Received March 25, 2019, accepted April 7, 2019, date of publication April 23, 2019, date of current version May 3, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2912897

Content Recommendation Algorithm for Intelligent Navigator in Fog Computing Based IoT Environment

FUHONG LIN[1], YUTONG ZHOU[2], ILSUN YOU[3], (Senior Member, IEEE), JIUZHI LIN[1], XINGSHUO AN[1], AND XING LÜ[1]

[1]School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China
[2]Institute of Education and Economy Research, University of International Business and Economics, Beijing 100029, China
[3]Department of Information Security Engineering, Soonchunhyang University, Asan-si 31538, South Korea

Corresponding author: Ilsun You (ilsunu@gmail.com)

This work was supported in part by the National Key R&D Program of China under Grant 2017YFC0820700, and in part by the Soonchunhyang University Research Fund.

ABSTRACT With the development of the Internet and mobile technologies, the Internet of Things (IoT) era has arrived. Vehicle networking technology can not only facilitate people's travel but also effectively alleviate traffic congestion. The development of fog computing technology provides unlimited possibilities for the Internet of Vehicles (IoV). Intelligent navigator is a very important part of human–computer interaction in IoV. It carries a large number of tasks of recommending content for users. In order to get more accurate recommendation content, we propose a weighted interest degree recommendation algorithm using association rules for intelligence in the IoV. First, the user data are analyzed to establish the association rule mining algorithm. Second, the user interest score is predicted by analyzing the relevance between user interests to recommend personalized service for the user. From the simulation results, we can see that the proposed algorithm can achieve higher recommendation accuracy.

INDEX TERMS Content recommendation, association rules, Internet of Vehicles, fog computing.

I. INTRODUCTION

We are embracing a new era of the Internet of Things (IoT) [1]. With the increasing of IoT devices, mass data is generated which may cause network congestion [2]. The traditional cloud computing paradigm cannot avoid such network congestion. Also many delay-sensitive services could not being guaranteed today's network architecture [3]. Then, a new computing paradigm called fog computing is proposed which can solve the above problems by migrating some services from the cloud server to the network edge. A typical fog computing architecture is composed by a user layer, a fog service layer, and a cloud service layer from the bottom up (Shown in Figure 1 [4]).

The application of fog computing in the IoT is very extensive [5]. For example, in the Internet of Vehicles (IoV), the user layer consists of on-board sensors and user equipment. The fog node consists of lightweight servers located at the network edge [6]. The cloud server is in charge of managing and monitoring the fog nodes and the entire network [7], [8]. In such an application scenario, the user and the in-vehicle navigation device generate a large amount of interaction, for example, the user may need the smart navigation device to personalize the content (bar, restaurant, and supermarket) that is of interest to the user [9]. In the fog computing, the workflow of the content recommendation of the vehicle intelligent navigator is shown in Figure 2.

When the fog node provides service, the massive retrieval of information is accompanied by an increase in user search and screening time [10], and a decrease in the precision rate. This causes the users are hardly to find the right data from a vast amount of network information resources [11]. The explosive growth of network data has made the problem of "information overload" more serious [12]. A large amount of unrelated redundant data information seriously interferes with people's choice of relevant useful information, making the cost of obtaining high-quality and valuable information higher and higher for users. The dataflow of the content

The associate editor coordinating the review of this manuscript and approving it for publication was Huan Zhou.

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

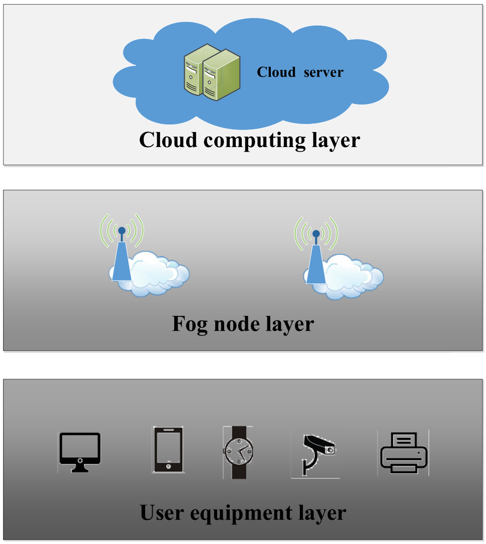


FIGURE 1. Fog computing architecture.

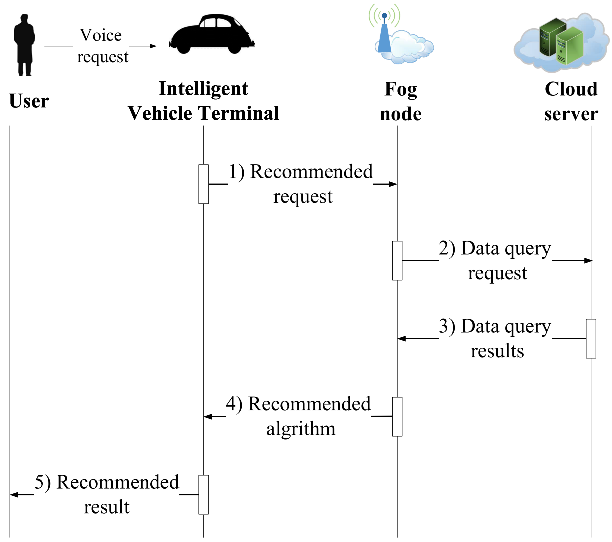


FIGURE 3. Dataflow of recommendation by the car smart navigator.

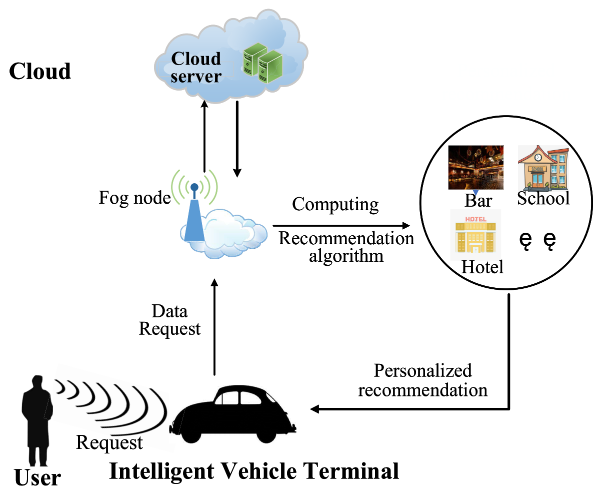


FIGURE 2. Workflow recommended by the car smart navigator.

recommendation of the vehicle intelligent navigator is shown in Figure 3.

