Computer Communications 169 (2021) 243–249

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

The prediction model for residence based on reliability in social network ☆

Bin Zhang *, Yanglan Fu, Jing Li, Zhou Ye

Blockchain Laboratory, Zhejiang University of Finance & Economics, Hangzhou, 310018, China

ARTICLE INFO

Keywords:
Online social
Predict residence
Untrustworthy executors
Trust model

ABSTRACT

Currently, the implementation judge often encounter such tough problems the person subject to enforcement usually change residence in secret to avoid execution. This problem has become a major obstacle for the courts in the process of law enforcement. Many research results indicated that the predictors based on the person's social network for simulate his behavior will be accurate. Consequently, this paper proposes a PLRU (predict Location Based on the Reliability of Social-online Users) model improving the execution efficiency of the court to the dishonest person subject to enforcement. We estimates and filters the trust degree of the relevant in social network, and then this paper presents the DDA (Density based late Differential Allocation) model to estimate the weight of the text in social conversation about the dimension of residence. Finally, the prediction range of the residence is obtained while combining with the person profile though the historical residence of the person subject to enforcement. The experiments reveal that the accuracy of the model is greatly improved compared with other models while it is well recognized by the executive judges of the court.

1. Introduction

At present, the difficulty of execution known to the whole society is a obstacle that exists in court system in China. And, the point is detecting the behavior trajectory of the person subject to enforcement. Therefore, this paper studies the efficient perception and collection method of intelligent multi-source heterogeneous social network data, probes the behavior characteristics of the person to enforcement and deduces the activity trajectory of the person subjected to execution based on artificial intelligence analysis technology, and provides an efficient model for the detection and discovery of the activity of the person subjected to execution by the court and other departments. So, this paper proposes a PLRU (Predict Location Based on the Reliability of Social-online Users) model of residential location estimation based on user social network reliability to predict of the hiding place of the court's untrustworthy executors and improve the efficiency. With the rapid development of mobile smart devices and wireless networks, the coverage of active users of social communication software is getting wider and wider. Currently, typical social software include WeChat, Instagram, Twitter, LINE, etc. [1]. In the process of social network, users will share or disclose geographic location information about their activities in their circle of Moments, which makes geographic location data gradually become a high-quality data mining target. Therefore, not only can information be disseminated through social networks, but also the reasons for various kinds of user behaviors can be analyzed [2]. Our social activities are formed in constant contact and interaction with others, so in many cases, human behavior can only be understood in the context of the situation. Therefore, based on the research topic of this project, this paper assumes that the trust value of the social network based on the social software is crucial to the construction of the activity trajectory of untrustworthy executors.

In this paper, the PLRU model based on user trust is mainly optimized in three aspects: user trust assignment, local word extraction and user profile of historical residence.

In summary, according to the traditional algorithm and the actual needs of this project in the course of court execution, we propose to design a PLRU model of residential location estimation based on user social network reliability to predict of the hiding place of the court's untrustworthy executors. By estimating the reliability of chat friends, select friends with high reliability, and then extract geographical keywords from their chat content to construct local words and emotional words. Finally, combining with the dimension score of the user profile of the historical residence, the residence prediction is used to deduce the activity trajectory of the subject. The model not only optimizes the real-time performance of the data, but also adds the emotional attitude to the geographical keywords, which effectively improves the execution rate and quality of the data processing. At the same time, combining with the residence user profile of untrustworthy executors, according to their own residential location characteristics, the geographical keywords are more targeted to score, and finally get the highest evaluation of the geographical location, that is, the predicted range of the residence.

☆ This work is supported by National Key Research and Development Program of China (2018YFC0830400).

* Corresponding author.
E-mail address: zhangbin@zufe.edu.cn (B. Zhang).

https://doi.org/10.1016/j.comcom.2021.01.030
Received 2 September 2020; Received in revised form 14 December 2020; Accepted 27 January 2021
Available online 2 February 2021

The major contributions of this work are three aspects:

A. A new model using social network trust to optimize the trajectory of the person being executed is proposed to accurately estimate the regional space of residence.

B. In the process of tracking the untrustworthy executor, the algorithm is more timely by using the social network information of them and focusing on digging geographical characteristics.

C. In combination with the social network user profile of the person subjected to execution, the predicted location information was optimized according to the residence preference of the untrustworthy executor to improve pertinence and accuracy.

2. Related work

2.1. Estimation of user trust in social networks

Firstly, the model needs to rank user reliability to obtain high-quality information and improve execution efficiency. Based on high-quality reliability, high-quality data related to geographical location of residence could be obtained from chat contents. This step will effectively narrow the scope for the next chat text extraction, improve the accuracy and speed up the operation efficiency. Therefore, A lot of previous research focus on the model of Flap also through social relations to calculate the degree of trust, and the co-addressing tendency among friends to calculate [3–6]. The all proposed a social tag recommendation algorithm based on friend trust and calculated the trust value between users by assigning different weights to different friends [7]. The score of trust is improved based on the implicit similarity of trust [7]. Based on sociological six degrees separation, the trusted users were calculated [8]. These efforts to explore trust intensity have well mined the implicit information in social relations, but they cannot deal with the sparse score and trust data well, thus greatly reducing the accuracy. In this paper, the deficiencies of these works will be optimized. In the case of sparse trust data, effective results can be obtained, and the trust strength between each group of users will be comprehensively considered.

