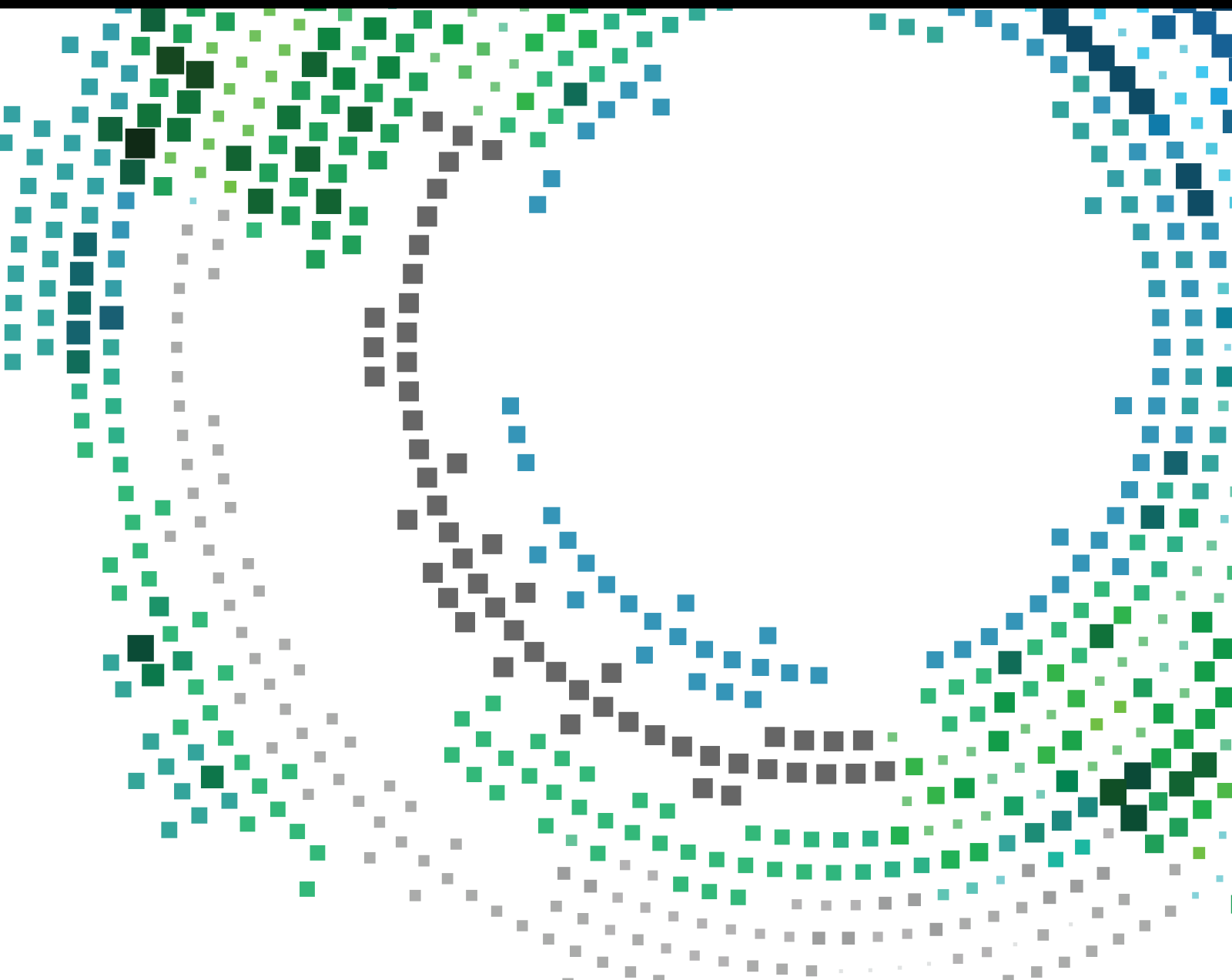


# Security Threats of Fog and Edge Computing Based Social Internet of Vehicles

Lead Guest Editor: Ke Gu

Guest Editors: Xiong Li, Kuo-Hui Yeh, Sheng Wen, and Fei Yu





---

# **Security Threats of Fog and Edge Computing Based Social Internet of Vehicles**

Mobile Information Systems

---

**Security Threats of Fog and Edge  
Computing Based Social Internet of  
Vehicles**

Lead Guest Editor: Ke Gu

Guest Editors: Xiong Li, Kuo-Hui Yeh, Sheng Wen,  
and Fei Yu



---

Copyright © 2021 Hindawi Limited. All rights reserved.

This is a special issue published in "Mobile Information Systems." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.







# Chief Editor

Alessandro Bazzi , Italy

## Academic Editors

Mahdi Abbasi , Iran  
Abdullah Alamoodi , Malaysia  
Markos Anastassopoulos, United Kingdom  
Marco Anisetti , Italy  
Claudio Agostino Ardagna , Italy  
Ashish Bagwari , India  
Dr. Robin Singh Bhadoria , India  
Nicola Biccocchi , Italy  
Peter Brida , Slovakia  
Puttamadappa C. , India  
Carlos Calafate , Spain  
Pengyun Chen, China  
Yuh-Shyan Chen , Taiwan  
Wenchi Cheng, China  
Gabriele Civitarese , Italy  
Massimo Condoluci , Sweden  
Rajesh Kumar Dhanaraj, India  
Rajesh Kumar Dhanaraj , India  
Almudena Díaz Zayas , Spain  
Filippo Gandino , Italy  
Jorge Garcia Duque , Spain  
Francesco Gringoli , Italy  
Wei Jia, China  
Adrian Kliks , Poland  
Adarsh Kumar , India  
Dongming Li, China  
Juraj Machaj , Slovakia  
Mirco Marchetti , Italy  
Elio Masciari , Italy  
Zahid Mehmood , Pakistan  
Eduardo Mena , Spain  
Massimo Merro , Italy  
Aniello Minutolo , Italy  
Jose F. Monserrat , Spain  
Raul Montoliu , Spain  
Mario Muñoz-Organero , Spain  
Francesco Palmieri , Italy  
Marco Picone , Italy  
Alessandro Sebastian Podda , Italy  
Maheswar Rajagopal, India  
Amon Rapp , Italy  
Filippo Sciarrone, Italy  
Floriano Scioscia , Italy

Mohammed Shuaib , Malaysia  
Michael Vassilakopoulos , Greece  
Ding Xu , China  
Laurence T. Yang , Canada  
Kuo-Hui Yeh , Taiwan

# Contents

---


## **Edge-Based Detection and Classification of Malicious Contents in Tor Darknet Using Machine Learning**

Runchuan Li, Shuhong Chen , Jiawei Yang, and Entao Luo  
Research Article (13 pages), Article ID 8072779, Volume 2021 (2021)



## **A Lightweight Authenticated Key Agreement Protocol Using Fog Nodes in Social Internet of Vehicles**

Tsu-Yang Wu , Xinglan Guo , Lei Yang , Qian Meng , and Chien-Ming Chen   
Research Article (14 pages), Article ID 3277113, Volume 2021 (2021)

## **Improving the Quality of Left-Behind Children's Participation in Sports through Wireless Network Monitoring**

Jinjin Zhao   
Research Article (10 pages), Article ID 3981893, Volume 2021 (2021)

## **An Intelligent Garbage Sorting System Based on Edge Computing and Visual Understanding of Social Internet of Vehicles**

Xuehao Shen, Yuezhong Wu , Shuhong Chen , and Xueming Luo  
Research Article (12 pages), Article ID 5231092, Volume 2021 (2021)

## Research Article

# Edge-Based Detection and Classification of Malicious Contents in Tor Darknet Using Machine Learning

Runchuan Li,<sup>1</sup> Shuhong Chen ,<sup>1</sup> Jiawei Yang,<sup>1</sup> and Entao Luo<sup>2</sup>

<sup>1</sup>School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China

<sup>2</sup>School of Electronics and Information Engineering, Hunan University of Science and Engineering, Yongzhou, Hunan 425199, China

Correspondence should be addressed to Shuhong Chen; shuhongchen@gzhu.edu.cn

Received 2 September 2021; Accepted 25 October 2021; Published 22 November 2021

Academic Editor: Ke Gu

Copyright © 2021 Runchuan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the increase of data in the network, the load of servers and communication links becomes heavier and heavier. Edge computing can alleviate this problem. Due to a sea of malicious contents in Darknet, it is of high research value to combine edge computing with content detection and analysis. Therefore, this paper illustrates an intelligent classification system based on machine learning and Scrapy that can detect and judge fleetly categories of services with malicious contents. Because of the nondisclosure and short survival time of Tor Darknet domain names, obtaining uniform resource locators (URLs) and resources of the network is challenging. In this paper, we focus on a network based on the Onion Router (tor) anonymous communication system. We designed a crawler program to obtain the contents of the Tor network and label them into six classes. We also construct a dataset which contains URLs, categories, and keywords. Edge computing is used to judge the category of websites. The accuracy of the classifier based on a machine learning algorithm is as high as 89%. The classifier will be used in an operational system which can help researchers quickly obtain malicious contents and categorize hidden services.

## 1. Introduction

The Darknet has a huge amount of data. Edge computing can process massive data on terminal devices and transfer the processed results to the server, which alleviates the computing pressure of servers and the load of communication links [1]. Tor (The Onion Router) Darknet [2], which is also known as onion network and dark web, is a network using anonymous communication technology [3]. It is hard to access hidden services and obtain resources from it without using specific software or a proxy agency. Their sites are not only not indexed by Google or other standard search engines but also invalidate quickly [4]. Due to the good concealment of Tor Darknet a lot of illegal contents exit from it, such as drugs, guns, and hacking technology. After the outbreak of COVID-19, many medical products and supplies also appeared in the Darknet market [5] that is not good for the stability of the society.

AlQahtani and El-Alfy [6] conducted extensive research on anonymous technology and the onion network. The strong concealment of Tor network was illustrated, but there were some defects in the design and implementation of Tor hidden service technology [7–9]. Furthermore, due to the special network structure and the characteristics of hiding identities on both sides of communication, the improved onion network technology was also applied to other applications such as the Internet of Vehicles (IoV) ad hoc network [10]. Because of a large number of high-quality resources and the difficulty of obtaining them from the dark web, the mining and analysis of Tor network resources has been a major research hotspot in the academic community.

There are many research methods for Tor network resources. Web crawler technology can improve the efficiency of obtaining network resources. Iliou et al. [11] and Monterrubio et al. [12] proposed a general crawling framework for automatically obtaining web resources in

the Tor network. Wang et al. [13] and He et al. [14] proposed a method to identify anonymous traffic in the Tor network. Biswas et al. [15] proposed a method for automatic recognition of services in the Tor network based on perceptual hashing, which identified services only through snapshots of services. In addition to the above methods, the analysis of web page text is also a commonly used research method [16]. It can adopt methods of data mining and machine learning to analyze the data in hacker forums and Darknet markets and quickly screen out relevant threat intelligence [17–19].

Before analyzing the content of the dark web, it is necessary to obtain a large number of URLs and classify the topic of the website, so as to conduct targeted research. Kan and Nguyen Thi obtained the theme of the website by analyzing URLs [20]. However, due to the particularity of the domain name in Tor Darknet, the same method cannot be used to directly determine the category of the website.

Al Nabki et al. [21] and Spitters et al. [22] comprehensively analyzed the hidden services in the Tor network based on the web page content and classified themes of websites. Biryukov et al. [23] obtained hidden service descriptors through port scanning and classified the content. They found that the content of Tor hidden services is diverse. Al Nabki et al. [24] divided the Tor network addresses into 26 categories manually and selected nine categories to be applied to training three different supervised classifiers. However, they did not integrate the crawling and analysis process to form a visual interactive system. Buldin and Ivanov [25] used the K-nearest neighbor algorithm to identify four categories of Darknet web pages. Graczyk and Kinningham [26] classified products in anonymous marketplaces based on the support vector machine model. However, the accuracy of their classifying models is about 79%.

This study focuses on six distinct categories of Tor Darknet web sites. They are “Counterfeit money,” “Counterfeit credit-cards,” “Cryptocurrency,” “Hacking,” and “Drugs,” respectively. These five categories do great harm to the society and are rich in resources in Tor Darknet. The rest of the classes are assigned to a 6th category which we called “Others.”

To obtain the content of the Tor website, a program based on the Scrapy framework was designed. Then, we created a dataset to train a classifier based on the K-nearest neighbor (KNN) technique using a web-text preprocessing algorithm to extract features from webpages that can highlight the main theme. The optimal parameter for the KNN model was chosen using cross validation and Grid Search.

Finally, we improved the accuracy of the classifier to 89%, which is 10% higher than that of the classifier based on the support vector machine (SVM) algorithm in [26]. Furthermore, we designed a system to automatically crawl and classify the content of websites that exist in Tor Darknet. It will help researchers more easily obtain lots of content and identify categories of websites in Tor Darknet.

## 2. Proposed Model for Tor Darknet Resource Detection in Edge Computing

*2.1. System Overview.* A system combining the functions of detecting and analyzing Tor Darknet resources is designed in our work. It enables researchers to easily acquire the contents of Tor Darknet websites and classify websites into unknown categories. As shown in Figure 1, the system is split into three modules.

The edge computing module is primarily responsible for detecting and preprocessing the Tor Darknet content. Traditional cloud computing systems are experiencing network latency and bandwidth congestion as a result of the enormous data created by Internet of Things (IOT) devices. Edge computing was born out of the need for huge IOT devices to network and the high demand for real-time performance of their applications. Major applications, services, and data storage are sunk to the network’s edge, bringing processing closer to the source of data. In the cloud computing paradigm, this will tackle issues such as excessive application delay and severe network load produced by enormous data being uploaded to the cloud data center for analysis. Web pages of URLs that are not categorized will be detected by the crawler. After preprocessing, the contents will be uploaded to a classifier, which relieves data processing pressure at the training module and decreases network transmission burden. The efficiency of accessing hidden services is limited due to the complexity of accessing the Tor network and the peculiarity of the routing protocol. Furthermore, the services generally have a short life cycle and are prone to malfunction. Therefore, web page content should be detected at edge devices, and the original data should then be processed into distinctive words that best describe the website category. The classification result may be returned more quickly and correctly after submitting the processed corpus to the classifier.

The classifier needs to be trained in the training module. To begin with, feature words in the dataset must be vectorized and weighted. The data is then split into two sets, a training set and a testing set, and they are used to train the classifier.

The output module is used to display the results. When unknown URLs are inputted, the crawler starts working and the data it collects is sent to the classifier for analysis. The system will display the classifier’s results and a performance report. Furthermore, users have the option of saving any object and generating word clouds.

*2.2. Domain Names Collection.* In Tor Darknet, a domain name’s complete format is “[digest].onion,” which is made up of two parts: the first [digest] is a random string of numbers mixed with English, and the second is a uniform suffix of Tor links, jsaljfslj4sfd5ad.onion, for example. It will not show any results when we search sites with the suffix “.onion.” Therefore, in order to classify the contents of Tor Darknet, domain names need to be obtained in various ways. In this paper, Tor domain names are collected in two ways: one is to collect them by Darknet directory sites; the other is

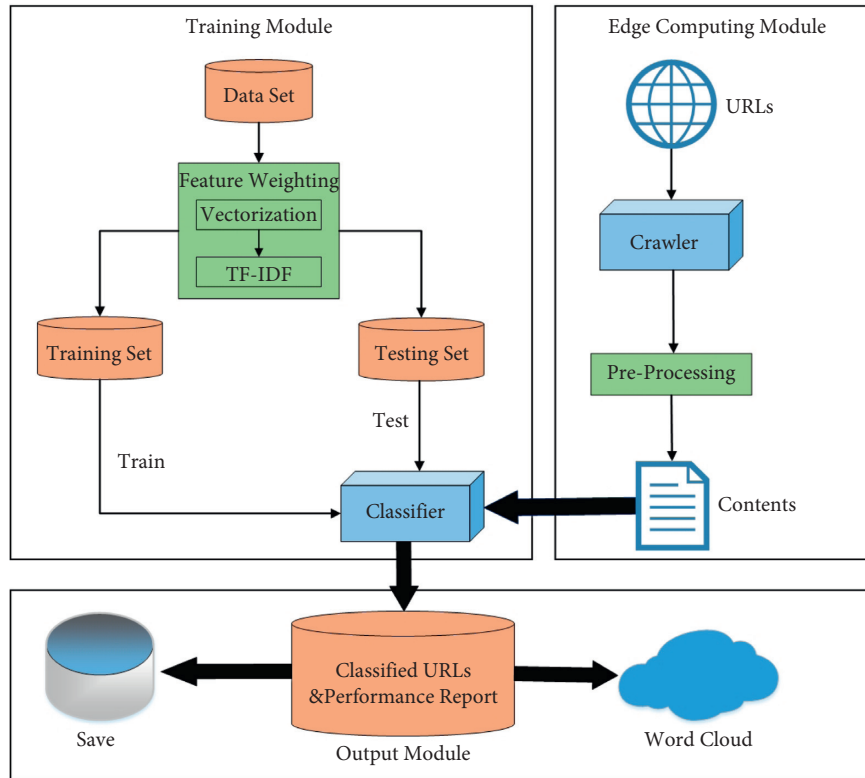


FIGURE 1: Tor Darknet detection and analysis system.

to use open datasets. URLs are mainly derived from a publicly available dataset called “Darknet Usage Text Addresses” (DUTA). It includes websites’ domain names, categories, and language types. However, because of the lack of text content related to the site, we were supposed to design crawlers to obtain them. The content would be treated as much as possible to be the words which can reflect web categories. Simultaneously, we built a new dataset for training a classification model.

2.3. *Communication Principle of Tor Darknet.* The term “hidden service” (HS) is used to describe a website that runs on Tor Darknet. Aside from the uniqueness of its domain name, the way it communities with others is also interesting. You need to run Tor agent software locally to access the network. Figure 2 and Algorithm 1 depict the specific communication procedure between a client and a hidden service in the network.

The network’s communication is established via a circuit made up of numerous routing nodes known as “Onion Routers” (OR). Following the launch of an HS, an Onion Agent (OA) will choose a router at random to serve as the Introduction Point Router (IPO), via which the hidden service will connect to the Tor network. The OA will then create a hidden service descriptor (HSD) including the IPO information, the timestamp, the HS public key, and other information, and upload it to a hidden service directory server (HSDS).

A user client connects to the network through the Onion Agent. The OA obtains the router information in the Tor

network from a directory server (DS) and chooses the best-performing routers to construct a communication circuit. By default, the client connects to the HSDS via 3-hop OR. According to the domain name address sent by the client, the server queries the corresponding hidden service descriptor and returns, and then, the client resolves the IPO address. The client’s OA chooses an OR as the Rendezvous Point Router (RPO) at random and transmits the RPO’s information to the HS through IPO. RPO will serve as the hub for anonymized data exchange between the client and the HS. After learning the information of RPO, HS builds a 6-hop link to form a communication circuit and starts data transmission with the client.

Network configuration is necessary before the client can access Tor Darknet. There are two methods to connect to the network: one is to use Tor browser and the other is to configure the proxy environment.

The socks protocol is used for network communication. However, some crawling modules do not support this protocol, so they cannot directly obtain and parse the response returned by the site.

In this experiment, combining with the scrapy framework, the original configuration was modified and the proxy conversion software Polipo was used to convert the socks protocol into the HTTP protocol, so as to achieve the acquisition of Tor Darknet resources. Figure 3 shows the agency conversion:

2.4. *Resource Detection for Tor Darknet.* We created a crawling program based on Scrapy framework for Tor

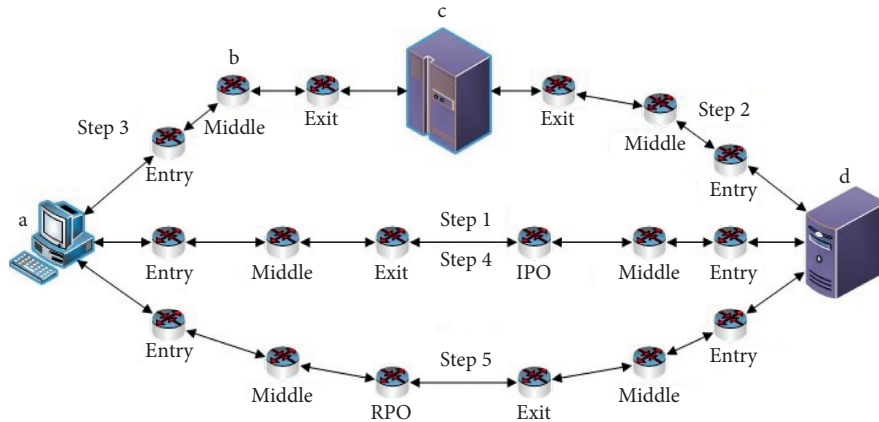


FIGURE 2: Communication between the client and HS. “a” is client, “b” is OR, “c” is HSDS, and “d” is HS. “Step  $n$ ” is the order of communication.

```

Input: URL
Output: STATUS
(1) Client sends request to DS
(2) Set  $R(OR[1], \dots, OR[N]) \leftarrow DS$  // DS returns the routing information to Client
(3) for  $i = 1$  to  $i = n$  do
(4) router whose performance is good will be selected
(5) end for
(6) STATUS = Client connects to HSDS and sends the URL to it
(7) if STATUS is failure then
(8) return false
(9) else
(10) Client get the information of IPO from HSDS
(11) end if
(12) for  $i = 1$  to  $i = n$  do
(13) Client selects a OR as RPO and sends its information to HS via IPO
(14) end for
(15) return true //data transmission can be started
    
```

ALGORITHM 1: Client connects to HS.

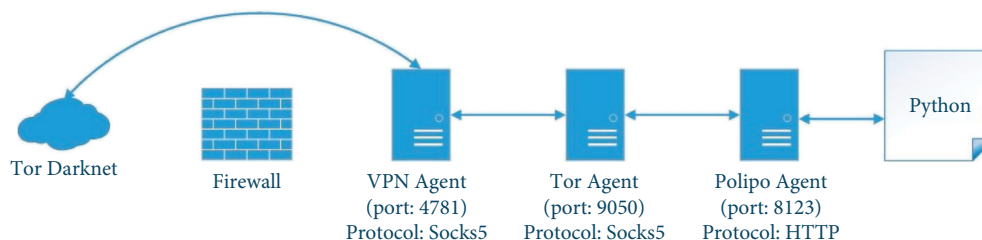


FIGURE 3: Agency conversion.

Darknet to accomplish automated access to resources due to the requirement to collect a huge number of URLs and the contents of each website.

The running process of the program is shown in Figure 4 and Algorithm 2.

A spider program reads all URLs in the list of “start\_urls” and sends them to the engine, the center of the whole framework, which handles the data flow between

components and triggers some operations. The engine will receive URLs to schedule, which will add URLs to the scheduling queue and wait for processing. After a URL is processed, the engine will receive a request sent by the schedule and trigger the downloader to work. After receiving a notification, the downloader will process a request according to the setting in downloader middleware and access Tor Darknet websites. The downloader will transmit a

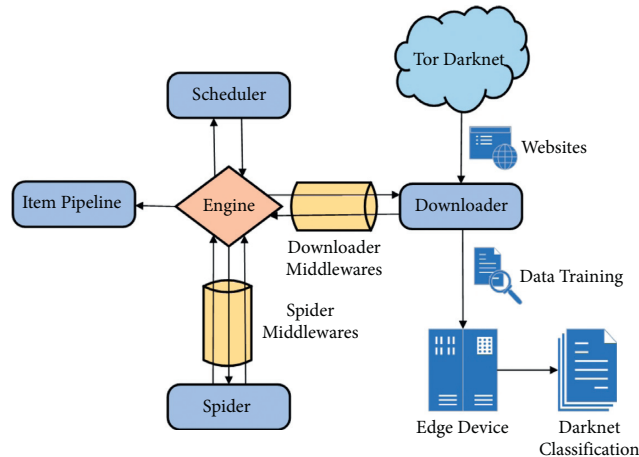


FIGURE 4: Framework of the crawling model.

processed request, namely, a response, to the spider for further processing. If the download fails, the engine will notify the schedule, then mark it, and reexecute the operation later.

When the spider receives a response, the “parse” function in the spider program will process the downloaded content according to the custom program. Then, items parsed in “parse” will be passed to the engine for further scheduling. The engine sends items to the item pipeline for data processing and storage in a file until the crawling of a URL is completed.

The URLs in the scheduling queue will then continue until all URLs have been processed. Scrapy does not repeat access to the site that has been visited, which solves the problem of multiple access to the same URL in the domain name set.

After the data is downloaded from Tor Darknet, the system will use the dataset to train a classification model and it will be downloaded to edge devices. Edge devices use the model to classify the content of unknown websites and finally obtain the categories of each website.

**2.5. Data Preprocessing.** The presentation of website page content and layout is realized through HTML code. After crawling a website to obtain texts, HTML elements in the texts need to be removed to filter out words. The data cleaning process is shown in Figure 5.

In Step 1, after the source code of a web page was downloaded, we parsed the HTML page content using “`lxml.html.document_from_string`” function from the LXML library. Then, “`lxml.html.clean.cleaner().clean_html()`” function was used to filter the script and HTML tags to obtain the text content displayed on the web page. There are many escape characters (ESC), carriage returns, tabs, and other meaningless characters in the content. Therefore, we combined them with Python’s operation on string types to replace them with a space.

In Step 4, the corpus had no html tags and looked more like regular text. Word formats, on the other hand, were inconsistent. This would reduce the effect of selecting feature words and increase the dimension of feature space. We used

the “`casefold()`” function to change all words to lowercase so that we could deal with them uniformly afterwards. All punctuation marks and Arabic numbers were filtered using an algorithm. We reduced inflection in words to their base forms using the Stemmer package, which helps to preprocess text, words, and documents for text normalization. We combine terms like “created” and “creates” into “create,” for example.

When all of the corpora’s formats were harmonized, they were expected to be processed further in Step 7. All the words that appear in the English stop word corpus were removed. These words, such as this, they, are, and so on, cannot represent any characteristics of websites. The last step was to remove strings that were longer than twelve or shorter than two. These characters are odd terms that do not correspond to the website’s subject.

The specific implementation method is shown in Algorithm 3. After the text content of each web page is cleaned, the corpus is integrated into a dataset containing URLs, categories, and key words. A machine learning algorithm, KNN, is then applied to such samples for the purpose of training a classifier in subsequent experiments.

### 3. Classification Model

For each type of website, there must be some words that highlight its characteristics. Therefore, after vectorizing the corpus, we combined the machine learning algorithm to train a classifier suitable for Tor Darknet websites.

**3.1. Text Vectorization.** Before the classification of web content, it is necessary to transform words based on text representation into a form that can be recognized and calculated. In other words, words need to be transformed into vectors and the calculation of similarity between text semantics is transformed into the calculation of distance between vectors.

If the correlation between words and web topic is measured directly according to word frequency, the

**Input:** Set *start\_urls* ( $u[1], \dots, u[n]$ )  
**Output:** Set *Content*

- (1) Queue  $Q = \text{Enqueue}(\textit{start\_urls})$
- (2) **while**  $Q$  not empty **do**
- (3)  $q = \text{Dequeue}(Q)$
- (4) **if**  $q$  has not been processed **then**
- (5)  $\textit{reponse} = \text{download}(q)$
- (6) **else**
- (7) CONTINUE
- (8) **end if**
- (9) **if** status of website = 200 **then**
- (10)  $\textit{items} \leftarrow \text{parse}(\textit{reponse})$
- (11) **else**
- (12) mark it and re-executed the operation later
- (13) CONTINUE
- (14) **end if**
- (15)  $\textit{Content} \leftarrow \textit{Content} \cup \{\textit{items}\}$
- (16) **end while**
- (17) **return** *Content*

ALGORITHM 2: Data crawling.