The development of the recommended system has so far had a history of more than 20 years. Due to its large application requirements, the recommendation system has received extensive attention. As a filtering mechanism, the recommendation system is an important means to solve information overload [13], [14]. The recommendation algorithm is one of the most important parts of the recommendation system. There are three types of traditional recommendation algorithms which are: the recommendation algorithm using content information [15], the recommendation algorithm using collaborative filtering [16], and the recommendation algorithm using both the above two aspects [17].

Therefore, this paper focuses on the content recommendation algorithm of the intelligent car navigation system in the fog computing environment to provide users with low-latency, high-accuracy content of interest to users. This paper's contributions are summarized as follows:

1) A weighted interest degree recommendation algorithm for vehicle network intelligent navigator is proposed.

2) The accuracy of the algorithm is verified by numerical simulation.

The rest of this paper is organized as follows. In Section 2, we will examine the work related to the study of recommended algorithms in the IoT. In Section 3, we will introduce some preliminary knowledge. In Section 4, we introduce a weighted interest recommendation algorithm based on association rules. Section 5 performs the simulation of the verification algorithm. The paper concludes in Section 6.

II. RELATED WORK

The fog computing environment empowers the Internet of Things, enabling many application services to be realized [18]. Mung et. al. in [19] summarized all kinds of merits and open vast vista about fog computing, mentioning the application of fog computing to the IoT, 5G and embedded artificial intelligence. Fog computing offers unlimited possibilities for the implementation of the Internet of Vehicles. Hou et al. in [20] proposed the concept of vehicular fog computing (VFC), it was a framework that VFC realize the computation and communication between all other devices by the cooperation of various terminal equipment or some devices that close to user which made better use of personal communication and computing resource of every car. Based on a publish/subscribe mode, Shin et. al. in [21] proposed a fog computing architecture in IoV. Authors described a traffic congestion control scenario using a smart traffic light system which operated on top of the proposed architecture. Truong et. al. in [22] proposed a brand new vehicular ad-hoc network (VANET), and this architecture used a prospective settle scheme which asked the combination of software defined network (SDN) and fog computing, both were emergent computing and network example. The architecture was decided by SDN contributes of dexterity, expansibility, programmability and global knowledge, at the same time fog computing could meet the requests of prospective VANETs scene [23].

The emerging IoT is used to connect physical devices with the Internet which can record the behaviors of the physical devices. To make the interactions efficiently between users and devices, the recommendation methods to be used. Yao et. al. in [24] proposed a unified probability factor-based framework for recommending new properties of devices' interest in the IoT. By integrating the relationships between different IoT entities, a hypergraph was developed to simulate things. The spatio-temporal association, based on this, could generate implicit object associations, establish an Internet of Things experimental platform, and verify the possibility and efficacy of the method. A large amount of IoT data from polyphase data initiators were converged and transmitted to online or offline systems, which resulted in increased intricacy of data storage and query, which was especially suitable for spatio-temporal data handling via online or offline systems. Cai et. al. in [25] designed a multi-level IoT database architecture. By combining spatial data with text, images and video, an IoT database model with multi-layer was gotten, which was composed by different kinds of nodes. They collaborated to promote data storage, indexing, and querying. At the same time, based on the proposed model, a search algorithm using pruning strategy was designed, which effectively reduced the complexity of the algorithm.

Cloud computing acts a prominent part in realizing the actual application of Industrial Internet of Things (IIoT) [26]. Luo et. al. were committed to optimizing the web or cloud-based services performance, based on kernel least mean square algorithm (KLMS), they proposed a data-driven IIoT QoS value prediction scheme in [27]. The scheme used correlation coefficients to search correlative QoS values for every known QoS entry based on homoplastic service customers and web service entries. The KLMS was used to analyze two types of QoS data, and calculated the deep relationship according to the highest similarity of these data. The final coefficients were used to predict the of the missing web service QoS value. Compared to the traditional method, the proposed prediction method had higher prediction accuracy in simulation.

Today, a lot of IoT services are available in the market, more and more IoT service provider are coming out. Recommending or pushing customized IoT services to users based on user-owned objects is critical to the success of the IoT [28], [29]. Mashal et. al. in [30] proposed a hypergraph model based IoT network. In this model, each superedge connected user, object and service, based on existing known indicators, analyzed and discussed different algorithms in IoT. Correlation between performances on service recommendations. The conclusion showed that using graph-based recommendation algorithm could make IoT recommendation system working with more efficiency. Based on the ranking-based extreme learning machine, Wang et. al. came up with a taxi driver recommended road model in [31]. The basic road network was divided into several clusters, and the characteristics of each cluster was processed according to the real data set. An order-based extreme learning machine model was introduced to assess users' discovery potential of each cluster. When the vehicle trajectory data was unavailable or incomplete, the taxi driver could obtain road clustering suggestions by using the model combined with the training cluster selection algorithm, and the actual recommendation effect of the model was good. KoHG et. al. in [32] introduced a multi-criteria matrix location and integration collaborative filtering method based on the advantages of multi-criteria grading in which users could use to dynamically identify possible tasks and use the IoT. User complexity was solved when various intelligent objects in the environment performed the required tasks.

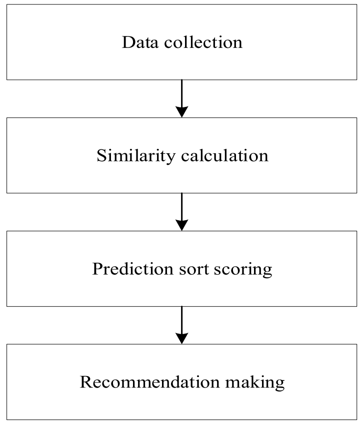


FIGURE 4. Recommended algorithm execution flow.

III. PREREQUISITE KNOWLEDGE

The core idea of collaborative filtering algorithm (CFA) is to find neighbors based on similarity, and then according to the prediction scores and recommendations, the execution process is shown in the Figure 4. First, the original data is collected. Second, the similarity is calculated. Third, the prediction is scored, and lastly recommendation is made. In the these steps, the key point is the user similarity calculation [33], [34].

A. SIMILARITY CALCULATION

The purpose of the CFA is to score the unrated items based on the similarity calculation. The computation process of the similarity among users becomes one of the keys to CFAs. Commonly used similarity measures include Euclidean distance, Pearson correlation coefficient, cosine similarity, modified cosine similarity, and so on.

The similarity is calculated in two ways. One uses the items' similarity, and the other one uses the user's similarity information. The choice between the two usually related to the users scale or items. The following uses the user-based collaborative filtering as an example to introduce the following similarity calculation methods.