2.2. Semantic feature extraction of social network

After selecting some trusted friends, extracting the semantic features of their relevant social records and spatial comments. Among the traditional methods of semantic feature extraction, there are LDA model, TF–IDF algorithm and keyword extraction algorithm based on co-occurrence words. A text retrieval method proposed based on kmeans clustering algorithm and LDA theme model and its validity verification, but it lacks the extraction of emotion words, which cannot reflect the attitude towards different geographical words [9–11]. The similar study improved the keyword extraction algorithm based on word co-occurrence, and proposed an algorithm superior to TF–IDF, but did not optimize the extraction of emotion words [12]. The theme of the SRC-LDA (semantic relation constrained-LDA) [13] used to implement the LDA fine-grained keywords extraction under the guidance of semantics, improve the granularity of key, and the extraction of the semantic relevance between emotional words, but does not address how to extract effectively the problem of geographic information in a text.

2.3. Dimensional analysis of residence based on behavior attributes

Finally, this paper uses the relevant geographical location information obtained from the social content as the training data of the user profile, divides the obtained geographic information words and emotional words into different dimensions according to the characteristics of the historical residence of the different faithless subjects, and finally filters out the effective prediction of the residence range. The user profile of residence is a process of extracting users' interest in different characteristics of residence based on their attributes and behaviors, and understanding users' preference for influencing factors of residence. The research analyzed the residential location from common factors and personality factors [14]. And similar research all generally divided the influencing factors into housing price, transportation, environment, employment accessibility, working place, family life cycle, residents' social attributes and income. According to the youth, the middle aged and the old, they have carried on the study of the factors of residence choice, and have obtained the important factors respectively [15–18]. Some researcher have studied the factors of residence choice of non-resident migrants and found that the housing price is the most important factor [19]. Therefore, the paper finally gets the highest rated location based on the evaluation of service facilities, transportation conditions, environmental conditions, housing price and sense of belonging.

3. The location prediction model based on reliability in social network

3.1. Problem settings

This paper is to obtain the prediction range of the place of residence combining the social content with friends. Therefore, there are mainly the following objectives: (i) How to accurately obtain reliable user trust weight through social network. (ii) How to improve the accuracy of geographic keyword and emotional word information obtained through social interaction. (iii) Accurate calculation of dimension weight of user profile of residence information.

In this section, This paper study the interactive information of the popular chat software of the studied object, and the reliability of chat friends is estimated. The chat content of friends with high reliability is selected to extract geographical keywords to construct local words and emotional words. Finally, combining the image of the users' previous residence of the untrustworthy person to screen and obtain the prediction range of the place of residence. The focus of this paper is to extract high quality emotional words and geographic words. The focus of this paper is to extract high quality emotional words and geographic words, but the premise is to improve the effectiveness of the collected data Chinese text extraction and the reliability of the extraction range. Therefore, in this paper, data pre-processing is carried out on the premise of user reliability calculation, which can effectively reduce the extraction range and improve the operation efficiency. Therefore, in this paper, data pre-processing is carried out on the premise of user reliability calculation, which can effectively reduce the extraction range and improve the operation efficiency.

3.2. Estimation of reliability based on social semantics

3.2.1. Estimation of direct friend algorithm

Firstly, this paper takes direct friends as an example to calculate the reliability. In order to establish the relationship between the users. We suppose that the trust relationship y is generated by a normal distribution of the trust vector, so the user Y can be written as

$$Y = \begin{bmatrix} y_{11}\ y_{12}\ y_{13}\ \cdots\ y_{1N} \\ y_{21}\ y_{22}\ y_{23}\ \cdots y_{2N} \\ \vdots \\ y_{N1}\ y_{N2}\ y_{N3}\ \cdots\ y_{NN} \end{bmatrix} \tag{1}$$

$$p(Y\,|\eta,\ \gamma, \sigma_F^2) = \frac{1}{\sigma_F\sqrt{2\pi}} e^{-\frac{(Y-g(\eta_i^T\gamma_j))^2}{2\sigma_F^2}} \tag{2}$$

Suppose that in the social network, each user with N friends, then the $c_{i,j}$ represents the friend relationship between user i and j, so we can obtain the user's trust matrix $Y = [Y_{i,j}]_{N\times N}$. $Y_{i,j}$ is the rating between user i and user j, $\eta \in T^{K\times M}$ and $\gamma \in T^{K\times M}$ represent the K-dimensional preference matrix for trusted users. And then the logistic regression $g(x) = \frac{1}{1+e^{-x}}$, Map $\eta^T\gamma$ to the range of [0, 1] to fit the value range of trust relationship Y. The formula indicates that the gaussian distribution with $g(\eta_i^T\gamma_j)$ as the mean and σ_F^2 as the variance is generated. Then You can get a logarithmic posterior probability distribution written as:

$$\begin{aligned} &Inp(Y\,|\eta,\gamma\;,\sigma_F^2) = \\ &-\frac{1}{2\sigma_F^2}\Big(\sum_{i=1}^{N}\sum_{j\in N}(Y_{i,j}-\eta^T\gamma)^2-\sum_{i=1}^{N}\eta_i^T\eta_i-\sum_{i=1}^{N}\gamma_j^T\gamma_j\Big) \\ &-\frac{1}{2}\sum_{i=1}^{N}\sum_{j\in N}In\sigma_F^2-(N\times K)In\sigma_F^2+C \end{aligned} \tag{3}$$