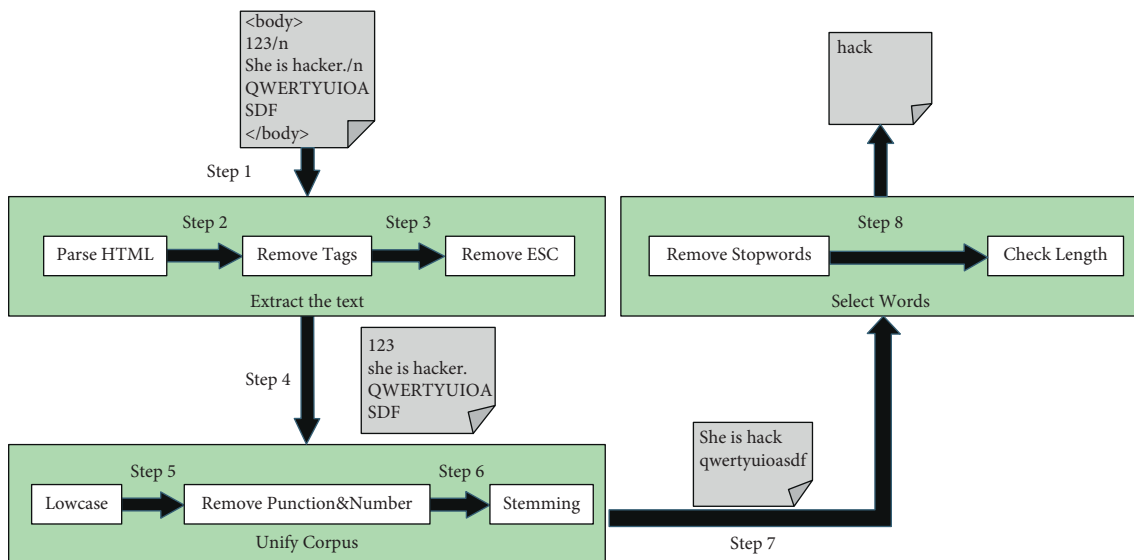


FIGURE 5: Data cleaning process.

**Input:** Web Set  $D(D[0], \dots, D[n])$   
**Output:** Corpus set  $W(W[0], \dots, W[n])$

- (1) **for**  $i = 0$  TO  $i = n$  **do**
- (2)  $\textit{content} = \text{obtain the HTML content of } D[i]$
- (3) use "lxml()" function to parse the content, then remove HTML tags, script and so on
- (4)  $\textit{text} = \text{preserve the text content displayed on the page}$
- (5) **for**  $ch$  in  $\textit{text}$  **do**
- (6) **if**  $ch = '\n'$  or  $ch = '\t'$  **then**
- (7)  $ch = ''$
- (8) **end if**
- (9) **end for**
- (10) Lowercase all English words
- (11) **for**  $\textit{temp}$  in  $\textit{text}$  **do**

ALGORITHM 3: Continued.



```

(12) if temp is a punctuation or a number then
(13) temp = ""
(14) end if
(15) end for
(16) PorterStemmer(text)//Unify all word formats
(17) word_list(wl[0], ..., wl[len1]) = text.split("")
(18) for i = 0 TO i = len1 do
(19) if word_list[i] ∈ stopwords(sw[0], ..., sw[m]) or 2 < len(word_list[i]) < 12 then
(20) delete_wordlist[i]
(21) end if
(22) end for
(23) SET W ← word_list(fw[0], ..., fw[len2]).join("")//Words are concatenated to strings
(24) end for
(25) return W

```

ALGORITHM 3: Data cleaning algorithm.

measurement results will be related to the size of the web page and sometimes the words with a high degree of correlation cannot be really calculated.

For example, “the” appears frequently on one page, but it also appears on other pages, so it is not of much importance. “Hack” appears relatively infrequently on the page, but it exists only on that page. Therefore, it is more important than “the” in the page.

Therefore, a better weighted method should be adopted to calculate the words that are more relevant to the topic of the web page. Combined with the idea of TF-IDF (term frequency-inverse document frequency), this paper adopts a new weighting method to calculate the value of word vectors.

After accessing all the domain names, we get contents of each website and then build a web page set  $W = (d_1, d_2, \dots, d_k)$ , where  $d_k$  represents the content contained in the  $k$ th web page.

Firstly, the frequency weighted method is used to count the occurrence times of all feature words in the corresponding web page text, as shown in formula (1), where  $n_{i,j}$  represents the occurrence times of feature words  $t_i$  in web page  $d_j$ . And,  $t_i$  represents the  $i$ th word in the web page.

$$F_{tor}(t_i, d_j) = n_{i,j}. \quad (1)$$

The number of web pages containing the feature word  $t$  is represented by  $G'_{tor}(t_i, d_j)$ .

According to formula (2), where  $k$  is the size of the web page set and  $\varepsilon$  denotes the smoothing factor which ensures that the denominator is not zero, we calculate the inverse document frequency (IDF) of the feature word  $t$  in the web page  $d$  which is represented by  $G_{tor}(t_i, d_j)$ .

$$G_{tor}(t_i, d_j) = 1 + \log \frac{\varepsilon + k}{\varepsilon + G'_{tor}(t_i, d_j)}. \quad (2)$$

Then, we multiply the two values of  $F_{tor}(t_i, d_j)$  and  $G_{tor}(t_i, d_j)$  to obtain the TF-IDF value of the feature word  $t$  in the web page  $d$ , as shown in

$$H_{tor}(t_i, d_j) = F_{tor}(t_i, d_j) \times G_{tor}(t_i, d_j). \quad (3)$$

Finally, the TF-IDF values of all feature words are normalized according to formula (4). The denominator is the square root of the sum of the weighted values of all words in web page  $d$ , and the numerator is the weighted values of all feature words.

$$H_{norm}(t, d) = \frac{H_{tor}(t, d)}{\sqrt{\sum_{i=1}^n H_{tor}^2(t_i, d)}}. \quad (4)$$

After calculating the weighted values of all feature words, generally, the higher the value, the better the feature of the web page.

As shown in Figure 6, a web page set has two samples. Each value in rectangles is the weighted value of a word. The content of page A is “the hack hack” and page B is “the drug drug.” According to the above formulas, the weighted value of each word in page A is calculated. Conspicuously, the weighted value of “hack” is higher than other words in the same page and “hack” is the word that can present the topic of the web page better.

### 3.2. Darknet Classification Model Based on KNN Algorithm.

KNN (K-nearest neighbor) is a well-known supervised machine learning classification method with a well-developed theory and intuitive reasoning. After calculating the distance between “data to be categorized” and “samples of known category,” the samples are classified by comparing the distance. Customize the K value, and choose K examples that are the most similar to the samples in the training set to be categorized. The proportion of the K sample categories is used to determine the target category of the sample to be categorized.

Assuming that the number of all feature words in web page set  $W$  is  $n$ , the text content of web page  $d_k$  is expressed as an  $n$ -dimensional vector  $(w_1, w_2, \dots, w_n)$ ,  $d_k \in W$ , where  $w_n$  represents the weight of the feature in the text. Calculate the distance between page  $x$  and page  $y$  according to

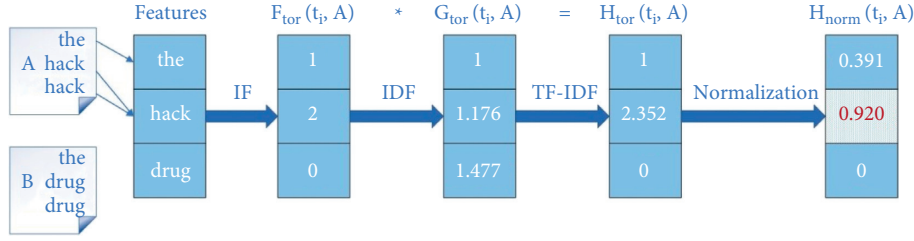


FIGURE 6: Text weighting process.

formula (5), where  $w_i^{(x)}$  represents the weight of the  $i$ th feature in the text content of page  $x$ .

$$\text{Distance}(x, y) = \sqrt{\sum_{i=1}^n (w_i^{(x)} - w_i^{(y)})^2}. \quad (5)$$

As shown in Figure 7, taking the two-dimensional space as an example, the square is a web page to be predicted. The triangle and circular are two different categories of websites. Calculate the distance of each sample point according to formula (5). Then, we take  $K = 3$  samples closest to the square. Because the number of triangles in the circular area is the largest, the predicted square category is consistent with the triangle.

The implementation method of applying KNN algorithm to Tor Darknet classification is shown in Algorithm 4.

**3.3. Model Optimization.** In the KNN classification model, the hyperparameter  $K$  of the algorithm will affect the prediction result. When we choose a value of  $K$  too small, it is easy to lead to overfitting of the classification model. On the contrary, too large may lead to underfitting. However, there is no proper theory to guide the selection of the  $K$  value, which can be selected by manual experience and then determined by the cross-validation method.

Grid Search is a method of adjusting parameters. It traverses all candidate parameters and tries every possibility. The best-performing parameter is the best parameter we want to apply to the classifier. As shown in formula (6),  $m_i$  represents the model whose value of  $K$  is  $i$ ,  $acc(m_i)$  represents the accuracy of model  $i$ , and  $M_{best}$  represents the highest accuracy of the best model.

$$M_{best} = \text{Max}(acc(m_1), acc(m_2), \dots, acc(m_i)). \quad (6)$$

Cross validation, also known as Cyclic Validation, is a method of model effect evaluation to avoid one-sided results caused by a single test. The concept of validation set is introduced to evaluate the accuracy of the model after tuning parameters, but it does not participate in training, which will more objectively evaluate the matching degree between the data outside the training set and the target attribute.

As shown in Figure 8, the dataset was divided into ten groups for ten model evaluations. Each time, one group of the dataset was taken as a validation set and the remaining nine groups were taken as the training set. As each validation set is different, the experiment will produce ten models and

their accuracy is  $E_i$ . The average accuracy of ten models  $acc(K)$  is the final accuracy of the classifier whose hyperparameter is  $K$ , as shown in formula (7).

Combined with the Grid Search and the 10-fold cross validation, the accuracy  $acc(K)$  is applied to the calculation of formula (6) to select the model with the highest accuracy.

$$acc(K) = \frac{1}{10} \sum_{i=1}^{10} E_i. \quad (7)$$

## 4. Experimental Analysis

Python is the best choice for data crawling and analysis. It is favored by many developers because of its simplicity and powerful features. Therefore, in order to achieve our design goal, our experiments were developed based on Python 3.7 under Windows 10 according to the actual situation.

**4.1. KNN Model Based on Frequency Weighting.** In this experiment, we utilized the Grid Search to find an appropriate  $K$  on the scale of one to seven. Then, the results of KNN classification models with different parameters were evaluated in combination with the cross validation of ten folds. As shown in Figure 9, for models with the same  $K$  value, different partitions of the dataset would result differently, so the mean value was finally needed for comparison.

As shown in Figure 10, the black bar is the result of frequency weighting. The white bar is the result of TF-IDF weighting. Each bar represents the accuracy of the model with different values of  $K$ . The average accuracy is the highest when the value of  $K$  is three. Table 1 shows the evaluation report of the classification model whose  $K$  is three. The classification accuracy is only 0.78. This accuracy is too low to be applied to the system, and there will be a large error, so the model still needs to be improved.

**4.2. KNN Model Based on TF-IDF Weighting.** Figure 11 is the accuracy of 10-fold cross validation after weighting the feature vectors with TF-IDF.

As shown in Figure 10, after being weighted by TF-IDF, the results of classification models with different  $K$  values are compared by Grid Search. When the  $K$  value is four, the average accuracy of the model is the highest. The accuracy of this classifier reached 0.89. Table 2 shows the report of this model whose hyperparameter is four.

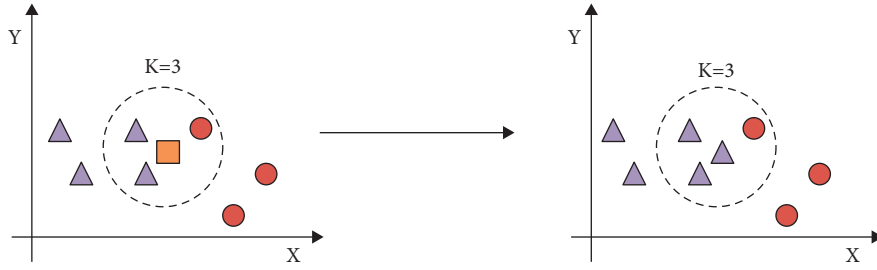


FIGURE 7: Classification principle of KNN algorithm.

```

Input: TRAINSET ( $T[1], \dots, T[n]$ ), Web page  $D$  to be predicted
Output: the category of  $D$ 
(1) for  $i = 1$  to  $i = n$  do
(2) Calculate the Euclidean distance between each sample in TRAINSET and  $D$ 
(3)  $distance = (d[1], \dots, d[n])$ 
(4) end for
(5)  $sorted\_dis = (distance)$  //Sort in ascending order
(6)  $sample(1, \dots, k) = \text{minimum}(sorted\_dis)$  //select  $k$  samples that have the shortest distance
(7) set  $classes = \text{count}(sample)$  //Count the number of each category in  $k$  samples
(8)  $category = \text{max}(classes)$  //The category with the largest output number is the predicted category
(9) return category
    
```

ALGORITHM 4: KNN classification algorithm.

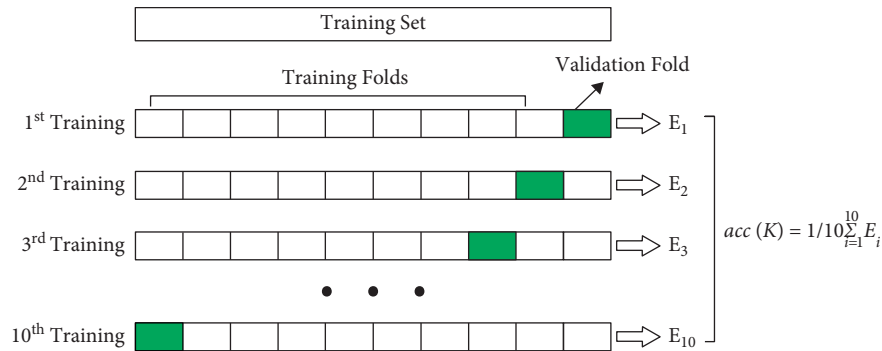


FIGURE 8: Ten-fold cross validation.

As shown in Table 3 and repeated experiments, the classification model after TF-IDF feature vectorization has higher accuracy and the performance is optimal when the hyperparameter  $K$  is four. Therefore, this model will be used in the subsequent system for its analysis function.

**4.3. Comparison of Relevant Research Methods and Results.** As shown in Table 4, the study in [24] manually marked 26 categories of domain names and selected 9 categories to apply in the training of the classification model. It adopted a Naive Bayes (NB) model based on TF-IDF weighting. Finally, the accuracy of cross validation of the model is 0.86.

The study in [25] categorized four types of illegal web pages based on the KNN algorithm. The accuracy for the test

sample is 0.80. The work in [26] designed a model to automatically categorized products for anonymous marketplaces. It extracted features using TF-IDF and then selected features from the text using principal component analysis. It categorized 12 product categories using the SVM method, with a model accuracy of up to 0.79.

In this paper, we categorized six types of websites and four was finally determined as the optimal hyperparameter of the classification algorithm by combining Grid Search with cross validation. The data was applied to train the KNN classification model based on this parameter, and we improved the accuracy to 0.89 finally.

**4.4. System Function Test.** A file is selected to store the domain names after the network environment has been set

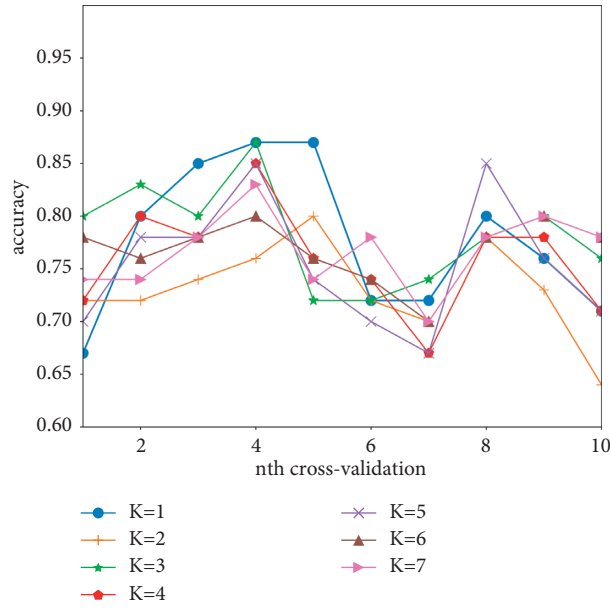


FIGURE 9: Ten-fold cross validation with frequency weighting.

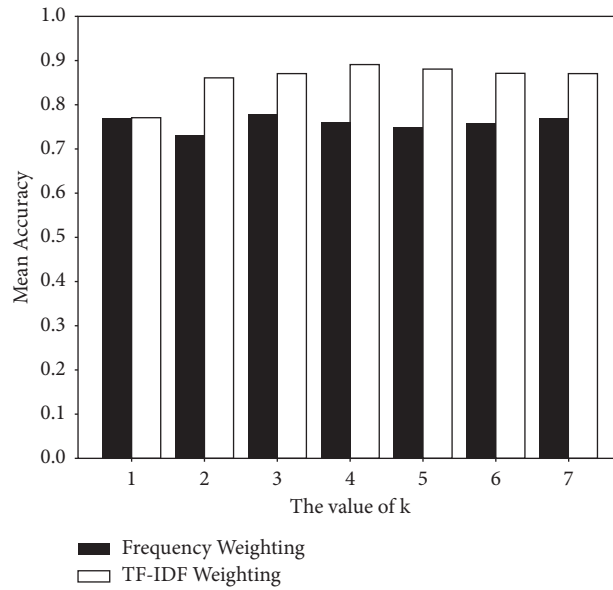


FIGURE 10: Cross-validation accuracy of model based on different weighting methods.

TABLE 1: Report of the KNN model which is weighted by frequency.

|                          | Precision | Recall | F1-score | Support |
|--------------------------|-----------|--------|----------|---------|
| Counterfeit credit-cards | 0.76      | 0.87   | 0.81     | 30      |
| Counterfeit money        | 1.00      | 0.83   | 0.91     | 12      |
| Cryptocurrency           | 0.93      | 0.79   | 0.85     | 47      |
| Drugs                    | 0.78      | 0.67   | 0.72     | 21      |
| Hacking                  | 0.30      | 0.70   | 0.42     | 10      |
| Other                    | 0.84      | 0.78   | 0.81     | 63      |
| Accuracy                 |           |        | 0.78     | 183     |
| Macro                    | 0.77      | 0.77   | 0.75     | 183     |
| Weighted avg             | 0.80      | 0.78   | 0.80     | 183     |

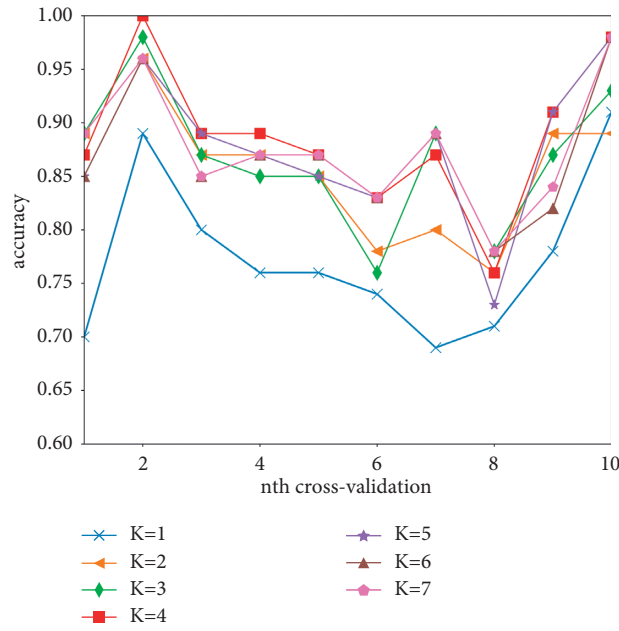


FIGURE 11: 10-fold cross validation with TF-IDF weighting.

TABLE 2: Report of the KNN model which is weighted by TF-IDF.

|                          | Precision | Recall | F1-score | Support |
|--------------------------|-----------|--------|----------|---------|
| Counterfeit credit-cards | 0.88      | 0.90   | 0.89     | 31      |
| Counterfeit money        | 0.33      | 0.50   | 0.40     | 4       |
| Cryptocurrency           | 1.00      | 0.89   | 0.94     | 38      |
| Drugs                    | 0.89      | 0.73   | 0.80     | 22      |
| Hacking                  | 1.00      | 0.77   | 0.87     | 13      |
| Other                    | 0.87      | 0.96   | 0.91     | 76      |
| Accuracy                 |           |        | 0.89     | 184     |
| Macro                    | 0.83      | 0.79   | 0.80     | 184     |
| Weighted avg             | 0.90      | 0.89   | 0.89     | 184     |

TABLE 3: Comparison of the best performance of KNN models based on different methods and sets.

|                | Frequency | TF-IDF |
|----------------|-----------|--------|
| Test set       | 0.80      | 0.89   |
| Validation set | 0.78      | 0.89   |

TABLE 4: Comparison of methods and results between this paper and others.

| Author                      | Accuracy | Method | Category |
|-----------------------------|----------|--------|----------|
| Al Nabki et al. [24]        | 0.86     | NB     | 9        |
| Buldin and Ivanov [25]      | 0.80     | KNN    | 4        |
| Graczyk and Kinningham [26] | 0.79     | SVM    | 12       |
| Li et al.                   | 0.89     | KNN    | 6        |

up. After that, it will be able to access all domain names. The deactivated and visited domain names will also not be used again.

Crawling contents will be showed in the display area of classification results, as shown in Table 5. The contents of data are showed in part.

TABLE 5: Partial result after the system completed the detection in the Darknet.

| Category                 | Url                               | Data   |
|--------------------------|-----------------------------------|--|
| Cryptocurrency           | http://grams7yngnpr5rzf.onion     | helix light gram learn lingo buy                           |
| Counterfeit credit-cards | http://aaaajqiyzj34rhjm.onion     | bring dream life low price                                 |
| Counterfeit money        | http://countervcgcdjp2y.onion/    | counterfeit usd qualiti usd                                |
| Counterfeit money        | http://fixedlwg3burzts.onion      | match european nation minor                                |
| Cryptocurrency           | http://22222222gib3ywf6.onion     | bitcoin quotation video onion site                         |
| Cryptocurrency           | http://5ifblitg2ywjjo2fgt.onion   | sourc bitcoin mixer quotation                              |
| Cryptocurrency           | http://3fjgpldvld7k7im.onion      | onion dir adult oniondir                                   |
| Cryptocurrency           | http://btce7mfhnhdkg5k.onion      | cost video xonion onion porn site                          |
| Drugs                    | http://napuzdankhadou3e.onion     | dank hank weed logo gram amnesia haze sample               |
| Drugs                    | http://smokerhxeb3tc6cy.onion/    | smoke finest organ cannabi buy weed                        |
| Hacking                  | http://hacker05xh5qtrrzmma.onion/ | rentahack hire hacker job imagin ddo                       |
| Hacking                  | http://beast7ruvpc3qjvh.onion     | itbazz hackasaserv hack ransomwar malwarew                 |
| Hacking                  | http://tinhat233xymse34.onion     | darknet message hack broken statist unpack introduct setup |
| Hacking                  | http://agenttoe2dlvxdei.onion     | price hack servic e-mail account devic                     |
| Hacking                  | http://underdj5ziov3ic7.onion     | tor commun perman move onion onion deprec                  |
| Other                    | http://a64r6sizrpegnggoj.onion    | cryptostorm commun forum cryptostorm forum                 |
| Other                    | http://darkw4u3xkeb5pzn.onion     | eva franco matt dark content dark content                  |
| Other                    | http://iz56hcijqh5uh5u.onion      | celebr underground celebr underground bitcoi               |
| Other                    | http://xujnrmw3lkpyj57r.onion     | share news casino game multiplay game poker                |



FIGURE 12: Word cloud.

As shown in Figure 12, any domain name can be selected on the system interface and the system will automatically generate and display the word cloud images of such websites.

This function is helpful for researchers to intuitively obtain website feature words and create better visual effects.

## 5. Conclusions

Hidden services generally only retain service status for a limited amount of time in order to avoid being tracked, resulting in frequent domain address changes and a short lifespan. Because it requires special software and data must be encrypted and decrypted as it goes via each node on the communication circuit, access to Tor Darknet is sluggish.

For the reasons stated above, manually classifying websites creates a significant amount of labor. Furthermore, the pertinence is so low that it is impossible to rapidly reach a certain domain name type.

As a result, we created a visual interactive system based on edge computing that may assist researchers in swiftly obtaining content and classifying website categories. The classifying model's accuracy is as high as 89%, which will increase researchers' efficiency in identifying illegal websites

and is of significant practical use. Furthermore, there are only six categories in this experiment due to the limited number of domain names already gathered. In the future, we will collect more domain names of other categories and obtain the feature words of different categories to expand the dataset which will contain more website categories.

## Data Availability

The dataset used in this experiment is DUTA-10k, which is a classic open-source dataset for collecting Darknet addresses and can be downloaded from related websites. The dataset used in this experiment is downloaded from the GVIS platform (<http://gvis.unileon.es/dataset/duta-darknet-usage-text-addresses-10k/>).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1005804, in part by the National Natural Science Foundation of China under Grants 61632009 and 62172159, in part by the Guangdong Provincial Natural Science Foundation under Grant 2017A030308006, and in part by the Natural Science Foundation of Hunan Province under Grant 2021JJ30294.

## References

- [1] J. Bellendorf and Z. Á. Mann, "Classification of optimization problems in fog computing," *Future Generation Computer Systems*, vol. 107, pp. 158–176, 2020.
- [2] S. Paul, D. Roger, and N. Mathewson, "Tor: the second-generation onion router," in *Proceedings of the Usenix Security Symposium*, pp. 303–320, San Diego, CA, USA, August 2004.

