KNN-based optimal acceptance loss function Eigenrank [22]. Markus et al. proposed a model based on CofiRank-maximum margin matrix factorization (MMMF) technique [23]. Yue et al. proposed a model xCLiMF to optimize the ranking learning evaluation index MRR [24].

The main idea of collaborative filtering is to find the similarities between users or items and the relations between users and items, basing on the interaction history between users and items. These similarities and relations mainly describe the user's interests and the item's properties. In the course recommendation task, people should consider whether the course will match the student's interests and consider whether the student will be good at this course. In our course recommendation algorithm, we first select some courses that match students' interests by using a user-based collaborative filtering algorithm. Using the ranking prediction algorithm discussed above, the final recommendation lists are obtained, which not only consider students' interests but also consider their predicted performances.

The rest of this paper is organized as follows: Section 2 discusses related work, followed by the models for energy-efficient optimization and makespan optimization designed in Section 3. The improved clonal selection algorithm for resource allocation is discussed in Section 4. Section 5 shows the simulation experimental results, and Section 6 concludes the paper with summary and future research directions.

2. Related Works

This section will talk about some existing works that are most related to our works, including neural networks [25], collaborative filtering, and neural network-based collaborative filtering [26].

2.1. Neural Networks. Neural networks are firstly inspired by neural science, and their initial objective is to simulate how

information transfers in the human brain. Neural networks consist of many connections and nodes. Information runs through these connections and nodes. In fact, most of the neural networks used in today's field of research can be regarded as a collection of many linear functions and nonlinear activation functions; for example, logistic regression can be regarded as an example of the simplest neural networks, which consists of one linear function and one nonlinear activation function (sigmoid).

Neural networks can also be regarded as a collection of many layers. These layers can be classified into three groups: input layer, hidden layer, and output layer. Information transfers through neural networks with a fixed direction. Neural networks can be used as an automatic solution for many tasks. Taking the classification problem as an example, after the original image being fed into neural networks, based on the characteristic of the input image, different neurons will be activated in different layers; outputs of deeper layers have more information for the classification mission. Normally, the number of neurons in the output layer is the same as the number of categories. The output value of each neuron in the output layer describes the predicted probability of the associated category.

In order to obtain a neural network for a specific task, we need to use data to train the neural network. When training a neural network, we need to provide the input and the output of the neural network, and the weights are optimized by the back propagation algorithm [27]. In fact, the main objective of a neural network is to simulate a function. Normally, this function is complicated, and it cannot be constructed by human analysis. However, it can be simulated by neural networks when given the input and output data.

Some previous works have used neural networks to predict student performance. In a very early study [28], neural networks are used to predict academic success in MBA programs. In this work, neural networks are compared with four other prediction methods: least square regression [29], stepwise regression, discriminate analysis [30], and logistic regression [31]. The object of this study is to help make the decision to accept students into the MBA program. In [32], an intelligent tutoring system based on neural networks is proposed. In order to provide an appropriate problem for a student, a neural network is trained at first to predict the number of errors that the student might make on a certain set of problems. Then, based on the prediction result, a suitable problem is decided for the student. Lykourentzou et al. used the students' prediction problem in an e-learning scenario [33]. The final grades are estimated based on the data collected before the middle of the course by a neural network-based model. Then, based on the predicted level of performance, students are clustered into two groups, and each group will be provided with suitable educational materials. One similar work [34] uses a neural network-based model to learn the interaction between students and courses to predict student performance.

2.2. Collaborative Filtering. The collaborative filtering algorithm [20] is the most used algorithm in the studies of recommendation system, and it has already been used in some real-life applications [35]. It was first introduced to recommend electronic documents to users [36]. The data, on which the collaborative filtering algorithm can be applied, can normally be represented by an interaction matrix which describes the interaction between the users and the items. In this interaction matrix, each row presents a user and each column presents an item; the values in this matrix presents the interaction between the related user and the related item. This interaction matrix is always sparse because many interactions between users and items are unknown. The main task of the collaborative filtering algorithm is to predict the unknown interactions basing on the existing interactions and, basing on the prediction results, make recommendations for users. In fact, the collaborative filtering algorithm is a collection of many algorithms. Most of today's collaborative filtering algorithms can be classified into two groups: similarity-based algorithm and latent factor algorithm.

There are two kinds of similarity-based algorithms: item-based collaborative filtering [20] and user-based collaborative filtering [37]. A similarity-based algorithm is a very intuitive algorithm, and it is mainly based on the similarities between users or the similarities between items. The similarities between users can be obtained by calculating the similarities between rows in the interaction matrix, since each row presents the interaction between the current user and all the items. Similarly, the similarities between items can be obtained by calculating the similarities between columns in the interaction matrix. A user-based collaborative filtering algorithm can be presented by

$$v_{a,j} = \sum_{i \in \text{neighbors}(a)} w_i v_{i,j}. \tag{1}$$

Here, we want to predict the interaction between user a and item j. The prediction is based on the interactions between item j and some users similar to user a. In this equation, w_i is proportional to the similarity between user I and current user a, which means that the users who are more similar to current user a will have more contribution on the prediction result.

The latent factor algorithm is an improvement of the content-based algorithm [38]. In the latent factor algorithm, we suppose that each user and each item have a latent representation. These latent representations can be used to describe the interaction between users and items. The recommendation is based on the similarity between the user's latent vector and the item's latent vector. Different from a content-based algorithm where the features of users and items are decided by human experts, in the latent factor algorithm, the latent vectors depend on history interactions between users and items and the interaction function, which measures the similarity between two latent vectors. One example of a latent factor algorithm is the matrix factorization method [35], as shown in equation (2). Here, the interaction between a user and an item is represented by the inner product of the user's latent vector and the item's latent vector. Many studies have focused on how to improve the latent representation and how to improve the interaction function:

$$v_{a,j} = v_a^T v_j, \tag{2}$$

where $v_{a,j}$ is the interaction between a user and an item, v_a the user's latent vector and v_j the item's latent vector.

Some previous works have applied a collaborative filtering algorithm to the course recommendation problem. In [39], collaborative filtering is combined with the Artificial Immune System. Students are first placed into several clusters using the Artificial Immune System clustering approach to calculate the affinities between different students in a training data pool. Then, collaborative filtering is applied to the data cluster to predict the rating for the course. In [40], collaborative filtering is combined with students' online learning style. Basing on their online learning styles, students are first clustered by the k-means algorithm. Then, item-based collaborative filtering algorithms and user-based collaborative filtering algorithms are applied to each cluster. In [41], a matrix factorization-based method is proposed to predict students' feedback ratings on courses. This work targets three problems: potential lack of rating data from students to courses, imbalance of the user-item matrix, and dependencies between courses.

There are also some works which use the collaborative filtering algorithm for students' performance prediction. In [42], both user-based collaborative filtering algorithms and item-based collaborative algorithms predict students' grades on elective courses. The objective of this work is to recommend elective courses for each student on which the student might have higher grades. Their experiments prove that the performances of user-based collaborative filtering algorithms and item-based collaborative filtering algorithms are similar in their data. The idea of this study is similar to the idea of our work. In a more recent work [43], a novel cross-user-domain collaborative filtering algorithm is designed to accurately predict the score of the optional course for each student by using the course score distribution of the most similar senior students and recommend the top t optional courses with the highest scores without time conflict. The difference is that, in our work, we consider not only the predicted performances of the student but also their interests in these courses. Based on this work's results, we also use a user-based collaborative filtering algorithm as one of our baseline methods for student performance prediction.

2.3. Neural Network-Based Collaborative Filtering. In fact, the neural network-based collaborative filtering algorithm [22] is a kind of latent factor algorithm, and it is also a popular direction of research in recent years [44]. The advantages of the neural network-based collaborative filtering algorithm is that its interaction function is based on neural networks which can be learned from the data, while in the previous algorithms, the interaction function is decided by a human, so if the chosen interaction function is not suitable for the current dataset, then the algorithm's performance will decrease. One of the most important neural network-based collaborative filtering works is neural collaborative filtering [22]. In this algorithm, each user and each item have two latent vectors. When calculating the interaction between the user's latent vectors and the item's latent vectors, the first result is obtained by the element-wise product between the first latent vector of the user and the first latent vector of the item. To get the second result, the second latent vector of the user and the second latent vector of the item are concatenated and used as input of a neural network, and the output of this neural network gives us the second result. The first and the second results are then concatenated, and another neural network is applied to this concatenated vector to get the final prediction of the interaction between the user and the item. The reason for using two latent vectors is that in this way, the interaction function can have both linear and nonlinear parts.

Few works have used neural collaborative filtering to predict the interactions between students and courses. In most recent works, Sun et al. used a multitask learning strategy to improve the neural collaborative filtering method and use it to predict student performance, and [45] used the instructor's identity as another input of the neural collaborative filtering model. The main idea of all these recent similar works is to use a neural network to the latent factor collaborative filtering method. Different from these works, in our work, we combine neural collaborative filtering with traditional similar-based collaborative filtering methods in order to make the prediction process more reasonable. And since neural collaborative filtering is the most used one in the previous works, we use it as one baseline method.

3. Course Recommendation Method

Our overall model is shown in Figure 1. It is mainly divided into two parts, namely, course recommendation and student ranking prediction. In the third and fourth sections, we will introduce the model design of these two parts, respectively.