B. EUCLIDEAN DISTANCE

Constructing a rating matrix based on user ratings, setting the row vectors to represent different users' ratings, and the column vectors to different users' ratings of the same item. The similarity is represented by equation (1):

$$sim = \frac{1}{1 + d_{AB}} \tag{1}$$

where *sim* indicates the degree of similarity, which ranges from 0 to 1, and d_{AB} represents the Euclidean distance of the score of the two items, and

$$d_{AB} = \sqrt{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 + \cdots + (\alpha_n - \beta_n)^2} \tag{2}$$

where $\alpha_i, \beta_i, (i \in [1, n])$ are the parameter values of the i'th dimension.

C. PEARSON CORRELATION COEFFICIENT

The Pearson correlation coefficient [35] can be used to measure the similarity between two vectors. Obviously the advantage of this method over Euclidean distance is that it is insensitive to user ratings, such as user a scoring 5 points for all items. While user b scores 1 for all items, the Pearson correlation coefficient considers the two vectors to be equal. Equation (3) shows the Pearson correlation coefficient similarity:

$$sim(A, B) = \frac{\sum_{i \in I_{AB}} \left(r_{A,i} - \overline{r_A}\right)\left(r_{B,i} - \overline{r_B}\right)}{\sqrt{\sum_{i \in I_{AB}} \left(r_{A,i} - \overline{r_A}\right)^2} \sqrt{\sum_{i \in I_{AB}} \left(r_{B,i} - \overline{r_B}\right)^2}} \tag{3}$$

where I_{AB} denotes a items' set that all users A and B jointly Scored, $r_{A,i}$,$r_{B,i}$ respectively denotes the users A and B's ratings for item i, and $\overline{r_A}$,$\overline{r_B}$ respectively denotes the average rating of user A and B.

D. COSINE SIMILARITY

Build a score matrix based on user ratings, set the row vector to represent different users' ratings, and the column vector indicate different users' ratings for the same item. The degree of similarity is measured by calculating the cosine of the two vectors' angle. If the vectors have the same direction, the similarity is 1.0; if the angle is 90 degrees, the similarity is 0. Equation (4) is shown at follows:

$$\cos\theta = \frac{A \cdot B}{\|A\| \, \|B\|} \tag{4}$$

where A and B represent two scoring vectors of two items and $\|A\|$ and $\|B\|$ represent the 2 norms of the vector.

E. MODIFIED COSINE SIMILARITY

The shortcoming of the cosine similarity measurement method is that it ignores the different users' rating criteria. So researchers propose a modified cosine similarity calculation method [36]. The calculation formula is shown in equation (5).

$$sim(A, B) = \frac{\sum_{i \in I_{AB}} \left(r_{A,i} - \overline{r_A}\right)\left(r_{B,i} - \overline{r_B}\right)}{\sqrt{\sum_{i \in I_A} \left(r_{A,i} - \overline{r_A}\right)^2} \sqrt{\sum_{i \in I_B} \left(r_{B,i} - \overline{r_B}\right)^2}} \tag{5}$$

where I_{AB} denotes a set of items that all users A and B jointly scored, I_A and I_B respectively denotes the collection of individual scoring items, $r_{A,i}$ and $r_{B,i}$ respectively denotes the rating of the user A and B for the project i, $\overline{r_A}$,$\overline{r_B}$ respectively denotes the average score of user A and B.

F. SKYLINE QUERY

Skyline queries extract data object sets from a data set that are not dominated by any other data object. This query method is essentially a multi-objective decision-making process. According to the rules of Skyline queries, we introduce the following two definitions.

Definition 1 (Domination): p is a D-dimensional data object generated by a fog node. The set of data objects is denoted as $S = \{p_1, \cdots, p_k\}$. For any two objects $p_i, p_j \in S$,p_j is dominated by p_i, i.e., $p_i \succ p_j$, if 1) any attribute value in any dimension of p_i is greater than or equal to p_j, and 2) there is at least one attribute value of p_i is greater than p_j. If p_i does not dominate p_j, it is denoted as $p_i \not\succ p_j$.

Definition 2 (Skyline set): Given a set S, the set of all the objects that are not dominated by any other objects in S is called the Skyline set. We denote it as $SKY(S) = \left\{p_i \,\middle|\, (\forall p_j \in S)(p_j \not\succ p_i)\right\}$.

In order to introduce our recommendation algorithm more clearly. An example will be introduced to explain the principle of Skyline query. We can use two-dimensional data in terms of the user rating and the distance in the vehicle intelligent navigator. The data distribution of each fog node is illustrated in Figure 5. It would be biased if the recommended results are determined using only a single dimension. Therefore, we need to find out the data objects that are not dominated by others through the Skyline query. In this example, $SKY(F) = \{f_2, f_3, f_8, f_9\}$.

IV. PROCESS OF THE CONTENT RECOMMENDATION

A. CONSTRUCTION OF USER INTEREST MODEL

A user interest model is a model representation of user information requirements and is very important in service recommendation system. The user model can usually show the user's interest in some specific topic information, which provides a basis for service providers to provide customers with more convenient services. The user model can usually be established in two ways. One is to directly obtain the user's interest and information demand tendency through the method provided by the user; the other method is to track and analyze the user's search, viewing, and other behavior records through the system to build a user profile.

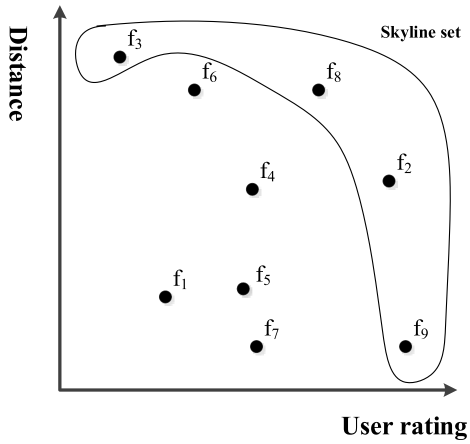


FIGURE 5. Skyline query in recommend of vehicle intelligent navigator.

In order to provide users with better services, this paper adopts a combination of active and passive acquisition to establish a user interest model. First of all, when actively acquiring user interests, the user can select the tags he is interested in from the list and perform the rating (the rating ranges from 1 to 5 stars) to express his/her own preference for the tag items, that is, the weight of interest. When the label in the table cannot meet the user's needs, the user can add keywords that he is interested in by manually inputting the information, and the user can also change the interest label and the score The user scoring matrix is constructed by this method of actively acquiring interest tags and interest scores.