- [3] M. Edman and B. Yener, "On anonymity in an electronic society," *ACM Computing Surveys*, vol. 42, no. 1, pp. 1–35, 2009.
- [4] G. Owenson, S. Cortes, and A. Lewman, "The darknet's smaller than we thought: the life cycle of tor hidden services," *Digital Investigation*, vol. 27, no. 17–22, 2018.
- [5] R. Broadhurst, M. Ball, and C. J. Jiang, "Availability of COVID-19 related products on tor darknet markets," *Australasian Policing*, vol. 12, no. 3, pp. 8–13, 2020.
- [6] A. A. AlQahtani and E.-S. M. El-Alfy, "Anonymous connections based on onion routing: a review and a visualization tool," *Procedia Computer Science*, vol. 52, pp. 121–128, 2015.
- [7] Q. Tan, Y. Gao, J. Shi, X. Wang, B. Fang, and Z. Tian, "Toward a comprehensive insight into the eclipse attacks of tor hidden services," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1584–1593, 2018.
- [8] A. Biryukov, I. Pustogarov, and R.-P. Weinmann, "Trawling for tor hidden services: detection, measurement, deanonymization," in *Proceedings of the 2013 IEEE Symposium on Security and Privacy*, pp. 80–94, IEEE, San Francisco, CA, USA, May 2013.
- [9] T. Elahi, K. Bauer, M. AlSabah, D. Roger, and I. Goldberg, "Changing of the guards: a framework for understanding and improving entry guard selection in tor," in *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society, WPES '12*, pp. 43–54, New York, NY, USA, October 2012.
- [10] M. Sayad Haghighi and Z. Aziminejad, "Highly anonymous mobility-tolerant location-based onion routing for vanets," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 2582–2590, 2019.
- [11] C. Iliou, K. George, T. Tsirikia, S. Vrochidis, and I. Kompatsiaris, "Hybrid focused crawling on the surface and the dark web," *EURASIP Journal on Information Security*, vol. 2017, no. 1, pp. 1–13, 2017.
- [12] S. Monterrubio, J. Naranjo, L. Lopez, and A. Caraguay, "Black widow crawler for tor network to search for criminal patterns," in *Proceedings of the 2021 Second International Conference on Information Systems and Software Technologies (ICI2ST)*, vol. 1, pp. 108–113, Los Alamitos, CA, USA, March 2021.
- [13] L. Wang, H. Mei, and V. S. Sheng, "Multilevel identification and classification analysis of tor on mobile and pc platforms," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1079–1088, 2020.
- [14] G.-F. He, M. Yang, J.-Z. Luo, and L. Zhang, "Online identification of tor anonymous communication traffic," *Ruanjian Xuebao/Journal of Software*, vol. 24, no. 3, pp. 540–556, 2013.
- [15] R. Biswas, V. González-Castro, E. Fidalgo, and E. Alegre, "Perceptual image hashing based on frequency dominant neighborhood structure applied to tor domains recognition," *Neurocomputing*, vol. 383, pp. 24–38, 2020.
- [16] P. Kaur, "Web content classification: a survey," *International Journal of Computer Trends and Technology*, vol. 10, no. 2, pp. 97–101, 2014.
- [17] S. Anna, A. Bessi, S. Damodaran, P. Shakarian, K. Lerman, and E. Ferrara, "Early warnings of cyber threats in online discussions," in *Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 667–674, IEEE, New Orleans, LA, USA, November 2017.
- [18] D. Rhumorbarbe, L. Staehli, J. Broséus, Q. Rossy, and P. Esseiva, "Buying drugs on a darknet market: a better deal? studying the online illicit drug market through the analysis of digital, physical and chemical data," *Forensic Science International*, vol. 267, pp. 173–182, 2016.
- [19] M. Ball and R. Broadhurst, "Data capture and analysis of darknet markets," 2021, <https://ssrn.com/abstract=3344936>.
- [20] M.-Y. Kan and H. O. Nguyen Thi, "Fast webpage classification using URL features," in *Proceedings of the 14th ACM International Conference on Information and Knowledge Management, CIKM '05*, pp. 325–326, New York, NY, USA, October 2005.
- [21] M. W. Al-Nabki, E. Fidalgo, E. Alegre, and L. Fernández-Robles, "ToRank: identifying the most influential suspicious domains in the Tor network," *Expert Systems with Applications*, vol. 123, pp. 212–226, 2019.
- [22] M. Spitters, S. Verbruggen, and M. Van Staalduinen, "Towards a comprehensive insight into the thematic organization of the tor hidden services," in *Proceedings of the 2014 IEEE Joint Intelligence and Security Informatics Conference*, pp. 220–223, IEEE, Washington, DC, USA, September 2014.
- [23] A. Biryukov, I. Pustogarov, F. Thill, and R.-P. Weinmann, "Content and popularity analysis of tor hidden services," in *Proceedings of the 2014 IEEE 34th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pp. 188–193, IEEE, Madrid, Spain, June 2014.
- [24] M. W. Al Nabki, E. Fidalgo, E. Alegre, and I. de Paz, "Classifying illegal activities on tor network based on web textual contents," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pp. 35–43, Valencia, Spain, April 2017.
- [25] I. D. Buldin and N. S. Ivanov, "Text classification of illegal activities on onion sites," in *Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pp. 245–247, Moscow, Russia, January 2020.
- [26] M. B. Graczyk and K. Kinningham, "Automatic product categorization for anonymous marketplaces," 2015, [http://cs229.stanford.edu/proj2015/184\\_report.pdf](http://cs229.stanford.edu/proj2015/184_report.pdf).



## Research Article

# A Lightweight Authenticated Key Agreement Protocol Using Fog Nodes in Social Internet of Vehicles

Tsu-Yang Wu <sup>1</sup>, Xinglan Guo <sup>1</sup>, Lei Yang <sup>1</sup>, Qian Meng <sup>1</sup> and Chien-Ming Chen <sup>2</sup>

<sup>1</sup>College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

<sup>2</sup>Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin, China

Correspondence should be addressed to Chien-Ming Chen; [chienmingchen@ieee.org](mailto:chienmingchen@ieee.org)

Received 2 September 2021; Accepted 27 October 2021; Published 17 November 2021

Academic Editor: Ke Gu

Copyright © 2021 Tsu-Yang Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, there has been rapid growth in the Internet of things, the Internet of vehicles, fog computing, and social Internet of vehicles (SIoV), which can generate large amounts of real-time data. Now, researchers have begun applying fog computing to the SIoV to reduce the computing pressure on cloud servers. However, there are still security challenges in SIoV. In this paper, we propose a lightweight and authenticated key agreement protocol based on fog nodes in SIoV. The protocol completes the mutual authentication between entities and generates the session key for subsequent communication. Through a formal analysis of the Burrows–Abadi–Needham (BAN) logic, real-oracle random (ROR) model, and ProVerif, the security, validity, and correctness of the proposed protocol are demonstrated. In addition, informal security analysis shows that our proposed protocol can resist known security attacks. We also evaluate the performance of the proposed protocol and show that it achieves better performance in terms of computing power and communication cost.

## 1. Introduction

With the popularization and development of the world wide web, the Internet of things (IoT) [1–3], which is a network of Internet extension and expansion, has emerged. With the continuous development of IoT applications, a “social network of intelligent objects” called social Internet of things (SIoT) [4] has been formed. Internet of vehicles (IoV) [5] is an extension of the concept of SIoT. IoV can realize network connections between the vehicle and vehicle (V2V), vehicle and infrastructure (V2I), and vehicle and pedestrian (V2P) and collect and share the key road information. With the rapid development of network and sensor technology, social connection in urban transportation systems is necessary, so social Internet of vehicles (SIoV) is produced [6–8]. SIoV is an application of SIoT in the field of vehicles and is a combination of vehicular ad hoc networks (VANET) and mobile networks, and it can generate a large amount of real-time data. In SIoV, intelligent vehicles can establish social relationships with other objects and form a specific social network.

For cloud computing processing of road real-time data, there are some problems associated with network delays, transmission efficiency, and others. Because the distance between the cloud computing server and vehicles is far, and the number of vehicles is increasing, the cloud server needs to process more real-time data, which increases the computing burden. Therefore, researchers have introduced fog computing to reduce the computational burden on cloud servers. The data, processing, and application of fog computing are stored on scattered and weak devices, almost outside the cloud, so the computing power is not strong. It can help the cloud server process some data that are not necessary or urgent at that moment. If it encounters data that it cannot process, it reports to the cloud server. Fog nodes can detect unsafe driving behavior in time, issue early warnings for the behavior, and provide the corresponding punishment when necessary. The application of fog node in IoT and IoV environments was mentioned in the articles [9–13]. In 2016, Azimi et al. [11] proposed a medical warning system in IoT based on fog computing. In 2019, Ismail et al. [12] proposed an implication of fog computing on the IoT.



In 2019, Ma et al. [10] proposed a protocol for fog-based IoV networks, which realized authenticated key agreement. In 2021, Eftekhari et al. [9] proposed a pairwise secret key agreement protocol using fog-based IoV, which was a three-part authentication protocol. The SIOV typical architecture based on fog nodes is shown in Figure 1.

However, in the SIOV environment based on fog nodes, there are still great risks related to security issues. For example, it is very challenging to ensure the confidentiality and privacy of data transmission based on ensuring the security of devices deployed on the network edge. The data transmitted through the public channel usually includes sensitive information such as the personal information of vehicle users, which needs to be kept secret. Recently, Ahmed et al. [6] researched a key agreement protocol for V2G in the SIOV environment, which was a two-party authentication protocol. The protocol [6] was based on an elliptic-curve (ECC) point multiplication and had a large computational cost. This shortcoming leads us to propose a more effective protocol.

We propose a lightweight and authenticated key agreement protocol based on three parties using fog nodes in an SIOV environment. In this protocol, vehicles and fog nodes authenticate each other with the help of a cloud server (CS) and establish a secure session key. Owing to the weak computing power of fog nodes, our protocol only uses lightweight primitives, such as hash function and XOR operation. Through formal analysis of the Burrows-Abadi-Needham (BAN) logic, real-oracle random (ROR) model, and ProVerif, the security, validity, and correctness of the proposed protocol are demonstrated. In addition, informal security analysis shows that our proposed protocol can resist known security attacks. We also evaluate the performance of the proposed protocol and show that it has better performance in terms of computing power and communication cost.

The rest of the paper is structured as follows: in Section 2, we review recent research results. The details of our proposed agreement are in Section 3. In Section 4, we use BAN, ROR, and ProVerif to verify the security, validity, and correctness of the proposed protocol. In addition, we conduct an informal security analysis. In Section 5, we compare our method with other protocols in terms of performance and security. Finally, we summarize this paper in Section 6.

## 2. Related Work

IoV is an open network environment, so this feature may threaten the identity information and relevant sensitive data of vehicle users. For many years, researchers have proposed many protocols to protect the privacy of vehicle users in IoV environments. In 2006, Raya et al. [14] proposed a vehicle communication protocol that stored multiple public and private key pairs and protected the privacy of vehicle users through the certificates stored in OBU. However, in 2008, Lu et al. [15] determined that the protocol [14] had high computing and storage cost because the key was changing at times and proposed a privacy protection protocol for vehicle

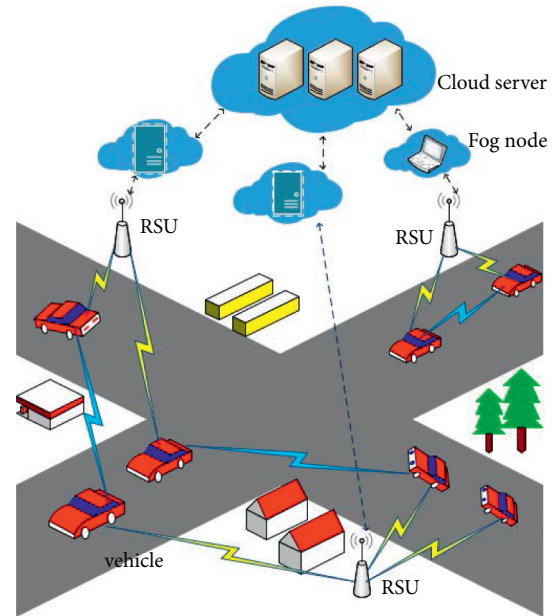


FIGURE 1: The SIOV architecture based on fog node.

communication. That same year, Zhang et al. [16] proposed an identity verification protocol for IoV. The protocol [16] realized privacy protection by the tamper-proof device to generate a random pseudonymity. In 2020, Cui et al. [17] researched a privacy-preserving scheme. The protocol [17] was based on edge computing and used lightweight primitives, such as elliptic-curve cryptography, instead of bilinear pairing-based primitives with high computational cost. Later, Hu et al. [18] proposed a privacy-preserving authentication scheme for IoV.

The protocols proposed by some researchers have high computing power. In 2014, Li et al. [19] proposed a protocol that provided PKC-based privacy protection for IoV and claimed that their protocol could resist replay and stolen smart card attacks. However, Amit et al. [20] revealed that Li et al.'s protocol [19] was susceptible to key compromise impersonation attacks and could not provide user anonymity. To reduce high computing cost caused by the use of PCK in the above protocol, the Trust-Extended Authentication Mechanism (TEAM) protocol was proposed [21]. In 2016, Kumari et al. [22] proposed an authentication protocol that also used TEAM. In 2017, Ying and Nayak [23] proposed an effective and lightweight protocol for an IoV environment, which could provide user anonymity. Chen et al. [5] demonstrated that [23] was vulnerable to replay and offline identity guessing attacks. Therefore, to solve the vulnerability of Ying and Nayak's protocol [23], Chen et al. [5] proposed a secure authentication scheme for IoV. However, the protocol [5] stored extensive data in the database, so it had high storage cost. In the same year, Mohit et al. [24] proposed an efficient authentication protocol for vehicular systems and deemed their protocol safe. However, Yu et al. [25] pointed out that the protocol [24] of Mohit et al. was susceptible to impersonation attacks and could not provide anonymity, traceability, and mutual authentication. Then, Yu et al. [25] proposed an authenticated protocol in

vehicular communications. In 2020, Sadri et al. [26] demonstrated that Yu et al.'s protocol [25] was susceptible to sensor capture attacks and impersonation attacks and could not provide traceability. Additionally, Sadri and Rajabzadeh Asaar [26] proposed a protocol in the IoV environment, which was based on lightweight primitives. In 2021, Wu et al. [27] proposed a protocol in IoV, and the protocol realized authentication key exchange (AKE).

There are increasingly more vehicles in the IoV environment, and data processing and transmission have become an inevitable challenge. Therefore, researchers began to apply cloud computing to IoV to solve the problem of processing a large amount of data to improve authentication efficiency. In an IoV environment, an authentication scheme based on cloud computing had been widely mentioned and applied in articles [28–31]. For an environment using cloud computing, problems such as network delay and transmission efficiency would exist, and the cloud server would need to process more data, which would increase the computing burden of the cloud server. Therefore, researchers have begun to introduce fog nodes for fog computing to share the pressure of cloud servers. In these papers [10–13], fog computing technology was applied. Ma et al.'s protocol [10] applied fog computing to IoV and proposed an authenticated key agreement protocol. They claimed that the protocol [10] was secure and efficient, but Eftekhari et al. [9] pointed out that Ma et al.'s protocol [10] was vulnerable to internal attacks, stolen smart card attacks, and known session-specific temporary information attacks. Therefore, Eftekhari et al. [9] proposed a more efficient authentication protocol. In 2021, Wu et al. [32] proposed a secure scheme using fog nodes in IoV, and the protocol realized AKE. In the same year, Maria et al. [33] proposed a blockchain-based anonymous authentication scheme, which used bilinear pairing. Some important related works are summarized in Table 1.

### 3. The Proposed Protocol

In this part, we introduce a lightweight and authenticated key agreement protocol using fog nodes in SIOV. Our protocol is based on the architecture of Figure 1. The protocol includes three entities: vehicle  $V_i$ , fog node  $FN_j$ , and CS. The symbols used in the protocol are shown in Table 2. The protocol has three phases: vehicle registration phase, fog node registration phase, and login authentication phase.

**3.1.  $V_i$  Registration Phase.** In the  $V_i$  registration phase,  $V_i$  registers with CS. The phase is shown in Figure 2, and the specific steps are as follows:

- (1) First,  $V_i$  selects its identity  $ID_i$ , password  $PSW_i$ , and a random number  $r_i$ , calculates its pseudoidentity  $PID_i = h(ID_i \| r_i)$ , and then transmits the  $PID_i$  to CS through the secure channel.
- (2) After receiving the message from  $V_i$ , CS calculates the value of  $HID_i = h(PID_i \| K_{CS})$ , initializes the

value of  $K_V$  to 0, and stores  $\{PID_i, K_V\}$  in its database. Finally, CS sends  $\{HID_i, K_V\}$  to  $V_i$ .

- (3) After receiving the message from CS,  $V_i$  calculates the value  $\alpha_i = HID_i \oplus h(PSW_i \| r_i)$ ,  $P_i = h(ID_i \| PSW_i \| r_i)$ , replaces  $HID_i$  with the value of  $\alpha_i$ , and stores the  $\{\alpha_i, P_i, r_i, K_V\}$  in its smart card.

**3.2.  $FN_j$  Registration Phase.** In  $FN_j$  registration phase,  $FN_j$  registers with CS. The phase is shown in Figure 3, and the specific steps are as follows:

- (1) First,  $FN_j$  selects its identity  $FID_j$  and a random number  $r_j$ , calculates its pseudoidentity  $PFID_j = h(FID_j \| r_j)$ , and then transmits  $\{PFID_j, FID_j\}$  to CS through the secure channel.
- (2) After receiving the message from  $FN_j$ , CS first selects a random number  $R_j$ , calculates the value of  $N_j = h(FID_j \| ID_{CS}) \oplus R_j$ ,  $K_{FN} = h(PFID_j \| K_{CS})$ ,  $HID_j = h(FID_j \| K_{CS})$ , and stores  $\{PFID_j, K_{FN}, FID_j\}$  in its database. Finally, CS sends  $\{K_{FN}, HID_j, N_j, ID_{CS}\}$  to  $FN_j$ .
- (3) After receiving the message from CS,  $FN_j$  calculates the value  $R_j = h(FID_j \| ID_{CS}) \oplus N_j$ ,  $\beta_j = HID_j \oplus h(R_j \| r_j)$ , and stores the  $\{K_{FN}, \beta_j, r_j, N_j\}$  in its database.

**3.3. Login and Authentication Phase.** In the login and authentication phase,  $V_i$ ,  $FN_j$ , and CS realize authentication and establish session key SK. This phase is shown in Figure 4, and the specific steps are as follows:

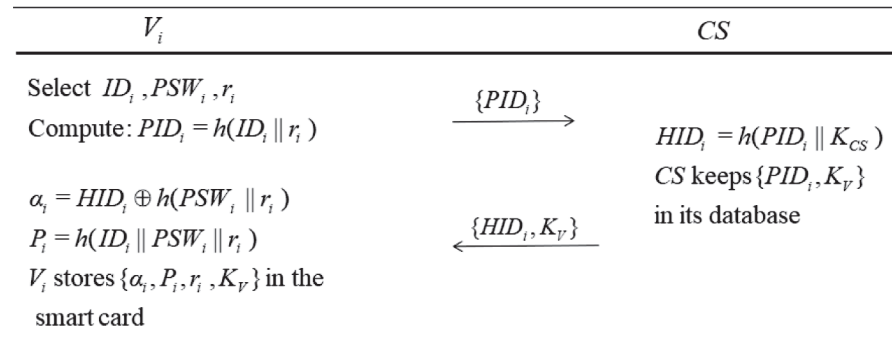
- (1) First,  $V_i$  inserts the smart card into the reader terminal, inputs its identity  $ID_i$ , password  $PSW_i$ , calculates the login authentication value  $P_i^* = h(ID_i \| PSW_i \| r_i)$ , and then compares  $P_i^* \stackrel{?}{=} P_i$ . If equal,  $V_i$  logs in successfully. Otherwise, the login fails. After successful login,  $V_i$  selects a random number  $N_1$  and calculates  $A_1 = h(ID_i \| r_i) \oplus N_1$ ,  $HID_i = \alpha_i \oplus h(PSW_i \| r_i)$ ,  $V_1 = h(HID_i \| K_V) \oplus N_1$ . Finally,  $V_i$  sends the login request  $M_1 = \{A_1, V_1, ID_{CS}, PID_i\}$  to  $FN_j$  through the common channel.
- (2) After receiving the message  $M_1$  from  $V_i$ ,  $FN_j$  first selects a random number  $N_2$  and then calculates  $A_2 = h(A_1 \| K_{FN} \| HID_j) \oplus N_2$ ,  $V_2 = h(A_2 \| K_{FN} \| V_1)$ , and finally  $FN_j$  transmits the message  $M_2 = \{PID_i, PFID_j, A_2, V_1, V_2\}$  to CS.
- (3) After receiving message  $M_2$  from  $FN_j$ , CS first indexes  $K_{FN}$  according to  $PFID_j$ , then calculates  $HID_i = h(PID_i \| K_{CS})$ ,  $N_1 = h(HID_i \| K_V) \oplus V_1$ ,  $V_1^* = h(HID_i \| K_V) \oplus N_1$ , and compares  $V_1^* \stackrel{?}{=} V_1$ . If it is equal, CS believes that  $V_i$  is legal. Otherwise, the authentication process is terminated. CS calculates  $V_2^* = h(A_2 \| K_{FN} \| V_1)$  and compares  $V_2^* \stackrel{?}{=} V_2$ . If it is equal, it means that CS believes that  $FN_j$  is legal. Otherwise, the authentication process is terminated. After authenticating  $V_i$  and  $FN_j$ , CS calculates  $A_1 = N_1 \oplus PID_i$ ,  $HID_j = h(FID_j \| K_{CS})$ ,  $N_2 = h(A_1 \| K_{FN} \|$

TABLE 1: The summary of authentication protocols.

| Protocols            | Cryptographic techniques   | Limitations   |
|----------------------|--|---|
| Li et al. [19]       | (1) Utilized digital signature<br>(2) Utilized asymmetric encryption<br>(3) Based on anonymous authentication      | (1) Does not resist key compromise impersonation attacks<br>(2) Does not provide user anonymity   |
| Ying and Nayak [23]  | (1) Utilized one-way hash function<br>(2) Based on Diffie–Hellman problem<br>(3) Based on anonymous authentication | (1) Does not resist replay attacks<br>(2) Does not resist offline identity guessing attacks   |
| Mohit et al. [24]    | (1) Utilized one-way hash function<br>(2) Based on smart card<br>(3) Two-factor                                    | (1) Does not resist impersonation attacks<br>(2) Does not provide anonymity and untraceability<br>(3) Does not provide mutual authentication                      |
| Yu et al. [25]       | (1) Utilized one-way hash function<br>(2) Based on smart card<br>(3) Two-factor                                    | (1) Does not resist sensor capture attacks<br>(2) Does not resist impersonation attacks<br>(3) Does not provide untraceability                                    |
| Ma et al. [10]       | (1) Utilized one-way hash function<br>(2) Based on smart card<br>(3) Utilized ECC                                  | (1) Does not resist internal attacks<br>(2) Does not resist stolen smart card attacks<br>(3) Does not resist known session-specific temporary information attacks |
| Wazid et al. [34]    | (1) Utilized one-way hash function<br>(2) Based on anonymous authentication<br>(3) Utilized ECC                    | —   |
| Eftekhari et al. [9] | (1) Utilized one-way hash function<br>(2) Based on anonymous authentication<br>(3) Utilized ECC                    | —   |
| Wu et al. [32]       | (1) Utilized one-way hash function<br>(2) Based on smart card<br>(3) Utilized ECC<br>(4) Two-factor                | —   |

TABLE 2: Notations used in the proposed protocol.

| Symbol                 | Description                        |
|------------------------|------------------------------------|
| $V_i$                  | The $i$ -th vehicle                |
| $FN_j$                 | The $j$ -th fog node               |
| CS                     | Cloud server                       |
| $ID_i, FID_j, ID_{CS}$ | Identities of $V_i, FN_j$ , and CS |
| $PSW_i$                | Password of the $V_i$              |
| $K_{FN}$               | Shared key of $FN_j$ and CS        |
| $K_{CS}$               | Secret key of CS                   |
| $K_V$                  | Counter value of $V_i$             |
| SK                     | Session key                        |

FIGURE 2:  $V_i$  registration phase.

| $FN_j$   | $CS$  |
|--|---|
| Select $FID_j, r_j$<br>$PFID_j = h(FID_j \  r_j)$  | choose $R_j$<br>$N_j = h(FID_j \  ID_{CS}) \oplus R_j$<br>$K_{FN} = h(PFID_j \  K_{CS})$  |
| $R_j = h(FID_j \  ID_{CS}) \oplus N_j$<br>$\beta_j = HID_j \oplus h(R_j \  r_j)$<br>$FN_j$ stores $\{K_{FN}, \beta_j, r_j, N_j\}$<br>in its database | $HID_j = h(FID_j \  K_{CS})$<br>$CS$ keeps $\{PFID_j, K_{FN}, FID_j\}$<br>in its database |
|  |   |

FIGURE 3:  $FN_j$  registration phase.

| $V_i$  | $FN_j$   | $CS$   |  |
|--|--|--|--|
| Enter $ID_i, PSW_i$<br>Compute: $P_i^* = h(ID_i \  PSW_i \  r_i)$<br>Check: $P_i^* = P_i$<br>Generate $N_1$<br>Compute: $A_1 = h(ID_i \  r_i) \oplus N_1$<br>$HID_i = \alpha_i \oplus h(PSW_i \  r_i)$<br>$V_1 = h(HID_i \  K_V) \oplus N_1$ | Generate $N_2$<br>Compute: $A_2 = h(A_1 \  K_{FN} \  HID_j) \oplus N_2$<br>$V_2 = h(A_2 \  K_{FN} \  V_1)$   | searches for $PFID_j$ and finds<br>$K_{FN} = h(PFID_j \  K_{CS})$<br>Compute: $HID_i = h(PID_i \  K_{CS})$<br>$N_1 = h(HID_i \  K_V) \oplus V_1$<br>$V_1^* = h(HID_i \  K_V) \oplus N_1$<br>Check: $V_1^* = V_1$<br>$V_2^* = h(A_2 \  K_{FN} \  V_1)$<br>Check: $V_2^* = V_2$<br>Compute: $A_1 = N_1 \oplus PID_i$<br>$HID_j = h(FID_j \  K_{CS})$<br>$N_2 = h(A_1 \  K_{FN} \  HID_j) \oplus A_2$ |  |
|  | $\{PID_i, PFID_j, A_2, V_1, V_2\}$   | Generate $N_3$<br>Compute:<br>$N_X' = h(HID_i \  N_1) \oplus N_2 \oplus N_3 \oplus HID_j$<br>$N_Y' = h(HID_j \  N_2) \oplus N_1 \oplus N_3 \oplus HID_i$<br>$SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$<br>$V_3^* = h(HID_j \  K_{FN} \  SK)$<br>Check: $V_3^* = V_3$  | $N_X' = h(HID_i \  N_1) \oplus N_2 \oplus N_3 \oplus HID_j$<br>$N_Y' = h(HID_j \  N_2) \oplus N_1 \oplus N_3 \oplus HID_i$<br>$SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$<br>$V_3 = h(HID_j \  K_{FN} \  SK)$<br>$V_4 = h(HID_i \  K_V \  SK)$<br>Update $K_V = K_V + 1$ |
| $\{A_1, V_1, ID_{CS}, PID_i\}$   | Compute:<br>$N_1 \oplus N_3 \oplus HID_i = h(HID_j \  N_2) \oplus N_Y'$<br>$SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$<br>$V_3^* = h(HID_j \  K_{FN} \  SK)$<br>Check: $V_3^* = V_3$ |  |  |
| Compute:<br>$N_2 \oplus N_3 \oplus HID_j = h(HID_i \  N_1) \oplus N_X'$<br>$SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$<br>$V_4^* = h(HID_i \  K_V \  SK)$<br>Check: $V_4^* = V_4$<br>Update $K_V = K_V + 1$                | $\{N_X', N_Y', V_3, V_4\}$   |  |  |
|  | $\{N_X', V_4\}$  |  |  |

FIGURE 4: Login and authentication phase.

$HID_j) \oplus A_2$ , selects a random number  $N_3$ , and calculates  $N_X' = h(HID_i \| N_1) \oplus N_2 \oplus N_3 \oplus HID_j$ ,  $N_Y' = h(HID_j \| N_2) \oplus N_1 \oplus N_3 \oplus HID_i$ ,  $SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$ ,  $V_3 = h(HID_j \| K_{FN} \| SK)$ ,  $V_4 = h(HID_i \| K_V \| SK)$ . Then, it updates  $K_V = K_V + 1$ , and finally,  $CS$  sends message  $M_3 = \{N_X', N_Y', V_3, V_4\}$  to  $FN_j$ .