The user-based collaborative filtering method is used to predict the elective courses which might match students' interests. The final recommendation list is a combination of the prediction of students' interests and the prediction of students' performance. The method for the prediction of students' performance is discussed in the previous section. In the prediction scenario of students' interest, we suppose that the students are from different years, noted as $Y_1, Y_2, \cdots, Y_m$ $(Y_1 < Y_2 < \cdots < Y_m)$. A student of year Y_1 means that the student begins his (her) study at university at year Y_1. To make recommendation for a target student of year Y_k and for courses in semester q, there are two steps: selection of similar students and selection of recommended courses.

The final output of the model is three lists. List 1 is the list of recommended courses that the student might be good at; lists 2 and 3 are the lists of recommended courses and required courses that the student might not be good at. $Y_i S_m$ is the Top-m most similar student from year Y_i. $Y_i S_m$ sim is the similarity between $Y_i S_m$ and the target student. $Y_i S_j C_k$ represents a course learned by $Y_i S_j$ but is not learned by the target student.

3.1. Selection of Students. Basing on the previous selected courses of the first $q-1$ semesters, we select N_0 most similar students from the students of each previous year

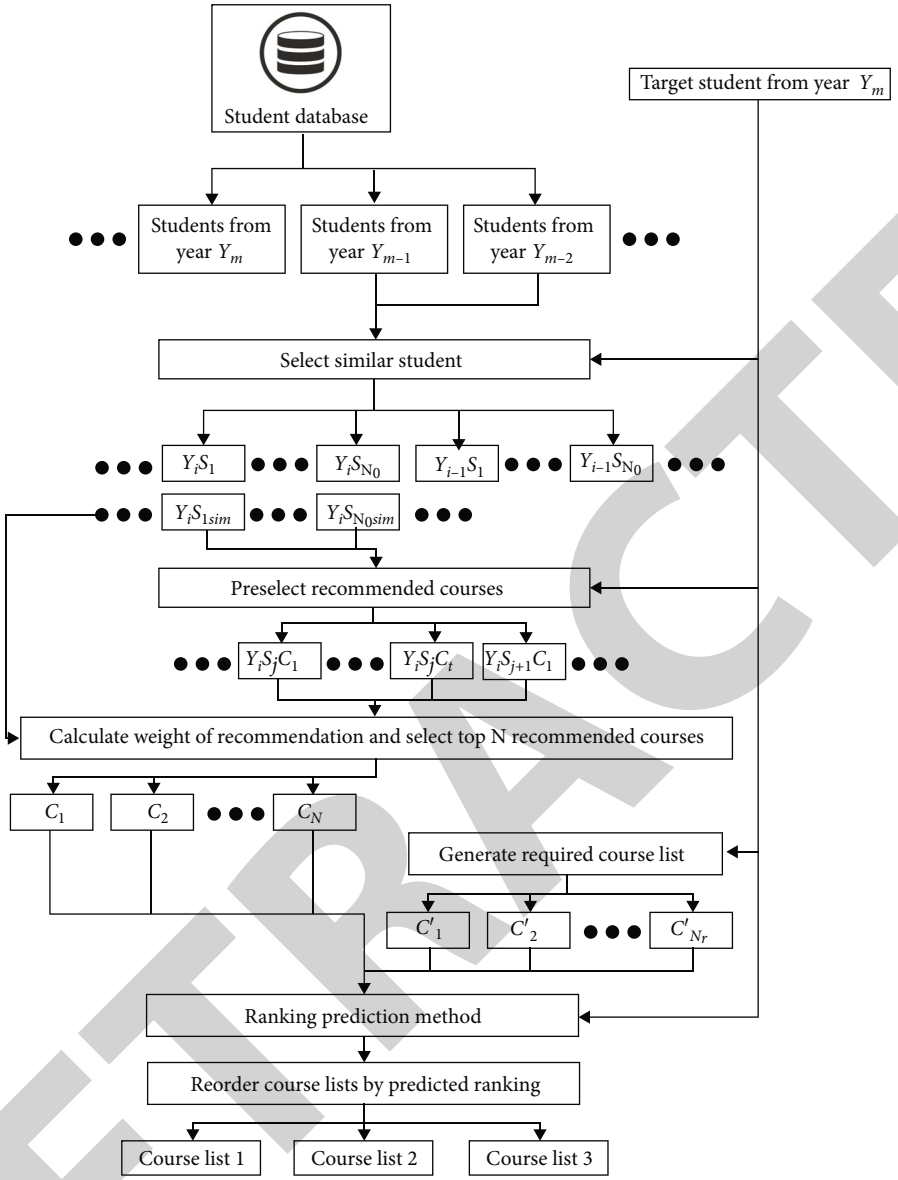


Figure 1: Visualization of the system model.

$(Y_1, Y_2, \cdots, Y_{k-1})$. The method to calculate the similarity is presented at the end of this part. These selected students are presented by set $\{Y_iS_j, i \in [1, k-1], j \in [1, N_0]\}$ which is named as the set of similar students. Meanwhile, each student in the set of similar student has a value of similarity compared to the target student; these values of similarities are presented by the set $\{Y_iS_j\text{sim}, i \in [1, k-1], j \in [1, N_0]\}$ in which $Y_iS_j\text{sim}$ presents the similarity between student Y_iS_j and the target student.

3.2. Selection of Recommended Courses. After obtaining the set of similar students, for each student in this set, we select the courses which he (she) has learned in the first q semesters and which the target student has not learned in the first $q-1$ semesters. These courses are presented by set $\{Y_iS_jC_t, i \in [1, k-1], j \in [1, N_0], t \in [1, Y_iS_jN]\}$ in which Y_iS_jN present the number of courses which student Y_iS_j has learned and the target student has not learned. In fact, in this set, different elements may be associated with the same course. This set is named the set of preselected courses. Then, we calculate the weight of recommendation of each course that appears in the set of preselected courses as shown in equation (3). In this equation, if $\text{course}_{\text{name}}$ is the same as $Y_iS_jC_t$, then the function $I(\text{course}_{\text{name}} = Y_iS_jC_t)$ will return to 1; otherwise, it will return to 0. The weight of recommendation describes quantitatively whether the course should be recommended, and this value can be used to sort the top N recommendation list:

$$\text{weight}_{\text{rec}}(\text{course}_{\text{name}}) = \sum_i \sum_j \sum_t Y_iS_j\text{sim} * I(\text{course}_{\text{name}} = Y_iS_jC_t). \tag{3}$$

In the next part, the method to calculate the similarity between courses and the method to measure the importance

of course are presented, in which the importance of course is an important indicator when calculating the similarity between courses.

3.3. Measure the Importance of Courses. Breese et al. proposed in [46]: similar ratings on some popular items do not represent a good indication that the two users have similar preferences, and similar ratings on niche items are more meaningful for reference.

The same idea can be used to optimize our model: The influences of different courses on the description of students' personality are different. The courses which are chosen by fewer students can better describe their personalities. The courses which are chosen by most of the students are usually some necessary courses for their major, and so these courses cannot describe students' personalities. Basing on this idea, we use equation (4) to measure the importance of the course. Therefore, we will reduce the weight of popular courses and increase the weight of minority courses. Here, N_{total} presents the total number of students in the current department and N_{course} presents the number of students who have chosen the current course:

$$\text{weight}(\text{course}) = \frac{N_{\text{total}}}{N_{\text{course}}}. \tag{4}$$

3.4. Measure the Similarity between Students. According to our method, the similarity between student S_i and student S_j is the similarity between the set of courses learned by student S_i: $\{C_{i,k}\}$ and the set of courses learned by student S_j: $\{C_{j,k}\}$. The method is shown in

$$\text{sim}(S_i, S_j) = \frac{\sum_{c \in \{C_{i,k}\} \cap \{C_{j,k}\}} \text{weight}(c)}{\sum_{c \in \{C_{i,k}\} \cup \{C_{j,k}\}} \text{weight}(c)}. \tag{5}$$

4. Ranking Prediction Method

There are many ways to describe a student's performance on a course. Two mainly used ways are students' score and ranking. A student's score can be influenced by many factors. Different courses may have different means and variances. The same course may also use different ways for evaluation on different semesters. Also, the score is uncertain. Even though the observed score is a fixed number, the ground truth score could be a distribution. On the contrary, students' ranking is a better choice for prediction. Firstly, the ranking will not be influenced by the mean and the variance of scores. Secondly, ranking can be used as an indicator that can directly describe whether the student is good at this course or not. Meanwhile, to better decrease the influence of uncertainty on the prediction results, we regroup the students into two categories. The first category contains students whose rankings are under 50%. The second category contains the students whose rankings are above 50%. In this way, the uncertainty of the score can only influence the students whose rankings are near 50%.

4.1. Course2Student. Our proposed method is named as Course2Student since our main idea is to use the characteristic of the previous courses and the associated performances to describe the characteristic of the student. Before introducing our method, we will first make a summary of the existing methods for students' performance prediction.