B. ASSOCIATION RULE MINING ALGORITHM

The function of association rules is to get meaningful internal relationships in large-scale data sets. These relationships can be expressed in two types: the frequent item-sets (FISs) and the association rules. FIS is the collection of items which often appear together, and the second type suggests that there may be strong associations between the two items. The most famous case in the association analysis is the "beer diapers case" in the supermarket "shopping basket" data. Through the analysis of "shopping basket" data, customers' consumption habits can be known [37], [38].

Assuming that $I = \{i_1, i_2, \ldots, i_k\}$ is a set consisting of k different data items, where the elements are called items and the set of items is called item-sets. The transaction database is $D = \{T_1, T_2, \ldots, T_v\}$, and each transaction is a subset of item-sets, then $|D|$ represents total number of transactions D.

According to [37], [38], the implicative formula of association rule is shown in equation (6).

$$R : X \Rightarrow Y \tag{6}$$

where $X \subset I, Y \subset I$, and $X \cap Y \neq \varnothing$, item-set X appears in a transaction, which leads to Y also appears in a transaction with a certain probability. Analysis of association rules for user interest tags can be measured by two criteria: support and credibility

Support represents the appearing probability of the item-set $\{X, Y\}$ in the total item-set [37], [38]. In other words, it is the percentage of the scale of items in the item set to the total scale of item-sets. The formula is shown in equation (7).

$$Support(X \Rightarrow Y) = P(XY) = \frac{\mathrm{sum}(\mathrm{X} \cup \mathrm{Y})}{|D|} \tag{7}$$

Confidence represents the probability of the item set containing Y in the item-set containing X [37], [38]. It can be written as in equation (8).

$$Confidence(X \Rightarrow Y) = P(X|Y) = \frac{\mathrm{sum}(\mathrm{X} \cup \mathrm{Y})}{\mathrm{sum}(\mathrm{X})} \tag{8}$$

C. MINING ALGORITHM OF FIS

In the specific application of association rule mining, the efficiency of the algorithm is undoubtedly very important, which is also the research focus of current data mining. During the data mining stage, the FIS generation process involves a huge amount of calculation, which is the most difficult and complexity part. Currently, there are three FISs mining algorithms commonly used: Apriori algorithm, Eclat algorithm, and FP-Growth algorithm. The following is a brief description of the Apriori algorithm as an example [39], [40].

Apriori principle: If a set of items is frequent, then all the subsets of it are also frequent. Using this principle, we can infer that assuming a set of items being an in-FIS, all its supersets are infrequent too, and through this principle, we can bring down the computation delay and increase the algorithm efficiency.

In the process of Apriori algorithm execution, firstly, it gets all the frequent 1-item sets by scanning all the databases. Then, it finds all the candidate 2-item sets using the Apriori_Gen algorithm, and then counts each item set to find all the frequent 2-item sets. This process is repeated. By analogy, frequent k-item sets are always found in an iterative manner.

D. WEIGHTED INTEREST MODEL BASED ON ASSOCIATION RULES

In the first part of this section, we obtain interest information by user-actively setting tags and scoring, and establish a user interest scoring matrix. However, in the process of actively setting the label, due to the inaccurate expression of the interest tag by the user, the simple and straightforward way cannot construct an accurate user model; therefore, based on this, we introduce an association rule algorithm to deeply study the relationship of user interest. Here we assume $I = \{I_1, I_2, \ldots, I_d\}$ is a collection of interest tags made up of d different tag items. And we assume $D = \{T_1, T_2, \ldots, T_n\}$ being a character database, in which each character T is a subset of I.

Here, we add various tags to each transaction, such as: sports. We set the following tags: basketball, football, tennis, table tennis, etc., through which we can better get more accurate the interest of users to build user interest model $\{(t_1, w_1), (t_2, w_2), \cdots (t_n, w_n)\}$, $t_1, t_2, \cdots, t_n$

represent different interest tags and $w_1, w_2, \cdots, w_n$ represents the weight of the corresponding interest label.

According to the user rating matrix, we can find the n tag items who has the biggest similarity to user A and generate the nearest neighbor set of users $B_A = \{b_1, b_2, \cdots, b_n\}$. According to equation (9), we can predict the user's prediction score for the interest label item.

$$p_{A,i} = \overline{R_A} + \frac{\sum_{B \in B_A} \left(R_{B,i} - \overline{R_B}\right) sim(A, B)}{\sum_{B \in N_A} |sim(A, B)|} \tag{9}$$

In the equation, $sim(A, B)$ represents the similarity between A and B, and $R_{B,i}$ denotes a score of users' interest tag i in its nearest neighbor set $B_A = \{b_1, b_2, \cdots, b_n\}$, $\overline{R_A}$ and $\overline{R_B}$ represent the average scores of users a and b their respective interests scores.

Based on the predicted value of interest labels and the users' interest model $\{(t_1, w_1), (t_2, w_2), \cdots (t_n, w_n)\}$obtained from association analysis. We can get a more accurate prediction score through equation (10).

$$f_{A,i} = w_i \cdot p_{A,i} = w_i \left(\overline{R_A} + \frac{\sum_{B \in B_A} \left(R_{B,i} - \overline{R_B}\right) sim(A, B)}{\sum_{B \in N_A} |sim(A, B)|} \right) \tag{10}$$

As can be seen from Figure 6, the algorithm execution flow of this paper contains a total of seven parts, namely user data information acquisition, user score matrix construction, search for neighbor user sets, association analysis to calculate neighbor interest scores, score prediction, and recommendations. And the user confidence feedback seven parts. The user feedback information helps the interest weight adjustment, so that the recommendation result is more in line with the user's needs.

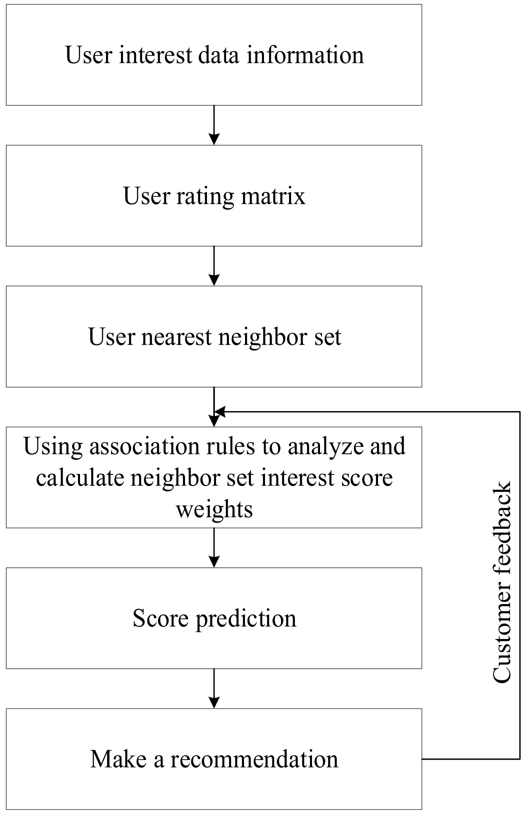


FIGURE 6. **Algorithm execution flow.**

E. ED-MAX FILTERING STRATEGY

In order to make the recommendation results more accurate, it is necessary to filter the final recommendation results. In this section, we implement this process with an Ed-Max filtering strategy (EMFS) based on Skyline query. Ed-Max algorithm was first proposed in [41] and the authors proved that the data filtered by this algorithm will not affect the global Skyline set. In addition, the filtering effect of this algorithm is better than that of MinMax algorithm. In this paper, Ed-max algorithm will be applied to filter the recommendation results again, in order to output more accurate and more user-friendly recommendation results.