- (4) After receiving message  $M_3$  from  $CS$ ,  $FN_j$  calculates  $N_1 \oplus N_3 \oplus HID_i = h(HID_j \| N_2) \oplus N_Y'$ ,  $SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$ ,  $V_3^* = h(HID_j \| K_{FN} \| SK)$ , and compares  $V_3^* = V_3$ . If it is equal, it means that  $FN_j$  believes that  $CS$  is legal. Otherwise, the authentication process is terminated. Finally,  $FN_j$  sends message  $M_4 = \{N_X', V_4\}$  to  $V_i$ .

- (5) After receiving message  $M_4$  from  $FN_j$ ,  $V_i$  calculates  $N_2 \oplus N_3 \oplus HID_j = h(HID_i \| N_1) \oplus N_X'$ ,  $SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_j \oplus HID_i)$ ,  $V_4^* = h(HID_i \| K_V \| SK)$ , and compares  $V_4^* = V_4$ . If equal, it means that  $V_i$  believes that  $FN_j$  and  $CS$  are legal. Otherwise, the authentication process is terminated. Finally,  $V_i$  updates  $K_V = K_V + 1$ .

## 4. Security Analysis

**4.1. BAN Logic.** BAN logic is a formal security analysis method [35]. In this part, we use BAN logic to prove that vehicles, fog nodes, and cloud servers share a session key  $SK$  and further prove the correctness of our protocol. The rules used in BAN logic are shown in the references.

#### 4.1.1. BAN Logic Rules

- (1) Message-meaning rule:  $(P \equiv P \stackrel{K}{\leftrightarrow} Q, P \triangleleft \{X\}_K) / (P \equiv Q \sim X)$
- (2) Freshness rule:  $(P \equiv \#(X)) / (P \equiv \#(X, Y))$
- (3) Nonce-verification rule:  $(P \equiv \#(X), P \equiv Q \sim X) / (P \equiv Q \equiv X)$
- (4) Jurisdiction rule:  $(P \equiv \#(X), P \equiv Q \sim X) / (P \equiv Q \equiv X)$
- (5) Belief rule:  $(P \equiv X, P \equiv Y) / (P \equiv (X, Y))$
- (6) Session key rule:  $(P \equiv \#(X), P \equiv Q \equiv X) / (P \equiv P \stackrel{K}{\leftrightarrow} Q)$

#### 4.1.2. Goals. **G1** $V_i \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$

- G2**  $V_i \equiv FN_j \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$
- G3**  $FN_j \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$
- G4**  $FN_j \equiv V_i \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$
- G5**  $CS \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$
- G6**  $CS \equiv V_i \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$
- G7**  $CS \equiv FN_j \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j$

#### 4.1.3. Idealizing Communication.

- M1**  $V_i \longrightarrow CS: \{PID_i, A_1, V_1\}$
- M2**  $FN_j \longrightarrow CS: \{PFID_j, A_2, V_2\}$
- M3**  $CS \longrightarrow FN_j: \{N'_Y, V_3\}$
- M4**  $CS \longrightarrow V_i: \{N'_X, V_4\}$

#### 4.1.4. Initial State Assumptions

- A1**  $V_i \equiv \#(N_1)$
- A2**  $FN_j \equiv \#(N_2)$
- A3**  $CS \equiv \#(N_3)$
- A4**  $CS \equiv V_i \stackrel{h(HID_i \| K_V)}{\equiv} CS$
- A5**  $CS \equiv \#(N_1)$
- A6**  $CS \equiv V_i \stackrel{h(A_1 \| K_{FN} \| HID_j)}{\equiv} N_1$
- A7**  $CS \equiv FN_j \stackrel{h(A_1 \| K_{FN} \| HID_j)}{\equiv} CS$
- A8**  $CS \equiv \#(N_2)$
- A9**  $CS \equiv FN_j \implies N_2$
- A10**  $CS \equiv HID_i$
- A11**  $CS \equiv HID_i \stackrel{h(HID_j \| N_2)}{\equiv} CS$
- A12**  $FN_j \equiv FN_j \stackrel{h(HID_j \| N_2)}{\equiv} CS$
- A13**  $FN_j \equiv CS \implies N_3$
- A14**  $FN_j \equiv CS \implies HID_j$
- A15**  $FN_j \equiv \#(N_1)$
- A16**  $FN_j \equiv \#(N_3)$
- A17**  $FN_j \equiv \#(HID_j)$
- A18**  $V_i \equiv V_i \stackrel{h(HID_i \| N_1)}{\equiv} CS$
- A19**  $V_i \equiv CS \implies N_3$

$$\mathbf{A20} \quad V_i \equiv CS \implies HID_j$$

$$\mathbf{A21} \quad V_i \equiv \#(N_2)$$

$$\mathbf{A22} \quad V_i \equiv \#(N_3)$$

$$\mathbf{A23} \quad V_i \equiv \#(HID_j)$$

4.1.5. *Detailed Steps.* By considering the message  $M1$  and using the seeing rule, we get

$$\mathbf{S1}: CS \triangleleft \left\{ V_1: \langle N_1 \rangle_{h(HID_i \| K_V)}, PID_i, A_1 \right\}. \quad (1)$$

Using S1, we get

$$\mathbf{S2}: CS \triangleleft \left\{ \langle N_1 \rangle_{h(HID_i \| K_V)} \right\} \quad (2)$$

Under the premise of assuming A4, using S2, and the message-meaning rule, we get

$$\mathbf{S3}: CS \equiv V_i \sim N_1. \quad (3)$$

In the case of conclusion S3, using assumption A5, the freshness rule, and the nonce-verification (N-V) rule, we get

$$\mathbf{S4}: CS \equiv V_i \equiv N_1. \quad (4)$$

In the case of conclusion S4, using assumption A6, and the jurisdiction rule, we get

$$\mathbf{S5}: CS \equiv N_1. \quad (5)$$

In addition, considering the message  $M2$ , we get

$$\mathbf{S6}: CS \triangleleft \left\{ PFID_j, A_2: \langle N_2 \rangle_{h(A_1 \| K_{FN} \| HID_j)}, V_2 \right\}. \quad (6)$$

Using S6, we get

$$\mathbf{S7}: CS \triangleleft \left\{ \langle N_2 \rangle_{h(A_1 \| K_{FN} \| HID_j)} \right\} \quad (7)$$

Under the premise of assuming A7, using S7, and the message-meaning rule, we get

$$\mathbf{S8}: CS \equiv FN_j \sim N_2. \quad (8)$$

In the case of conclusion S8, using assumption A8, the freshness rule, and the nonce-verification (N-V) rule, we get

$$\mathbf{S9}: CS \equiv FN_j \equiv N_2. \quad (9)$$

In the case of conclusion S9, using assumption A9, and the jurisdiction rule, we get

$$\mathbf{S10}: CS \equiv N_2. \quad (10)$$

Because  $SK = h(N_1 \oplus N_2 \oplus N_3 \oplus HID_i \oplus HID_j)$ , according to the conclusions A10, A11, S10, and S5 and the belief rule, we get

$$\mathbf{S11}: CS \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j, \quad (\mathbf{G5}). \quad (11)$$

Using A5, S11, and the SK rule, we get

$$\mathbf{S12}: CS \equiv V_i \equiv V_i \stackrel{SK}{\leftrightarrow} FN_j, \quad (\mathbf{G6}). \quad (12)$$



Using A8, S11, and the SK rule, we get

$$\text{S13: } \text{CS} \mid \equiv \text{FN}_j \mid \equiv V_i \stackrel{\text{SK}}{\leftrightarrow} \text{FN}_j, \quad (\text{G7}). \quad (13)$$

By considering the message  $M3$  and using the seeing rule, we get

$$\text{S14: } \text{FN}_j \triangleleft \left\{ V'_Y: \langle N_1, N_3, \text{HID}_i \rangle_{h(\text{HID}_i \| N_2)}, V_3 \right\}. \quad (14)$$

Using S14, we get

$$\text{S15: } \text{FN}_j \triangleleft \left\{ \langle N_1, N_3, \text{HID}_i \rangle_{h(\text{HID}_i \| N_2)} \right\}. \quad (15)$$

Under the premise of assuming A12, using S15, and the message-meaning rule, we get

$$\text{S16: } \text{FN}_j \mid \equiv \text{CS} \mid \sim (N_1, N_3, \text{HID}_i). \quad (16)$$

In the case of conclusion S16, using assumptions A13 and A14, the freshness rule, and the nonce-verification (N-V) rule, we get

$$\text{S17: } \text{FN}_j \mid \equiv \text{CS} \mid \equiv (N_1, N_3, \text{HID}_i). \quad (17)$$

Applying this for each component, we get

$$\begin{aligned} \text{S18: } \text{FN}_j \mid \equiv \text{CS} \mid \equiv N_1, \\ \text{S19: } \text{FN}_j \mid \equiv \text{CS} \mid \equiv N_3, \\ \text{S20: } \text{FN}_j \mid \equiv \text{CS} \mid \equiv \text{HID}_i. \end{aligned} \quad (18)$$

In the case of conclusion S18, using assumption A15, and the jurisdiction rule, we get

$$\text{S21: } \text{FN}_j \mid \equiv N_1. \quad (19)$$

In the case of conclusion S22, using assumption A16, and the jurisdiction rule, we get

$$\text{S22: } \text{FN}_j \mid \equiv N_3. \quad (20)$$

In the case of conclusion S23, using assumption A17, and the jurisdiction rule, we get

$$\text{S23: } \text{FN}_j \mid \equiv \text{HID}_i. \quad (21)$$

Because  $\text{SK} = h(N_1 \oplus N_2 \oplus N_3 \oplus \text{HID}_i \oplus \text{HID}_j)$ , according to the conclusions S21, S22, and S23 and the belief rule, we get

$$\text{S24: } \text{FN}_j \mid \equiv V_i \stackrel{\text{SK}}{\leftrightarrow} \text{FN}_j, \quad (\text{G3}). \quad (22)$$

Using A15, S24, and the SK rule, we get

$$\text{S25: } \text{FN}_j \mid \equiv V_i \mid \equiv V_i \stackrel{\text{SK}}{\leftrightarrow} \text{FN}_j, \quad (\text{G4}). \quad (23)$$

By considering the message  $M4$  and using the seeing rule, we get

$$\text{S26: } V_i \triangleleft \left\{ V'_X: \langle N_2, N_3, \text{HID}_j \rangle_{h(\text{HID}_i \| N_1)}, V_4 \right\}. \quad (24)$$

Using S26, we get

$$\text{S27: } V_i \triangleleft \left\{ \langle N_2, N_3, \text{HID}_j \rangle_{h(\text{HID}_i \| N_1)} \right\}. \quad (25)$$

Under the premise of assuming A18, using S27, and the message-meaning rule, we get

$$\text{S28: } V_i \mid \equiv \text{CS} \mid \sim (N_2, N_3, \text{HID}_j). \quad (26)$$

In the case of conclusion S28, using assumption A19 and A20, the freshness rule, and the nonce-verification (N-V) rule, we get

$$\text{S29: } V_i \mid \equiv \text{CS} \mid \equiv (N_2, N_3, \text{HID}_j). \quad (27)$$

Applying this for each component, we get

$$\begin{aligned} \text{S30: } V_i \mid \equiv \text{CS} \mid \equiv N_2, \\ \text{S31: } V_i \mid \equiv \text{CS} \mid \equiv N_3, \\ \text{S32: } V_i \mid \equiv \text{CS} \mid \equiv \text{HID}_j. \end{aligned} \quad (28)$$

In the case of conclusion S30, using assumptions A21, and the jurisdiction rule, we get

$$\text{S31: } V_i \mid \equiv N_2. \quad (29)$$

In the case of conclusion S31, using assumptions A22, and the jurisdiction rule, we get

$$\text{S32: } V_i \mid \equiv N_3. \quad (30)$$

In the case of conclusion S32, using assumptions A23, and the jurisdiction rule, we get

$$\text{S33: } V_i \mid \equiv \text{HID}_j. \quad (31)$$

Because  $\text{SK} = h(N_1 \oplus N_2 \oplus N_3 \oplus \text{HID}_i \oplus \text{HID}_j)$ , according to the conclusions S31, S32, and S33 and the belief rule, we get

$$\text{S34: } V_i \mid \equiv V_i \stackrel{\text{SK}}{\leftrightarrow} \text{FN}_j, \quad (\text{G1}). \quad (32)$$

Using A21, S34, and the SK rule, we get

$$\text{S35: } V_i \mid \equiv \text{FN}_j \mid \equiv V_i \stackrel{\text{SK}}{\leftrightarrow} \text{FN}_j, \quad (\text{G2}). \quad (33)$$

**4.2. Formal Security Analysis.** In this part, we use the ROR model to formally prove the security of our proposed protocol. The ROR model judges the security of the protocol by calculating the session key SK probability of an ordinary situation [36, 37].

**4.2.1. ROR Model.** The protocol consists of three entities: vehicle, fog node, and cloud server. In the ROR model, we use  $\Pi_{V_i}^x$ ,  $\Pi_{\text{FN}_j}^y$ , and  $\Pi_{\text{CS}}^z$  to represent the  $x$ -th communication of the  $V_i$ , the  $y$ -th communication of the  $\text{FN}_j$ , and the  $z$ -th communication of the CS, respectively. We also define that the attacker  $A$  can have the following query capabilities, where  $Z = \{\Pi_{V_i}^x, \Pi_{\text{FN}_j}^y, \Pi_{\text{CS}}^z\}$ .

Execute ( $Z$ ): by performing this query operation,  $A$  can intercept the messages transmitted on the public channel.

Hash (string): by performing this query operation,  $A$  can obtain the hash value of the input string.

Send ( $Z, M$ ): by performing this query operation,  $A$  can send message  $M$  to  $Z$  and receive the response from  $Z$ .

Corrupt ( $Z$ ): by performing this query operation,  $A$  can obtain a party's secret values, such as some values in the smart card, long-term key, or temporary information.

Test ( $Z$ ): by performing this query operation,  $A$  flips a coin  $c$ . If  $c = 1$ ,  $A$  can obtain an accurate session key; if  $c = 0$ ,  $A$  can obtain a random string of the same length as the session key.

**4.2.2. Theorem.** In the ROR model, assume  $A$  can perform execute, hash, send, corrupt, and test queries. Then, the probability that  $A$  can break the proposed protocol  $P$  in polynomial time is  $\text{Adv}_A^P(\xi) \leq (q_{\text{send}}/2^{l-1}) + (3q_{\text{hash}}^2/2^l) + 2 \max\{C' \cdot q_{\text{send}}', (q_{\text{send}}/2^l)\} + ((q_{\text{exe}} + q_{\text{send}})/p)$ , where  $q_{\text{hash}}$  represents the number of times hash queries are executed,  $q_{\text{send}}$  represents the number of times send queries are executed,  $q_{\text{exe}}$  represents the number of times execute queries are executed,  $l$  represents the bits of biological information, and  $C'$  and  $s$  are constants in Zipf's law.

**4.2.3. Proof.** We played five rounds of games, which were expressed as follows:  $\text{GM}_0$  to  $\text{GM}_6$ .  $\text{Succ}_A^{\text{GM}_i}(\xi)$  represents the event that  $A$  can win in the game  $\text{GM}_i$ .  $\text{Succ}_P^{A, \text{GM}_i} = \Pr[\text{Succ}_A^{\text{GM}_i}]$  represents the advantage of  $A$  for winning  $\text{GM}_i$ .  $\Pr[Z]$  is the probability of event  $Z$ .  $\text{Adv}_P^A$  represents the advantage  $A$  has in breaking the security of SK for protocol  $P$ . The specific steps of  $\text{GM}_i$  are as follows:

$\text{GM}_0$ :  $\text{GM}_0$  is the first-round game in the ROR model and a real attack. We choose a coin  $c$  to start the round. Therefore, in  $\text{GM}_0$ , we can obtain the probability that  $A$  can successfully break  $P$  as

$$\text{Adv}_A^P = \left| 2\Pr[\text{Succ}_A^{\text{GM}_0}] - 1 \right|. \quad (34)$$

$\text{GM}_1$ :  $\text{GM}_1$  adds an execute query to  $\text{GM}_0$ . In  $\text{GM}_1$ ,  $A$  can only obtain the messages transmitted on the public channel. After  $\text{GM}_1$ ,  $A$  will query the session key SK through the test, but  $A$  cannot obtain five values  $\{N_1, N_2, N_3, \text{HID}_i, \text{HID}_j\}$ , so the probability that  $\text{GM}_0$  is equal to that of  $\text{GM}_1$  is

$$\Pr[\text{Succ}_A^{\text{GM}_1}] = \Pr[\text{Succ}_A^{\text{GM}_0}]. \quad (35)$$

$\text{GM}_2$ :  $\text{GM}_2$  adds a send query to  $\text{GM}_1$ . According to Zipf's law [38], we obtain

$$\left| \Pr[\text{Succ}_A^{\text{GM}_2}] - \Pr[\text{Succ}_A^{\text{GM}_1}] \right| \leq (q_{\text{send}}/2^l). \quad (36)$$

$\text{GM}_3$ :  $\text{GM}_3$  adds the hash query to  $\text{GM}_2$ . The maximum probability of text collision in transmission is  $(q_{\text{exe}} + q_{\text{send}})^2/2p$ , and we can obtain

$$\begin{aligned} & \left| \Pr[\text{Succ}_A^{\text{GM}_3}] - \Pr[\text{Succ}_A^{\text{GM}_2}] \right| \\ & \leq \frac{(q_{\text{exe}} + q_{\text{send}})^2}{2p} + \frac{q_{\text{hash}}^2}{2^{l+1}}. \end{aligned} \quad (37)$$

$\text{GM}_4$ : in this round, we verify the security of the session key SK using two events. One is to obtain the long-term key of  $\Pi_{\text{CS}}^z$  to verify the perfect forward security, and the other is to obtain temporary information to verify that the protocol can resist the known session-specific temporary information attacks.

- (1) Perfect forward security: using  $\Pi_{\text{CS}}^z$ ,  $A$  attempts to obtain the private key  $K_{\text{CS}}$  of CS, or  $A$  uses  $\Pi_{V_i}^x$  or  $\Pi_{\text{FN}_j}^y$  to obtain some secret values in the registration phase.
- (2) Known session-specific temporary information attacks:  $A$  uses one of  $\Pi_{V_i}^x$  or  $\Pi_{\text{FN}_j}^y$  or  $\Pi_{\text{CS}}^z$  to attempt to obtain temporary information.

For the first event, if  $A$  obtains the private key  $K_{\text{CS}}$  of CS, or the secret value of  $\Pi_{V_i}^x$  and  $\Pi_{\text{FN}_j}^y$  in the registration phase, but  $A$  cannot get the random number  $N_1, N_2, N_3, \text{HID}_j$ , it cannot calculate session key SK, where  $\text{SK} = h(N_1 \oplus N_2 \oplus N_3 \oplus \text{HID}_i \oplus \text{HID}_j)$ . For the second event, if  $A$  can obtain  $N_1$ , but the values of  $N_2$  and  $N_3$  are confidential, the SK cannot be calculated. Similarly, if  $N_2$  and  $N_3$  are leaked, SK cannot be calculated by  $A$ . Therefore, the probability of this round is

$$\left| \Pr[\text{Succ}_A^{\text{GM}_4}] - \Pr[\text{Succ}_A^{\text{GM}_3}] \right| \leq \frac{q_{\text{hash}}^2}{2^{l+1}}. \quad (38)$$

$\text{GM}_5$ : in this round of the game,  $A$  uses the corrupt query to obtain the parameter  $\{\alpha_i, P_i, r_i, K_V\}$  stored in the smart card, so  $A$  wants to conduct the offline key guessing attacks.  $V_i$  uses random numbers and passwords for registration, so  $A$  must guess  $P_i = h(\text{ID}_i \parallel \text{PSW}_i \parallel r_i)$ , but the probability of guessing a random number is  $1/2^l$ , which can be ignored. Using Zipf's law [38], we can obtain

$$\left| \Pr[\text{Succ}_A^{\text{GM}_5}] - \Pr[\text{Succ}_A^{\text{GM}_4}] \right| \leq \max \left\{ C' \cdot q_{\text{send}}', \frac{q_{\text{send}}}{2^l} \right\}. \quad (39)$$

$\text{GM}_6$ : this round of the game is to verify that protocol  $P$  can resist the impersonation attacks,  $A$  uses  $h(N_1 \oplus N_2 \oplus N_3 \oplus \text{HID}_i \oplus \text{HID}_j)$  to query, and the game is terminated. Therefore, the probability that  $A$  can guess SK is

$$\left| \Pr[\text{Succ}_A^{\text{GM}_6}] - \Pr[\text{Succ}_A^{\text{GM}_5}] \right| \leq \frac{q_{\text{hash}}^2}{2^{l+1}}. \quad (40)$$

Because the probability of success and failure of the  $\text{GM}_6$  is  $1/2$ ,

$$\begin{aligned}
\frac{1}{2} \text{Adv}_A^P &= \left| \Pr[\text{Succ}_A^{\text{GM}_0}] - \frac{1}{2} \right| \\
&= \left| \Pr[\text{Succ}_A^{\text{GM}_0}] - \Pr[\text{Succ}_A^{\text{GM}_6}] \right| \\
&= \left| \Pr[\text{Succ}_A^{\text{GM}_1}] - \Pr[\text{Succ}_A^{\text{GM}_6}] \right| \\
&\leq \sum_{i=0}^5 \left| \Pr[\text{Succ}_A^{\text{GM}_{i+1}}] - \Pr[\text{Succ}_A^{\text{GM}_i}] \right| \quad (41) \\
&= \frac{q_{\text{send}}}{2^l} + \frac{3q_{\text{hash}}^2}{2^{l+1}} + \max \\
&\quad \cdot \left\{ C' \cdot q_{\text{send}}', \frac{q_{\text{send}}}{2^l} \right\} + \frac{(q_{\text{hash}} + q_{\text{send}})^2}{2p}.
\end{aligned}$$

Finally, we can obtain

$$\begin{aligned}
\text{Adv}_A^P &\leq \frac{q_{\text{send}}}{2^{l-1}} + \frac{3q_{\text{hash}}^2}{2^l} + 2 \max \{ C' \cdot \\
&\quad \cdot q_{\text{send}}', \frac{q_{\text{send}}}{2^l} \} + \frac{(q_{\text{exe}} + q_{\text{send}})^2}{p}. \quad (42)
\end{aligned}$$

**4.3. ProVerif.** ProVerif is a formal automatic verification tool, which can verify confidentiality, identity, anonymity, and so on [39, 40]. In this paper, we use the ProVerif code to achieve vehicle registration, fog node registration, and authentication between the two parties and the CS and verify the security and effectiveness of our proposed protocol through ProVerif.

ProVerif demonstrates that the specific operation works as follows. Our protocol includes three entities: vehicle, fog node, and cloud server. The symbols and operation definitions used in ProVerif are shown in Figure 5.

The proof contains six events, as shown in Figure 6. The six events are `veclestarted()`, `vecleauthored()`, `cloudserveracvehicle()`, `cloudserveracfognode()`, `fognodeaccloudserver()`, and `vecleaccloudserver()`, indicating that the vehicle starts certification, the vehicle completes certification, the cloud server completes the vehicle certification, and the cloud server completes the fog node certification, respectively. The fog node completed the certification of the cloud server, and the vehicle completed the certification of the cloud server.

Then, we use ProVerif to query whether  $A$  can calculate the session key  $SK$  through the data transmitted on the common channel. The query operation is shown in Figure 7.

Finally, we get the verification result using the ProVerif tool, as shown in Figure 8. The result shows that  $A$  cannot calculate the session key  $SK$  of the  $V_i$ ,  $FN_j$ , and CS.

**4.4. Informal Security Analysis.** This part is an informal security analysis of our proposed agreement. We have proved that the protocol can meet common security requirements. The specific proof is as follows.

**4.4.1. Mutual Authentication.** In the authentication phase, with the help of CS, mutual authentication between  $V_i$  and  $FN_j$  is realized.  $V_1$  in message  $M_1$  is the value CS uses to authenticate  $V_i$ ,  $V_2$  in message  $M_2$  is the value CS uses to authenticate  $FN_j$ , and  $V_3$  and  $V_4$  in message  $M_3$  are the values CS uses to authenticate  $FN_j$  and  $V_i$ , respectively. Therefore, the mutual authentication among  $V_i$ ,  $FN_j$ , and CS is realized in the authentication phase.

**4.4.2. Replay Attacks.** In this protocol, we use cumulative value  $K_V$  to resist replay attacks. In the  $V_i$  registration phase, we initialize  $K_V$  to 0. As the session progresses, it carries out +1 operation on the value  $K_V$ , saves it to its database after CS authenticates  $V_i$ ,  $FN_j$ , and carries out the necessary calculation. After CS authenticates  $V_i$  and generates the session key, it also carries out the +1 operation on the value  $K_V$  and saves it to the smart card. In this manner,  $K_V$  on both sides is synchronous and equal, and the session process is completed smoothly. If  $A$  repeatedly sends message  $M_1$  intercepted in the public channel, CS continues to calculate the value  $K_V + 1$  in the authentication phase. Value  $V_4$  generated using  $K_V$  is not equal to value  $V_4$  calculated by  $V_i$  using  $K_V$  stored in its smart card, because the value  $K_V$  in the smart card of  $V_i$  cannot keep up with the CS update speed, so the authentication fails. Thus, our protocol can resist replay attacks.

**4.4.3. Man-in-the-Middle Attacks.** Suppose that  $A$  can intercept the message  $M_1 = \{A_1, V_1, \text{ID}_{\text{CS}}, \text{PID}_i\}$  transmitted on the public channel between  $V_i$  and  $FN_j$ . Since  $A$  cannot obtain the information  $\{\alpha_i, K_V, r_i\}$  in the smart card and the identity  $\text{ID}_i$  of  $V_i$ ,  $A$  cannot calculate the values  $\{K_V, \text{HID}_i, N_1\}$  required for  $V_1$ , where  $V_1 = h(\text{HID}_i \| K_V) \oplus N_1$ . Therefore, after  $A$  tampers with  $M_1$ , it cannot pass the authentication of  $FN_j$ . Similarly, because the privacy value is unknown,  $A$  cannot calculate the authentication value  $V_2$ ,  $V_3$ , or  $V_4$  and cannot complete the verification after intercepting the information  $M_2$ ,  $M_3$ , or  $M_4$ . Therefore, our protocol can resist man-in-the-middle attacks.

**4.4.4. User Anonymity.** The real identities of  $V_i$  and  $FN_j$  are transmitted on the secure channel and are protected by pseudoidentity  $\text{PID}_i$  and  $\text{PFID}_j$  in the authentication phase. The anonymity of  $V_i$  and  $FN_j$  is ensured. Therefore, our protocol can provide user anonymity.

**4.4.5. Untraceability.** If  $A$  wants to trace the  $V_i$ , it intercepts the messages  $\{M_1, M_2, M_3, M_4\}$  transmitted on the common channel. Since the random numbers  $\{N_1, N_2, N_3\}$  are used, this means that messages  $\{M_1, M_2, M_3, M_4\}$  are different during each session. In addition,  $A$  cannot obtain the random numbers  $\{N_1, N_2, N_3\}$ , so  $A$  cannot be traced back to  $V_i$ . Therefore, our protocol can provide untraceability.