Most of the methods for student performance prediction can be regarded as a function. The inputs of the function are some indicators of students which are correlated with students' performance. The output is the students' predicted performance. These methods can be divided into two types: parametric method and nonparametric method. For the nonparametric method, a rule is proposed. Basing on this rule and the inputs of the prediction model, the correlated information is selected from the data, and then, they are used to predict the final result. Some commonly used nonparametric methods include K-nearest neighborhood and collaborative filtering [20]. For parametric methods, there will be some parameters in the model, and the data optimize these parameters. We can consider that the nonparametric model directly takes data as memories and the parametric model first learn something from the data and then forget about the original data.

Course2Student is a parametric method. It is based on neural collaborative filtering and linear regression. Linear regression is one of the earliest methods used for students' performance prediction [47], and even in recent years, it is still a commonly used method [48]. Equation (6) is an example of linear regression, in which S presents the student's performance which we want to predict, X_i presents one indicator of students which can be used to predict the student's performance, and w_i presents the associated weight of indicator X_i. In previous works, many indicators have been studied for the prediction of student performance [48] Generally, the most correlated indicators to students' performance prediction are the students' performances on his (her) previous courses. So, in our work, we choose students' performances on his (her) previous courses as indicators for performance prediction:

$$S = \sum_i w_i X_i. \tag{6}$$

One advantage of linear regression is that based on the weights associated with each indicator, we can easily understand the importance of each indicator on the prediction result. One disadvantage of linear regression is that, after training, the model can only be used in the current mission, which means that it can only be used to predict the performance of one course. In a real-life application, we need to predict the results of many courses, so we have to train separately many independent linear regression models. Also, since the weights are correlated with the indicators, if one indicator in a model is changed, then we have to train a new model. To solve the above problems, one possible solution is to reuse the weights, as shown in equation (7). Since in our method the rankings of new courses are predicted by previous courses' rankings, we use X to present both the prediction results and the indicators. In equation (7), X_j presents the ranking of the current student on course j, $w_{i,j}$ describes the influence of course i on the ranking prediction of course j, or we can say that it describes the

similarity between course i and course j on the ranking prediction problem. Set A presents the set of courses that the current student has learned. By this method, after obtaining all the weights in $\{w_{i,j}\}$, given one student and one course, the student's ranking on that course can be predicted basing on his (her) rankings of previous courses:

$$X_j = \sum_{i \in A} w_{i,j} X_i. \tag{7}$$

The above solution makes the prediction model more flexible, and it still could be further improved. By this method, supposing that there are N courses, then $\{w_{i,j}\}$ will have N^2 parameters (supposing that the similarity between two courses is not symmetric which means that $w_{i,j}$ is not necessarily equal to $w_{j,i}$ when $i \neq j$). But in fact most of the weights in $\{w_{i,j}\}$ are not independent, so it is not necessary to have N^2 parameters. For example, if the ranking of course k only depends on the ranking of course i as $X_k = w_{i,k} X_i$, and the ranking of course j only depends on the ranking of course i as $X_j = w_{i,j} X_i$, then when predicting the ranking of course k basing on the ranking of course j, we will have $X_k = w_{j,k} X_j = w_{i,k}/w_{i,j} X_j$, which means that $w_{j,k} = w_{i,k}/w_{i,j}$; these three parameters are not independent. In order to further improve the prediction method, we propose a new method as shown in equation (8), which is our proposed Course2Student method. In this equation, embedding$(\cdot)$ is a function which projects the identity of the course to the latent vector of the course. sim$(\cdot)$ is a function basing on neural networks which measures the similarity between two latent vectors of courses in ranking prediction problems:

$$X_j = \sum_{i \in A} \text{sim}(\text{embedding}(i), \text{embedding}(j)) X_i. \tag{8}$$

The function embedding$(\cdot)$ has MN parameters where M presents the length of the latent vector of the course. Since there are relatively many courses $M < N$, this method has fewer parameters compared with the previously discussed method. Besides, this method can keep the transitivity of similarity, which means that if course i is similar to course j and course j is similar to course k, then by using the Course2Student method, we can obtain that course i is also similar to course k, which is logical. The idea of using neural networks to measure the similarity comes mainly from the work of neural collaborative filtering. Since it is hard for human to choose the best way of the measure of similarity basing on the given data, it is better to use neural networks to learn a measure of similarity directly from the data. The Course2Student method is also illustrated in Figure 2.

4.2. Selection of Correlated Courses. In our prediction scenario, normally, a student will have a lot of previous courses; it will take a lot of time if we use all these previous courses to train the Course2Student model. So, before training the Course2Student model, for each target course for prediction, N_0 most correlated courses are selected. The correlation between two courses is calculated by Pearson's correlation coefficient, as shown in equation (9), in which the correlation between course x and course y is calculated. In this equation, n presents the number of students who have learned both course x and course y, and x_i and y_i present, respectively, the score on course x and the score on course y of the i^{th} student who has learned both course x and course y. Considering that some courses might have very few learners, in that case, Pearson's correlation coefficient may not be able to describe the correlation between these courses properly. Therefore, when calculating the correlation, if the number of students who have learned both courses is under 30, then the correlation between these two courses is considered 0:

$$\text{correlation}(x, y) = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \sqrt{\sum_i^n (y_i - \bar{y})^2}}. \tag{9}$$

4.3. Hybrid Prediction Method. Considering that different prediction results have different uncertainties, but one determinist prediction model can only give one result, basing on this result, it is hard to measure the uncertainty of this prediction. In order to better estimate the uncertainty of prediction and filter out the uncertain prediction results to improve the overall accuracy, we use a hybrid prediction method by combining Course2Student and a user-based collaborative filtering method.

We choose these two methods because Course2Student is a parametric method and user-based collaborative filtering is a nonparametric method, so these two methods will have very different decision boundaries. If these methods give the same prediction result, then we consider that this prediction result has high confidence. In the hybrid prediction method, we only consider the result which has high confidence as an available prediction result. Basing on this hybrid method, we can divide each course list (list of required courses or list of courses that the student might like) into two sublists. The first sublist is the list of courses that the student might be good at (available ranking prediction result is in top 50%), and the second sublist is the list of courses that the student might not be good at (available ranking prediction result is not in 50%).

The user-based collaborative filtering method for ranking prediction is shown in equation (10). In order to predict the ranking of student a on course j which is presented by $X_{a,j}$, we first select a set of students who are most similar to student a and who have also learned course j, then use their rankings on course j and their similarities with student a to predict $X_{a,j}$. Function sim(i, a) measures the similarity between two students, $C(i)$ presents the set of courses which student i has learned, and card (A) presents the number of elements in set A:

$$X_{a,j} = \sum_{i \in \text{neighbors}(a)} \text{sim}(i, a) X_{i,j},$$

$$\text{sim}(i, a) = \frac{1}{\sum_{j \in C(i) \cap C(a)} \left(X_{i,j} - X_{a,j}\right)^2 / \text{card}\,(C(i) \cap C(a))}. \tag{10}$$

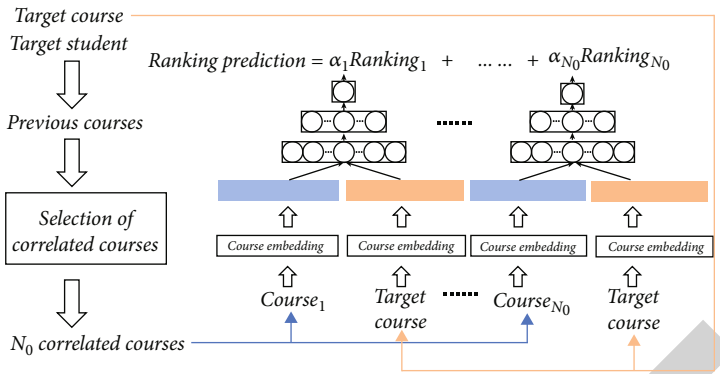


Figure 2: Course2Student ranking prediction method.

Table 1: Recommendation results.

Method	User-based collaborative filtering			
Department	Computer science			
Year	2016			2017
Semester	5	7	8	5
Recall				
$N = 10$	0.1164	0.1421	0.1392	0.1711
$N = 20$	0.1674	0.196	0.2187	0.3061
$N = 30$	0.2179	0.2342	0.2888	0.3809
Precision				
$N = 10$	0.0269	0.0484	0.0329	0.0432
$N = 20$	0.0182	0.0324	0.0262	0.0365
$N = 30$	0.0161	0.0258	0.0233	0.0303
Department	Honors college			
Year	2016			2017
Semester	5	7	8	5
Recall				
$N = 10$	0.539	0.4637	0.3977	0.4667
$N = 20$	0.6288	0.553	0.5208	0.5289
$N = 30$	0.6741	0.5912	0.5775	0.568
Precision				
$N = 10$	0.2957	0.3657	0.2862	0.2796
$N = 20$	0.1755	0.2234	0.1885	0.1598
$N = 30$	0.1374	0.1634	0.1404	0.1155

Table 2: Train/test data split options.