Assuming that the data set in the recommendation result is $D = \{D_1, D_2, \cdots D_n\}$, and the data dimension is M. First, we need to store the maximum value of each dimension in the data set in memory.

The set of $D_{\max} = \{D_{1-\max}, \cdots, D_{M-\max}\}$ represents the tuple set of the biggest values between the first dimension and the Mth-dimension. $D_{\max}$ can be calculated directly by traversing the database once.

Next, we need to calculate the data object with the maximum Euclidean distance in the data space.$(D_n)_i$ is used to represent the value in ith dimension of the tuple D_n. Based on the definition of Euclidean distance, we can calculate the $Ed(D_n)_i$ as follows:

$$Ed(D_n)_i = \sqrt{\sum_{i=1}^{M} [(D_n)_i]^2} \tag{11}$$

There is n $Ed(D_n)_i$ after calculating which stores in memory. Then the maximum value of $Ed(D_n)_i$ can be calculated. In the data set $(D_n)_i$, the max $Ed(D_n)_i$ denotes the maximum Euclidean distance value (except for the tuple of $D_{\max}$), and its tuple is written using $D_{\max Ed}$.

If the data space is an M-dimensional coordinate one and every tuple is mapped to a space point, the EMFS focuses on an area encircled by three parameters which are $D_{\max}$, $D_{\max Ed}$, and the coordinates. A two-dimensional information is used to illustrate the EMFS procedures in Figure 7.

Figure 7 shows the filtration effect of EMFS. In Figure 7, there are 11 tuples in total. After EMFS processing, there are only five tuples left. A large part of the redundant data is reduced.

In the following, we introduce the content recommendation algorithm pseudocode in Algorithm 1. In this algorithm, the step 3 is similar to the fog node filtering strategy (FNFS)

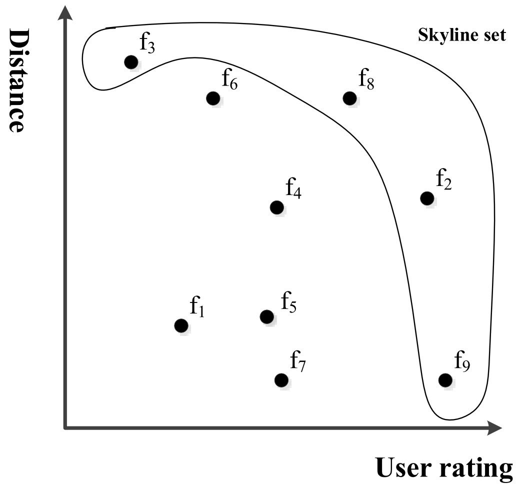


FIGURE 7. Ed-Max filtering strategy.

TABLE 1. Tag data conversion table.

Basketball	Tennis	Football	Pingpong	…	Action	War
1	1	0	1		1	1

proposed in [41], for both of them need to find the optimal Euclidean distance.

V. SIMULATION AND ANALYSIS

Since this article uses the combination of active and passive methods to obtain the user's interest score, but most of the current rating information in the Internet is the score after shopping and watching. It is not applicable to the experimental test of this article. Therefore, this article obtains some students' interest ratings for various news through questionnaire survey. The content of the questionnaire includes military, finance, entertainment, technology, digital, history, sports, movies, etc, with a score range of 1-5, and detailed classification of each content. Such as sports are divided into: basketball, tennis, football, table tennis and so on. In the refinement label of various news items in the questionnaire, the setting of the check mark selected by the experimenter is 1, and the unmarked value is 0. The conversion effect is shown in Table 1.

A. ANALYSIS OF RESULTS

From Figure 8 we can see that boys who like to watch action movies usually prefer to focus on basketball, and girls who like action movies usually prefer badminton.

In this experiment, the accuracy of the proposed algorithm is compared with Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) [42], [43]. MAE is a commonly used measurement method for measuring the accuracy and comparison of statistics and can accurately reflect the quality of recommendation. It can be used to measure the predicted user rating and the actual user rating. The value of MAE

Algorithm 1 The Content Recommendation Algorithm

Step1: Data collection, building user scoring matrix A
Calculate user similarity:
def User_Similarity_Calculating (TR):
  S = dict()
    for s in TR.keys():
      for i in TR.keys():
        ifs == i:
          continue
        S [s][i] = len(TR[u] & TR[v])
        S [s][i] = /= math.sqrt(len(TR[u]) ∗ len(TR[v]) ∗ 1.0)
      end for
    end for
return S
Step2: Build a nearest neighbor set,
Use the mining algorithm of FIS to get frequent item attribute sets. Further, find correlation between interests
C_1 = get_frequent_1-itemsets(P);
for (k=2;$C_{k-1} \neq \varnothing$;k++)
      M_k = appriori_gen(L_{k-1},Min_Supp);
        for each transaction t∈P
          M_t = subset(M_k,t);
      end for
end for
for each candidate m∈M_t
    m.count++;
end for
L_k ={m∈M_k|c.count≥Min_Supp};
return L=$S_k L_k$;
function appriori_gen(C_{k-1},Min_Supp)
for each itemset l_1 ∈C_{k-1}
      for each itemset l_2 ∈L_{k-1}
        if((c_1[1]∈c_2[1])^(c_1[2]∈c_2[2])^…(c_1[k-1]∈c_2[k-2])=(c_1[k-1]∈c_2[k-1]))
          m=c_1&c_2
        else
          add m to M_k;
        end if
      return;
    end for
end for
for each (k-1)-subset of m
        if s not belong to $C_{(k-1)}$
          return the result is TRUE
        else
          return the result is FALSE
end for
Step 3: Optimal Euclidean distance is processed
Ed-Max = []
for projects in the list
    DIS = Eucl(projects)
    Ed.append(DIS)
end for