## 5. Security and Performance Comparisons

In this part, we compare our protocol with those of Ma et al. [10], Wazid et al. [34], Eftekhari et al. [9], and Wu et al. [32] in terms of security, computational cost, and communication cost.



```

(* channel*)
free ch :channel. (* public channel *)
free sch: channel [private]. (* secure channel, used for registering *)
(* shared keys *)
free SKv : bitstring [private].
free SKf : bitstring [private].
free SKc : bitstring [private].
(* constants *)
free Kcs:bitstring [private].
free P:bitstring.
free B:bitstring.
free yj:bitstring.
(* functions & reductions & equations *)
fun h(bitstring) :bitstring. (* hash function *)
fun mult(bitstring,bitstring) :bitstring. (* scalar multiplication operation *)
fun add(bitstring,bitstring):bitstring. (* Addition operation *)
fun sub(bitstring,bitstring):bitstring. (* Subtraction operation *)
fun mod(bitstring,bitstring):bitstring. (* modulus operation *)
fun con(bitstring,bitstring):bitstring. (* concatenation operation *)
reduc forall m:bitstring, n:bitstring; getmess(con(m,n))=m.
fun xor(bitstring,bitstring):bitstring. (* XOR operation *)
equation forall m:bitstring, n:bitstring; xor(xor(m,n),n)=m.
fun Gen(bitstring):bitstring. (* Generator operation *)
fun Rep(bitstring,bitstring):bitstring.

```

FIGURE 5: The definition in the ProVerif tool.

*5.1. Security Comparisons.* When comparing protocol security, we use ✓ to indicate that the protocol can resist the attacks and × to indicate that the protocol cannot resist the attacks. The results of comparing protocol security are shown in Table 3. It can be seen that our protocol can resist known attacks and have better security. Ma et al.'s protocol [10] cannot provide user anonymity and untraceability and is vulnerable to impersonation attacks and known session-specific temporary information attacks. The protocols in [9, 32, 34] and our protocol are secure.

*5.2. Performance Comparison.* Performance analysis is conducted from the aspects of computational cost and communication cost. We analyze and compare the computational cost from the login authentication phase of each protocol. The computational cost of XOR and join operations is negligible. The computational cost comparison is shown in Table 4. It is obvious that the protocols of Ma et al.

[10], Wazid et al. [34], Eftekhari et al. [9], and Wu et al. [32] perform point multiplication, Wazid et al. [34] and Wu et al. [32] perform fuzzy extraction, and Wazid et al.'s protocol [34] and Eftekhari et al. [9] also perform ECC point addition. Only our proposed protocol performs the hash operation, so its computational cost is less.

Here,  $T_{pm}$  represents the time taken to perform a point multiplication operation,  $T_{pa}$  represents the time taken to execute an ECC point addition,  $T_f$  represents the time taken to execute a fuzzy extraction function, and  $T_h$  represents the time taken to execute a hash operation.

In the comparison of communication cost, we assume that the length of the identity and the random number are 160 bits, the length of the timestamp is 32 bits, the length of the one-way hash function is 256 bits, and the length of ECC point is 320 bits. Therefore, based on our assumption, the communication costs of the protocols of Ma et al. [10], Wazid et al. [34], Eftekhari et al. [9], and Wu et al. [32] are 4512 bits, 3488 bits, 4416 bits, and 4448 bits. Here, we illustrate our

```

(* -----Vehicle's process----- *)
let ProcessVehicle=new IDi : bitstring; (* the Vehicle's ID *)
  new PSWi : bitstring;
  new ri : bitstring;
  let PIDi=h(con(IDi,ri)) in
  out(sch, (PIDi));
  in(sch, (xHIDi:bitstring,xKv:bitstring));
  let ai=xor(xHIDi,h(con(IDi,ri))) in
  let Pi=h(con(con(IDi,PSWi),ri)) in
  ! (event VehicleStarted());
  let Pi'=h(con(con(IDi,PSWi),ri)) in
  if Pi=Pi' then
  new N1:bitstring;
  let A1=xor(h(con(IDi,ri),N1)) in
  let HIDi=xor(ai,h(con(PSWi,ri))) in
  let V1=xor(h(con(HIDi,xKv)),N1) in
  new IDcs:bitstring;
  out(ch,(A1,V1,IDcs,PIDi)); (*-----authentication----- *)
  event VehicleAuthed();
  in(ch, (xNx':bitstring,xV4:bitstring));
  (*let xor(xor(N2,N3),HIDj)=xor(h(con(HIDi,N1)),xNx') in *)
  let P=xor(h(con(HIDi,N1)),xNx') in
  let SKv=h(xor(xor(xor(P,HIDi),N1),HIDi)) in
  let V4=h(con(con(HIDi,xKv),SKv)) in
  if V4=xV4 then event VehicleAcCloudServer();

(* -----FNj process----- *)
let ProcessFNj=new FIDj:bitstring;
  new rj:bitstring;
  let PFIDj=h(con(FIDj,rj)) in
  out(sch, (FIDj,PFIDj));
  in(sch, (yKFN:bitstring,yHIDj:bitstring,yNj:bitstring,yIDcs:bitstring));
  let Rj=xor(h(con(FIDj,yIDcs)),yNj) in
  ! (in(ch, (yA1:bitstring,yV1:bitstring,yIDcs:bitstring,yPIDi:bitstring));
  new N2:bitstring;
  let A2=xor(h(con(con(yA1,yKFN),yHIDj)),N2) in
  let V2=h(con(con(A2,yKFN),yV1)) in
  out(ch, (yPIDi,PFIDj,A2,yV1,V2));
  in(ch, (yNx':bitstring,yNY:bitstring,yV3:bitstring,yV4:bitstring));
  let Q=xor(h(con(yHIDj,N2)),yNY) in
  let SKf=h(xor(xor(Q,N2),yHIDj)) in
  let V3=h(con(con(yHIDj,yKFN),SKf)) in
  if V3=yV3 then event FogNodeACcloudServer();
  out (ch, (yNx',yV4));

(* -----CloudServer's process----- *)
let VehicleReg= in(sch, (zPIDi:bitstring));
  let HIDi=h(con(zPIDi,Kcs)) in
  new Kv:bitstring;
  out(sch, (HIDi,Kv));
  0;

let FogNodeReg= in(sch, (zPFIDj:bitstring,zFIDj:bitstring));
  new Rj:bitstring;
  new IDcs:bitstring;
  let Nj=xor(h(con(zFIDj,IDcs)),Rj) in
  let KFN=h(con(zPFIDj,Kcs)) in
  let HIDj=h(con(zFIDj,Kcs)) in
  out(sch, (KFN,HIDj,Nj,IDcs));
  0;

let
CloudServerAuth= in(ch, (zPIDi:bitstring,zPFIDj:bitstring,zA2:bitstring,zV1:bitstring,zV2:bitstring));
  let KFN=h(con(zPFIDj,Kcs)) in
  let HIDi=h(con(zPIDi,Kcs)) in
  new Kv:bitstring;
  let N1=xor(h(con(HIDi,Kv)),zV1) in
  let V1'=xor(h(con(HIDi,Kv)),N1) in
  if V1'=zV1 then event CloudServerAcVehicle();
  let V2'=h(con(con(zA2,KFN),zV1)) in
  if V2'=zV2 then event CloudServerAcFogNode();
  let A1=xor(N1,zPIDi) in
  new FIDj:bitstring;
  let HIDj=h(con(FIDj,Kcs)) in
  let N2=xor(h(con(con(A1,KFN),HIDj)),zA2) in
  new N3:bitstring;
  let P=xor(xor(N2,N3),HIDj) in
  let Q=xor(xor(N1,N3),HIDi) in
  let Nx'=xor(h(con(HIDi,N1)),P) in
  let NY=xor(h(con(HIDj,N2)),Q) in
  let SKc=h(xor(xor(xor(N1,N2),N3),HIDj),HIDi) in
  let V4=h(con(con(HIDj,KFN),SKc)) in
  let V4=h(con(con(HIDi,Kv),SKc)) in
  out(ch, (Nx',NY,V3,V4));
  0;

let ProcessCloudServer= VehicleReg | FogNodeReg | CloudServerAuth.
(* ----- main----- *)
process
  (!ProcessVehicle | !ProcessFNj | !ProcessCloudServer )

```

FIGURE 6: Process in the ProVerif tool.

```

(* queries *)
query attacker(SKv).
query attacker(SKf).
query attacker(SKc).
query inj-event(VehicleAuthed()) ==> inj-event(VehicleStarted()).
query inj-event(CloudServerAcFogNode()) ==> inj-
event(CloudServerAcVehicle()).
query inj-event(VehicleAcCloudServer()) ==> inj-
event(FogNodeACcloudServer()).
(* event *)
event VehicleStarted().
event VehicleAuthed().
event CloudServerAcVehicle().
event CloudServerAcFogNode().
event FogNodeACcloudServer().
event VehicleAcCloudServer().

```

FIGURE 7: Queries and events in the ProVerif tool.

protocol as an example to show the specific analysis. In our protocol, the messages transmitted in the login authentication phase are  $M_1 = \{A_1, V_1, ID_{CS}, PID_i\}$ ,  $M_2 = \{A_2, V_1, V_2, PFID_j, PID_i\}$ ,  $M_3 = \{N'_X, N'_Y, V_3, V_4\}$ , and  $M_4 = \{N'_X, V_4\}$ , where  $\{A_1, A_2, N'_X, N'_Y\}$  are random strings,  $ID_{CS}$  is an identity, and  $\{V_1, V_2, V_3, V_4, PFID_j, PID_i\}$  are hash values. Therefore, the total communication cost of our proposed

protocol is 2336 bits. The comparison of communication cost is shown in Table 5. Obviously, the communication cost of our proposed protocol is less.

In the security comparison, we found that Ma et al.'s protocol [10] cannot provide user anonymity and untraceability and is vulnerable to impersonation attacks and known session-specific temporary information attacks.

Verification summary:  
 Query not attacker(SKv[]) is true.  
 Query not attacker(SKf[]) is true.  
 Query not attacker(SKc[]) is true.  
 Query inj-event(VehicleAuthenticated) ==> inj-event(VehicleStarted) is true.  
 Query inj-event(CloudServerAcFogNode) ==> inj-event(CloudServerAcVehicle) is true.  
 Query inj-event(VehicleAcCloudServer) ==> inj-event(FogNodeACcloudServer) is true.

FIGURE 8: Results in the ProVerif tool.

TABLE 3: Comparisons of security.

| Security properties                                  | [10] | [34] | [9] | [32] | Ours |
|--|------|------|-----|------|------|
| Mutual authentication                                | ✓    | ✓    | ✓   | ✓    | ✓    |
| User anonymity                                       | ×    | ✓    | ✓   | ✓    | ✓    |
| Perfect forward secrecy                              | ✓    | —    | ✓   | ✓    | ✓    |
| Man-in-the-middle attacks                            | ✓    | —    | —   | ✓    | ✓    |
| Impersonation attacks                                | ✓    | ✓    | ✓   | ✓    | ✓    |
| Known session-specific temporary information attacks | ×    | —    | ✓   | ✓    | ✓    |
| Untraceability                                       | ×    | ✓    | ✓   | ✓    | ✓    |
| Offline password guessing attacks                    | ✓    | ✓    | —   | ✓    | ✓    |
| Replay attacks                                       | ✓    | ✓    | ✓   | ✓    | ✓    |

TABLE 4: Computational cost comparison.

| Protocol             | $V_i$                      | $FN_j$                     | CS                          | Total                              |
|----------------------|----------------------------|----------------------------|-----------------------------|------------------------------------|
| Ma et al. [10]       | $3T_{pm} + 4T_h$           | $4T_{pm} + 4T_h$           | $10T_{pm} + 11T_h$          | $17T_{pm} + 19T_h$                 |
| Wazid et al. [34]    | $3T_{pm} + 2T_f + 22T_h$   | $2T_{pm} + T_{pa} + 14T_h$ | $3T_h$                      | $5T_{pm} + 2T_{pa} + 2T_f + 43T_h$ |
| Eftekhari et al. [9] | $3T_{pm} + T_{pa} + 11T_h$ | $3T_{pm} + T_{pa} + 12T_h$ | $3T_{pm} + 2T_{pa} + 15T_h$ | $9T_{pm} + 4T_{pa} + 38T_h$        |
| Wu et al. [32]       | $2T_{pm} + T_f + 8T_h$     | $4T_{pm} + 5T_h$           | $4T_{pm} + 13T_h$           | $10T_{pm} + T_f + 26T_h$           |
| Ours                 | $7T_h$                     | $5T_h$                     | $11T_h$                     | $23T_h$                            |

TABLE 5: Communication cost comparison.

| Protocol             | Round | Communication cost (bits) |
|----------------------|-------|---------------------------|
| Ma et al. [10]       | 4     | 4512                      |
| Wazid et al. [34]    | 3     | 3488                      |
| Eftekhari et al. [9] | 4     | 4416                      |
| Wu et al. [32]       | 4     | 4448                      |
| Ours                 | 4     | 2336                      |

Although the protocols of [9, 32, 34] can resist known security attacks, the overhead in the aspect of computational cost and communication cost is much more than that of our proposed protocol. Therefore, our protocol is better in terms of security and performance.

## 6. Conclusions

In this paper, we first review the AKE protocol in IoV and SIOV, and then, we propose a lightweight and authenticated key agreement protocol using fog nodes. The security analysis of the protocol is conducted by using BAN, ROR, and ProVerif. The comparison of security and performance shows that the protocol achieves higher performance in terms of computing power and communication cost compared with

other protocols. In future research, we will focus on improving the security and performance of the protocol in SIOV.

## Data Availability

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

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This article was supported by the Guangxi Key Laboratory of Trusted Software (no. KX202033).

## References

- [1] S. Arasteh, S. F. Aghili, and M. Hamid, "A new lightweight authentication and key agreement protocol for internet of things," in *Proceedings of the 2016 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC)*, pp. 52–59, IEEE, Tehran, Iran, September 2016.



- [2] M. Azroul, J. Mabrouki, A. Guezzaz, and Y. Farhaoui, "New enhanced authentication protocol for internet of things," *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 1–9, 2021.
- [3] X. Hu, Y. Wu, C. Jin, and S. Kumari, "Efficient and privacy-preserving authentication protocol for heterogeneous systems in iiot," *IEEE Internet of Things Journal*, vol. 7, no. 12, pp. 11713–11724, 2020.
- [4] K. Park, Y. Park, A. K. Das, S. Yu, J. Lee, and Y. Park, "A dynamic privacy-preserving key management protocol for v2g in social internet of things," *IEEE Access*, vol. 7, pp. 76812–76832, 2019.
- [5] C.-M. Chen, B. Xiang, Y. Liu, and K.-H. Wang, "A secure authentication protocol for internet of vehicles," *IEEE Access*, vol. 7, pp. 12047–12057, 2019.
- [6] S. Ahmed, S. Kumari, M. A. Saleem, K. Agarwal, K. Mahmood, and M.-H. Yang, "Anonymous key-agreement protocol for v2g environment within social internet of vehicles," *IEEE Access*, vol. 8, pp. 119829–119839, 2020.
- [7] T. A. Butt, R. Iqbal, K. Salah, M. Aloqaily, and Y. Jararweh, "Privacy management in social internet of vehicles: review, challenges and blockchain based solutions," *IEEE Access*, vol. 7, pp. 79694–79713, 2019.
- [8] L. Zhang, Z. Zhao, Q. Wu, H. Zhao, H. Xu, and X. Wu, "Energy-aware dynamic resource allocation in uav assisted mobile edge computing over social internet of vehicles," *IEEE Access*, vol. 6, pp. 56700–56715, 2018.
- [9] S. A. Eftekhari, M. Nikooghadam, and M. Rafiqhi, "Security-enhanced three-party pairwise secret key agreement protocol for fog-based vehicular ad-hoc communications," *Vehicular Communications*, vol. 28, Article ID 100306, 2021.
- [10] M. Ma, D. He, H. Wang, N. Kumar, and K.-K. R. Choo, "An efficient and provably secure authenticated key agreement protocol for fog-based vehicular ad-hoc networks," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8065–8075, 2019.
- [11] I. Azimi, A. Anzanpour, A. M. Rahmani, P. Liljeberg, and T. Salakoski, "Medical warning system based on internet of things using fog computing," in *Proceedings of the 2016 international workshop on big data and information security (IWBIS)*, pp. 19–24, IEEE, Jakarta, Indonesia, October 2016.
- [12] B. Ismail, A. Sari, and P. Österberg, "Security implications of fog computing on the internet of things," in *Proceedings of the 2019 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–6, IEEE, Las Vegas, NV, USA, January 2019.
- [13] X. Jia, D. He, N. Kumar, and K.-K. R. Choo, "Authenticated key agreement scheme for fog-driven iot healthcare system," *Wireless Networks*, vol. 25, no. 8, pp. 4737–4750, 2019.
- [14] M. Raya, P. Papadimitratos, and J.-P. Hubaux, "Securing vehicular communications," *IEEE wireless communications*, vol. 13, no. 5, pp. 8–15, 2006.
- [15] R. Lu, X. Lin, H. Zhu, P.-H. Ho, and X. Shen, "ECP: efficient conditional privacy preservation protocol for secure vehicular communications," in *Proceedings of the IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, pp. 1229–1237, IEEE, Phoenix, AZ, USA, April 2008.
- [16] C. Zhang, R. Lu, X. Lin, P.-H. Ho, and X. Shen, "An efficient identity-based batch verification scheme for vehicular sensor networks," in *Proceedings of the IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, pp. 246–250, IEEE, Phoenix, AZ, USA, April 2008.
- [17] J. Cui, L. Wei, H. Zhong, J. Zhang, Y. Xu, and L. Liu, "Edge computing in vanets-an efficient and privacy-preserving cooperative downloading scheme," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1191–1204, 2020.
- [18] X. Hu, J. Chen, M. Qian, and Y. Zhao, "Conditional privacy-preserving authentication protocol with dynamic membership updating for vanets," *IEEE Transactions on Dependable and Secure Computing*, 2020.
- [19] J. Li, H. Lu, and M. Guizani, "Acpn: a novel authentication framework with conditional privacy-preservation and non-repudiation for vanets," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 4, pp. 938–948, 2014.
- [20] D. Amit, N. Kumar, A. K. Das, and W. Susilo, "Secure message communication protocol among vehicles in smart city," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4359–4373, 2017.
- [21] M.-C. Chuang and J.-F. Lee, "Team: trust-extended authentication mechanism for vehicular ad hoc networks," *IEEE Systems Journal*, vol. 8, no. 3, pp. 749–758, 2013.
- [22] S. Kumari, M. Karupiah, X. Li, F. Wu, A. K. Das, and V. Odelu, "An enhanced and secure trust-extended authentication mechanism for vehicular ad-hoc networks," *Security and Communication Networks*, vol. 9, no. 17, pp. 4255–4271, 2016.
- [23] B. Ying and A. Nayak, "Anonymous and lightweight authentication for secure vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10626–10636, 2017.
- [24] P. Mohit, R. Amin, and G. P. Biswas, "Design of authentication protocol for wireless sensor network-based smart vehicular system," *Vehicular Communications*, vol. 9, pp. 64–71, 2017.
- [25] S. Yu, J. Lee, K. Lee, K. Park, and Y. Park, "Secure authentication protocol for wireless sensor networks in vehicular communications," *Sensors*, vol. 18, no. 10, p. 3191, 2018.
- [26] M. J. Sadri and M. Rajabzadeh Asaar, "A lightweight anonymous two-factor authentication protocol for wireless sensor networks in internet of vehicles," *International Journal of Communication Systems*, vol. 33, no. 14, Article ID e4511, 2020.
- [27] T.-Y. Wu, Z. Lee, L. Yang, and C.-M. Chen, "A provably secure authentication and key exchange protocol in vehicular ad hoc networks," *Security and Communication Networks*, vol. 2021, Article ID 9944460, 17 pages, 2021.
- [28] S. Bitam, A. Mellouk, and S. Zeadally, "Vanet-cloud: a generic cloud computing model for vehicular ad hoc networks," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 96–102, 2015.
- [29] Q. Jiang, J. Ni, J. Ma, L. Yang, and X. Shen, "Integrated authentication and key agreement framework for vehicular cloud computing," *IEEE Network*, vol. 32, no. 3, pp. 28–35, 2018.
- [30] M. A. Salahuddin, A. Al-Fuqaha, M. Guizani, and S. Cherkaoui, "RSU cloud and its resource management in support of enhanced vehicular applications," in *Proceedings of the 2014 IEEE Globecom Workshops (GC Wkshps)*, pp. 127–132, IEEE, Austin, TX, USA, December 2014.
- [31] Y. Rong, X. Huang, J. Kang et al., "Cooperative resource management in cloud-enabled vehicular networks," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7938–7951, 2015.
- [32] T.-Y. Wu, Z. Lee, L. Yang, J.-N. Luo, and R. Tso, "Provably secure authentication key exchange scheme using fog nodes in vehicular ad hoc networks," *The Journal of Supercomputing*, vol. 77, no. 7, pp. 6992–7020, 2021.
- [33] A. Maria, V. Pandi, J. Deborah Lazarus, M. Karupiah, and M. S. Christo, "BBAAS: blockchain-based anonymous authentication scheme for providing secure communication in

- vanets,” *Security and Communication Networks*, vol. 2021, Article ID 6679882, 11 pages, 2021.
- [34] M. Wazid, P. Bagga, A. K. Das, S. Shetty, J. J. P. C. Rodrigues, and Y. Park, “AKM-IOV: authenticated key management protocol in fog computing-based internet of vehicles deployment,” *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8804–8817, 2019.
- [35] R. M. Needham, M. Burrows, and M. Abadi, “A logic of authentication,” *Proceedings of the Royal Society of London. Mathematical and physical sciences Series*, vol. 426, no. 1871, pp. 233–271, 1989.
- [36] C. Ran, O. Goldreich, and S. Halevi, “The random oracle methodology, revisited,” *Journal of the ACM*, vol. 51, no. 4, pp. 557–594, 2004.
- [37] T.-Y. Wu, L. Yang, Z. Lee, C.-M. Chen, J.-S. Pan, and S. K. Hafizul Islam, “Improved ECC-based three-factor multiserver authentication scheme,” *Security and Communication Networks*, vol. 2021, Article ID 6627956, 14 pages, 2021.
- [38] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, “Zipf’s law in passwords,” *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2776–2791, 2017.
- [39] B. Blanchet, V. Cheval, X. Allamigeon, and B. Smyth, “Proverif: cryptographic protocol verifier in the formal model; 2012,” 2019, <https://prosecco.gforge.inria.fr/personal/bblanche/proverif/>.
- [40] Y. Luo, W. M. Zheng, and Y.-C. Chen, “An anonymous authentication and key exchange protocol in smart grid,” *Journal of Network Intelligence*, vol. 6, no. 2, pp. 206–215, 2021.

## Research Article

# Improving the Quality of Left-Behind Children's Participation in Sports through Wireless Network Monitoring

Jinjin Zhao 

Jilin Agricultural University, Changchun 130118, China

Correspondence should be addressed to Jinjin Zhao; zhaojinjin@jlau.edu.cn

Received 30 July 2021; Accepted 24 August 2021; Published 8 September 2021

Academic Editor: Ke Gu

Copyright © 2021 Jinjin Zhao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Urbanization is the process that people shift from rural to urban areas, which has led to large numbers of left-behind children in China. The left-behind children stay in rural regions of China while their parents work in urban areas. The left-behind children have few opportunities to participate in sports due to the lacking of concern, and it is not of high quality even though they participate in sports. Therefore, it is necessary to improve the quality of left-behind children's sports participation through wireless network monitoring. Wireless network monitoring transmits high-definition (HD) video streaming in real time to facilitate feedback timely. This paper studies the two-dimensional (2D) integer discrete cosine transform (DCT) and analyzes the reason for image distortion, then an improved DCT coefficient quantization approach is proposed for long-distance real-time transmission of HD video streaming, and a noise processing with a zero-mean noise processing is added in optimized approach to solve the image distortion problem. The experimental results show that the proposed improved approach has a good performance in reducing the blocking artifacts, and within the image reconstruction, the proposed approach improves the subjective video quality.

## 1. Introduction

The problem of left-behind children in China has become prominent over the past decades. More and more young and middle-aged farmers have worked in urban areas with the rapid advance in China's economy, and large numbers of left-behind children have emerged in the rural areas. According to related surveys, there are more than 10 million left-behind children in rural areas in China. Left-behind children have attracted the attention of the whole society. At present, the problems of left-behind children in health and education are more prominent, and most of them lack physical exercise, so it is necessary to improve the quality of left-behind children's sports participation through wireless network monitoring [1].

Wireless network monitoring transmits high-definition video streaming in real time to facilitate feedback timely [2]. The HD video streaming consists of moving images by frame one after another [3, 4]. When the images are continuously displayed at a certain rate, people will have a sense of video

due to the persistence of vision by eyes. The processing of video streaming can be roughly divided into the following three steps: (i) the original HD video streaming information is processed digitally, including quantization, coding, and other processes; (ii) the digital HD video streaming is transmitted to the receiver; and (iii) receiver will receive digital HD video streaming for antidigital processing, including inverse quantization, decoding, and other processes. Video monitoring is completed by the above steps in order to observe the quality of left-behind children's participation in sports.

High-Definition Multimedia Interface (HDMI) is a digital video/audio interface technology for video transmission; at the same time, the receiver can receive audio signals [5]. HDMI transmission distance is up to 30 meters, so it is not suitable for long-distance transmission. The HDMI cable can be seen as a low-pass filter. If the digital signals of HD video streaming are transmitted over long distances through the related interface, the high-frequency components will inevitably be severely attenuated. In

addition, the transmission rate of digital signals of HD video streaming is very fast and more likely to cause intersymbol interference, so that the receiver is difficult to decode and display. Therefore, in order to realize real-time and long-distance transmission of HD video streaming, it is necessary to compress the signal source. The improved DCT quantization approach can reserve the raw data to the greatest extent, which means eliminating redundant information as far as possible. As a result, it is necessary to compress the HD video streaming of HDMI.

Accordingly, the main contributions of this paper are summarized as follows: (i) an improved DCT coefficient quantization approach for long-distance real-time transmission of HD video streaming is proposed and (ii) a noise processing with a zero-mean noise processing is added in an optimized approach to solve the image distortion problem.

The remainder of this paper is organized as follows. Section 2 reviews the related work. In Section 3, the 2D DCT approach is studied. In Section 4, an improved DCT coefficient quantization approach is proposed. The experimental results are shown in Section 5. Section 6 concludes this paper.

## 2. Related Work

Many strategies of DCT for video or image processing have been proposed. In [6], a new multiobjective optimization algorithm was proposed to search for an efficient integer DCT matrix, which had the coding performance as close as possible to the transform in high-efficiency video coding but implemented with reduced hardware and power. In [7], an efficient hybrid image fusion method was proposed which is suitable for visual sensor networks based on the integer lifting wavelet transform and the DCT. In [8], a reconfigurable transform architecture was presented to flexibly support the reusability of different transform sizes. The proposed architecture maximally reused the hardware resources by rearranging the order of input data for different transform sizes while still exploiting the butterfly property. In [9], a new approximation for the 8-point discrete tchebichef transform was proposed with a higher power-and compression efficiency by exploring coefficient truncation. In [10], an image-dependent optimum nonnegative integer bit allocation algorithm was proposed, which was then mapped into desired image-independent solution via utilization of a prepared combined image and proposed modified step size mapping technique. In [11], the authors presented a hardware architecture for  $8 \times 8$  2D DCT and inverse DCT using Taylor-series expansion of trigonometric functions. In [12], the authors presented an efficient and low complexity integer approximation of the DCT for image compression. Their new approach involved replacing the bit shift elements of a variant of the signed DCT transform by zeros, in order to eliminate the bit shift operations. In [13], an efficient hardware implementation was proposed for high-speed vector-radix decimation-in-frequency three-dimensional DCT with an optimum area and power consumption. In [14], a novel algorithm was proposed to determine the minimum number of low-frequency DCT

coefficients required for transform and quantization block in high-efficiency video coding. In [15], a novel semifragile watermarking technique using integer wavelet transform and DCT for tamper detection and recovery to enhance enterprise multimedia security was proposed. In [16], a new model of inverse DCT kernel for high-efficiency video coding was proposed. In [17], the authors presented a blind and robust scheme using YCbCr color space, integer wavelet transform, and DCT for color image watermarking.