The maximized logarithmic posteriori probability distribution equivalent to the minimization loss function if the variance parameter is kept constant which can be written as

$$L_F(Y,\eta,\gamma)=\frac{1}{2}\Big(\lambda_F\sum_{i=1}^{N}\sum_{j=1}^{N}(Y_{i,j}-\eta^T\gamma)^2+\lambda_\eta\sum_{i=1}^{N}\eta_i^T\eta_i+\lambda_\gamma\sum_{i=1}^{N}\gamma_j^T\gamma_j\Big) \tag{4}$$

The trust vector η and the trusted vector γ are obtained by gradient descent method. However, the implied similarity degree of both user i and user j is $S_{i,j}^{\eta}=\eta_i^T\eta_j$, the constant trust value of the trusted is $S_{i,j}^{\gamma}=\gamma_i^T\gamma_j$. The score similarity $S_{i,j}^{\xi}$ between user i and user j is calculated from the score vectors of user to obtain $S_{i,j}^{\xi}=I_i^TI_j$. We assume that the trust strength matrix among users is S, which is generated by the normal distribution of the score similarity degree $S_{i,j}^{\xi}$, and the implicit similarity degree $S_{i,j}^{\eta}$ and $S_{i,j}^{\gamma}$, so S can be written as

$$p\Big(S\,\Big|S_{i,j}^{\xi},\;S_{i,j}^{\eta},S_{i,j}^{\gamma},\sigma_s^2\Big)=\frac{1}{\sigma_s\sqrt{2\pi}}e^{-\frac{(Y-g(S_{i,j}^{\xi}+S_{i,j}^{\eta}+S_{i,j}^{\gamma}))^2}{2\sigma_s^2}} \tag{5}$$

The σ_s^2 represents the variance of the gaussian distribution, which can also be understood as the noise condition of the estimated value. We get a logarithmic posterior probability distribution which can be written as

$$Inp\Big(S\,\Big|S_{i,j}^{\xi},\;S_{i,j}^{\eta},S_{i,j}^{\gamma},\sigma_s^2\Big)=In\frac{1}{\sigma_S\sqrt{2\pi}}-\frac{(Y-g(S_{i,j}^{\xi}+S_{i,j}^{\eta}+S_{i,j}^{\gamma}))^2}{2\sigma_S^2} \tag{6}$$

In order to maximize formula (6), we need to minimize the loss function which can be written as

$$L_S(S,I,\eta,\gamma)=\frac{\lambda_S}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{S}_{i,j}-g(I_i^TI_j+\eta_i^T\eta_j+\gamma_i^T\gamma_j))^2 \tag{7}$$

Obviously, we can obtain the minimum value of 0 when the $\hat{S}_{i,j}$ equals $g(I_i^TI_j+\eta_i^T\eta_j+\gamma_i^T\gamma_j)$. However, considering that the implicit similarity of score and trust relation is combined, the estimation of trust intensity has noise.

3.2.2. Estimation of indirect friend algorithm

In addition to the above algorithm for friends, the reliability of indirect friends can be estimated by the following formula, and the reliability generated by different paths can be added, but more than one. The user k is the direct friend of user i. And ϖ_i, ϖ_j and ϖ_k respectively represent the collection of friends in the social circle of user i, j and k. The party i originally did not know actor j, but the friend actor k knew actor j, so that party i would have contact with actor j in the hiding process. When the relationship between party i and actor j is called indirect connection, and the degree of trust should also be calculated. By summing up the indirect trust degree of its related actors, the trust degree is not higher than one. And $j\in(\omega_k-\varpi_j)$, through the formula, we finally get the trust ranking. So it can be written as

$$Y_{ij}=\min\left(\sum_{j\in(\varpi_k-\varpi_j)}\frac{|\varpi_i\bigcap\varpi_k|}{|\varpi_i\bigcup\varpi_k|}\cdot\frac{|\varpi_j\bigcap\varpi_k|}{|\varpi_j\bigcup\varpi_k|},1\right) \tag{8}$$

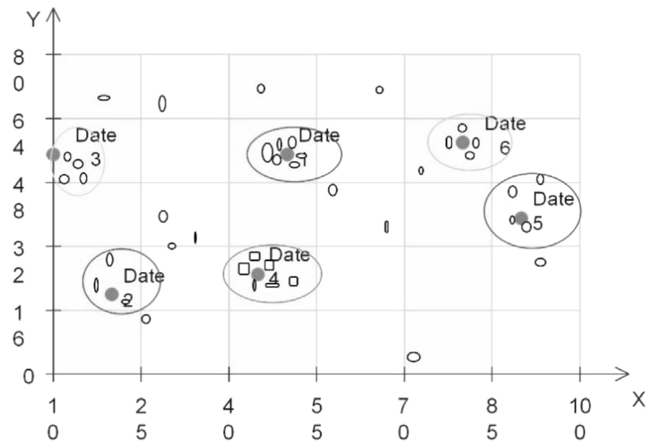

Fig. 1. The data clustering of geographical location.