Year	Computer science and honors college				Data split 1	Data split 2	Data split 3
	2015	2016	2017	2018			
Semester	1				Train	Train	Train
	2						
	3						
	4	1					
	5	2					
	6	3					
	7	4	1				
	8	5	2				
	9	6	3				
	10	7	4	1	Test	Test	Test
	11	8	5	2			
	12	9	6	3			

5. Experiment

In order to evaluate our methods, we collect our own dataset. This dataset contains students of years 2015, 2016, 2017, and 2018 from the department of computer science and honors college at our university. The data of the department of computer science contains 844 students and 715 courses. The data of honors college contains 977 students and 1457 courses. According to the education system of our university, a bachelor's degree takes 4 years. Each year has three semesters, the semester of autumn, the semester of spring, and the semester of summer. The semester of summer is a short semester, and only a part of the students will take this semester. So before obtaining the bachelor's degree, a student will have totally 12 semesters. Until these data are collected, we get data of 12 semesters for the students 2015, 9 semesters for the students 2016, 6 semesters for the students 2017, and 3 semesters for the students 2018. In order to protect students' privacy, the identity of each student is replaced by a unique string. In the original data, there is no information about whether the course is required or elective. On the other hand, as the students from different years might have different required course lists, for now, we have not got all the required course lists for all students. In order to know whether a course is an elective course or not, we check the number of students who have learned the course. If more than half of the students have learned the course, then the course will be treated as a required course. Otherwise, it will be treated as an elective course.

5.1. Advice on Equation Results of Course Recommendation. In this part of the experiments, we make a recommendation for the courses of semesters 5, 7, and 8 of the students of year

Table 3: Prediction results of student ranking.

Method	Department	Option	Accuracy
Neural collaborative filtering [26]	Computer science	1	0.6765
		2	0.6458
		3	0.6391
	Honors college	1	0.7025
		2	0.7055
		3	0.6895
User-based collaborative filtering	Computer science	1	0.6751
		2	0.6581
		3	0.6609
	Honors college	1	0.7116
		2	0.7214
		3	0.6949
Course2Student	Computer science	1	**0.6938**
		2	**0.6635**
		3	**0.6616**
	Honors college	1	**0.7208**
		2	**0.7250**
		3	**0.7088**

2016 and for the courses of semester 5 of the students of year 2017. The reason why we choose these semesters is that there are relatively more elective courses, so it is better for the evaluation of recommendation methods. The recommendation lists are generated based on the previous choices of courses. We use recall and precision to evaluate the recommendation results, and different lengths of recommendation lists are used: 10, 20, and 30. Table 1 shows the results by using the user-based collaborative filtering method. Here, we are more interested in the results of recalls because in this mission, we want that the recommendation list can contain all the courses that the student might choose and at the same time provide some courses that the student has not noticed but may catch his (her) interests.

Precision is also used since it is commonly used to evaluate the recommendation results. We can see that the results on honors college are much better than the results on the department of computer science. It may be caused by the fact that the students from the honors college will choose their options at the end of the second year, so from their choices of elective courses, we can see some personal characteristics. From this point of view, we can see that this recommendation method is more suitable for the recommendation scenarios where students have more elective courses to choose from and where students have their own options to choose from. Considering the values of recall, the average value of recalls of semester 5 of the students from the honors college of year 2016 is above 0.5 even though $N = 10$, which is a result that could prove that the current method can be used in real life. In future work, we will pay more attention on the departments which education systems are similar to the education system of the honors college and explore new recommendation methods for the departments in which students have relatively fewer elective courses. For example, we can combine natural language processing methods or knowledge graph [42] to add some prior knowledge about the courses.

Table 4: Prediction results of student ranking by hybrid method.

Method	Department	Option	Accuracy	Coverage
Hybrid	Computer science	1	0.7213	0.8219
		2	0.7109	0.7775
		3	0.7155	0.7478
	Honors college	1	0.7585	0.8344
		2	0.7618	0.8551
		3	0.7459	0.8225

5.2. Results of Ranking Prediction. In order to make our experiments more similar to the real prediction scenarios, we use the time node to separate the training and testing data. For example, for the students of year 2018, in order to get the ranking prediction after the second semester, we only use the data which can be collected before the beginning of the second semester. To better evaluate the generalization of our methods, we use three different time nodes to separate the training and testing data; for a student of year 2016, it is the time nodes before 8th, 7th, and 5th semesters. The details about the training/testing data split are shown in Table 2. In the second and third data split options, since all the data of students of year 2018 are in the testing data, we only consider the data of students of years 2015, 2016, and 2017.

We treat the students' ranking prediction problem as a binary classification problem. The rankings which are in the

Table 5: Prediction results of student ranking.

Method	Department	Option	Ranking improvement	Improvement percentage
Hybrid	Computer science	1	0.1677	0.7894
		2	0.1011	0.7739
		3	0.1537	0.7672
	Honors college	1	0.2138	0.7468
		2	0.2307	0.7318
		3	0.2405	0.7864

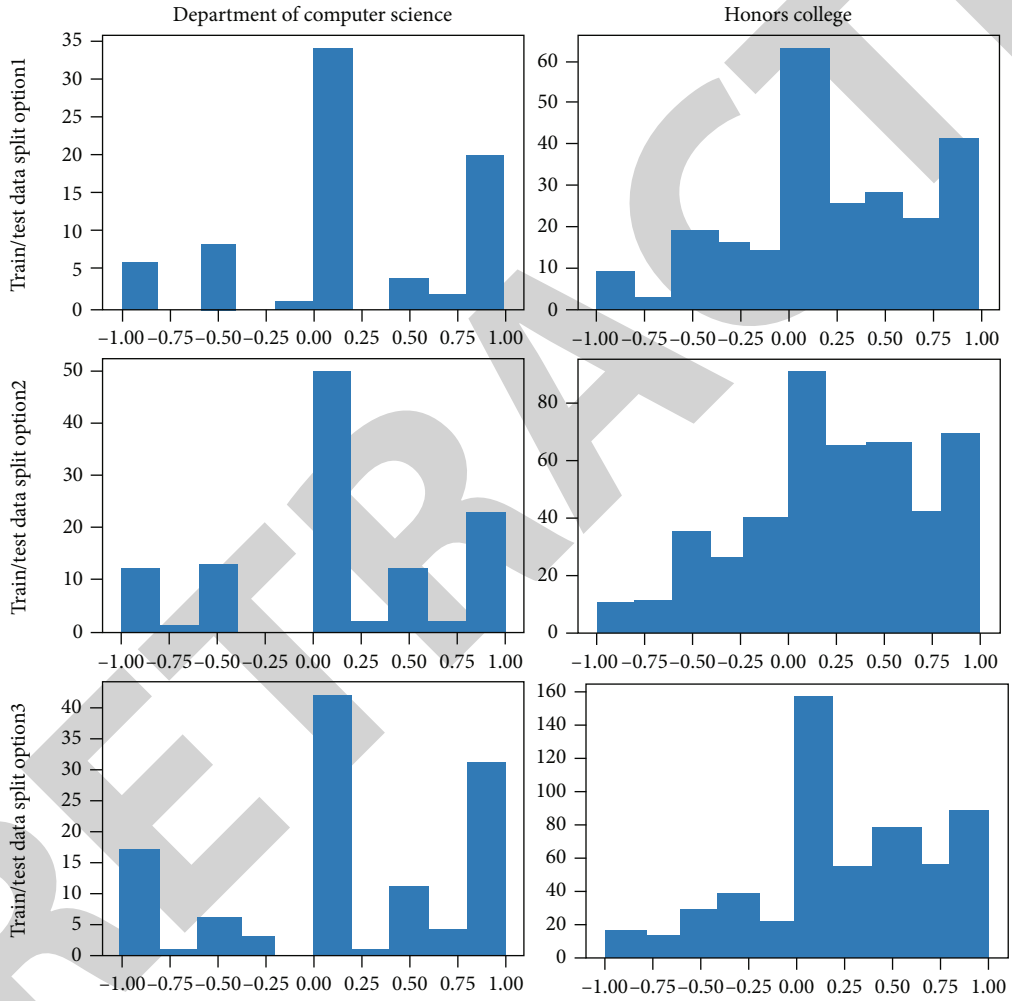


Figure 3: Distribution of ranking improvement.

first 50% are regarded as the first category, and the rankings which are not in the first 50% are regarded as the second category. In order to evaluate our proposed Course2Student method, its performance is compared with the performances of two other methods: neural collaborative filtering and user-based collaborative filtering, which are the two most used methods for similar problems. The accuracy of the testing data is used to evaluate the performance of each method. Table 3 shows the results.

For classification model f and test data set D with size n, accuracy is defined as

$$\text{Accuracy}(f; D) = \frac{1}{n}\sum_{i=1}^{n}(f(x_i) = \text{label}_i). \tag{11}$$

We bold the highest accuracy of each data split option.

Results show that our method achieves the highest accuracies on all data and all training/testing split options.

To further improve the accuracy and the confidence of prediction results, we use the hybrid method. The prediction results of user-based collaborative filtering and Course2Student are combined. Only the same prediction results are regarded as available prediction results and used. The other results are regarded as results with high uncertainties. Table 4 shows the results of the hybrid method, which includes the accuracies and the percentages of the available results. We can see that the hybrid method can indeed further improve prediction accuracy, which means that this estimation of uncertainty is logical.