Some studies on blocking artifacts have also been proposed. In [18], the authors presented a deep network to eliminate image compression artifacts (usually denoted by image deblocking) based on image fusion in a multiscale manner. In [19], a novel dual-residual network was proposed to reduce compression artifacts caused by lossy compression codecs. In [20], a new method of image upscaling along with deblocking of compressed images was proposed. In [21], a wavelet transform based on the Meyer algorithm with edge-angle tracking capability was proposed for edge and blocking artifact reduction of an image, during image compression processes. In [22], a blocking artifact detection method was proposed for motion-compensated frame-rate upconversion algorithms. In [23], the authors presented the design of a partial overlapping block using exact Legendre moment computation for gray-level image reconstruction. There have also been several researches with respect to image compression [24–27].

## 3. Image Compression

Improving the quality of left-behind children's participation in sports is mainly achieved through wireless network monitoring, and how to improve the quality of video has become the key to research. The large volume of HD video streaming is well beyond the Ethernet transmission ability, and there are strict requirements in the real-time performance of HD video, so it needs to be a more appropriate compression approach to perform the compression of real-time HD video streaming. Thus, the real-time transmission of peer to peer can be realized by using HDMI through Ethernet. Theoretically, HDMI can send both video and audio streaming at a data transfer rate up to 4.5 Gb/s. According to the above, taking the resolution of  $1920 \times 1080$  HD video streaming as an example, which requires that the transfer rate must be 2 Gb/s. Considering the fact that no IP packet loss occurs when HD video streaming is transmitted over 100 M-Ethernet, the transfer rate on the physical link should be less than 100 Mb/s. Consequently, the compression ratio should be 25:1 to 34:1.

*3.1. DCT.* The original HD video streaming is highly correlated, which also has enough redundancy, while using the compression approach is to eliminate redundancy and reserve the original information as much as possible. That is, by reducing the correlation of the HD video streaming sequence, less bits are used to quantify the HD video streaming in order to compress the HD video streaming. There are three aspects in information redundancy that are



interframe redundancy, intraframe redundancy, and information entropy redundancy.

Under such context, this paper has relatively strict requirements on real-time and compression complexity.

Accordingly, I use a DCT-based video compression algorithm while 2D DCT is defined as follows:

$$F(u, v) = \frac{2}{\sqrt{AB}} c(u)c(v) \left[ \sum_{i=0}^{A-1} \sum_{j=0}^{B-1} f(i, j) \cos \frac{(i+0.5)u\pi}{A} \cos \frac{(j+0.5)v\pi}{B} \right], \quad (1)$$

$$F(i, j) = \frac{2}{\sqrt{AB}} \left[ \sum_{u=0}^{A-1} \sum_{v=0}^{B-1} c(u)c(v)F(u, v) \cos \frac{(i+0.5)u\pi}{A} \cos \frac{(j+0.5)v\pi}{B} \right],$$

where  $u = 0$  and  $v = 0$ , and  $c(u) = c(v) = (1/\sqrt{2})$ ; when  $u, v \neq 0$ ,  $c(u) = c(v) = 1$ .

**3.2. 2D DCT Approach.** The characteristics of human vision are not sensitive to the distortion of the image caused by the loss of high-frequency component packets in HD video streaming. In addition, there are a small number of high-frequency components in the image. Hence, the distributed energy in the image can be concentrated in the middle and low-frequency region by DCT. In other words, it makes the most of the energy of the image in the upper left of the DCT coefficients, and this can be transmitted to the receiver using a wireless network. Then, the receiver will realize the inverse transformation of the energy concentrated in the low-frequency parts of DCT so as to recover the original image information. However, from the DCT transformation formula, it can be seen that the computation of float is still huge, which conflicts with the requirements of low latency and low complexity in this paper; moreover, in the receiver decoding process, it will inevitably affect the accuracy. For this reason, the 2D integer DCT is used as the intraframe compression algorithm in this paper.

A 2D DCT can be realized by conducting one-dimensional DCT (1D DCT) twice, in which a 1D DCT is performed in the row first and then performed in the column. The 2D DCT can be represented by the matrix as shown in equation (2). Figuratively speaking, it is to figure out which 2D cosine waves constitute the image.

$$F = AfA^T, \quad (2)$$

where  $F$  is the coefficients of the transformation,  $f$  is the pixels of the image, and  $A$  is the transformation matrix.

In the process of 2D integer DCT, in order to reduce the computation and complexity, all the elements of the transformation matrix  $A$  are modified to integers and the scale factor matrix is added in the quantization process so as to realize the transformation and quantization. Meanwhile, the coding rate is also greatly improved. 2D integer DCT transformation matrix is defined as follows:

$$F = (A_{\text{int}} f A_{\text{int}}^T) \otimes E, \quad (3)$$

where  $E$  is the scale factor and  $A_{\text{int}}$  is the transformation matrix.

In the video compression algorithm, the single frame from the video is composed of many  $8 \times 8$ -pixel blocks. To realize the domain transformation of the image content, 2D integer DCT is used to transform the  $8 \times 8$ -pixel blocks, respectively, and then 64 DCT coefficients can be obtained. This paper uses the 2D integer DCT to effectively avoid the error of precision in the computation of float in inverse DCT. The computation is greatly reduced to addition and subtraction and shift operation by means of 2D integer DCT, and the division is further avoided. Moreover, the data within the pixel block is highly correlated. In this way, a large number of 0 coefficients will appear in the lower right after 2D integer DCT, and then “z” font will be used for scanning and coding, so that the volume of data will be much smaller. This lays a foundation for the realization of real-time transmission with low latency.

After using the above transformation processing, the next step is to use the conventional quantization method to operate the 64 DCT coefficients of each pixel block obtained by transformation. Quantization can not only adjust the low energy coefficient to zero which realizes certain compression of the image content but also reduce the transformed DCT coefficients, which provides convenience for the subsequent 8b/10b code transmission. According to the characteristics of human vision, they are not sensitive to the loss of high-frequency components, so high-frequency information can be lost to compress large amounts of data. However, the adjacent pixel blocks ( $8 \times 8$ ) in the content of a single frame image are quantized separately, resulting in no relation between the quantization errors of adjacent blocks. For the boundary pixels of adjacent blocks, if the quantization error between adjacent blocks has a jump, the smooth texture in the original image content will change greatly on the boundary region between adjacent blocks, so that the image distortion gets more serious. I take a frame of left-behind children participating in sports from the video for analysis. The comparison of the 2D integer DCT compression effects is shown in Figure 1, and the comparison of local magnification of 2D integer DCT compression effects is shown in Figure 2.





FIGURE 1: The comparison of the 2D integer DCT compression effects.

It can be seen from Figure 2 that the visual effect of local magnification of reconstructed images is obviously worse because there is a contradiction between the extremely limited bandwidth and the real-time transmission of big data. Therefore, it is necessary to compress the data source of HD video streaming effectively to realize the transmission of HD video streaming with low latency. And in the process of image domain transformation, DCT also greatly destroys the correlation between pixel blocks. In addition, the conventional quantization processing of transformed DCT coefficients will cause serious distortion to the reconstructed image at the receiver.

#### 4. Improved Approach

**4.1. The Reason for Image Distortion.** In this paper, the 2D integer DCT is used as the intraframe compression algorithm, but the distortion of the image reconstructed by the receiver is still serious which is caused by the coefficient quantization in 2D integer DCT. The coefficient quantization process and reconstruction process of classical integer DCT are defined as follows:

$$F_{Q_s}(u, v) = \text{round}\left(\frac{F_D(u, v)}{Q_s(u, v)}\right), \quad (4)$$

$$F_R(u, v) = F_{Q_s}(u, v)Q_s(u, v),$$

where  $F_{Q_s}(u, v)$  is the 2D integer DCT coefficients after quantization,  $F_D(u, v)$  is the 2D integer DCT coefficients before quantization,  $Q_s(u, v)$  is the quantization step size of the quantization process,  $F_R(u, v)$  is the reconstructed 2D integer DCT coefficients, and round is the round function. Then the quantization error of the 2D integer DCT coefficients is defined as follows:

$$\sigma(u, v) = F_R(u, v) - F_D(u, v). \quad (5)$$

There is still a strong correlation between adjacent pixel blocks in the original single frame image content before 2D integer DCT is used, but the DCT process broke the correlation between adjacent pixel blocks. Moreover, all pixel blocks ( $8 \times 8$ ) in the image content are quantified simultaneously and independently, so that the quantization errors are not correlated. If there are quantization errors between adjacent pixel blocks, this will cause the original smooth texture to fluctuate at the boundary between adjacent blocks of pixels, that is, blocking artifacts. As shown in Figure 2, the boundary fluctuation of the reconstructed image at the

receiver gives a worse visual effect to the human eye. Therefore, it is necessary to use the necessary quantization approach to reserve the original information of the image as much as possible.

#### 4.2. Improved DCT Coefficient Quantization Approach.

The monitoring of left-behind children requires low latency and clear pictures, so as to handle some exceptions in time through monitoring. There are many approaches for the reduction of blocking artifacts. As can be seen from Figure 3, there are four boundaries around pixel block 1 [28]. If the quantization error of adjacent pixel blocks on one of the boundaries is discontinuous, this will cause the blocking artifacts.

According to the spatial redundancy of the image, it can be seen that there is a little change in a certain part of the image for the pixel quantization of a pixel block with the same color, and from the characteristics of human vision, the blocking artifacts are derived from the boundary of a single frame image. When the quantization error of pixel 1 is compared with the quantization error of the other four pixel blocks, if it is smaller than the mean value of the quantization error of the other four pixel blocks, then it is not processed. On the contrary, it indicates that the boundary error of the pixel block is relatively high and the blocking artifacts may easily occur in this region. The size of  $n$  of the sampling matrix (i.e., the number of nonzero values of sampling matrix) is determined by equation (6) for requantization of a pixel block so as to reduce the quantization error of the current pixel block and ensure the continuity of boundary quantization error. With the consideration of the low latency requirement of long-distance transmission, the complexity of the improved approach cannot be increased, and the source of HD video streaming is directly lost after conventional quantization compression, which leads to serious image distortion. In this paper, noise processing with zero-mean noise processing is added in an optimized approach to solve the image distortion problem [29]. Given this, an improved DCT coefficient quantization approach is proposed for long-distance real-time transmission of HD video streaming.

$$n = \left\lceil \frac{4 \times \sigma_1(u, v)}{\delta \times (\sigma_2(u, v) + \sigma_3(u, v) + \sigma_4(u, v) + \sigma_5(u, v))} \right\rceil, \quad (6)$$

where  $\sigma_1(u, v)$ ,  $\sigma_2(u, v)$ ,  $\sigma_3(u, v)$ ,  $\sigma_4(u, v)$ , and  $\sigma_5(u, v)$  are the quantization error of each pixel block, and  $\delta$  is the coefficient.



FIGURE 2: The comparison of local magnification of 2D integer DCT compression effects. (a) Original image. (b) Reconstructed image.

The optimization process is summarized as follows:

- Step 1: using integer DCT and compressing DCT coefficients.
- Step 2: using the boundary detection algorithm to detect the boundary of the whole reconstructed image.
- Step 3: dividing the DCT coefficients into two regions (upper left and lower right), as described, most of the energy in the image is concentrated at the DCT coefficients in the upper left. Therefore, the quantization approach of DCT coefficients in the upper left is readjusted. The new quantization equations (7) and (8) are defined as follows:

$$F_{Qs,\min}(u, v) = \frac{F_D(u, v)}{Qs(u, v)}, \quad (7)$$

$$F_{Qs,\max}(u, v) = \frac{F_D(u, v)}{Qs(u, v)} + 1. \quad (8)$$

Thus, each DCT coefficient in the upper left will have two possible quantized values which constitute  $2^n$  different quantization coefficients matrix. Each quantization coefficient matrix is reconstructed, and the boundary errors of four boundaries are calculated. When the sum of squares of the four boundary error is minimized, the quantization matrix value is the quantization value of the integer DCT, which is defined as follows:

$$\sigma'(u, v) = \sum_{k=1}^4 (F_R'(u, v) - F_D(u, v))_k^2. \quad (9)$$

- Step 4: repeating step 3 for each pixel block at the boundary of the image. A large number of tests show that although the blocking artifacts are

reduced, integer DCT belongs to unitary transformation and has the property of energy conservation. At the same time, the signal-to-noise ratio of the image may be slightly decreased after inverse DCT, and DCT coefficient value of  $F(0, 0)$ ,  $F(0, 1)$ ,  $F(1, 0)$ ,  $F(0, 2)$ , and  $F(2, 0)$  has a great impact on the blocking artifacts, so I select  $n$  value of equation (6) no less than 5.

The optimized DCT coefficient quantization is used to reduce the blocking artifacts of the image boundary, so that the video streaming received at the receiver and the original are consistent as much as possible, and people cannot detect the compression processing of the HD video streaming. Figure 4 shows the comparison of 2D integer DCT compression and improved quantization compression effects (the 30th frame image in the video) as the quantization step is 20. Figure 5 shows the comparison of local magnification of 2D integer DCT compression and improved quantization compression effects. It can be seen that the improved DCT coefficient quantization can get a better subjective visual effect.

## 5. Experiment and Results Analysis

**5.1. Setup.** The data source depends on the left-behind children of QingGangYuan hope primary school. QingGangYuan is a village in the city of Sinan, Guizhou Province, one of the poorest provinces in China. Among a hundred and ten students enrolled in QingGangYuan hope primary school, approximately a total of them are left-behind children or children whose parents are off to work in larger cities and rarely come home. Wireless network monitoring is used to improve the quality of left-behind children in sports. I take the first 100 frames of monitoring video to verify the improved approach with good performance. The first 100 frames of the Foreman, Claire, Carphone, and News (QCIF format) standard video streaming sequences are encoded in intraframe,

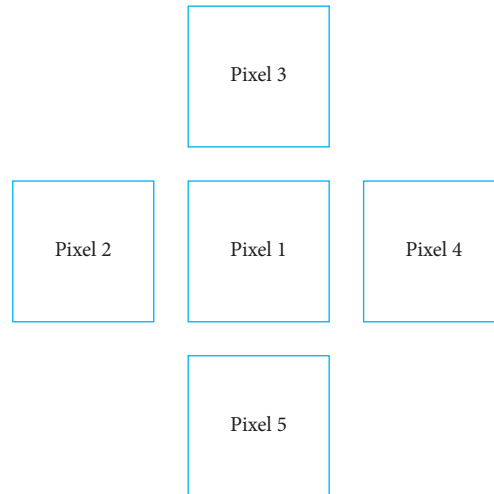


FIGURE 3: The distribution between pixel block 1 and adjacent pixel blocks.



FIGURE 4: The comparison of 2D integer DCT compression and improved quantization compression effects. (a) Original image. (b) Reconstructed image (improved quantization).

and H.264 video encoder is used. The quantization step is 10–25. According to blocking artifacts measurement [30], the approach with improved quantization and the approach without improved quantization are for comparison in reducing the blocking artifacts.

**5.2. Comparison Analysis.** In Figure 6,  $\delta = 0.9$ , and Figures 6(a)–6(d) show the blocking artifact size of reconstructed images. It can be seen that the improved quantization approach can effectively reduce the blocking artifacts; in particular, when the quantization step size is large, the improved approach can be well reflected.

In Figure 7, when  $\delta = 0.75$  and quantization step size is 20, the image is reconstructed at the 10th frame of the video. Among them, Figure 6(a) is the reconstructed

image without the improved quantization; that is, the high-frequency information is lost through direct compression, so the reconstructed image has the worst effect. Figure 6(b) is a reconstructed image using the improved quantization approach, and the image distortion problem is solved by adding zero-mean noise processing which has greatly improved the image reconstruction effect. Figure 6(c) uses a neural network-based deblocking method to reconstruct images [31]. Figure 6(d) uses a novel frame-wise filtering method to reconstruct images [32]. The compression effect of Figure 6(a) is the worst by the naked eye, while the approach proposed in this paper seems to be the best.

Figure 8 shows the local magnification comparison of each reconstructed image in Figure 7. It can be clearly seen that the improved quantization approach has a good effect





FIGURE 5: The comparison of local magnification of 2D integer DCT compression and improved quantization compression effects. (a) Original image. (b) Reconstructed image (improved quantization).

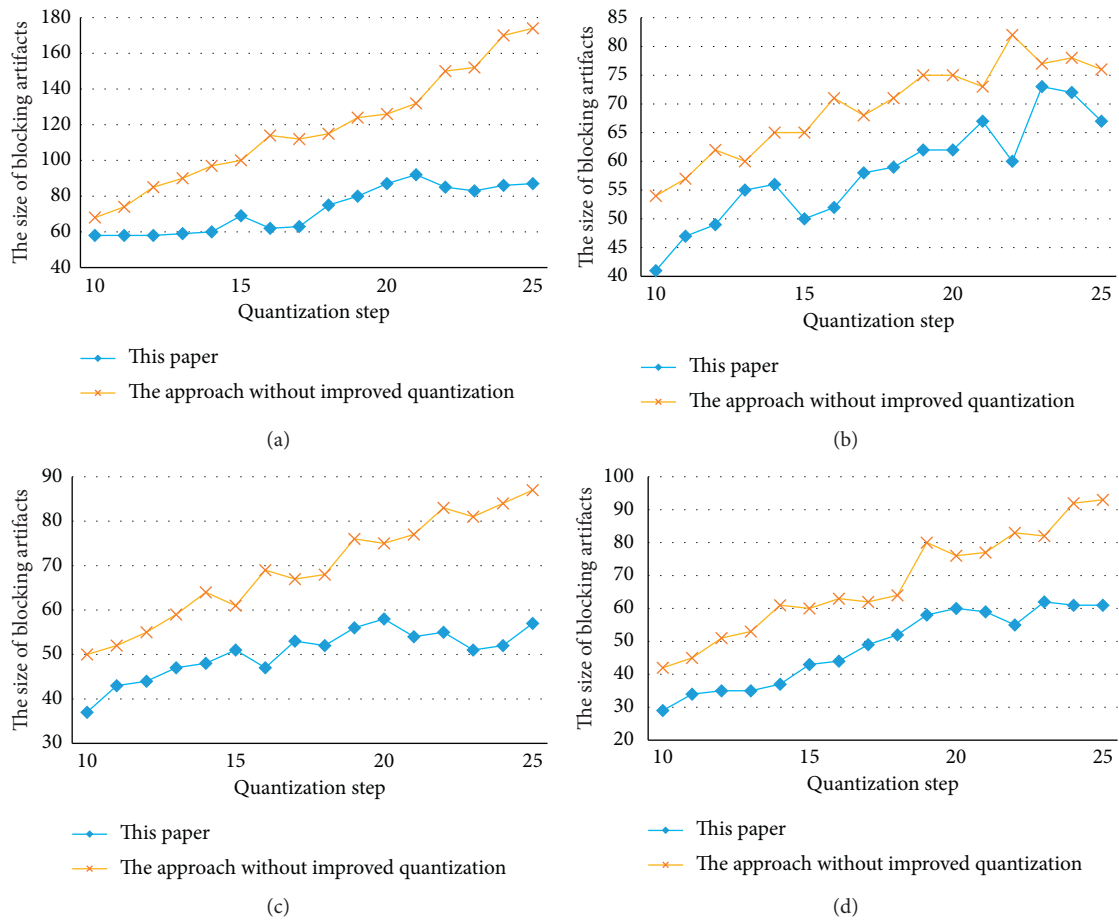


FIGURE 6: The comparison of reducing blocking artifacts. (a) Foreman sequence. (b) Claire sequence. (c) Carphone sequence. (d) News sequence.



FIGURE 7: The comparison of reconstructed images at the 10th frame of the video.



FIGURE 8: Continued.



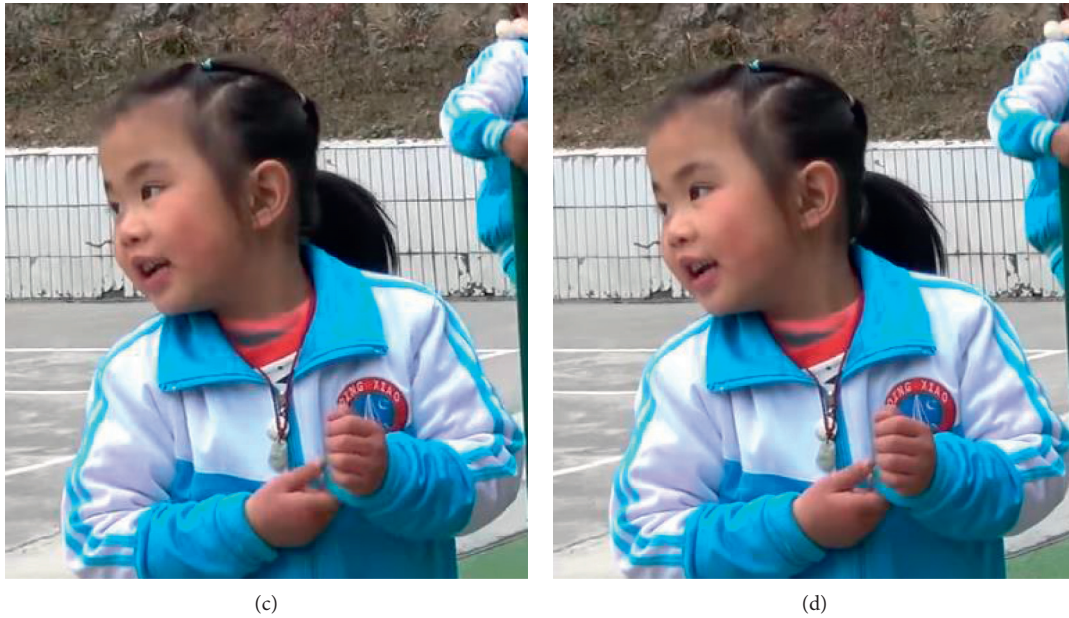


FIGURE 8: The comparison of local magnification of reconstructed images at the 10th frame of the video.

on the naked eye. The proposed approach can greatly reduce the blocking artifacts and improve the subjective video quality.

## 6. Conclusions

The increasing number of left-behind children in China has dramatically become a problem. They seldom take part in sports due to the lacking of monitoring. To improve the quality of left-behind children's participation in sports based on wireless network monitoring, an improved DCT coefficient quantization approach is proposed for long-distance real-time transmission of HD video streaming, and the image distortion problem is solved by adding zero-mean noise processing which has greatly improved image reconstruction effect. The experimental results demonstrate that the proposed improved approach has a good performance in reducing the blocking artifacts, and when the quantization step size is large, the improved quantization approach can be well reflected. However, there are some limitations of this paper. At first, in the process of transmitting video stream over a wireless channel, the performance of the video stream is greatly disturbed by external interference, so the visual effect of the receiver needs to be improved. Then, it can get better results by using field-programmable gate array high-speed parallel processing. The approach complexity will be optimized to improve the compression efficiency for further research.

## 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 no conflicts of interest.

## Acknowledgments

This work was supported by the National Social Science Foundation: 2016 "13th Five-Year Plan" Education Youth Project, grant no. CLA160197.

## References

- [1] C. Bi, D. Oyserman, Y. Lin, J. Zhang, B. Chu, and H. Yang, "Left behind, not alone: feeling, function and neurophysiological markers of self-expansion among left-behind children and not left-behind peers," *Social Cognitive and Affective Neuroscience*, vol. 15, no. 4, pp. 467–478, 2020.
- [2] J. Zhou, "Artificial intelligence driven wireless network remote monitoring based on Diffie-Hellman parameter method," *Computer Communications*, vol. 160, pp. 132–138, 2020.
- [3] M. Taha, A. Canovas, J. Lloret, and A. Ali, "A QoE adaptive management system for high definition video streaming over wireless networks," *Telecommunication Systems*, vol. 77, no. 1, pp. 63–81, 2021.
- [4] S. Xie, Y. Xu, Q. Shen, Z. Ma, and W. Zhang, "Modeling the perceptual quality of viewport adaptive omnidirectional video streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 9, pp. 3029–3042, 2020.
- [5] P.-L. Chen, "A fully synthesizable ultra-N audio frequency multiplier for HDMI applications," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 10, pp. 2134–2138, 2020.
- [6] J. Chen, S. Liu, G. Deng, and S. Rahardja, "Hardware efficient integer discrete cosine transform for efficient image/video compression," *IEEE Access*, vol. 7, pp. 152635–152645, 2019.
- [7] B. Latreche, S. Saadi, M. Kiouss, and A. Benziane, "A novel hybrid image fusion method based on integer lifting wavelet and discrete cosine transformer for visual sensor networks," *Multimedia Tools and Applications*, vol. 78, no. 8, pp. 10865–10887, 2019.