Algorithm 1 *(Countinued.)* The Content Recommendation Algorithm

```
M = max (Ed).position
max_list.append(list[M])
#Filtering
for project in max_list
    for project2 in list
        if project2 <= project
            list.pop(project2)
        end if
    end for
end for
```

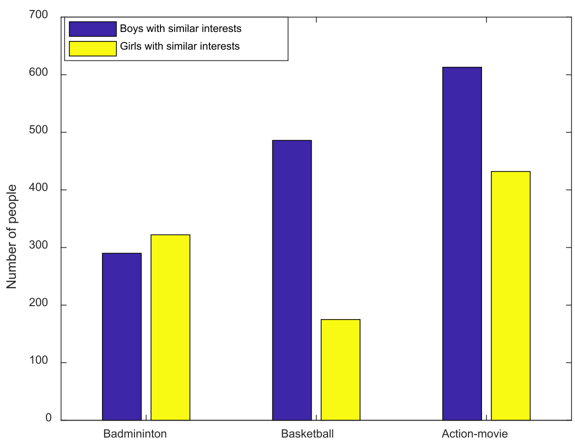


FIGURE 8. Interest association diagrams.

and final accuracy is opposite. If the MAE value decreases, the recommendation accuracy increases. Conversely, accuracy of the recommended algorithm gets worse. The calculation method is shown in equation (12).

$$MAE = \frac{\sum_{i=1}^{n} \left| R_{A,i} - f_{A,i} \right|}{n} \tag{12}$$

In the formula, $R_{A,i}$ denotes the actual user A's score for interest tag i, and $f_{A,i}$ is the predicted user A's score for the same interest tag calculated by the proposed algorithm.

RMSE was proposed to measure the deviation with the observed and real value. It has the advantage of better reflecting the real situation. The RMSE value and the accuracy of the recommendation also is opposite. If the RMSE value decreases, the recommendation accuracy increases. Conversely, the accuracy of the recommended performance gets worse. The calculation method is shown in equation (13).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(R_{A,i} - f_{A,i}\right)^2}{n}} \tag{13}$$

In formula (13), $R_{A,i}$ also denotes the actual user A's score for interest tag i, and $f_{A,i}$ is the predicted user A's score for the same interest tag calculated by the proposed algorithm.

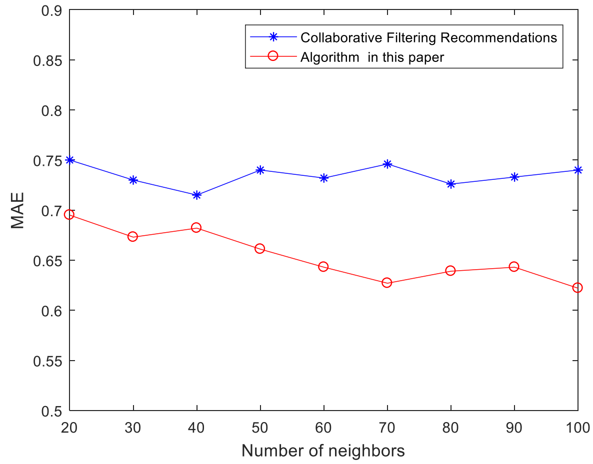


FIGURE 9. MAE value comparison between the proposed algorithm and collaborative filtering recommendations.

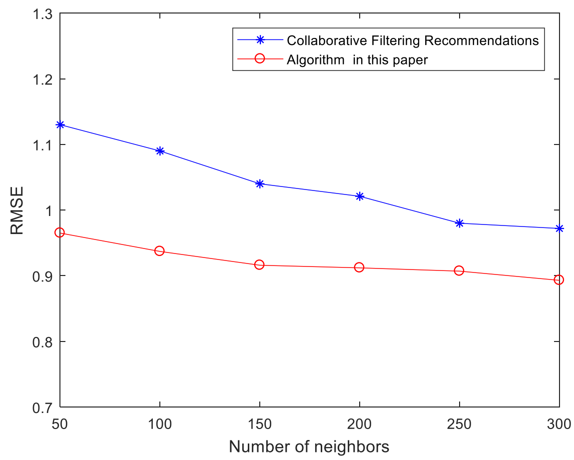


FIGURE 10. RMSE value comparison between the proposed algorithm and collaborative filtering recommendations.

The MAE value comparison between the presented algorithm and the conventional recommendation algorithm is shown in Figure 9. The vertical axis denotes the MAE value, and the horizontal axis indicates the number of users of the neighboring matrix used when predicting the score.

From Figure 9, we can see that when the neighbor matrix users increases, the MAE value of the proposed scheme is constantly decreasing slowly, and the overall MAE value of the proposed scheme is smaller than the conventional recommendation algorithm. This means that the proposed algorithm is outperformance compared with the conventional recommendation algorithm.

The following conclusion can be drawn from Figure 10. When the neighbor users increases, the RMSE value decreases. However, comparing with the conventional collaborative filtering recommendation, the proposed algorithm has lower RMSE value and better recommendation effect.

In this experiment, the prediction experiment is carried out by randomly selecting unrated items of 100-500 users, and the numerical simulation results are given in Figure 11. It states

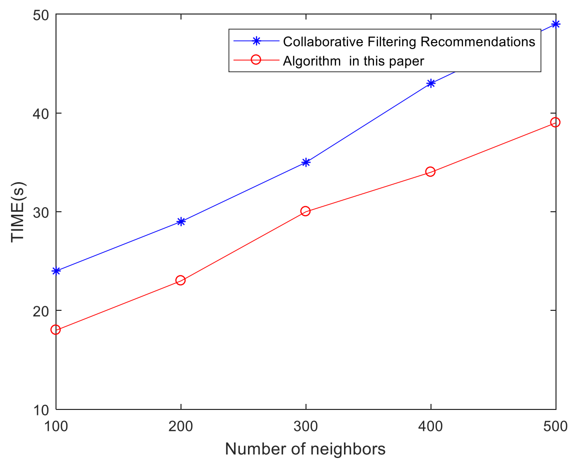


FIGURE 11. **Algorithm execution time varies with user number.**

that when the number of users increases, the execution delay of the algorithm also increase. However, compared with the conventional CFA, the proposed algorithm takes less time, that is, the proposed algorithm is time efficient.