5.3. Results of Ranking Prediction. To evaluate whether this method can improve students' ranking, based on the results of the hybrid prediction method, we separate the list of elective courses into two sublists for each student in the testing data. The first sublist contains the courses which have a predicted ranking in the first 50%, and the second sublist contains the courses which have a predicted ranking not in the first 50%. For each student, we use the difference between the mean of the real rankings of the courses in the first sublist and the second sublist to describe the ranking improvement of that student by only choosing the elective courses that the student is predicted to be good at. Here, we use 0 and 1 to present the ranking. 1 means that it is in the top 50%, and 0 means that it is not in the top 50%. The ranking improvement is described by the mean of ranking improvements of all associated students. Table 5 shows the results. We can see that under different training and testing split options, there are always over 70% of students whose rankings are improved according to this method of evaluation. Figure 3 presents the details about the distribution of ranking improvement. One point that needs to be discussed is that now we only use the history data to evaluate our methods, but when these methods are applied in a real-life situation, will the result of the student be changed when he (she) knows the predicted result? In our future studies and applications, we will concentrate on this point.

5.4. Advantages of the Proposed Method. Our proposed method has four major advantages.

Firstly, the framework of the proposed ranking prediction method is based on the traditional method so that it can always have acceptable prediction results; on the contrary, when using the end-to-end machine learning methods, like the ones used in previous works, sometimes we could get some extremely abnormal result; for example, the predicted value can be out of the interval of possible results.

Secondly, the proposed ranking prediction method makes it possible for us to know the influences of each historical course on the ranking prediction of the target course, which can help us understand how the prediction result is made and as a result has a better vision of the resulting model. And this vision and understanding of the model can help us debug the model during the training process much easier compared with an end-to-end model.

Thirdly, as the recommended courses are firstly selected according to the students' preference and then reordered by the ranking prediction result, all the recommended courses are acceptable considering the student's interest. Some previous works use the performance prediction results directly to recommend courses, and in that way, some recommended results may be totally irrelevant to students' future plans.

Fourthly, when recommending courses only based on the performance prediction result, we need to predict students' performance on each course in the database, which is an extremely time-consuming process. In our method, it is only necessary to predict the performance of the student on the courses in the preselected list (selected by students' preference) which takes much less time.

6. Conclusions

In this work, we propose a new method that can automatically generate personal advice about courses in the next semester. Particularly, we explore the application of deep learning methods on students' ranking prediction problem and propose a new method that combines neural networks with traditional methods. The results of experiments prove that our methods can indeed recommend courses for students that can match their interests, and students have a high possibility to improve their average rankings when they choose the elective courses which, according to the prediction result, they might be good at. For now, our studies are mainly based on the existing data. In future work, we will concentrate more on the changes in students' behaviors when they know the prediction results. For example, students will change their learning strategies and achieve higher scores when they have known that they might not be good at certain courses. The relevant machine learning methods and course recommendation methods mentioned in this paper can be widely applied in IOT and industrial vocational learning systems. So in our next steps, we will work with the associated departments and try to apply our methods to a real-life system.

Data Availability

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

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2019YFB2102200) and National Science Foundation of China (Grant No. 61977003).

References

[1] L. Huang, Q. Fu, M. He, D. Jiang, and Z. Hao, "Detection algorithm of safety helmet wearing based on deep learning," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 13, article e6234, 2021.

[2] D. Jiang, G. Li, C. Tan, L. Huang, Y. Sun, and J. Kong, "Semantic segmentation for multiscale target based on object recognition using the improved Faster-RCNN model," *Future Generation Computer Systems*, vol. 123, pp. 94–104, 2021.

[3] S. Liao, G. Li, H. Wu et al., "Occlusion gesture recognition based on improvedSSD," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 6, p. e6063, 2021.

[4] Q. Gao, J. G. Liu, and Z. J. Ju, "Robust real-time hand detection and localization for space human robot interaction based on deep learning," *Neurocomputing*, vol. 390, pp. 198–206, 2020.

[5] Q. Gao, J. G. Liu, Z. J. Ju, and X. Zhang, "Dual-hand detection for human-robot interaction by a parallel network based on hand detection and body pose estimation," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9663–9672, 2019.

[6] X. Ma and Z. Zhou, "Student pass rates prediction using optimized support vector machine and decision tree," in *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)*, pp. 209–215, Las Vegas, NV, USA, 2018.

[7] H. Hamsa, S. Indiradevi, and J. Kizhakkethottam, "Student academic performance prediction model using decision tree and fuzzy genetic algorithm," *Procedia Technology*, vol. 25, pp. 326–332, 2016.

[8] M. Ashenafi, G. Riccardi, and M. Ronchetti, "Predicting students' final exam scores from their course activities," in *2015 IEEE Frontiers in education conference (FIE)*, pp. 1–9, El Paso, TX, USA, 2015.

[9] D. Petkovic, M. Sosnick-Pérez, K. Okada et al., "Using the random forest classifier to assess and predict student learning of software engineering teamwork," in *2016 IEEE Frontiers in education conference (FIE)*, pp. 1–7, Erie, PA, USA, 2016.

[10] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: a brief review," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 7068349, 13 pages, 2018.

[11] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [review article]," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[12] Y. D. Li, Z. B. Hao, and H. Lei, "Survey of convolutional neural network," *Journal of Computer Applications*, vol. 36, no. 9, pp. 2508–2515, 2016.

[13] P. Kamavisdar, S. Saluja, and S. Agrawal, "A survey on image classification approaches and techniques," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 1, pp. 1005–1009, 2013.

[14] P. N. Druzhkov and V. D. Kustikova, "A survey of deep learning methods and software tools for image classification and object detection," *Pattern Recognition and Image Analysis*, vol. 26, no. 1, pp. 9–15, 2016.

[15] M. Ludewig, N. Mauro, S. Latif, and D. Jannach, "Performance comparison of neural and non-neural approaches to session-based recommendation," in *Proceedings of the 13th ACM conference on recommender systems*, pp. 462–466, Copenhagen, Denmark, 2019.

[16] A. Pigeau, O. Aubert, and Y. Prié, *Success Prediction in MOOCs A Case Study*, International Educational Data Mining Society, 2019.

[17] K. Bhumichitr, S. Channarukul, N. Saejiem, R. Jiamthapthaksin, and K. Nongpong, "Recommender systems for university elective course recommendation," in *The 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 1–5, NakhonSiThammarat, Thailand, 2017.

[18] Z. D. Zhao and M. S. Shang, "User-based collaborative-filtering recommendation algorithms on Hadoop," in *2010 third international conference on knowledge discovery and data mining*, pp. 478–481, Phuket, Thailand, 2010.

[19] Z. A. Pardos, Z. Fan, and W. Jiang, "Connectionist recommendation in the wild: on the utility and scrutability of neural networks for personalized course guidance," *User Modeling and User-Adapted Interaction*, vol. 29, no. 2, pp. 487–525, 2019.

[20] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, pp. 285–295, Minneapolis, MN, 2001.

[21] N. N. Liu, Z. Min, and Y. Qiang, "Probabilistic latent preference analysis for collaborative filtering," in *Proceedings of the 18th ACM conference on Information and knowledge management (CIKM '09)*, pp. 759–766, New York, NY, USA, 2009.

[22] M. Chen, Y. Ma, B. Hu, and L. Zhang, "A ranking-oriented hybrid approach to QoS-aware web service recommendation," in *2015 IEEE International Conference on Services Computing*, pp. 578–585, New York, NY, USA, 2015.

[23] W. Markus, K. Alexandros, Q. V. Le, and S. Alex, "COFI-RANK maximum margin matrix factorization for collaborative ranking," in *Proceedings of the 20th International Conference on Neural Information Processing Systems*, pp. 1593–1600, Red Hook, NY, USA, 2007.

[24] S. Yue, K. Alexandros, B. Linas, L. Martha, and H. Alan, "XCLiMF: optimizing expected reciprocal rank for data with multiple levels of relevance," in *Proceedings of the 7th ACM conference on Recommender systems*, pp. 431–434, New York, NY, USA, 2013.

[25] B. Hidasi and A. Karatzoglou, "Recurrent neural networks with top-k gains for session-based recommendations," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 843–852, New York, NY, USA, 2018.

[26] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. S. Chua, "Neural collaborative filtering," in *Proceedings of the 26th International Conference on World Wide Web*, pp. 173–182, Republic and Canton of Geneva, Switzerland, 2017.

[27] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[28] B. C. Hardgrave, R. L. Wilson, and K. A. Walstrom, "Predicting graduate student success: a comparison of neural networks and traditional techniques," *Computers & Operations Research*, vol. 21, no. 3, pp. 249–263, 1994.

[29] D. J. Hamilton, *Multiple Regression Analysis and Prediction of GPA upon Degree Completion*, College Student Journal, 1990.

[30] W. R. Klecka, G. R. Iversen, and W. R. Klecka, *Discriminant Analysis*, Sage, 1980.

[31] S. J. Press and S. Wilson, "Choosing between logistic regression and discriminant analysis," *Journal of the American Statistical Association*, vol. 73, no. 364, pp. 699–705, 1978.

[32] T. Wang and A. Mitrovic, "Using neural networks to predict student's performance," in *International Conference on Computers in Education, 2002. Proceedings*, pp. 969–973, Auckland, New Zealand, 2002.