- [8] M. Zheng, J. Zheng, Z. Chen, L. Wu, X. Yang, and N. Ling, "A reconfigurable architecture for discrete cosine transform in video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 810–821, 2020.
- [9] G. Paim, L. M. G. Rocha, G. M. Santana, L. B. Soares, E. A. C. da Costa, and S. Bampi, "Power-, area-, and compression-efficient eight-point approximate 2-D discrete tchebichef transform hardware design combining truncation pruning and efficient transposition buffers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 2, pp. 680–693, 2019.
- [10] V. S. Thakur, K. Thakur, S. Gupta, and K. R. Rao, "Image-independent optimal non-negative integer bit allocation technique for the DCT-based image transform coders," *IET Image Processing*, vol. 14, no. 1, pp. 11–24, 2020.
- [11] D. Mukherjee and S. Mukhopadhyay, "Hardware efficient architecture for 2D DCT and IDCT using Taylor-series expansion of trigonometric functions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2723–2735, 2020.
- [12] N. Brahimi, T. Bouden, T. Brahimi, and L. Boubchir, "A novel and efficient 8-point DCT approximation for image compression," *Multimedia Tools and Applications*, vol. 79, no. 11–12, pp. 7615–7631, 2020.
- [13] V. Arunachalam, A. N. Joseph Raj, and S. Deepika, "Performance improvement of vector-radix decimation-in-frequency 3D-DCT/IDCT using variable word length," *Circuits, Systems, and Signal Processing*, vol. 40, no. 4, pp. 1818–1831, 2021.
- [14] A. Singhadia, P. Bante, and I. Chakrabarti, "A novel algorithmic approach for efficient realization of 2-D-DCT architecture for HEVC," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 3, pp. 264–273, 2019.
- [15] N. Sivasubramanian and G. Konganathan, "A novel semi fragile watermarking technique for tamper detection and recovery using IWT and DCT," *Computing*, vol. 102, no. 6, pp. 1365–1384, 2020.
- [16] S. Chatterjee and K. Sarawadekar, "WHT and matrix decomposition-based approximated IDCT architecture for HEVC," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 6, pp. 1043–1047, 2019.
- [17] R. Sinhal, D. K. Jain, and I. A. Ansari, "Machine learning based blind color image watermarking scheme for copyright protection?" *Pattern Recognition Letters*, vol. 145, pp. 171–177, 2021.
- [18] C.-H. Yeh, C.-H. Lin, M.-H. Lin, L.-W. Kang, C.-H. Huang, and M.-J. Chen, "Deep learning-based compressed image artifacts reduction based on multi-scale image fusion," *Information Fusion*, vol. 67, pp. 195–207, 2021.
- [19] J. Li, D. Li, C. Chen, Q. Yan, and X. Lu, "A dual-residual network for JPEG compression artifacts reduction," *Signal, Image and Video Processing*, vol. 15, no. 3, pp. 485–491, 2021.
- [20] A. Singh and J. Singh, "A content adaptive method of deblocking and super-resolution of compressed images," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11095–11131, 2021.
- [21] M.-T. Wu, "Wavelet transform based on Meyer algorithm for image edge and blocking artifact reduction," *Information Sciences*, vol. 474, pp. 125–135, 2019.
- [22] N.-U. Kim and Y.-L. Lee, "Blocking-artifact detection in frequency domain for frame-rate up-conversion," *Personal and Ubiquitous Computing*, vol. 22, no. 1, pp. 173–184, 2018.
- [23] Z. Bahaoui, K. Zenkouar, H. E. Fadili, H. Qjidaa, and A. Zarghili, "Blocking artifact removal using partial overlapping based on exact Legendre moments computation," *Journal of Real-Time Image Processing*, vol. 14, no. 2, pp. 433–451, 2018.
- [24] K. Al-Khayyat, I. Al-Shaikhli, and M. Al-Hagery, "Second compression for pixelated images under edge-based compression algorithms: JPEG-LS as an example," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 6, pp. 10661–10669, 2021.
- [25] S. Kudo, S. Orihashi, R. Tanida, S. Takamura, and H. Kimata, "GAN-based image compression using mutual information for optimizing subjective image similarity," *IEICE Transactions on Information and Systems*, vol. E104.D, no. 3, pp. 450–460, 2021.
- [26] A. M. Sandu, C. A. Mifale, M. A. A. Ungureanu, and E. I. Scarlat, "Case comparison between direct image compression and hologram compression," *University Politehnica of Bucharest Scientific Bulletin-Series A-Applied Mathematics and Physics*, vol. 83, no. 1, pp. 235–246, 2021.
- [27] D. Tellez, G. Litjens, J. van der Laak, and F. Ciompi, "Neural image compression for gigapixel histopathology image analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 2, pp. 567–578, 2021.
- [28] H. Zhang, B. Zheng, and H. Zu, "Application of SVGA video real-time transmission technology in music education information communication," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 2, pp. 1733–1744, 2020.
- [29] X. Xue and J. Zhang, "Matching large-scale biomedical ontologies with central concept based partitioning algorithm and adaptive compact evolutionary algorithm," *Applied Soft Computing*, vol. 106, pp. 1–11, 2021.
- [30] P. He, X. Jiang, T. Sun, and S. Wang, "Detection of double compression in MPEG-4 videos based on block artifact measurement," *Neurocomputing*, vol. 228, pp. 84–96, 2017.
- [31] Y. Zhang, E. Salari, and S. Zhang, "Reducing blocking artifacts in JPEG-compressed images using an adaptive neural network-based algorithm," *Neural Computing & Applications*, vol. 22, no. 1, pp. 3–10, 2013.
- [32] H. Huang, I. Schiopu, and A. Munteanu, "Frame-wise CNN-based filtering for intra-frame quality enhancement of HEVC Videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 6, pp. 2100–2113, 2021.

## Research Article

# An Intelligent Garbage Sorting System Based on Edge Computing and Visual Understanding of Social Internet of Vehicles

Xuehao Shen,<sup>1</sup> Yuezhong Wu ,<sup>1</sup> Shuhong Chen ,<sup>2</sup> and Xueming Luo<sup>3</sup>

<sup>1</sup>College of Railway Transportation, Hunan University of Technology, Zhuzhou 412007, China

<sup>2</sup>School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China

<sup>3</sup>Scholl of Computer Science, Hunan University of Technology, Zhuzhou 412007, China

Correspondence should be addressed to Yuezhong Wu; wuyuezhong@hut.edu.cn and Shuhong Chen; shuhongchen@gzhu.edu.cn

Received 10 July 2021; Revised 6 August 2021; Accepted 21 August 2021; Published 31 August 2021

Academic Editor: Ke Gu

Copyright © 2021 Xuehao Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to enable Social Internet of Vehicles devices to achieve the purpose of intelligent and autonomous garbage classification in a public environment, while avoiding network congestion caused by a large amount of data accessing the cloud at the same time, it is therefore considered to combine mobile edge computing with Social Internet of Vehicles to give full play to mobile edge computing features of high bandwidth and low latency. At the same time, based on cutting-edge technologies such as deep learning, knowledge graph, and 5G transmission, the paper builds an intelligent garbage sorting system based on edge computing and visual understanding of Social Internet of Vehicles. First of all, for the massive multisource heterogeneous Social Internet of Vehicles big data in the public environment, different item modal data adopts different processing methods, aiming to obtain a visual understanding model. Secondly, using the 5G network, the model is deployed on the edge device and the cloud for cloud-side collaborative management, aiming to avoid the waste of edge node resources, while ensuring the data privacy of the edge node. Finally, the Social Internet of Vehicles devices is used to make intelligent decision-making on the big data of the items. First, the items are judged as garbage, and then the category is judged, and finally the task of grabbing and sorting is realized. The experimental results show that the system proposed in this paper can efficiently process the big data of Social Internet of Vehicles and make valuable intelligent decisions. At the same time, it also has a certain role in promoting the promotion of Social Internet of Vehicles devices.

## 1. Introduction

Social Internet of Vehicles (SIOV) is considered to be the core component of the future intelligent transportation system, and it is also one of the most promising practical technologies for 5G vertical applications [1]. Since the US Department of Transportation issued the “Intelligent Transport System (ITS) Strategic Plan” in 2015, SIOV technology has been vigorously developing around the two themes of intelligence and information sharing [2]. On November 11, 2020, the World Intelligent Connected Vehicle Conference released the “Intelligent Connected Vehicle Technology Roadmap 2.0,” which further pointed out the direction for the development of intelligent vehicle networking [3]. It is a feasible solution to use vehicle as edge

node device to provide computing services and offload tasks to the edge of the network. In recent years, with the development of social economy and the improvement of material consumption, the appearance of domestic waste has become more and more diversified and complicated, and the global waste production has also shown a cliff-like growth compared with previous years. In response to this, my country has successively introduced a series of policies. In December 2016, the fourteenth meeting of the Central Finance and Economics Leading Group hosted by General Secretary Xi Jinping proposed that “it is necessary to accelerate the establishment of a garbage disposal system for classified release, classified collection, classified transportation, and classified treatment, and form a waste treatment system based on the rule of law, promoted by the



government, and a garbage classification system with participation of the whole people, urban and rural planning, and local conditions, and strive to increase the coverage of the garbage classification system.” The latest revision of the “Law of the People’s Republic of China on the Prevention and Control of Environmental Pollution by Fixed Wastes” in 2020 requires that local people’s governments at or above the county level should speed up the establishment of a domestic waste management system for classified release, classified recycling, transportation, and treatment. At present, China’s economy is developing rapidly, people’s quality of life has been greatly improved, public environmental issues have become the focus of attention, and one of the key factors affecting the public environment is the garbage issue.

At this stage, garbage classification is mainly concentrated in the outdoor public environment for fixed groups of people to deal with according to categories. There are problems such as high labor intensity, low sorting efficiency, and poor working environment, people’s low awareness of classification, and a wide variety of garbage. Therefore, from the perspectives of practicability, environmental protection, and intelligence, it is of great significance to study and design an intelligent garbage sorting vehicle system based on edge computing of SIoV. In recent years, with the rapid development of artificial intelligence and robotics, service robots have attracted widespread attention. At present, there have not been public reports about the work carried out by service robots for autonomous garbage detection and classification. At the same time, service robots are one of the edge devices in SIoV. Therefore, it is of great practical significance to implement garbage classification and detection algorithms on service robots. However, only using the detection and classification model can only realize the identification and positioning of garbage, and the degree of intelligence is not high. If you want to make robots have the ability to recognize and discriminate objects in a public environment like humans, for example, humans can understand what they see. The items in the scene can be inferred and classified based on the association and imagination based on these items, so they should not only rely on the appearance and geometric characteristics of the items, but also rely on the guidance and reasoning of the high-level prior knowledge of the items. Public environmental goods information has the characteristics of diversity, semantics, and relevance, so it can be considered to use knowledge graphs to express and store this rich prior knowledge in a structured form. Then, the effective garbage detection and classification algorithm is used to complete the identification and positioning of items. The use of 5G networks to realize collaborative computing between edge device nodes and the cloud and intelligent decision-making of garbage classification on the big data of items in the scene are the key issues studied in this paper. The main innovations and contributions of this paper include the following: First, build a new visual understanding model. This model uses the knowledge graph to uniformly characterize and store the multimodal information of items in the public environment and combines the YOLOv4 detection algorithm to identify and locate items in the scene; the second is to propose the use of cloud-side collaborative

computing. Using the 5G network, the visual understanding model deployed on the edge device and the cloud are used for cloud-side collaborative management. The cloud is used to store a large amount of data, while avoiding the waste of edge node resources, while ensuring the data privacy of the edge node; the third is to build an intelligent garbage sorting system based on edge computing and visual understanding of SIoV. Common items can be detected, identified, and classified by edge devices; abnormal items can be connected to the cloud for identification, reasoning, and decision-making.

The organizational structure of the rest of this paper is as follows. Related work will be discussed in the second part. The third part introduces the overall design of the system in detail. The fourth part mainly introduces the experimental setup and result analysis. The fifth part is the summary of this paper.

## 2. Related Work

*2.1. Deep Learning.* Deep learning is an important breakthrough in the field of artificial intelligence in the past decade, and it is widely used in target detection and classification. Target detection can be divided into two categories in terms of methods. One is the regression-based one-stage algorithm represented by the YOLO [4] series and the SSD series, and the other is based on the candidate region two-stage algorithm represented by Fast R-CNN and Faster R-CNN. In recent years, many researchers have used deep learning technology in the research of garbage identification, classification, and detection. Literature [5] proposed a method of automatic identification of garbage types combined with ResNet-50 and multilevel SVM. GCNet [6] consists of a feature extractor and a classifier. The feature extractor uses ResNet101 as the backbone network, which contains 5 Bottlenecks, and each Bottleneck adds an attention mechanism, and then the extracted different features are merged. Literature [7] proposed an improved network’s robustness framework DNN-TC for junk image recognition. The network adds two fully connected layers after the output layer and the global average pooling layer of the classification latitude to reduce model parameter redundancy. The last layer uses the log softmax function to calculate the confidence of each label. Literature [8] uses DenseNet with higher detection accuracy and at the same time modifies the connection mode between dense blocks to achieve the purpose of improving the detection speed. Ma Wen et al. [9] used ResNet50 to replace the original VGG16 basic network in Faster R-CNN and Soft-NMS instead of the original NMS, which improved accuracy and reduced time. Wang Mingjie [10] used K-mean++ to determine the size of the prior frame and then used migration learning to complete the location and classification of garbage using YOLOv3 and classified the waste into recyclable items, dry garbage, wet garbage, and hazardous garbage. However, with the increasing demand for garbage classification by mobile edge devices and considering the accuracy and real-time performance, this paper uses YOLOv4 as the basic network. However, in order to distinguish it from the existing YOLO

detection algorithm, the model not only relies on a large amount of labeled data to train and fit a large number of parameters for prediction, but also considers the role of prior knowledge to guide the reasoning of the model. Therefore, it is planned to add a knowledge graph on the basis of the YOLOv4 algorithm to further enhance the intelligent level of the system.

*2.2. Knowledge Graph.* The knowledge graph can use the knowledge triples composed of nodes and relationships to intuitively represent the association relationships between items in the scene, and it can be stored in a structured form at the same time. Therefore, in a public environment, it is a very effective method to use knowledge graph to express rich visual associate on information and prior knowledge of objects. Knowledge Graph (KG) [11] is a technical method that uses graph models to describe the relationship between knowledge and modeling the world. KG was first applied to improve the capabilities of search engines; after that, KG showed greater application value in assisting intelligent question and answering, natural language processing, big data analysis, recommendation calculations, and interpretable artificial intelligence [12–14]. Marino et al. [15] studied the application of structured prior knowledge in the form of knowledge graphs in image classification. Jiang et al. [16] proposed a hybrid knowledge routing module in view of the current detection algorithm ignoring the semantic association information of the target in the scene and the long tail phenomenon of the sample size distribution of different categories. Chen et al. [17] introduced statistical target objects and their possible coexistence of prior knowledge to constrain the relationship prediction space to improve the accuracy of the model in fewer categories. Wang et al. [18] introduced the prior knowledge of the association between the characters in the scene and the surrounding objects and performed explicit reasoning based on knowledge. Wu et al. [19] proposed a visual question and answer method, which constructs a textual representation of the semantic content of an image and merges it with the textual information from the knowledge base to achieve a deeper understanding of the scene.

*2.3. Edge Computing.* As a new paradigm, edge computing can sink the computing functions and services of the cloud to the network edge devices, providing real-time data analysis and intelligent processing nearby, which can effectively solve the problems of network congestion and network delay caused by the transmission and processing of massive data. At present, edge computing accelerates the transformation and upgrading of the economy by providing key-capabilities such as computing, network, and intelligence nearby. It has gradually become a new direction of the computer system and a new format in the information field and has received extensive attention from academia and industry. With the popularization and development of products and application scenarios such as smart phones, smart homes, and smart connected cars, artificial intelligence is gradually migrating from the cloud to the embedded

end of the edge, and intelligent edge computing has emerged from this [20]. The concept based SIOV is an extension of the Internet of Things. Real-time collection of vehicle operating data is achieved through on board sensing units, roadside acquisition modules, vehicle-to-road communication units, and other equipment, then builds a data platform for monitoring large-scale vehicle real-time operating information, and provides various data service [21]. In the era of the Internet of everything, the massive amounts of data generated by various smart devices have put forward higher requirements on computing, storage, and network service capabilities. In order to relieve the computing pressure on the cloud and at the same time improve the computing power and operating efficiency of the mobile side, some services are deployed at the network edge close to the mobile side to build a mobile edge computing system [22–24]. The introduction of edge computing based SIOV is an inevitable trend. However, edge computing is deployed near the network infrastructure. On the one hand, it is vulnerable to attacks from edge vehicles and network infrastructure such as counterfeiting, privacy theft, and false information; on the other hand, unauthorized internal attackers may also access and steal storage at the edge. Sensitive information is in the data center [25–27]. Therefore, the efficient processing and valuable intelligent decision-making [28] based SIOV big data on the edge device side can protect the privacy and safety of vehicles in the edge computing of SIOV.

*2.4. SLAM.* Simultaneous localization and mapping (SLAM) is a process in which the robot uses its own vision, laser, and other sensors to complete its own positioning while constructing environmental maps and path planning [29]. The SLAM system that uses the camera to collect image information as the source of environmental perception information is called Visual SLAM (VSLAM). Compared with other SLAM systems, VSLAM can perceive richer colors, textures, and other environmental information. With the rapid development of deep learning technology, it has very successful applications in various fields of Computer Vision (CV). For example, SLAM technology is playing an increasingly important role in the fields of service robots and driverless cars. In this context, in recent years, more and more SLAM researchers use deep learning-based methods to extract environmental semantic information to obtain high-level scene perception and understanding and apply it in the VSLAM [30] system to assist VSLAM. The system improves positioning performance and map visualization, thereby giving robots more efficient human-computer interaction capabilities. The combination of deep learning and SLAM improves the limitations caused by manual design features and potentially improves the learning ability and intelligent level of the robot [31]. VSLAM can construct a 3D map of the surrounding environment and calculate the position and direction of the camera. The combination of deep learning and SLAM is a hot research direction in recent years. Among them, the semantic SLAM combining VSLAM and deep learning obtains environmental geometric information during the mapping process and at the same time recognizes

independent objects in the environment and obtains semantic information such as their positions, poses, and individual contours, which expands the research content of traditional SLAM problems. Integrate some semantic information into SLAM research to cope with the requirements of complex scenarios [32]. Therefore, this paper applies SLAM on the edge device to further enhance the autonomy of the system.

### 3. Overall System Design

*3.1. Overall Design.* This paper experimentally designs an intelligent garbage sorting system based on edge computing and visual understanding of SIOV. The overall architecture of the system is shown in Figure 1. This architecture diagram is divided into three parts:

- (1) Visual understanding model: First, the KG is used to uniformly characterize and store the multimodal information knowledge of items in the public environment. The YOLOv4 detection algorithm is used to identify and locate the items in the scene, and the constructed multimodal KG is combined with the YOLOv4 visual detection method. Construct a visual understanding model, as shown in Figure 2. Then put the model on the cloud server for large-scale data training and then deploy it to the NVIDIA Jetson Nano development board after the training is completed. Finally, when the edge device is in a public scene, use the camera to collect pictures in real time and transmit the pictures wirelessly to the development board. The development board uses the trained model to detect whether there is garbage in the picture, and if it exists, it sends it to the driver board STM32 through the serial port. STM32 controls the robotic arm to grab the garbage and put it into the corresponding garbage bin according to the recognition and classification results. The judgment of this process is to realize the intelligent decision-making of items through intelligent question and answer technology and realize the goal of garbage classification.
- (2) Cloud edge collaboration: What is used here is an edge-oriented cloud-side collaborative computing form. In this form, the cloud is only responsible for the initial training work, and the model is downloaded to the edge after the training is completed. While performing computing tasks at the edge, it will also use real-time on-site data to perform subsequent calculations on the model. This model can meet individual application requirements, make better use of local data, and avoid waste of edge node resources to ensure the edge data privacy of the node.
- (3) Edge device applications: The mapping navigation unit is based on the ROS distributed framework, uses lidar to collect the environmental information of the cleaning area, realizes the SLAM function based on the scanning matching algorithm, and uses the optimal path algorithm to autonomously plan and

traverse the cleaning area. During the traversal process of the edge sorting device, the target detection algorithm detects and classifies the real-time images obtained by the camera and obtains the coordinates and angle information of the target as the input information of the sorting control unit and controls the robotic arm to perform the garbage capture task.

In order to enable edge device to achieve intelligent and automated garbage classification in a public environment, while avoiding network congestion caused by simultaneous access to a large amount of data, and considering the combination of mobile edge computing and SIOV, in order to give full play to the high bandwidth and low latency characteristics of mobile edge computing, this paper builds an intelligent garbage sorting system based on edge computing and visual understanding of SIOV (Figure 2). First of all, according to the existence of two modalities of video and image in the public environment, the YOLOv4 detection algorithm is used to extract the location of the entity category; use BLSTM-LCRF and PCNN-BLSTM-Attention proposed by Wang Huan et al. [33] to extract entities and relationships from text modalities. The open source structured data collected from the Internet and the entity relationship extracted above form a knowledge triple. Secondly, the knowledge triples are used to uniformly represent and store the semantic description information, attribute information, and spatial location information of the items in the scene using the KG. Finally, when detecting and classifying garbage items in a public environment, the YOLOv4 detection algorithm will perform real-time detection to obtain its location and category information. At the same time, it will use OCR technology to obtain the description information of the item's outer packaging, and the previously constructed KG will be further developed through VQA technology. The auxiliary detection model matches and determines whether the item is garbage and what type of garbage in the form of intelligent question and answer and makes further intelligent decision-making.

*3.2. Multimodal Knowledge Graph.* With the continuous popularization of the Internet and media technologies, information from different sources such as text, images, video, audio, etc. collectively portrays the same or related content, presenting complex and multilevel semantic relationships, forming "multimodal" information. First of all, multimedia data containing different modalities present internally synchronized semantic associations, while information from different sources and modalities across media presents dynamic, complex, multilayered temporal and semantic associations. Secondly, the cross-media forms are heterogeneous, the content is diverse, and the distribution is complex. The traditional analysis and processing methods are mostly based on the assumption of independent and identical distribution, and it is difficult to effectively utilize and learn the massive and complex cross-media information. Finally, the application scenarios involved in cross-media are more extensive, such as cross-media content

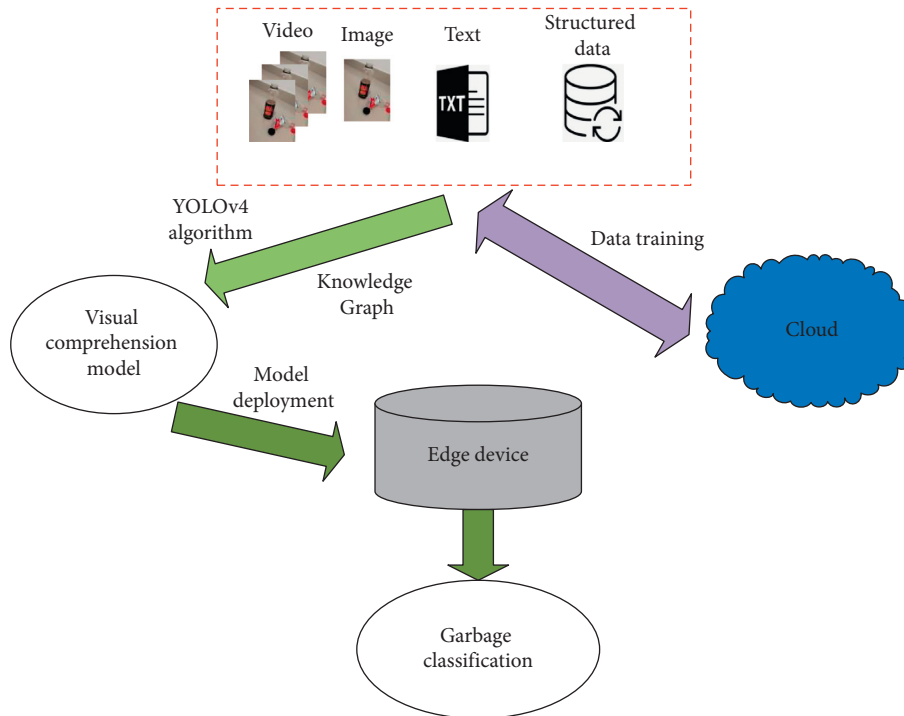


FIGURE 1: System overall architecture diagram.

search, recommendation, question and answer, etc. Abundant item data has been accumulated in public scenarios, such as “item name,” “item packaging,” “item category,” and “item production information.” However, the data are not related to each other, failing to form effective knowledge. Constructing the item data into a KG and developing upper-level applications can effectively enable garbage classification in public scenarios, solve the pain points of the scenario, and give full play to the value of knowledge. The construction and application process of the KG is shown in Figure 3.

The process of building knowledge graphs of objects in public scenarios: ontology design, knowledge extraction, knowledge mapping, knowledge fusion, and disambiguation. Ontology design is a process of knowledge modeling. This process requires the realization of ontology design by summarizing the knowledge in the field. Generally speaking, RDFS, OWL, and other languages can be used for modeling. The semiautomated form realizes the paper knowledge modeling in public scenarios and quickly completes the paper KG ontology design. The ontology includes concepts such as item names, item attributes, and associations between items. Knowledge extraction completes the extraction of knowledge triples based on relevant algorithms by collecting relevant data. For structured data, after simple conversion, triples can be generated. For unstructured data, the document format needs to be converted first to obtain easy-to-handle text formats such as txt and docx; literature [30] proposed BLSTM-LCRF and PCNN-BLSTM-Attention models to extract entities and relationships from text data. Knowledge mapping, the triples obtained in the information extraction stage, needs to map the extraction results to the ontology through knowledge mapping, in order to generate

a KG. If the amount of data is large, the process can be accelerated with the help of big data technologies such as Hadoop. Finally, it is necessary to integrate and disambiguate heterogeneous data under a unified standard for the knowledge from different data sources to complete the creation of the KG. For example, an item with the same apple name can actually refer to both fruits and apple mobile phones. Typical applications included in the paper KG based on public scenarios include intelligent question answering, intelligent reasoning, and intelligent decision-making.

**3.3. YOLOv4 Network.** Based on the YOLOv3 algorithm, Alexey B officially launched YOLOv4 in 2020 (shown in Figure 4). Its network structure is divided into input, backbone feature extraction network (CSPDarknet53), enhanced feature extraction network (Spatial Pyramid Pooling, SPP, and Path Aggregation Network, PANet), and detection network (YOLO Head). The principle of YOLOv4 model detection is to divide the image into an  $S \times S$  grid, where each cell in the grid is responsible for predicting  $B$  location bounding boxes and conditional probabilities belonging to  $C$  categories, and finally output whether the bounding box contains the target and the bounding box confidence information. YOLOv4 uses CSPDarknet53 instead of Darknet53 in its backbone network. The main changes are as follows: first, the introduction of CSPNet (Cross Stage Partial Network) network structure. The gradient variability is integrated and mapped to the feature map from beginning to end, and then the feature map is divided into two parts: one part performs residual stacking operation, and the other part is directly combined with the last convolution result, which effectively solves the gradient information factor in

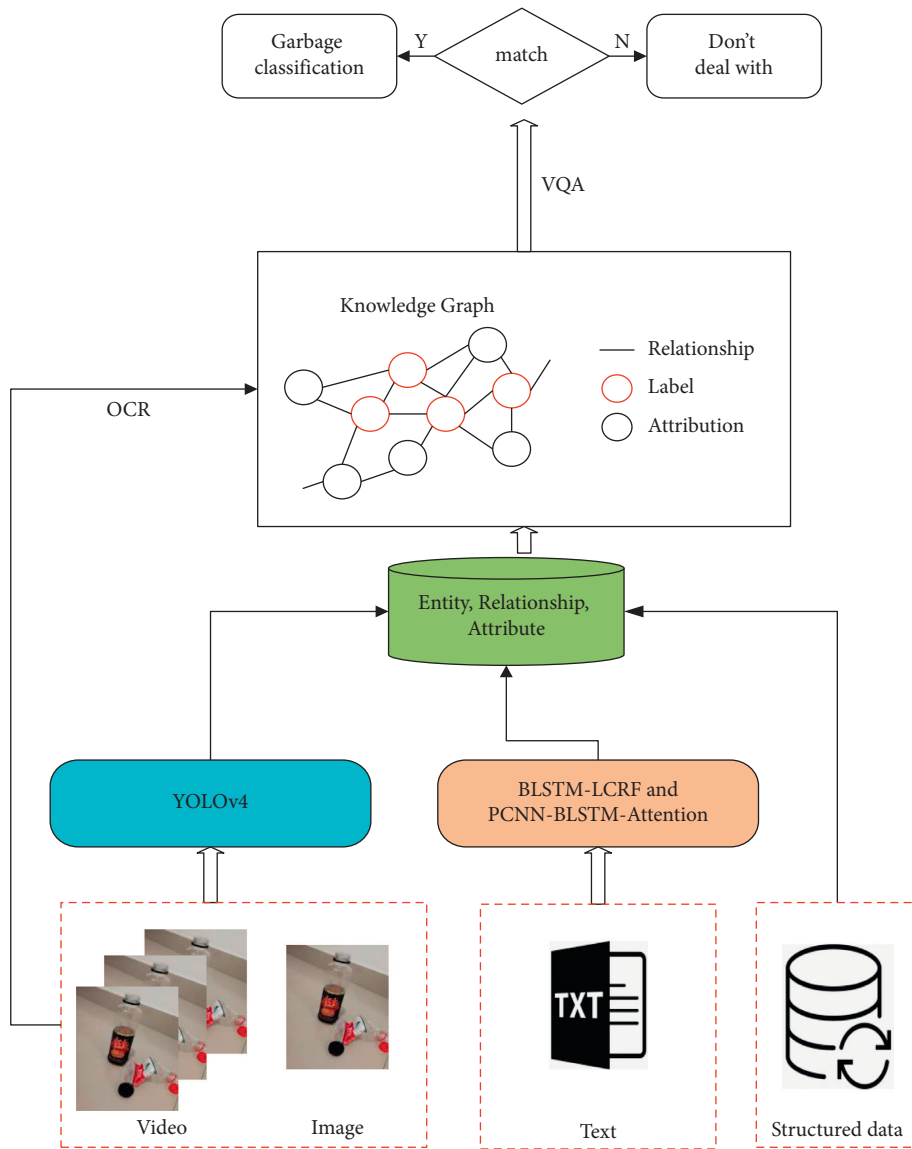


FIGURE 2: Visual comprehension model.