3.3. Extraction of geographic feature words and emotion words based on semantics

The several friends with high reliability ranking were obtained through early processing, geographic keywords can be extracted from their social documents. In this paper, such words are called local word set R, where n represents the number of words $R=\{r_1,r_2,\ldots,r_n\}$. The most typical local words among them, are the places or landmark nouns that directly indicate the location, such as city, street, landmark building, etc. And the set of words that appear in the chat text about the emotional description of geographical location, where n represents the number of words $Z=\{z_1,z_2,\ldots,z_n\}$.

In this paper, a DDA model based on DBSCAN clustering algorithm and LDA subject model is used to extract the local word R from text retrieval and obtain the relevant emotion word Z to display the relevant attitude, so as to obtain the main geographical location of emotion preference. Firstly, relevant text data are obtained from the text data set. Then, by setting the neighborhood radius χ and density threshold DT, the points satisfying the domain radius are classified into a center group, and the points outside the density distance can be de-noised effectively. And they can be written as $\chi=\frac{2\sum_{i=1}^{n}dis_{ij}}{n(n-1)}$, $DT=\min(dis_{ij},\vartheta)$. (See Fig. 1.)

The dis_{ij} represents the straight-line distance obtained according to the map of two geographical locations and ϑ represents the maximum density threshold of the input, in order to prevent the clustering after the generation of a large size of the clustering group and increase the search range.

By calculating the term frequencies of each word in each social document and the document frequency used to calculate the particularity of word will result in different weight calculation results. In order to solve the influence of the length of the document on the weight calculation of words, as well as the missing words which are omitted in the document content, the formula can be written as

$$W_{ij}=\sqrt{\frac{tf_{ij}}{\max T}\times\log(\frac{N}{df_i}\times n_j)} \tag{9}$$

W_{ij} is the weight of the word f_j in the article after correction, tf_{ij} is the number of times the word appears in the chat content, N is the number of set chat text, and nj is the number of letters.max T is the word frequency of the total number of words with the maximum word frequency in the file set, which is used to solve the problem of higher word frequency.

After the main geographic keywords are obtained and the corresponding weights are calculated, the feature words and emotion words are matched to obtain the must-link semantics. Combining with the following formula, the semantic relationship strength of emotion words to feature words is calculated, which is ST. Where, η_i is the frequency threshold of local feature words, and θ_2 is the frequency threshold of local emotion words. $f_c(v_i, v_j)$ is the co-occurrence frequency of candidate feature word v_i and candidate emotion word v_j in sentences. The formula can be written as

$$ST(v_i, v_j) = \left| \frac{Inf_c(v_i, v_j)}{Inw_{ij}^2 - f_c(v_i, v_j)} \right| \quad (10)$$
$$f(v_i) < \theta_1, f(v_j) < \theta_2$$

The feature semantic matrix can be written as

$$\lambda = \left\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n\right\}, \ \mu = \left\{\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n\right\} \quad (11)$$

Among them, ε represents the emotional preference expressed by the untrustworthy executor to the geographical location mentioned before, such as convenience, remoteness, suitability, Satisfies or not. λ represents different evaluation dimensions of geographical location, such as location, environment, transportation, housing price, and consumption.

3.4. The weight model based on fuzzy comprehensive analysis

After the algorithm mentioned above, we will calculate several major geographical locations where the untrustworthy executor often appears and has emotional preference according to the selected trusted users, namely the friends and relatives in the execution case and the chat contents between them, which will become an important location in the experimental prediction of residence range. At the same time, we also need to find the multidimensional information of the untrustworthy executor, such as: historical residence data including id card address, property ownership card address, rental information address and Taobao receiving place, so as to effectively provide the residence user portrait for the executive judge.

In this paper, the dimension of historical user portrait of the residence of the person who has broken the promise is defined as U, and I is the dimension number, $U = \{u_1, u_2, \dots, u_l\}$. Many scholars have studied the factors affecting the choice of residential areas from different perspectives, such as age group, migrant population, etc. This paper adopts the fuzzy comprehensive evaluation, which can avoid the inherent subjectivity of target selection by experience and make the merger and acquisition decision more scientific and reasonable by taking advantage of the advantages of this method, and calculates the weight of five dimensions of U, including service facilities, traffic arrangement, environmental conditions, housing price and sense of belonging. Similarly, the obtained emotional word Z is divided into the above five dimensions according to the same classification criteria. According to the ratio of the word number tf_{Zn} of each dimension to the word number af of the total dimension, the word frequency of each dimension can be obtained respectively, the weight can be written as $W_{Zn} = \frac{tf_{Zn}}{af}$, and z_n is the classification of dimensions obtained from the text. δ stands for sensitivity, that is, the sensitivity to different dimensions shown in the chat content. The higher the sensitivity, the higher the reliability of W_{Zn} it represents. In combination with the dimension weight score W_{bn} obtained from its historical residence, relevant weight score can be obtained by using nearby or similar residence assessment report, and the final weight W_n can be obtained by weighting. They can be written as

$$\delta = \frac{1}{\sqrt{2\pi}} e^{\frac{-ST^2(v_i, v_j)}{2}}, \ W_n = \frac{\delta \times W_{zn} + (1-\delta) \times W_{bn}}{10} \quad (12)$$

According to the dimensionality scale we obtained, several key geographical locations were scored, so as to get the geographical location with the highest score, that is, the prediction point of the final residence.

