H K . N C C P
International Journal of Applied Mathematics and Soft Computing
Volume 4, Issue 1, February, 2018

Location Tracking and Prediction Method for Social Network Users based on Data Mining

Jian Wang[1,2]
[1]Shandong University, Jinan, 276700, China
[2]Qilu Institute of Technology, Jinan, 250200, China

Abstract: In order to improve the ability of social network user behavior analysis and trajectory tracking prediction, optimize social network construction, based on data mining and behavior analysis, the user characteristics of social network are analyzed, and the regularity of social network users is found. A prediction method of social network user location tracking is proposed based on data mining. An association topology model of social network user location distribution is constructed, and segmental feature extraction method is used to extract the relevant feature of social network user location. The Parallel Sets argument axis sorting method is used to schedule social network users' locus storage structure, the implicit pattern of data set is found by fuzzy partition clustering method, and the fuzzy C-means clustering method is used to mine the data. The prediction of social network user location tracking is realized. The simulation results show that the proposed method has high accuracy and high precision of data mining.

Keywords: Data mining; Social network; User location tracking prediction; Feature extraction; Topological model

1. Introduction

With the rapid development of Web2.0, social network is not only a tool but also an interactive platform. It has a profound impact on the establishment of online social relations and interaction. Social network has gradually become the main platform and way of people's emotional communication, relationship maintenance and information communication. The social network has become the focus of the industry in the academic world today. At present, the research and utilization of social networks has been very hot, such as user analysis, relationship analysis, social search, network structure, user privacy and so on, among them, users research is a very important direction, exploring the location and behavior of social network users, combining data mining and behavioral analysis to analyze the user characteristics of social networks and discover the regularity of social network users[1].

The data in the social network is very huge, it is extracted by the related technology, and the location of the network user is located and tracked and forecasted by combining the data mining results. In order to improve the ability of user behavior analysis and promote the optimization and upgrading of social network, the network resource scheduling and optimal allocation are carried out based on the prediction results of user location locus, so as to improve the ability of user behavior analysis of social network. It has great practical significance to study the methods of user location mining and tracking prediction[2-4]. The first step to realize the locus prediction of social network users is time series analysis and feature recombination of big data information flow, and adaptive learning method is used to search and cluster data relevance, such as K-means algorithm, fuzzy C-means algorithm and hierarchical segmentation clustering mining method. The fuzzy C-means clustering method is easy to fall into the local optimal solution in the clustering processing of social network user location locus data mining, and the convergence is not good in the feature extraction of dynamic data. However, the association mining of big data based on hierarchical segmentation clustering method is affected by segmentation threshold, which leads to greater sensitivity to the initial clustering center. The related literatures have improved the algorithms of locus tracking and data mining for social network users. In reference [5], a method of locus tracking prediction for social network users based on adaptive learning of frequent terms is proposed. In order to improve the classification ability of social network user location trajectory data mining, the segmentation prewhitening matching detection algorithm is used to suppress interclass closed frequent item interference. However, with the increase of data association characteristic disturbance, the accuracy of the mining method is not high. In reference [6], a cloud user locus data mining method is proposed based on the fusion of neighbor propagation and density, and the clustering center is adjusted repeatedly to optimize the location trajectory of social network users. The density information of the data is extracted from the nearest neighbor node to carry out the feature clustering, and then predict the location of the users in the social network. This method has a lag in the trajectory prediction[7].

In order to solve the above problems, a method of location tracking and prediction is proposed based on data mining for social network users, and constructs an associated topology model of the location distribution of social network users. The segmented feature extraction method is used to extract the relevant features of the user location of social network, and the Parallel Sets argument axis sorting method is used to schedule the location of social network user locus storage structure. Combined with the fuzzy partition clustering method to discover the hidden pattern in the data set, the fuzzy C-means clustering method is used to mine the data and realize the locus tracking prediction of social network users. Finally, the performance tests are carried out through simulation experiments to demonstrate the superior performance of the proposed method in improving the accuracy of user location trajectory prediction.