VI. SUMMARY AND OUTLOOK

Fog computing enables the IoT application scenario such as intelligent vehicle networking. Intelligent navigator is the core component of Human-Computer Interaction in Vehicle Networking. Accurate content recommendation is needed in the use of navigators. This paper proposes a weighted interest user model based on association rules. It uses the association rule to mine the user's interest label to calculate its weight value, and uses the scoring matrix to accurately calculate the user's nearest neighbor set. Through the above methods, we get more accurate interest score and improve the recommendation accuracy. The numerical simulations indicate that the performance of the proposed algorithm is excellent.

ACKNOWLEDGMENT

(Fuhong Lin and Yutong Zhou are co-first authors.)

The authors would like to express our sincere thanks to the referees and the editors for their valuable suggestions.

REFERENCES

[1] Z. Ai, Y. Zhou, and F. Song, "A smart collaborative routing protocol for reliable data diffusion in IoT scenarios," *Sensors*, vol. 18, no. 6, p. 1926, 2018.
[2] F. Song and E. al, "Smart collaborative caching for information-centric IoT in fog computing," *Sensors*, vol. 17, no. 11, p. E2512, Nov. 2017.
[3] F. Lin, Y. Zhou, G. Pau, and M. Collotta, "Optimization-oriented resource allocation management for vehicular fog computing," *IEEE Access*, vol. 6, pp. 69294–69303, 2018.
[4] X. An *et al.*, "Hypergraph clustering model-based association analysis of DDOS attacks in fog computing intrusion detection system," *J. Wireless Commun. Netw.*, to be published.
[5] F. Lin, Y. Zhou, X. An, I. You, and K. Choo, "Fair resource allocation in an intrusion-detection system for edge computing: Ensuring the security of Internet of Things devices," *IEEE Consum. Electron. Mag.*, vol. 7, no. 6, pp. 45–50, Nov. 2018.
[6] J. Pan *et al.*, "α-fraction first strategy for hirarchical wireless sensor neteorks," *J. Internet Technol*, vol. 19, no. 6, pp. 1717–1726, 2018.
[7] Y. Zhang *et al.*, "Constructing chaotic systems from a class of switching systems," *Int. J. Bifurcation Chaos*, vol. 28, no. 2, pp. 18500321–1850032-9, 2018.
[8] F. Song *et al.*, "Smart collaborative connection management for identifier-based network," *IEEE Access*, vol. 5, pp. 7936–7949, 2017.
[9] F. Song *et al.*, "Modeling space-terrestrial integrated networks with smart collaborative theory," *IEEE Network*, vol. 33, no. 1, pp. 51–57, Feb. 2019.
[10] X. An *et al.*, "A novel differential game model-based intrusion response strategy in fog computing," *Secur. Commun. Netw.*, vol. 2018, pp. 1–10, Aug. 2018.
[11] Z. Ai, Y. Liu, F. Song, and H. Zhang, "A smart collaborative charging algorithm for mobile power distribution in 5G networks," *IEEE Access*, vol. 6, pp. 28668–28679, 2018.
[12] H. Zhou, H. Wang, X. Chen, X. Li, and S. Xu, "Data offloading techniques through vehicular ad hoc networks: A survey," *IEEE Access*, vol. 6, pp. 65250–65259, 2018.
[13] C. Huang, Y. Chen, S. Xu, and H. Zhou, "The vehicular social network (VSN)-based sharing of downloaded geo data using the credit-based clustering scheme," *IEEE Access*, vol. 6, pp. 58254–58271, 2018.
[14] M. Ruan, X. Chen, and H. Zhou, "Centrality prediction based on K-order Markov chain in mobile social networks," *Peer-To-Peer Netw. Appl.*, vol. 16, pp. 1–11, Mar. 2019.
[15] H. Zhang, Y. Qi, H. Zhou, J. Zhang, and J. Sun, "Testing and defending methods against DoS attack in state estimation," *Asian J. Control*, vol. 19, no. 3, pp. 1–11, 2017.
[16] Y. Cai, H.-F. Leung, Q. Li, H. Min, J. Tang, and J. Li, "Typicality-based collaborative filtering recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 3, pp. 766–779, Mar. 2014.
[17] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Model. User-Adapted Interact.*, vol. 12, no. 4, pp. 331–370, Nov. 2002. doi: 10.1023/A:1021240730564.
[18] X. An *et al.*, "Sample selected extreme learning machine based intrusion detection in fog computing and MEC," *Wirel. Commun. Mob. Comput.*, vol. 2018, Aug. 2018, Art. no. 7472095.
[19] Chiang, Mung, and Z. Tao, "Fog and IoT: An overview of research opportunities," *IEEE Trans. Softw. Eng.*, vol. 3, no. 6, pp. 854–864, Dec. 2017.
[20] Hou, Xueshi, "Vehicular fog computing: A viewpoint of vehicles as the infrastructures," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3860–3873, Sep. 2016.
[21] Shin, Sangjin, "A pub/sub-based fog computing architecture for Internet-of-vehicles," in *Proc. IEEE Int. Conf. Cloud Comput. Technol. Sci.*, Dec. 2017, pp. 1–10.
[22] N. B. Truong *et al.*, "Software defined networking-based vehicular adhoc network with fog computing," in *Proc. IEEE Int. Symp. Integr. Netw. Manage.*, Jul. 2015, pp. 25–36.
[23] F. Song, R. Li, and H. Zhou, "Feasibility and issues for establishing network-based carpooling scheme," *Pervasive Mobile Comput.*, vol. 24, pp. 4–15, Dec. 2015.
[24] L. Yao *et al.*, "Things of interest recommendation by leveraging heterogeneous relations in the internet of things," *ACM Trans. Internet. Technol.*, vol. 16, no. 2, p. 9, 2016.
[25] H. Cai *et al.*, "A multi-layer Internet of things database schema for online-to-offline systems," *Int. J. Distrib. Sens. Netw.*, vol. 12, no. 8, 2018, Art. no. 1550147716664248.
[26] F. Lin, X. Lü, I. You, and X. Zhou, "A novel utility based resource management scheme in vehicular social edge computing," *IEEE Access*, vol. 6, pp. 66673–66684, 2018.
[27] X. Luo *et al.*, "A large-scale web QoS prediction scheme for the Industrial Internet of Things based on a kernel machine learning algorithm," *Comput. Netw.*, vol. 101, pp. 81–89, May 2016.
[28] F. Song *et al.*, "Smart collaborative distribution for privacy enhancement in moving target defense," *Inf. Sci.*, to be published. doi: 10.1016/j.ins.2018.06.002.
[29] F. Song *et al.*, "An optimization-based scheme for efficient virtual machine placement," *Int. J. Parallel Program.*,vol. 42, no. 5, pp. 85–872, 2014.
[30] I. Mashal, O. Alsaryrah, and T. Y. Chung, "Analysis of recommendation algorithms for Internet of Things," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–6.
[31] R. Wang *et al.*, "TaxiRec:Recommending road clusters to taxi drivers using ranking-based extreme learning machines," in *Proc. Sigspatial Int. Conf. Adv. Geographic Inf. Syst.*, 2015, pp. 125–136.
[32] H. G. Ko, I. Y. Ko, and D. Lee, "Multi-criteria matrix localization and integration for personalized collaborative filtering in IoT environments," *Multimedia Tools Appl.*, vol. 77, no. 4, pp. 4697–4730, 2018.