Table 1
Reliability rating of respondent's friends.

Respondent's friends	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}
Trust implied similarity	0.91	0.89	0.89	0.87	0.84	0.79	0.71	0.65	0.64	0.55
Total reliability score	0.95	0.92	0.91	0.91	0.89	0.83	0.74	0.69	0.67	0.59

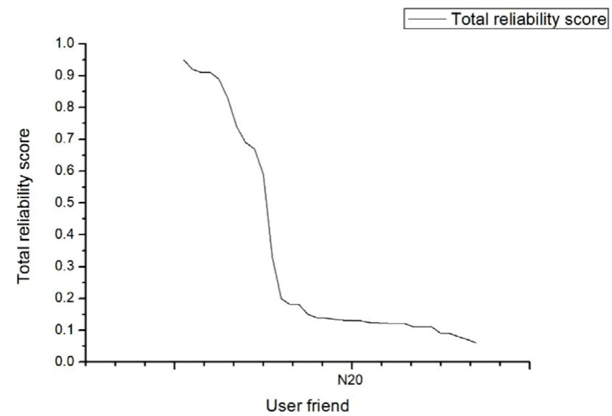

Fig. 2. The total trust rate of respondent.

4. Results and discussion

The main purposes of the experiment are as follows. Firstly, we obtain a reliable user trust weight. Secondly, we extraction of key local words. Thirdly, we obtain the dimension weight of user portrait of reasonable residence. This experiment uses WeChat, QQ and other popular chat information as the data source to get the ranking of friends with a high degree of trust. Secondly, the selected friends were selected for chat text to extract keywords of geographic information and related emotional words. Then, the user profile is made with the historical residence information of the untrustworthy executor, and the weight of each dimension is obtained. Finally, through the calculation of the extracted geographical locations, a high-quality sorted list of geographical locations is generated to obtain the prediction range of the place of residence.

4.1. Estimation of user reliability based on social network

This experiment estimates the trustworthiness of the social friends of the untrustworthy executor based on the experimental data collected from the social network, and the data acquisition estimation of chat frequency, content and time, the trust value obtained is (0, 1). The users are divided into direct friends and indirect friends. The formula is used to calculate the user's trust degree respectively, and the trust degree generated by different paths of indirect friends is summed. (See Table 1.)

On the premise of a large sample size, we can find that the total score of users' trust decreases in a half curve with the increase of the number of friends. According to the results in Fig. 2, we can see that the selection of high-quality users' friends is concentrated in the first n, so we can select the first n with effective reliability and high score as effective screening results for the obtained experiment result. (See Fig. 3.)

4.2. Extraction of local words and emotional words

Through the calculation of experiment 1, select the top n friends who are trusted by users. On this basis, the chat content between the untrustworthy executor and his friends, as well as the spatial dynamics and comments of the friends are processed in data, and the geographical keywords and emotional words are extracted, and the

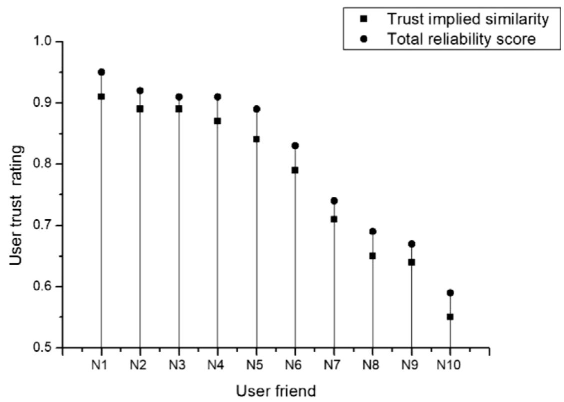

Fig. 3. The trust rate of respondent social network.

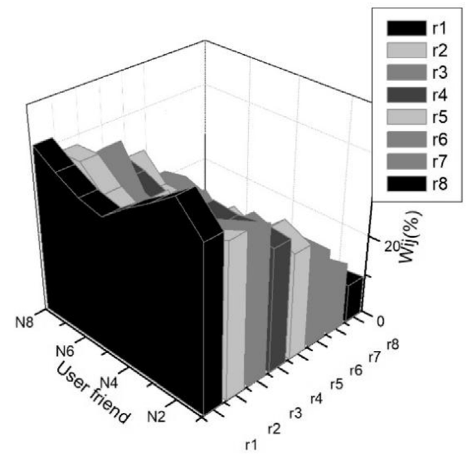

Fig. 4. The weight of the local word R.