2. Associated Topological Structure Model and Feature Extraction Preprocessing

2.1. User locus topology model of social network

In the design of social network information resource transmission system, the user topology of social network is represented by matrix C, where C is the two-dimensional matrix of $N \times N$. N is the number of nodes in the network[8]. For user node i, assuming that N_i^1 is the first adjacent node, the multivariate neighbor node with user locus distribution of social network is obtained as follows:

$$N_i^2 = N_i^1 \cup \left(\bigcup_{j \in N_i^1} N_j^1 \right) \tag{1}$$

Suppose that a frame is composed of M time slots, and the allocation of time slots is represented by matrix X, where $X = x_{mi}$ is a matrix of $M \times N$ -dimension. Where: when the social network user location distribution node i can be sent at the m time slot of the time frame, the value of m is 1, otherwise, the value of x_{mi} is 0, in different initial clustering centers, the location of social network user locus is stored in the $G_1 = \left(M^a{}_1, M^b{}_1, Y_1\right)$, $G_1 = \left(M^a{}_1, M^b{}_1, Y_1\right)$. The attribute distribution is represented by the subgraph, then it satisfies the $G_1 \subseteq G_2 \Leftrightarrow Y_1 \subseteq Y_2$. Let $A = \{a_1, a_2, ..., a_n\}$ be the amplitude of the user trajectory distribution, and the state characteristic of the user locus storage node in the social network is expressed as $p(i,j) = lim_{t\to\infty} p\{ a_t = i, b_t = j \}$, then the total average number of child nodes in the topology tree of the locus distribution of social network users is the ratio of the sum is:

$$\overline{D} = \sum_{l_i=1}^{M-1} \left| D_{l_i} \right| \Big/ \sum_{j=1}^{M-1} \left| L_j \right| \tag{2}$$

The data user node $a(i,j)$ is inputted into the link layer, and the time series of user locus of social network collected by critical path node is x_n, and the expected response is d_n. Thus the user location distribution locus topology of social network is constructed[10]. The structure diagram is shown in figure 1.

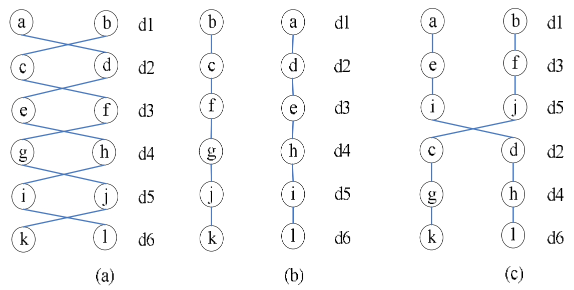

Figure 1. User locus topology of social networks

2.2. Feature extraction of user location trajectory tracking

On the basis of constructing the associated topology model of user location distribution in social network, the feature extraction of user locus tracking is carried out, and the association feature extraction of user location in social network is carried out by using segmental feature extraction method[11-13]. The location tracking step of any non-root node in the user position locus transfer tree is:

$$CW_{\min}^{l_{i+1}} = \begin{cases} CW_{\min}^{0} \times (1+\overline{D}_{l_0})^c, i = 0 \\ CW_{\min}^{l_i} \times (1+\overline{D}_{l_i})^c, 0 < i < M \end{cases} \tag{3}$$

The weighted progressive coefficients of the decision tree for user trajectory tracking are expressed as follows:

$$c = \log_{(1+\overline{D})^M} A \big/ CW_{\min}^{0} \tag{4}$$

In the locus tracking step, the weighted control structure is the same among each step[14]. The Parallel Sets variable axis sorting method is used to schedule the location locus storage structure of social network users. The scheduling model is shown as follows.

According to the average number of sub-nodes between the user layer and the layer in the social network, the feature extraction and the distributed network design are carried out, which provides the data input basis for the user position trajectory tracking of the social network.

$$\begin{cases} \max U = u_1 + u_2 + \ldots + u_n \\ u_i = p_i \\ \sum_i^n p_i = 1, 0 < p_i < 1 \\ \dfrac{p_1/(1-p_1)}{w_1} = \dfrac{p_i/(1-p_i)}{w_i} = .. = \dfrac{p_n/(1-p_n)}{w_n} = \dfrac{1}{K} \end{cases} \tag{5}$$

3. Prediction Algorithm Optimization

3.1. Association data mining for the location of social network users

On the basis of constructing the associated topological structure model of user location distribution and feature extraction of social network, this paper designs an improved algorithm for tracking and predicting user location. In this paper, a new method for social network based on data mining is proposed. Based on the method of tracking and predicting the location of the user, the method of segment feature extraction is used to mine the relevant data of the user's position in social network[15], and the characteristic values of the location of the user are obtained as follows:

$$\begin{aligned} CW_{\min}^{l_M} &= CW_{\min}^{l_{M-1}} \times (1+\overline{D}_{l_{M-1}})^c \\ &= CW_{\min}^{0} \times (1+\overline{D}_{l_0})^c \times (1+\overline{D}_{l_1})^c \times \cdots \times (1+\overline{D}_{l_{M-1}})^c \\ &= CW_{\min}^{0} \times [(1+\overline{D}_{l_0}) \times (1+\overline{D}_{l_1}) \times \cdots \times (1+\overline{D}_{l_{M-1}})]^c \end{aligned} \tag{6}$$