[33] Y. Bergner *et al.*, "Model-based collaborative filtering analysis of student response data: Machine-learning item response theory," in *Int. Educ. Data Mining Soc.*, Aug. 2012, p. 8.

[34] M. Y. H. Al-Shamri and N. H. Al-Ashwal, "Fuzzy-weighted similarity measures for memory-based collaborative recommender systems," *J. Intell. Learn. Syst. Appl.*, vol. 6, no. 1, pp. 1–10, Aug. 2014.

[35] G. Pirlo and D. Impedovo, "Cosine similarity for analysis and verification of static signatures," *IET Biometrics*, vol. 2, no. 4, pp. 151–158, Aug. 2013.

[36] X. He and Y. Luo, "Mutual information based similarity measure for collaborative filtering," in *Proc. IEEE Int. Conf. Prog. Inform. Comput.*, Sep. 2011, pp. 1117–1121.

[37] M. Narvekar and S. F. Syed, "An optimized algorithm for association rule mining using FP tree," *Procedia Comput. Sci.*, vol. 45, pp. 101–110, Jun. 2015.

[38] A. Mangalampalli and V. Pudi, "Fuzzy association rule mining algorithm for fast and efficient performance on very large datasets," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Aug. 2009, pp. 1–8.

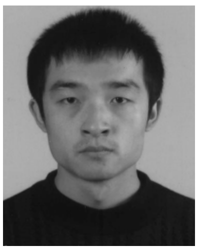
[39] K. Lin *et al.*, "A frequent itemset mining algorithm based on the Principle of Inclusion-Exclusion and transaction mapping," *Inf. Sci.*, vol. 5, pp. 278–289, Jul. 2014.

[40] F. Zhang *et al.*, "Accelerating frequent itemset mining on graphics processing units," *J. Supercomput.*, vol. 66., no. 1, pp. 94–117,2013.

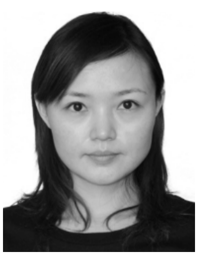
[41] X. An, F. Lin, L, Yang, "Node state monitoring scheme in fog radio access networks for intrusion detection," *IEEE Access*, vol. 7, pp. 21879–21888, 2019. doi: 10.1109/ACCESS.2019.2899017.

[42] X. Yu and M. Q. Li, "Effective hybrid collaborative filtering algorithm for alleviating data sparsity," *J. Comput. Appl.*, vol. 29, no. 6, pp. 1590–1593, Aug. 2014.

[43] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)," *Geosci. Model Develop.*, vol. 7, no. 3, pp. 1247–1250, 2014.



FUHONG LIN received the M.S. and Ph.D. degrees in electronics engineering from Beijing Jiaotong University, Beijing, China, in 2006 and 2010, respectively. He is currently an Associate Professor with the Department of Computer and Communication Engineering, University of Science and Technology Beijing, China. His research interests include edge/fog computing, network security, and big data. He received the Provincial and Ministry of Science and Technology Progress Award, in 2017. He received the track Best Paper Award from the IEEE/ACM ICCAD 2017. Two of his papers have listed in the "Top 100 of the most influential international academic papers in China" in 2015 and 2016, respectively. He has served as the Co-Chair of the first and the third IET International Conference on Cyberspace Technology, and the General Chair of the second IET International Conference on Cyberspace Technology. He was the leading Guest Editor of the Special Issue "Recent Advances in Cloud-Aware Mobile Fog Computing" for Wireless Communications and Mobile Computing. Currently, he also serves as a Reviewer for more than 20 international journals, including the IEEE Transactions on Industrial Informatics, the IEEE Access, *Information Sciences*, the IEEE Internet of Things Journal, *The Computer Journal*, and *China Communications*.



YUTONG ZHOU is currently an Associate Researcher with the Institute of Education and Economy Research, University of International Business and Economics. She is also a Columnist for multiple publications and has published over 30 research articles in journals and magazines. Her current research interests include the Internet-based education, the Internet-assisted economic development, and poverty relief networks.

ILSUN YOU (SM'13) received the M.S. and Ph.D. degrees in computer science from Dankook University, Seoul, South Korea, in 1997 and 2002, respectively, and the second Ph.D. degree from Kyushu University, Japan, in 2012. From 1997 to 2004, he was with THIN Multimedia Inc., Internet Security Co., Ltd., and Hanjo Engineering Co., Ltd., as a Research Engineer. He is currently an Associate Professor with the Department of Information Security Engineering, Soonchunhyang University. He is a Fellow of the IET. He has served or is currently serving as a Main Organizer of international conferences and workshops, such as MIST, MobiWorld, MobiSec, and so forth. He is the Editor-in-Chief of the *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications* (JoWUA). He is on the Editorial Board of *Information Sciences* (INS), *Journal of Network and Computer Applications* (JNCA), IEEE Access, *Intelligent Automation & Soft Computing* (AutoSoft), *International Journal of Ad Hoc and Ubiquitous Computing* (IJAHUC), *Computing and Informatics* (CAI), and *Journal of High Speed Networks* (JHSN). Especially, he has focused on 4/5G security, security for wireless networks and the mobile internet, the IoT security, and so forth, while publishing more than 180 papers in these areas.

JIUZHI LIN received the master's degree, in 2019. He is currently pursuing the degree in electronics and communication engineering with the University of Science and Technology Beijing. His research interests include edge computing and fog computing.

XINGSHUO AN was born in Shandong, China, in 1988. He received the master's degree from the University of Science and Technology Beijing, China, in 2014, where he is currently pursuing the Ph.D. degree with the School of Computer and Communication Engineering. His research interests include fog computing and network security.

XING LÜ received the Ph.D. degree in computer science and technology from the Beijing University of Posts and Telecommunications, Beijing, China, in 2012. He is currently a Professor with the Department of Communication Engineering, School of Computer and Communication Engineering, University of Science and Technology Beijing, China. His research interests include fog computing, optical soliton communication, and big data.