Table 2
The weight of the local word R.

W_{ij} (%)	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
N_1	40.6	38.1	37.0	30.7	26.5	22.5	16.1	9.7
N_2	47.2	35.1	34.4	31.2	29.4	22.3	14.5	8.8
N_3	42.6	31.7	31.7	30.0	23.5	20.4	12.7	6.7
N_4	38.4	38.1	37.0	30.7	26.5	22.5	16.1	9.7
N_5	33.5	30.1	27.2	22.3	20.9	16.5	10.1	8.2
N_6	34.9	33.1	29.0	26.7	23.5	20.8	11.1	9.4
N_7	37.9	38.1	39.0	29.3	23.5	22.9	16.9	7.7
N_8	41.6	37.7	33.3	30.7	29.5	21.5	12.8	6.4

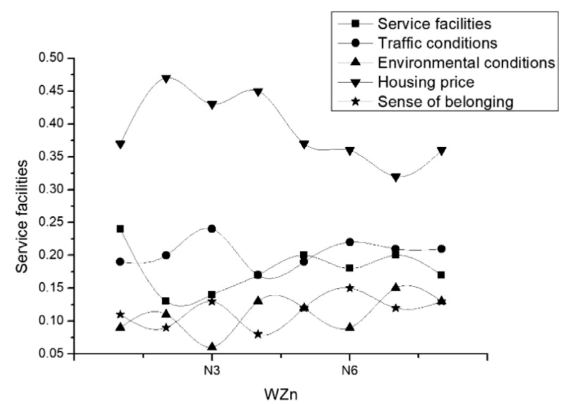

Fig. 5. The weight of the user profile.

Table 3
The weight of the user profile.

W_{Zn}	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Service facilities	0.24	0.13	0.14	0.17	0.20	0.18	0.20	0.17
Traffic conditions	0.19	0.20	0.24	0.17	0.19	0.22	0.21	0.21
Environmental conditions	0.09	0.11	0.06	0.13	0.12	0.09	0.15	0.13
Housing price	0.37	0.47	0.43	0.45	0.37	0.36	0.32	0.36
Sense of belonging	0.11	0.09	0.13	0.08	0.12	0.15	0.12	0.13

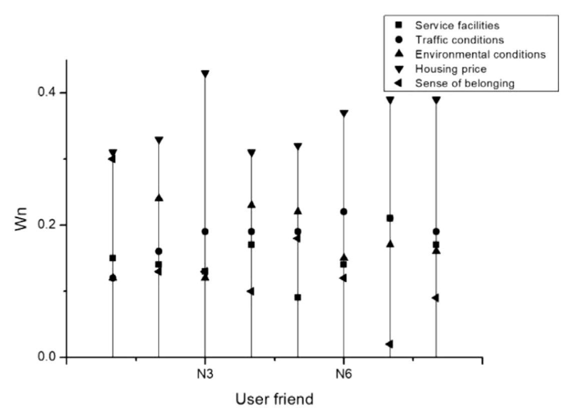

Fig. 6. The weight of the user profile of residence.

Table 4
The weight of the user profile of residence.

W_n	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Service facilities	0.15	0.14	0.13	0.17	0.09	0.14	0.21	0.17
Traffic conditions	0.12	0.16	0.19	0.19	0.19	0.22	0.21	0.19
Environmental conditions	0.12	0.24	0.12	0.23	0.22	0.15	0.17	0.16
Housing price	0.31	0.33	0.43	0.31	0.32	0.37	0.39	0.39
Sense of belonging	0.30	0.13	0.13	0.10	0.18	0.12	0.02	0.09

weight ratio is finally obtained. In the result list of experiment 2, the first m local words r_i with high weight were selected, including the residential addresses of users' friends, to obtain the weight table of local words R according to the formula. (See Table 2.)

The obtained emotional words were divided into five dimensions of service facilities, traffic conditions, environmental conditions, housing price and sense of belonging according to the attributes of user profile in the residential area, and the total word frequency of the dimensions was calculated to obtain the dimension weight W_{Zn} of user profile in the residential area. (See Table 3, Figs. 4 and 5.)

4.3. Weight model based on fuzzy comprehensive evaluation

Through experiment 2, the dimension weight W_{Zn} table of user portrait based on emotion words is obtained, and then the rating table of user portrait of residence in this paper is calculated by combining with the dimension weight obtained from the user portrait of the historical residence of the person who has lost faith. (See Table 4 and Fig. 6.)

After obtaining the above data, it is put into the residential location evaluation index system and weight table, and the selected geographical locations are scored in a form. The total score is 100, and the highest score is the predicted location range of the final residence. (See Table 5, Figs. 7 and 8.)

4.4. Discussion

This paper uses the archived data of historical cases of the court to carry out the corresponding experimental test. There are 708,920 pieces of chat content with location in relevant cases among them,

Table 5
The table of residential location evaluation index system and weight.