Where

$$\begin{aligned} &[(1+\overline{D}_{l_0}) \times (1+\overline{D}_{l_1}) \times \cdots \times (1+\overline{D}_{l_{M-1}})] \\ &\le ((1+\overline{D}_{l_0} + 1 + \overline{D}_{l_1} + \cdots + 1 + \overline{D}_{l_{m-1}}) / M)^M \\ &= (1+\overline{D})^M \end{aligned} \tag{7}$$

Then

$$\begin{aligned} CW_{\min}^{l_m} &\le CW_{\min}^{0} \times [(1+\overline{D})^M]^c \\ &= CW_{\min}^{0} \times A / CW_{\min}^{0} = A \end{aligned} \tag{8}$$

The cardinality of social network user locus set satisfies $p_0 = \dfrac{1}{m}$, $m = r_1 + r_2 + \ldots + r_i + \ldots + r_n$, and the optimized feature extraction of social network user locus association rule is as follows:

$$p_i = \frac{r_i}{m} = \frac{r_i w_0}{m w_i} = \frac{w_i}{w_1 + \ldots + w_i + \ldots + w_n} \tag{9}$$

According to the data mining results, the trajectory tracking cluster analysis is carried out according to the data mining results.

3.2. Implementation of fuzzy clustering and user location tracking prediction

Combined with the fuzzy partition clustering method to find the hidden pattern in the data set, the asynchronous progressive weighting coefficient of the user's locus track digging amount is obtained as follows:

$$CW_{\min}^{n} = \begin{cases} ((1-B_{l(n)})e^{1-a_n} + B_{l(n)}) \times CW_{\min}^{l(n)}, a_n > 1 \\ CW_{\min}^{l(n)}, \quad 0 \le a_n \le 1 \end{cases} \tag{10}$$

$$B_{l(n)} = (1/(1+\overline{D}_{l(n)-1}))^c \tag{11}$$

$$a_n = \begin{cases} |D_n| / (\overline{D}_{l(n)}), \overline{D}_{l(n)} \ne 0 \\ 0, \quad \overline{D}_{l(n)} = 0 \end{cases} \tag{12}$$

Where, a_n is the asynchronous iterative coefficient of child nodes and $B_{l(n)}$ is the adjustment coefficient of track tracking of network users. At the same time, for any network user node n, because:

$$(1+\overline{D})^M > 1 \tag{13}$$

$$A / CW_{\min}^{0} > 1 \tag{14}$$

Therefore, the center adjustment coefficient of fuzzy subdivision clustering is obtained:

$$c > 0 \tag{15}$$

While $(1+\overline{D}_{l_i})^c > 1$, the output eigenvalues of data mining with user position locus are satisfied:

$$CW_{\min}^{l_{i+1}} > CW_{\min}^{l_i} \tag{16}$$

Then

$$0 < (1-B_{l(n)})e^{1-a_n} < 1 - B_{l(n)} \tag{17}$$

$$B_{l(n)} < (1-B_{l(n)})e^{1-a_n} + B_{l(n)} < 1 \tag{18}$$

Thus:

$$((1-B_{l(n)})e^{1-a_n} + B_{l(n)}) \times CW_{\min}^{l(n)} > CW_{\min}^{l(n)-1} \tag{19}$$

While $a_m > 1$, Then

$$CW_{\min}^{l(n)} > CW_{\min}^{l(m)} > CW_{\min}^{m} \tag{20}$$

Finally, the prediction correlation matrix of user location trajectory tracking is obtained:

$$R = \begin{bmatrix} r(V_1, V_1) & \mathbf{L} & r(V_1, V_{k-1}) \\ \mathbf{L} & \mathbf{L} & \mathbf{L} \\ r(V_{k-1}, V_1) & \mathbf{L} & r(V_{k-1}, V_{k-1}) \end{bmatrix} \tag{21}$$

Thus, the decision criteria of data association mining are as follows:

$$p(Q_s) = \frac{1}{\sqrt{2p}\, s_s} \exp\left[-\frac{(Q_s - \langle Q_s \rangle)^2}{2 s_s^{\,2}} \right] \tag{22}$$

$$\int_{-\infty}^{\infty} p(Q_s)\, dQ_s = 1 \tag{23}$$

If, the absolute error between Q_0 and $\langle Q_s \rangle$ exceeds the threshold Q_c, such as:

$$p(|Q_0 - \langle Q_s \rangle| > Q_c) \le 0.05 \tag{24}$$

On the basis the above analysis, the trajectory prediction of user's position is realized, and the schematic diagram of trajectory distribution is obtained as shown in figure 2.