[33] I. Lykourentzou, I. Giannoukos, G. Mpardis, V. Nikolopoulos, and V. Loumos, "Early and dynamic student achievement

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7072849, 7 pages
https://doi.org/10.1155/2021/7072849

Research Article

Research on News Recommendation System Based on Deep Network and Personalized Needs

Weijia Zhang[1] **and Feng Ling**[2]

[1]*City Institute, Dalian University of Technology, Dalian, Liaoning 116600, China*
[2]*Dalian Daily, Dalian, Liaoning 116600, China*

Correspondence should be addressed to Weijia Zhang; zwj1978@m.fafu.edu.cn

Received 2 September 2021; Revised 10 October 2021; Accepted 22 October 2021; Published 2 November 2021

Academic Editor: Yinghui Ye

In order to solve the problems of poor performance of the recommendation system caused by not considering the needs of users in the process of news recommendation, a news recommendation system based on deep network and personalized needs is proposed. Firstly, it analyzes the news needs of users, which is the basis of designing the system. The functions of the system module mainly include the network function module, database module, user management module, and news recommendation module. Among them, the user management module uses the deep network to set the user news interest model, inputs the news data into the model, completes the personalized needs of the news, and realizes the design of the news recommendation system. The experimental results show that the proposed system has good effect and certain advantages.

1. Introduction

With the continuous development of the Internet, the mode or speed of information transmission has undergone earth-shaking changes. Therefore, great changes have taken place in the way of obtaining information [1]. Coupled with the rapid development of mobile Internet and intelligent devices, people get information more quickly and can get news information anytime and anywhere. The value of news lies in real-time, so news communication is more valuable in the form of Internet development [2]. New media occupy a more and more important position in the modern media industry because of its rich forms, strong interaction, wide channels, high coverage, accurate arrival, high cost performance, and convenient promotion. The "new" of news media is reflected in the following three points: all-round digitization, interactivity, and personalization. The personalized information service provided by this news media enables the audience to obtain personalized customized services. The news media based on the Internet not only brings convenience for users to obtain information but also brings the problem of information overload, and this problem is more serious than the traditional media. In the face of a large amount of news information, users cannot quickly find the information they need. The existing Mengguang website provides the same news information to different users and also provides a certain retrieval function, which meets the needs of users to a certain extent, but still cannot meet the deep-seated needs of users [3]. Under a large amount of news information, how to quickly classify and mine massive information and then push it to different users has become an urgent need to be solved. A recommendation system is a system that mainly solves the problem of information overload on the Internet. It can filter the information on the Internet, obtain more user interest preferences, and push the information similar to the user's interest to the user, so as to effectively control the accumulation of ineffective information in the user's field of vision [4]. Recommendation system is divided into a nonpersonalized recommendation system and a personalized recommendation system. The nonpersonalized recommendation system provides recommendations with the same content to all users. The personalized recommendation system can recommend thousands of people and thousands of faces according to users [5]. At

the same time, users can be grouped through user clustering, and people with the same hobbies can be placed in the same group. Because the effect of personalized recommendation is very good, it has been widely used in existing Internet products. Therefore, in order to improve the efficiency of news recommendation, a lot of research has been done on news recommendation system in this field [6].

Literature [7] proposed a personalized news recommendation based on event ontology. Aiming at the problems of cold start, sparse data, lack of semantics, and low recommendation accuracy in the traditional recommendation system, a recommendation algorithm based on event ontology is proposed. The event ontology is constructed by combining the news classification structure and news corpus, the elements of the news browsed by users are extracted, and the user interest model is constructed. The classification structure based on event ontology calculates the similarity between news events. The user interest similarity is calculated through the user interest model, and the relevant news events are found according to the semantic radius of the nonhierarchical structure of the event ontology. The news personalized recommendation results are obtained from the three aspects of event ontology similarity, user interest similarity, and nonhierarchical structure similarity. The experimental results show that the recommendation results of this algorithm are better than the collaborative filtering recommendation algorithm and content-based recommendation algorithm. This recommendation method does not consider the personalized needs of users, and there are more uninterested information in the recommended news information. Literature [8] proposed a hybrid recommendation algorithm based on incremental collaborative filtering and latent semantic analysis. The hybrid recommendation algorithm dynamically adjusts the recommendation list by incrementally updating the item similarity list in the item-based incremental update collaborative filtering module. The experimental results show that the proposed IULSACF algorithm is better than the traditional recommendation method in various evaluation indexes. This method takes less consideration of news recommendation and has some limitations. Literature [9] proposed the design and implementation of the personalized news recommendation system based on collaborative filtering. The key research contents include proposing a personalized news recommendation model, improving traditional collaborative filtering algorithm, and realizing a personalized news recommendation system combined with a mobile platform. The personalized news recommendation model includes four modules, namely, news classification, user interest analysis, user clustering, and recommendation result generation. The improvement of collaborative filtering algorithm includes filling the default value with the average user score and the popularity of items to solve the problem of data sparsity and setting the time weight of user score data with the time attenuation function to solve the problem of user interest migration. The implementation of a personalized news recommendation system includes analyzing the requirements of the system, introducing the design ideas and implementation methods of the system, integrating the improved recommendation algorithm into the system, and finally realizing the development of a personalized news recommendation system. This paper proposes a personalized news recommendation model and improves the traditional collaborative filtering algorithm. Finally, a personalized news recommendation system based on collaborative filtering is designed and implemented on the basis of the Android system. The personalized news recommendation system designed and implemented in this paper is in line with the era background of mobile Internet, can bring users a personalized service experience, and has high research value and application significance. Literature [10] made an in-depth study of the literature from 2001 to 2019 and listed 81 relevant studies, which were roughly divided into six categories for discussion and solved some of the many challenges in the field of journalism. Literature [11] proposed a new personalized news recommendation framework, the Hybrid Personalised News Recommendation (HYPNER). In Literature [12], the author proposes a context-mixing, deep-learning-based method for session-based news recommendation, which can utilize multiple information types.

The Internet of Things (IoT) is an information carrier based on the Internet, traditional telecommunication networks, etc. It enables all ordinary physical objects that can be independently addressed to form interconnected networks. The application of the Internet of Things involves all aspects. The application of the Internet of Things in the fields of industry, agriculture, environment, transportation, logistics, security, and other infrastructure effectively promotes the intelligent development of these aspects, makes the limited resources more rationally used and allocated, and thus improves the efficiency and benefits of the industry. In the household, medical and health care, education, finance and service industry, tourism and other fields closely related to life, service scope, service mode, and service quality have been greatly improved, greatly improving people's quality of life. With the progress of science and technology in human society, the development of the Internet of Things has become an irresistible trend.

In order to solve the shortcomings of the above recommendation, this paper designs a news recommendation system based on deep network and personalized needs. Firstly, it analyzes the news needs of users, which is the basis of designing the system. The functions of the system module mainly include network function module, database module, user management module, and news recommendation module. Among them, the user management module uses the deep network to set the user news interest model, inputs the news data into the model, completes the personalized needs of the news, and realizes the design of the news recommendation system. The experimental results show that the proposed system has a good effect and certain advantages.

Our contribution includes the following three points:

(1) In order to solve the problems of poor performance of the recommendation system caused by not considering the needs of users in the process of news recommendation, a news recommendation system

based on the deep network and personalized needs is proposed

(2) We use the deep network to set up the user's news interest model, input the news data into the model, complete the personalized needs of news, and realize the design of the news recommendation system

(3) The experimental results show that the proposed system has good effect and certain advantages

2. Research on News Recommendation System Based on Depth Network and Personalized Needs

2.1. Demand Analysis of the News Recommendation System. The content application of news shows all the news in the platform to users according to category and time. Users can retrieve the content according to their favorite news categories and the navigation displayed by the news system and search directly for news with specific keywords or clear content. However, the problem brought by this news system is that all users see the same news information list. It does not process news information for different users and allows the news system to customize their own news homepage for each user [13, 14]. The personalized news recommendation system will obtain the user's list of news of interest according to the user's demographic information and usage records on the platform. In this paper, the collaborative filtering algorithm is used for personalized news content recommendation. It recommends relevant news content to users according to the correlation between users or the correlation between news. The advantage of this method is that it can quickly build the similarity between users and recommend the news information viewed by similar users to each other.

The main requirements of the personalized news recommendation system are as follows:

(1) The system can recommend real-time news for users according to their current browsing behavior, so as to ensure that users can quickly obtain real-time news

(2) The news recommendation list can be updated in time, and the user recommendation list can be calculated according to the way that the news system collects Xinli, so as to ensure that users can obtain the latest recommendation list after logging in the Xinyang system every time

(3) The system should better solve the problem of news recommendation for new users and the problem of user cold start

(4) The newly added news also needs to be quickly recommended to users to solve the long tail effect in news information and better solve the cold start problem of the project

(5) The system tracks the accuracy and recall rate of the recommendation system [15], ensures that the news recommendation has a high accuracy and recall rate, and makes users satisfied with the current recommendation

(6) The system should be able to adapt to a large number of users and new news on the platform and can quickly calculate the push, save, and proofread type

(7) The system has good expansibility and meets the system's ability to deal with SeaView data in terms of capacity and computing power

(8) The real-time computing H-frame is used to calculate the personalized recommendation model, so as to improve the actual push and storage capacity of the whole platform

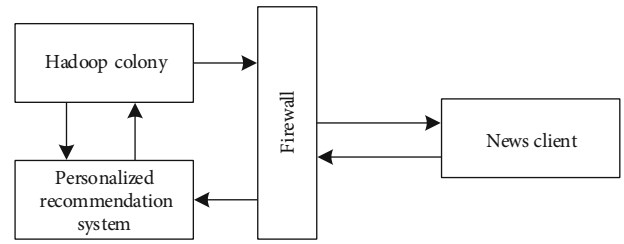


Figure 1: Network function design of the news recommendation system.