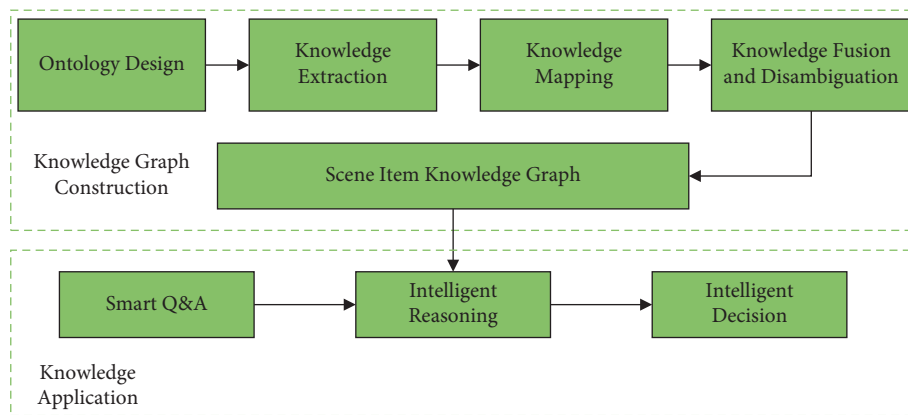


FIGURE 3: Knowledge graph construction and application flow chart.

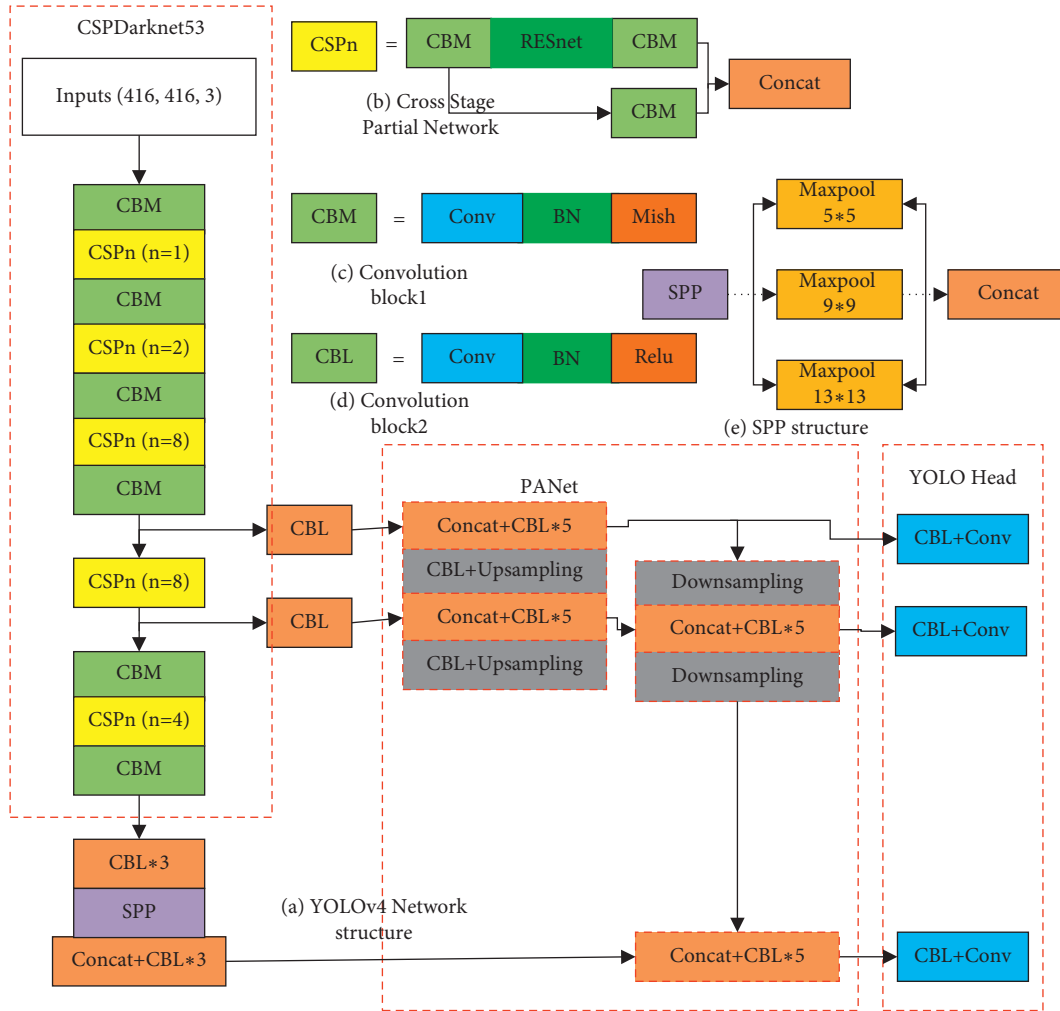


FIGURE 4: YOLOv4 network structure diagram.

the deep network processing; repeated learning causes the problem of increased calculation. The second is to replace the activation function LeakyReLU with the Mish activation function, so that the overall network detection has a higher accuracy rate. Introduce the Spatial Pyramid Pooling (SPP) layer and use the maximum pooling of different pooling core sizes to pool the input feature layer, using the maximum pooling method of  $1 \times 1$ ,  $5 \times 5$ ,  $9 \times 9$ , and  $13 \times 13$ , respectively. This structure can use different dimensions of the same image as input to obtain pooling features of the same length. The SPP network can effectively increase the range of feature acceptance and extract more important contextual features without reducing the network operation speed. PANet is created on the basis of FPN (Feature Pyramid Networks) in YOLOv3. It is a bottom-up path enhancement designed to promote information flow and use accurate low-level positioning information to enhance the overall feature level, thereby shortening the information path between the low-level and top-level features. This structure makes full use of feature fusion and changes the previous additive fusion method to multiplicative fusion, so that the network has higher detection accuracy. The detection part still uses YOLO Head detection using the YOLOv3 algorithm.

## 4. Experimental Setup and Result Analysis

**4.1. Experimental Environment Configuration.** The experiment in this paper is completed under the Ubuntu system. CUDA is a general parallel computing architecture launched by NVIDIA; CUDNN is a GPU acceleration library for deep neural networks. The data is trained through the cooperation of the two, and the experimental environment is configured, as shown in Table 1.

First, the collected image data is cleaned to remove some invalid images; then the processed valid data set is labeled according to the PASCAL VOC data set format; then the migration learning method is adopted, and the pretraining of yolov4.pt provided on the official website is used. The weight is used as the initial parameter of network training; finally, after 70 epochs of iteration, the loss function reaches a minimum and tends to a balanced state, and the training is stopped to obtain the best weight file. The model training loss function curve in this paper is shown in Figure 5. Some network parameter descriptions are shown in Table 2.

The model training process in this paper: calculate and predict the input data through forward propagation to obtain the loss function of the network model, then use the

TABLE 1: Experiment environment configuration.

| Project                  | Experimental environment |
|--------------------------|--------------------------|
| System                   | Ubuntu18.04              |
| Programming environment  | Pycharm                  |
| GPU                      | NVIDIA TITAN RTX         |
| Memory                   | 24 GB                    |
| Pytorch version          | Pytorch1.6               |
| Python version           | Python3.6                |
| CUDA version             | CUDA10.1                 |
| CUDNN version            | CUDNN7.6                 |
| Data bases               | Neo4j4.2.2               |
| Java runtime environment | JDK15.0.1                |

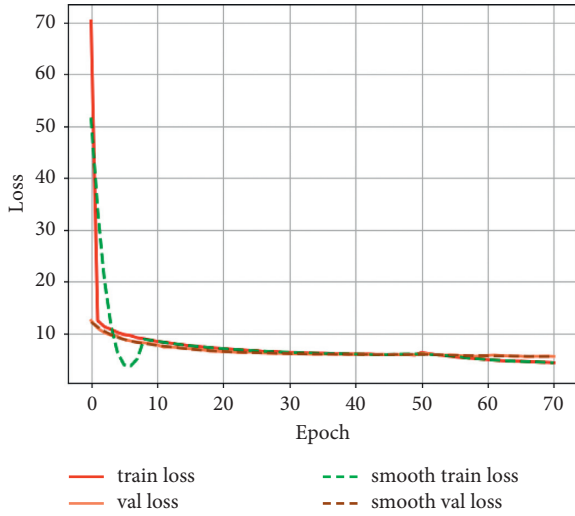


FIGURE 5: Loss curve of our model.

TABLE 2: Network parameter description table.

| Parameter name | Parameter values |
|----------------|------------------|
| Learning rate  | 0.001            |
| Momentum       | 0.9              |
| Decay          | 0.0005           |
| Batch size     | 64               |

combination of backpropagation and gradient descent to find the direction of gradient descent, and update the model weight parameters, and the whole process is repeated iteratively, and a network model with better detection effect is finally obtained. The loss function used in this network model is as follows:

$$\begin{aligned}
\text{Loss} = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B W_{ij}^{\text{obj}} [(x_i - x'_i)^2 + (y_i - y'_i)^2] \\
& + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B W_{ij}^{\text{obj}} \left[ (\sqrt{w_i} - \sqrt{w'_i})^2 + (\sqrt{h_i} - \sqrt{h'_i})^2 \right] \\
& + \sum_{i=0}^{S^2} \sum_{j=0}^B W_{ij}^{\text{obj}} (C_i - C'_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B W_{ij}^{\text{noobj}} (C_i - C'_i)^2 \\
& + \sum_{i=0}^{S^2} W_{ij}^{\text{obj}} \sum_{c \in \text{classes}} [(p_i(c) - p'_i(c))]^2.
\end{aligned} \tag{1}$$

It can be seen from formula (1) that the entire loss function is composed of three parts: the first two lines in the formula represent the error of the center coordinates and width and height of the prediction box, which is the first part, where  $(x_i, y_i, w_i, h_i)$  represents the center coordinates and width and height of the prediction box, which means marking the center coordinates and width and height of the box. The third row represents the confidence error, which is the second part. The left half of the third row represents the confidence error that the prediction box contains the target, and the right half represents the confidence error that the prediction box does not contain the target, where  $C_i$  indicates that the prediction box contains the object confidence, where  $C'_i$  represents the IOU of the predicted box and the marked box. The fourth row represents the category error of the target, which is the third part, where  $p_i(c)$  represents the marked category value and  $p'_i(c)$  represents predicted category value. In formula (1),  $W_{ij}^{\text{obj}}$  means that the target is detected by the  $j$  prediction frame in the first  $i$  grid, and  $W_{ij}^{\text{noobj}}$  means that the target is not detected, and  $W_i^{\text{obj}}$  means that the target is in the first  $i$  grid.

**4.2. Data Collection.** The data set used in this paper has a total of 15,000 domestic garbage pictures, most of which come from the data set in the garbage classification competition held by Alibaba Cloud Tianchi and some pictures of domestic garbage collected by the author. The data set can be divided into four categories in general, namely, recyclable garbage, kitchen waste, hazardous garbage, and other garbage. Each category contains multiple objects. Among them are recyclable trash: power bank, bag, wash supplies, plastic toy, plastic utensils, plastic hangers, glassware, metalware, courier bags, plug wire, old clothes, ring-pull can, pillow, plush toy, shoes, cutting board, carton, wine bottle, metal food can, ironware, wok, edible oil drum, drink bottle, paper books; harmful trash: dry battery, unguentum, expired drugs; other trash: disposable snack box, stained plastic, butt, toothpick, flowerpot, chinaware, chopsticks, stained paper; use the LabelImage tool to label the items in the picture, and divide the data set into a training set and a test set at a ratio of 8 : 2.

**4.3. Experimental Results and Analysis.** The evaluation criteria of the results of this experiment are mainly Precision (P), Recall (R), Mean Average Precision (MAP), and detection speed Frames Per Second (FPS). Among them, P represents the ratio of the real samples in the recognized positive samples, namely,

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{2}$$

Among them,  $R$  represents the proportion of correctly identified positive samples in the total number of samples, namely,

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{3}$$



TABLE 3: Algorithm detection result.

| Algorithm | P (%) | R (%) | MAP (%) | FPS |
|-----------|-------|-------|---------|-----|
| YOLOv4    | 84.3  | 81.9  | 72.5    | 26  |
| Our model | 84.6  | 82.3  | 73.8    | 24  |

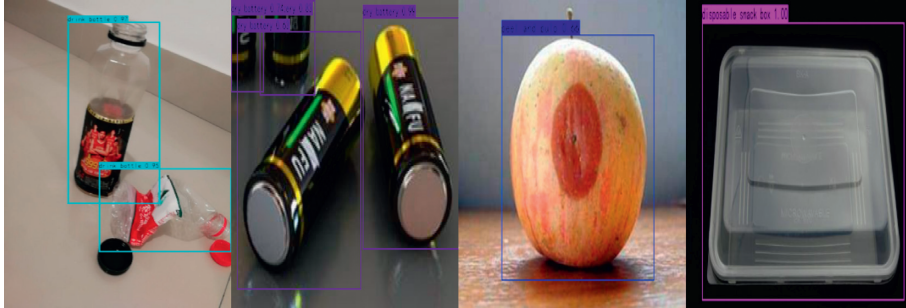


FIGURE 6: Visualization result graph.

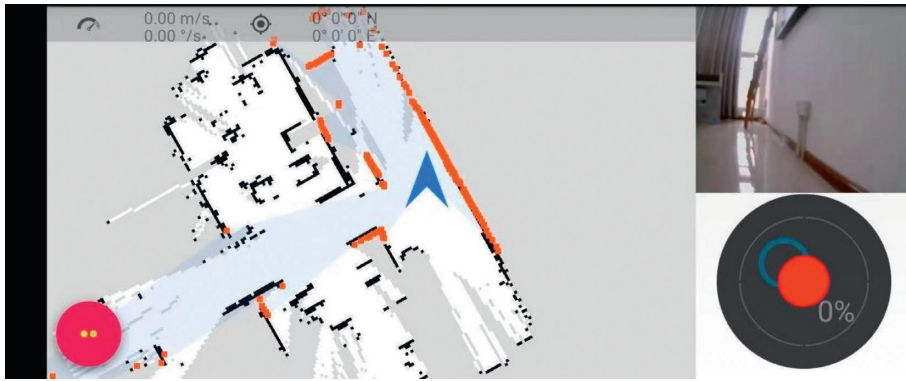


FIGURE 7: Build map.



FIGURE 8: Navigation.

TP is the number of positive samples that correctly classify the target, FP is the number of positive samples that are incorrectly classified as the target, and FN refers to the number of positive samples that are incorrectly classified as negative samples.

In order to verify the effectiveness of this test method, the visual understanding model and YOLOv4 proposed in this paper were trained and verified with different algorithms according to the network parameters in Table 2. The  $P$ ,  $R$ ,  $MAP$ , and  $FPS$  are shown in Table 3.





FIGURE 9: Recognition result.

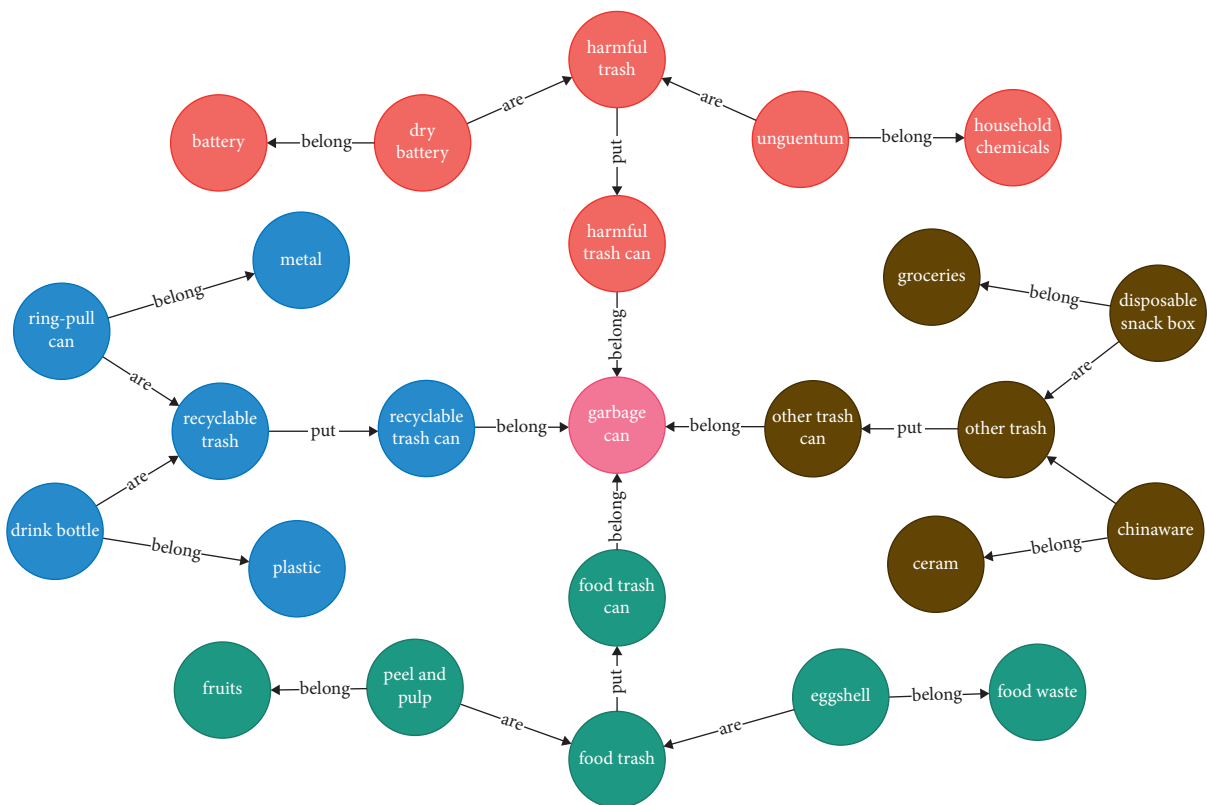


FIGURE 10: Local knowledge graph.

From the results in Table 3, the model proposed in this paper has better performance than YOLOv4. Under the premise of equivalent detection speed, the MAP can reach nearly 74%. Figure 6 shows part of the visualization results of the model detection proposed in this paper.

Taking into account the limited equipment in the actual public environment, this paper uses a trolley to simulate an

indoor scene. The results of mapping, navigation, and recognition results are shown in Figure 7, Figure 8 and Figure 9.

Figure 10 shows the constructed knowledge graph of the local scene. Different colors represent different types, among which red is harmful trash, blue is recyclable trash, green is food trash, and gray is other trash. Through the YOLOv4 algorithm, the entity name and location information of the



FIGURE 11: Classification result.

item in the scene can be obtained, the identified entity target object is matched with the entity in the KG, and the location relationship or attribute information between the entities is used to make further intelligent decisions and judgments. If the item is garbage, determine the specific type of garbage and return it to the corresponding garbage bin to facilitate the next step of sorting by the edge device; if it is not garbage, the model will not process it. Figure 11 is the result of querying the target entity in the KG and performing garbage classification. For example, a deformed beverage bottle on the ground is a recyclable trash and should be placed in a recyclable trash can.

### 5. Conclusions

This paper builds an intelligent garbage sorting system based on edge computing and visual understanding of SIoV. Experimental results show that the system can provide efficient and valuable intelligent decision-making, free human hands, reduce back-end waste processing, and improve work efficiency; at the same time, through cloud-side collaborative computing, the use of edge devices can be fully utilized, which can avoid cloud network congestion. It also guarantees the data privacy of edge nodes. In addition, due to the limited data set currently collected, the constructed KG is not complete, and the classification effect of some

uncommon or severely defaced items is relatively poor. On the one hand, future research work can add multimodal data analysis experiments to visual understanding; on the other hand, it can also consider using 5G networks to achieve collaborative work between multiple edge devices and increase related algorithm analysis and experiments in SIoV.

### Data Availability

The data used to support the findings of this study are not applicable because the data interface cannot provide external access temporarily.

### Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

### Acknowledgments

This work was supported in part by the National Key R&D Program of China under Grant nos. 2019YFE0122600 and 2018YFB1700200, in part by the Hunan Provincial Key Research and Development Project of China under Grant no. 2019GK2133, in part by the Natural Science Foundation of Hunan Province under Grant no. 2021JJ50050, and in part

by the Scientific Research Project of Hunan Provincial Department of Education under Grant no. 19B147.

## References

- [1] Y. Xiao, H. Liu, and X. H. Cheng, "Key technologies of internet of vehicles and their development trends and challenges," *Communications Technology*, vol. 54, no. 1, pp. 1–8, 2021.
- [2] Y. Wei, "Analysis on the development of internet of vehicles based on 5G communication technology," *Application of Mechanics-electronics Technology*, vol. 54, no. 3, pp. 223–225, 2019.
- [3] F. Gao, "Interpretation of intelligent connected vehicle technology Roadmap 2.0," *Internet of Things Technology*, vol. 11, no. 1, pp. 3–4, 2020.
- [4] A. Bochkovski and C. Y. Wang, "YOLOv4: optimal speed and accuracy of object detection," 2020, <https://arxiv.org/abs/2004.10934v1>.
- [5] O. Adedeji and Z. Wang, "Intelligent waste classification system using deep learning convolutional neural network," *Procedia Manufacturing*, vol. 35, pp. 607–612, 2019.
- [6] Z. Y. Dong and W. G. Han, "Classification algorithm of garbage images based on convolutional neural network," *Computer Systems & Applications*, vol. 29, no. 8, pp. 199–204, 2020.
- [7] A. H. Vo, L. Hoang Son, M. T. Vo, and T. Le, "A novel framework for trash classification using deep transfer learning," *IEEE Access*, vol. 7, pp. 178631–178639, 2019.
- [8] C. Bircanoglu, M. Atay, F. Beser, M. A. Kizrak, and O. Genc, "RecycleNet: intelligent waste sorting using deep neural networks," in *Proceedings of the 2018 IEEE International Conference on INnovations in Intelligent Systems and Applications (INISTA)*, July 2018.
- [9] W. Ma, J. Yu, and J. Y. Chen, "Garbage detection and classification based on improved Faster R-CNN," *Computer Engineering*, vol. 47, no. 8, pp. 294–300, 2020.
- [10] M. J. Wang, "Automatic garbage location and classification based on YOLOV3," *Wireless Internet Technology*, vol. 16, no. 20, pp. 110–112, 2019.
- [11] A. Singhal, "Official Google Blog: Introducing the Knowledge Graph: things, not strings," 2012, <http://googleblog.blogspot.pt/2012/05/introducing-knowledge-graph-things-not.html>.
- [12] M. Li, Z. Sun, S. Zhang, and W. Zhang, "Enhancing knowledge graph embedding with relational constraints," *Neurocomputing*, vol. 429, pp. 77–88, 2021.
- [13] Z. Li, H. Liu, Z. Zhang, T. Liu, and J. Shu, "Recalibration convolutional networks for learning interaction knowledge graph embedding," *Neurocomputing*, vol. 427, pp. 118–130, 2021.
- [14] F. Gong, M. Wang, H. Wang, S. Wang, and M. Liu, "SMR: medical knowledge graph embedding for safe medicine recommendation," *Big Data Research*, vol. 23, 2021.
- [15] K. Marino, R. Salakhutdinov, and A. Gupta, "The more you know: using knowledge graphs for image classification," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20–28, 2017.
- [16] C. Jiang, H. Xu, X. Liang, and L. Lin, "Hybrid knowledge routed modules for large-scale object detection," in *Proceedings of the In Advances in Neural Information Processing Systems*, pp. 1552–1563, Montreal Convention Centre, Montreal, Canada, December 2018.
- [17] T. Chen, W. Yu, R. Chen, and L. Lin, "Knowledge-embedded routing network for scene graph generation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019*, pp. 6163–6171, Long beach california, June 2019.
- [18] Z. X. Wang, T. S. Chen, S. J. Jimmy, W. Yu, H. Cheng, and L. Lin, "Deep reasoning with knowledge graph for social relationship understanding," 2018, <https://arxiv.org/abs/1807.00504v1>.
- [19] Q. Wu, P. Wang, C. Shen, D. Anthony, and A. van den Hengel, "Ask me anything: free-form visual question answering based on knowledge from external sources," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4622–4630, Las Vegas, NV, USA, June 2016.
- [20] Z. Wang, "Development status and trend outlook of edge computing," *Automation Panorama*, vol. 38, no. 2, pp. 22–29, 2021.
- [21] C. C. Juan, Z. Sherali, and G. Juan Antonio, "Internet of vehicles: architecture, protocols, and security," *IEEE Internet of Things Journal*, vol. 5, pp. 3701–3709, 2017.
- [22] W. S. Shi, X. Z. Zhang, Y. F. Wang, and Q. Zhang, "Edge computing: state-of-the-art and future directions," *Journal of Computer Research and Development*, vol. 56, no. 1, pp. 73–93, 2019.
- [23] Z. Liao, Y. Ma, J. Huang, J. Wang, and J. Wang, "HOTSPOT: a UAV-assisted dynamic mobility-aware offloading for mobile-edge computing in 3-D space," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10940–10952, 2021.
- [24] J. Zhang, J. Cui, H. Zhong, Z. Chen, and L. Liu, "PA-CRT: Chinese remainder theorem based conditional privacy-preserving authentication scheme in vehicular ad-hoc networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 722–735, 2021.
- [25] R. Roman, J. Lopez, and M. Mambo, "Mobile Edge Computing, Fog et al.: a Survey and Analysis of Security Threats and Challenges," *Future Generation Computer Systems*, vol. 78, pp. 680–698, 2018.
- [26] J. Zhou, X. L. Dong, and Z. F. Cao, "Research advances on privacy preserving in recommender systems," *Journal of Computer Research and Development*, vol. 56, no. 10, pp. 2033–2048, 2019.
- [27] Y. Z. Wu, Q. Liu, R. R. Chen, C. Li, and Z. Peng, "A group recommendation system of network document resource based on knowledge graph and LSTM in edge computing," *Security and Communication Networks*, vol. 2020, Article ID 8843803, , 2020.
- [28] D. Liu and Y. Luo, "A consensus reaching process based on the concordance correlation measure of intuitionistic fuzzy sets in multi-criteria decision making," *Journal of Intelligent & Fuzzy Systems*, pp. 1–16, 2021.
- [29] C. Cadena, L. Carlone, H. Carrillo et al., "Past, present, and future of simultaneous localization and mapping: toward the robust-perception age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [30] Y. Zhao, G. L. Liu, G. H. Tian et al., "A survey of visual SLAM based on deep learning," *Robot*, vol. 39, no. 6, pp. 889–896, 2017.
- [31] S. P. Li and T. Zhang, "A survey of deep learning application in visual SLAM," *Aerospace Control and Application*, vol. 45, no. 2, pp. 1–10, 2019.
- [32] R. J. Liu, X. S. Wang, and C. Zhang, "A survey on visual SLAM based on deep learning," *Journal of System Simulation*, vol. 32, no. 7, pp. 1244–1256, 2020.
- [33] H. Wang, W. Q. Zhu, Y. Z. Wu, P. J. He, and L. J. Wan, "Named entity recognition based on equipment and fault field of CNC machine tools," *Journal of Engineering Science*, vol. 42, no. 4, pp. 476–482, 2020.