W_n	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
Service facilities	88	77	63	77	69	74	71	77
Traffic conditions	79	86	79	79	69	72	61	79
Environmental conditions	46	74	82	83	82	65	87	76
Housing price	90	73	83	81	72	87	79	89
Sense of belonging	86	83	73	80	88	82	82	79

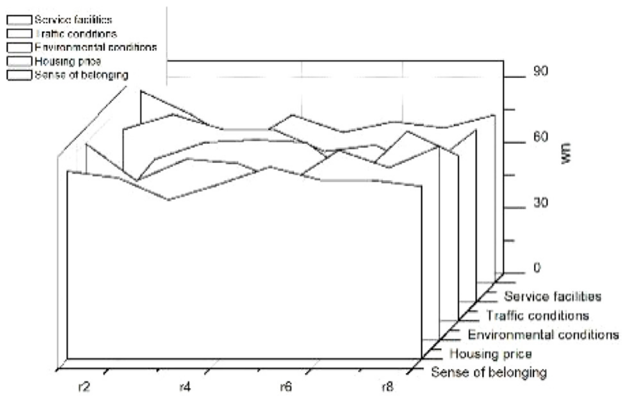

Fig. 7. The residential location evaluation index system.

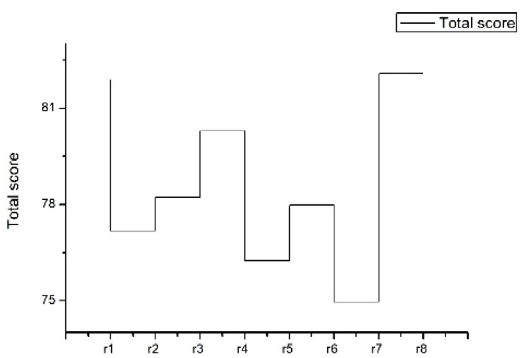

Fig. 8. The score of forecast land of residence.

Table 6
Comparison of experimental results of user location prediction.

Methods	Mean migration distance (km)	Migration distance within 15 km (%)
Flap	29.7	50.3
PLRU	17.1	73.4
UGC-LI	20.1	69.6

accounting for 33% of all data. The PLRU model proposed in this paper was used to conduct big data experiment, and the experimental results were compared with Flap model and UGC-LI model.

The accuracy of the results was found to be improved and running time was also reduced by 20% and 13% respectively. The validity of the prediction model is validated. (See Table 6, Figs. 9 and 10.)

Among the experimental results, some of the experimental results were selected for accuracy comparison, and it was found that the experimental accuracy of this model was higher. Meanwhile, the Flap model and the UGC-LI model were used as the benchmark to compare the execution efficiency of the our PLRU model. Based on the research and analysis of the main influence dimensions of residence choice in different age groups and whether or not they are aliens, this paper makes a more targeted analysis of the dimension classification and weight rating of the residence prediction of different categories of untrustworthy executors. Combined with their historical residential user profiles, this paper introduces their personal living habits, and further scores the selected geographic texts, so as to obtain a number of points with a high total score and obtain the final prediction range of the residence. On the basis of combining the personal user portrait and narrowing the text search scope according to the user's trust, the accuracy of the experimental results is not only improved, but also the execution efficiency is improved.

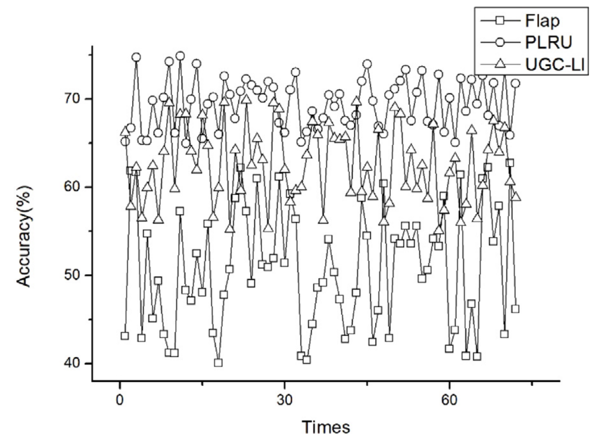

Fig. 9. The accuracy of three algorithms.

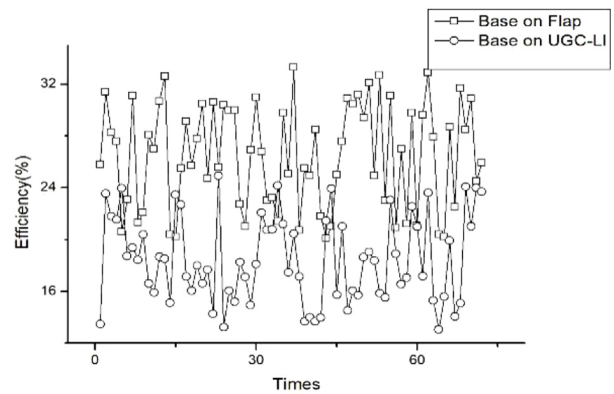

Fig. 10. The execution efficiency of two algorithms.