2.2. Function Module Design of News Recommendation System. In order to realize the effectiveness of the news recommendation system, based on the existing hardware, the system designs a network function module, database module, user management module, and news recommendation module.

2.2.1. Network Function Module of the News Recommendation System. The network is a key module of its implementation, and the fast and slow response of the network response reflects the performance of the system. Therefore, this system network takes advantage of the computing power and storage power of the Hadoop cluster [16] and stores the final computing results in a NoSQL database with a high concurrency power like HBase. The personalized news recommendation system modifies the relevant task parameters and submits the task to the Hadoop platform for computation, while the platform needs to monitor whether the entire recommended calculation is working properly. The Hadoop platform is responsible for the calculation and storage of the recommendation models and regularly conducts the recommendation model calculation to ensure the normal operation of the recommendation model calculation and finally stores the calculation results in [17] in HBase. The news client of the recommendation system is displayed by pulling the relevant user news recommendation list from HBase. The personalized network deployment diagram is shown in Figure 1.

2.2.2. News Recommendation System Database Module. The database in the news recommendation system is its recommended core function module. Therefore, the recommended

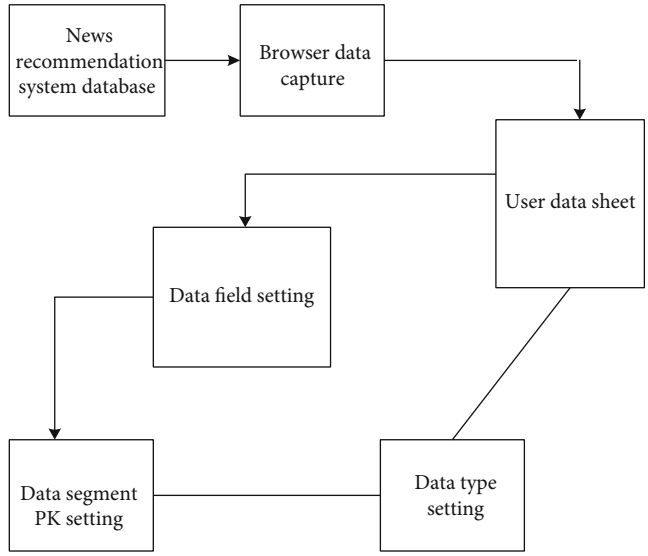


Figure 2: Recommended system database structure.

Table 1: User data.

The field name is called	Data type	Do you have the main key	Field description
ID	INT	Y	User number
Account	VARCHAR11	N	Login ID
Password	VARCHAR32	N	Login password
Salt	VARCHAR20	N	Password salt
Name	DATETIME	Y	Name of user
Type	TINYTIME	N	Role type
Create time	DATETIME	Y	Creation time

system provides a detailed design of the database module. All data in this system is stored as [18] via MySQL, with hot data and user data successfully logged in in Redis, for easy user access. The database structure of the system is shown in Figure 2.

The news recommendation system database mainly includes user data, permission data, and news information data; user data and news data are shown in Tables 1 and 2.

2.2.3. User Management Module Based on a Deep Network. In the user management module of the recommendation system, the personalized needs of users are mainly managed, and their interests are analyzed, so as to be the recommended object of the recommendation system. In user management, the news word vector is trained first, needing to divide the news text and remove the stop words included therein first. This paper uses the three deactivated word tables of HIT, the Baidu deactivated word table, and the machine intelligence laboratory of Sichuan University, combined to obtain a comprehensive deactivated word table. We place words in news text into the deword table to match. After removing the stop words, word vectors were trained using the Word2Vec model in Gensim. In order not to miss keywords, the minimum word frequency of the model was set to 1, the number of iterations to 5, and the word vector dimension to 200, and the model was trained to obtain the word vector model [19], and the calculated word vectors can be described as

$$V = [v_1, v_2, \cdots v_{200}]. \tag{1}$$

Every user will have a certain interest tendency; manifested in the news system is the click and reading tendency of the news content. How to measure users' interest is the key to recommendation algorithms. This article treats the news content that the user has viewed as inner ci of interest, generating three vectors of interest for each user through records of the user browsing, i.e.,

$$I = \frac{\sum_I^N W_i}{N}, \tag{2}$$

where I represents the number of set elements of all words included in the user's reading record, N represents the set of word vectors of all words included in the user's reading record, and W_i represents word vectors.

Based on this, this paper manages and trains user data with the help of a deep neural network. The structure of the deep network is composed of a number of restricted Boltzmann machines, in which the network layer can be divided into two layers: visual layers, and follows the rules of no connection and connection between layers. Usually, the hidden layer unit is mainly used to train the higher-order correlation characteristics shown by extracting visual layer data, as shown in Figure 3.

The average vector was calculated as the user vector of interest based on the deep network structure determined above. We enter the user vector of interest data into a deep network, convert the one-dimensional user vector into an

Table 2: News data.

The field name is called	Data type	Do you have the main key	Field description
ID	INT	Y	News number
Title	VARCHAR45	N	Head(ing) to come
Content	VARCHAR32	N	News content
Author	VARCHAR20	N	Author name
Image-url	DATETIME	Y	Figure connection
Browse	TINYTIME	N	Browse number
Create time	DATETIME	Y	Creation time

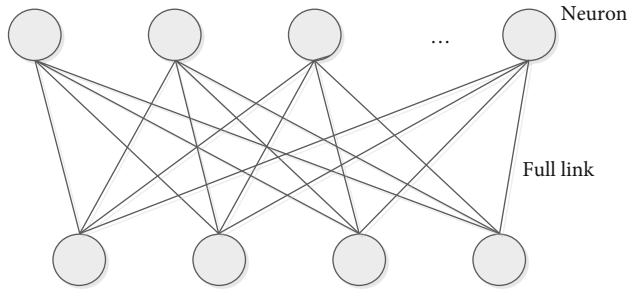


Figure 3: Basic structure of the deep network.

ordinary dimensional vector, and obtain it by using the convolutional operation in the deep network:

$$C_i = f\left(w \times X_{i,j,h-1} + b\right). \tag{3}$$

Among these, C_i represents the eigenvalues in the eigenmap, w represents the deep network kernel function, and b represents the sliding window.

At this time, the acquired feature vector is

$$C = [c_1, c_2, \cdots c_{n-h-1}]. \tag{4}$$

According to the determined user interest feature vector, the user news interest determination model is constructed to obtain

$$C_P = [c_1, c_2, \cdots c_{n-h-1/2}]. \tag{5}$$

It is then trained through a deep network, yielding the following:

$$L = -\sum_{1}^{N} Y_i \log\left(y_i\right). \tag{6}$$

Among them, N represents the number of news and y_i represents the degree of user love.

2.2.4. News Personalized Recommendation Module. According to the above analysis, the subject model was established and trained, the new input news is subject classified, and the content of the news is preprocessed; then, through the subject model classification, the feature vectors of the news are stored into the file system HDFS. The personalized recommended news list is completed, filtering by collaborative filtering methods based on two aspects of the user and item. Finally, the obtained news list is filtered and integrated, the processed structure is newly ranked, and the news list to the user personalized recommendation module is shown in Figure 4.

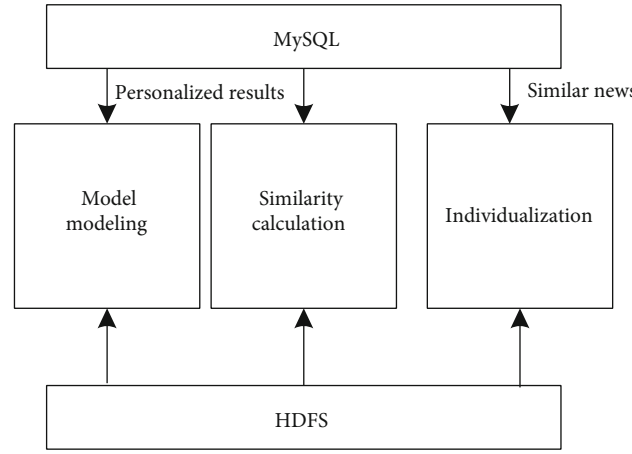


Figure 4: User-personalized recommended module structure.

3. Experimental Analysis

3.1. Experimental Scheme. To test the validity of the proposed system, an experimental analysis was performed. The hardware and software environment for the experimental test are shown in Table 3.

The network architecture of the experimental test system is shown in Figure 5.

3.2. Analysis of Experimental Results. In order to highlight the effectiveness of the method, it is compared against the accuracy of the system in [7] and the system in [8], and the results are shown in Figure 6.