5. Conclusion

This paper proposes a PLRU model to predict the location of the person's residence based on the trust of the user's social network, when the court cannot use GPS to locate the person or the person whose residence is dishonest cannot be found by hiding his identity. The model effectively integrates user trust and geographical location frequency in social networks. In the process of inferring the residence location of the untrustworthy executor, not only the trust between users and chat content are considered, but also the user profile of the residence is introduced to screen the location points of the possible residence, which effectively improves the accuracy of the results. In the process of tracking the person whose word has been broken, the user friends of the untrustworthy executor will publish real-time dynamic information, which can help us update the local word database in time and make it more timely. Any one's daily behavior will leave more traces in the network with the continuous development of the network information age. Therefore, the court can obtain more reliable and effective data sources, supplement the algorithm factor in this paper, and improve the accuracy of the algorithm.

CRediT authorship contribution statement

Bin Zhang: Conception or design of the work or the acquisition, Analysis, or interpretation of data for the work, Editing and Writing assistance, Read and contributed to the manuscript. **Yanglan Fu:** Conception or design of the work or the acquisition, Analysis, or interpretation of data for the work, Editing and Writing assistance, Read and contributed to the manuscript. **Jing Li:** Conception or design of the work or the acquisition, Analysis, or interpretation of data for the work, Editing and Writing assistance, Read and contributed to the manuscript. **Zhou Ye:** Conception or design of the work or the acquisition, Analysis, or interpretation of data for the work, Editing and Writing assistance, Read and contributed to the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors acknowledge the support by National Key Research and Development Program of China (2018YFC0830400).

References

[1] C. Hsing-Chung, Collaboration IoT-based RBAC with trust evaluation algorithm model for massive IoT integrated application, Mobile Netw. Appl. 24 (3) (2019) 839–852.

[2] H. Wu, K. Yue, Y. Pei, et al., Collaborative topic regression with social trust ensemble for recommendation in social media systems, Knowl.-Based Syst. 97 (Apr.1) (2016) 111–122.

[3] Shuo Wang, Richard O. Sinnott, Surya Nepal, Privacy-protected statistics publication over social media user trajectory streams, Future Gener. Comput. Syst. 87 (2018) 792–802.

[4] Y. Liao, W. Lam, L. Bing, et al., Joint modeling of participant influence and latent topics for recommendation in event-based social networks, ACM Trans. Inf. Syst. 36 (3) (2018) 1–31.

[5] P. Lv, X. Meng, Y. Zhang, FeRe: Exploiting influence of multi-dimensional features resided in news domain for recommendation, Inf. Process. Manage. 53 (5) (2017) 1215–1241.

[6] F.Z. Lebib, H. Mellah, H. Drias, Enhancing information source selection using a genetic algorithm and social tagging, Int. J. Inf. Manage. 37 (6) (2017) 741–749.

[7] Zafran Khan, Naima Iltaf, Hammad Afzal, Haider Abbas, DST-HRS: A topic driven hybrid recommender system based on deep semantics, Comput. Commun. 156 (2020) 183–191.

[8] M.H. Alam, W.J. Ryu, S.K. Lee, Joint multi-grain topic sentiment: Modeling semantic aspects for online reviews, Inform. Sci. 339 (2016) 206–223.

[9] H. Ma, M. Jia, D. Zhang, et al., Combining tag correlation and user social relation for microblog recommendation, Inform. Sci. 385386 (2017) 325–337.

[10] Y. Gong, Q. Zhang, X. Huang, Hashtag recommendation for multimodal microblog posts, Neurocomputing 272 (jan.10) (2018) 170–177.

[11] X. Li, M. Jiang, H. Hong, et al., A time-aware personalized point-of-interest recommendation via high-order tensor factorization, ACM Trans. Inf. Syst. 35 (4) (2017) 1–23.

[12] Sergio G. Burdisso, Marcelo Errecalde, Manuel Montes-y-Gómez, A text classification framework for simple and effective early depression detection over social media streams, Expert Syst. Appl. 133 (2019) 182–197.

[13] M. Goudjil, M. Koudil, M. Bedda, et al., A novel active learning method using SVM for text classification, Int. J. Autom. Comput. 15 (03) (2018) 44–52.

[14] G. Liu, Y. Liu, A. Liu, et al., Context-aware trust network extraction in large-scale trust-oriented social networks, World Wide Web 21 (3) (2018) 713–738.

[15] J.C. Bailón-Elvira, M.J. Cobo, E. Herrera-Viedma, A.G. López-Herrera, Latent Dirichlet Allocation (LDA) for improving the topic modeling of the official bulletin of the spanish state (BOE), Procedia Comput. Sci. 162 (2019).

[16] Kaveh Bastani, Hamed Namavari, Jeffrey Shaffer, Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints, Expert Syst. Appl. 127 (2019) 256–271.

[17] Ali Mohsenzadeh, Amir Jalaly Bidgoly, Yaghoub Farjami, A novel reward and penalty trust evaluation model based on confidence interval using Petri net, J. Netw. Comput. Appl. 154 (C) (2020).

[18] Praveen S. Challagidad, Mahantesh N. Birje, Multi-dimensional dynamic trust evaluation scheme for cloud environment, Comput. Secur. 91 (2020).

[19] A. Mohsenzadeh, H. Motameni, M.J. Er, Retraction note to: A new trust evaluation algorithm between cloud entities based on fuzzy mathematics, Int. J. Fuzzy Syst. 21 (6) (2019) 1988.