The experimental results in Figure 6 show that with the continuous change of analysis time, the accuracy of user demand data recommendation using this system against the literature [7] system and the literature [8] system is different in news recommendation. Among them, the accuracy of the system of user demand data in news recommendation is about 98%. The best accuracy of the user demand data recommendation in the news recommendation is at least about 89%. The accuracy of user demand data recommendation is about 80%. By contrast, the accuracy of the system in the news recommendation is 9% higher than the literature [7] system and 18% above the literature [8] system. In contrast, the system recommended better results. This is due to the user management module using a deep network setting user news interest model in this system design. The news data is entered into the model, the personalized needs for news are completed, the design of the news recommendation system is realized, and the effectiveness of the system is improved.

Table 3: Test environment settings.

Hardware settings	Dell, Intel i7, internal storage, internal memory, EMS memory: 8 GB
Software settings	CentOS 6.4, MySQL
Database	SQL
The amount of news data (GB)	2

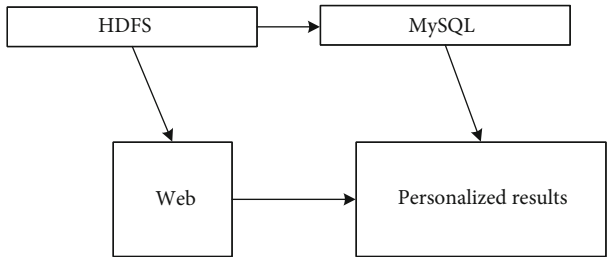


Figure 5: Test system environment.

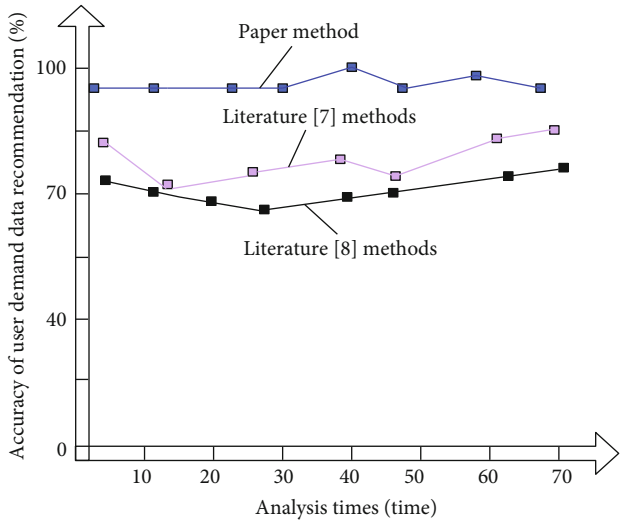


Figure 6: Accuracy analysis of user demand data recommendations in news recommendations.

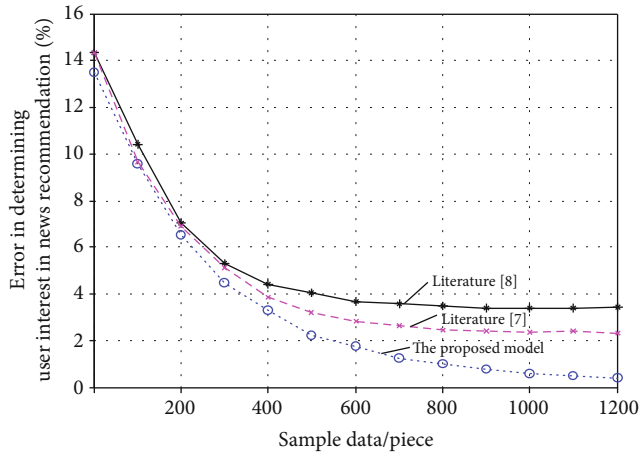


Figure 7: Error analysis of user interest determination in the news recommendations.

The experiment compares the errors of user interest in this system, literature [7] system, and literature [8] system in news recommendation. The results are shown in Figure 7.

The experimental results in Figure 7 show that the errors determined in the user interest of the paper system, literature [7] system, and literature [8] system are different. When the sample data size is constantly changing, the error is determined by user interest in news recommendations. Among them, the error of user interest in the news recommendation is about 0.1%, 2.1% in the literature [7] system, and about 3.9% in the news recommendation. This is the result of the analysis in the system design and verifies the effectiveness of the system.

4. Conclusion

In order to solve the problem of a deep network and personalized demand, resulting in the poor performance of the news recommendation system, a news recommendation system is proposed. First, we analyze the user news needs as the basis of the system design; the system module functions mainly include the network function module, database module, user management module, and news recommendation module. Among them, the user management module uses the deep network to set up the user news interest model, inputs the news data into the model, completes the personalized needs of the news, and realizes the design of the news recommendation system. The experimental results show that the proposed system is highly effective and has certain advantages. Although our method has achieved good experimental results, with the development of deep learning today, we should integrate the deep learning method into the algorithm of this paper. In the future, we will devote ourselves to this endeavor.

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 conflict of interest.

References

[1] K. Han, "Personalized news recommendation and simulation based on improved collaborative filtering algorithm," *Complexity*, vol. 2020, Article ID 8834908, 2020.

[2] H.-S. Sheu, Z. Chu, D. Qi, and S. Li, "Knowledge-guided article embedding refinement for session-based news

recommendation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 99, pp. 1–7, 2021.

[3] J. Fergie, A. Howard, L. Huang, and A. Srivastava, "Implementation experience with meningococcal serogroup B vaccines in the United States: impact of a nonroutine recommendation," *The Pediatric Infectious Disease Journal*, vol. 40, no. 3, pp. 269–275, 2021.

[4] C. Song, W. Liu, Z. Liu, and X. Liu, "User abnormal behavior recommendation via multilayer network," *PLoS One*, vol. 14, no. 12, article e0224684, 2019.

[5] A. Gcs and B. Ml, "Stacked DeBERT: all attention in incomplete data for text classification," *Neural Networks*, vol. 136, no. 1, pp. 87–96, 2021.

[6] M. A. Ibrahim, M. U. Ghani Khan, F. Mehmood, M. N. Asim, and W. Mahmood, "GHS-NET a generic hybridized shallow neural network for multi-label biomedical text classification," *Journal of Biomedical Informatics*, vol. 116, no. 1, p. 103699, 2021.

[7] Z. H. U. Wenyue, L. I. U. Wei, and L. I. U. Zongtian, "News personalized recommendation based on event ontology," *Computer Engineering*, vol. 45, no. 6, pp. 267–272, 2019.

[8] L. I. U. Hui, W. A. N. Cheng-feng, and W. U. Xiao-hao, "A hybrid recommendation model based on incremental collaborative filtering and latent semantic analysis," *Computer Engineering and Science*, vol. 41, no. 11, pp. 2033–2039, 2019.

[9] L. I. U. Jin-hui, C. U. I. Xiang-yang, Y. A. N. G. Fan, and L. I. U. Li-yan, "Design and implementation of a combined recommendation system based on spring boot and user portrait," *Electronic Component and Information Technology*, vol. 3, no. 5, pp. 24–29, 2019.

[10] C. Feng, M. Khan, A. U. Rahman, and A. Ahmad, "News recommendation systems-accomplishments, challenges & future directions," *IEEE Access*, vol. 8, pp. 16702–16725, 2020.

[11] A. Darvishy, H. Ibrahim, F. Sidi, and A. Mustapha, "HYPNER: a hybrid approach for personalized news recommendation," *IEEE Access*, vol. 8, pp. 46877–46894, 2020.

[12] G. D. S. P. Moreira, D. Jannach, and A. M. Da Cunha, "Contextual hybrid session-based news recommendation with recurrent neural networks," *IEEE Access*, vol. 7, pp. 169185–169203, 2019.

[13] N. J. Krishna, H. Purohit, and H. Rangwala, "Diversity-based generalization for neural unsupervised text classification under domain shift," vol. 15, no. 24, pp. 1145–1152, 2020, http://arxiv.org/abs/2002.10937.

[14] I. Chalkidis, M. Fergadiotis, S. Kotitsas, P. Malakasiotis, N. Aletras, and I. Androutsopoulos, "An empirical study on large-scale multi-label text classification including few and zero-shot labels," vol. 45, no. 2, pp. 123–128, 2020, http://arxiv.org/abs/2010.01653.

[15] I. Baldini, D. Wei, K. N. Ramamurthy, M. Yurochkin, and M. Singh, "Your fairness may vary: group fairness of pretrained language models in toxic text classification," vol. 45, no. 1, pp. 15–21, 2021, http://arxiv.org/abs/2108.01250.

[16] R. Asgarnezhad, M. Soltanaghaei, and S. A. Monadjemi, "An application of MOGW optimization for feature selection in text classification," *The Journal of Supercomputing*, vol. 16, no. 15, pp. 154–160, 2020.

[17] M. Oleynik, A. Kugic, Z. Kasáč, and M. Kreuzthaler, "Evaluating shallow and deep learning strategies for the 2018 n2c2 shared task on clinical text classification," *Journal of the American Medical Informatics Association*, vol. 26, no. 11, pp. 1247–1254, 2019.

[18] T. Igamberdiev and I. Habernal, "Privacy-preserving graph convolutional networks for text classification," vol. 7, no. 1, pp. 868–871, 2021, http://arxiv.org/abs/2102.09604.

[19] Q. I. U. Ning-jia, G. A. O. Peng, W. A. N. G. Peng, and T. A. O. Yue, "Research on ACO-WNB classification algorithm based on improved information gain," *Computer Simulation*, vol. 36, no. 1, pp. 295–299, 2019.