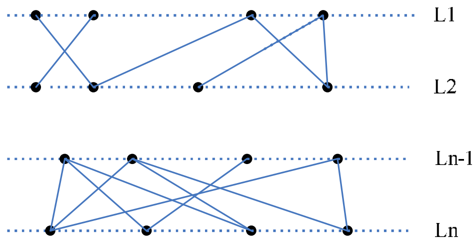

Figure 2. Trajectory tracking predictive output for social users

4. Simulation Experiment and Result Analysis

In order to test the performance of this method in realizing the locus tracking prediction of social network users, the simulation experiment is carried out. The software environment of the experiment is Matlab 7, the number of user nodes of social network is 100, the size of partition is 2400M. The distribution layer number is 25, the embedding dimension is 5, the correlation coefficient is 1.2, the discrete sampling rate of user locus sampling is $f_s = 10 * f_0 \mathrm{Hz} = 10\mathrm{KHz}$, and the training sample scale is 2000. According to the above simulation environment and parameter setting, the user locus of social network is divided. The mining results are shown in figure 3.

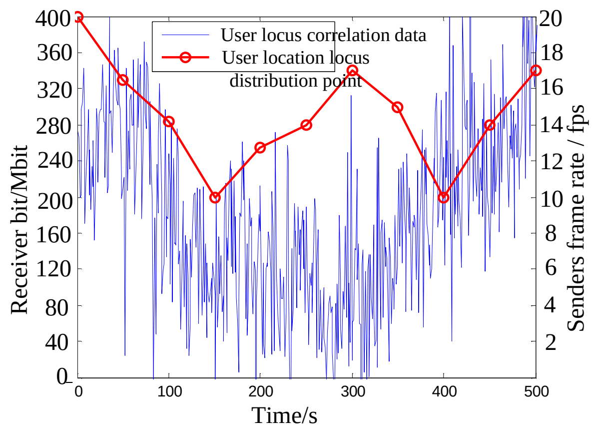

Figure 3. Association data mining for user locus distribution of social networks

According to the data mining results in figure 3, the user location trajectory tracking prediction is carried out, and the prediction accuracy is compared with different methods. The comparison results are shown in figure 4, and the analysis figure 4 shows that the method of this paper is used for social networking. The precision of tracking and predicting the user's position is high.

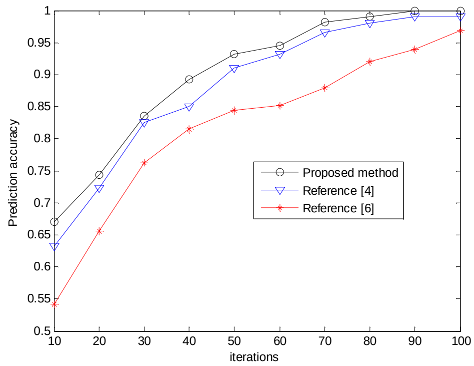

Figure 4. Prediction accuracy comparison

5. Conclusions

In this paper, a data mining based location tracking prediction method for social network users is proposed. The associated topological structure model of the location distribution of social network users is constructed, and the segmented feature extraction method is used to carry out the social network. Based on the feature extraction of user location, the Parallel Sets argument axis sorting method is used to schedule the location of social network user locus storage structure. Combined with the fuzzy partition clustering method, the implicit pattern of data set is found, and the fuzzy pattern is adopted. The C-means clustering method is used for data mining to realize the locus tracking prediction of social network users. The simulation results show that the proposed method has high accuracy and high precision of data mining. It has good practical value in social network user locus mining and behavior analysis.

References

[1] LIU Yazhou, WANG Jing, PAN Xiaozhong, FU Wei. Evolutionary game considering node degree in social network. Journal of Computer Applications, 2018, 38(4): 1029-1035.

[2] ZHU G H, JIANG G P, XIA L L. Rumor spreading model considering conformity phenomena in complex social networks[J]. Computer Science, 2016, 43(2):135-139.

[3] YANG Fengrui, ZHENG Yunjun, ZHANG Chang. Hybrid recommendation algorithm based on probability matrix factorization. Journal of Computer Applications, 2018, 38(3): 644-649.

[4] DURAO F, DOLOG P. Improving tag-based recommendation with the collaborative value of wiki pages for knowledge sharing[J]. Journal of Ambient Intelligence Humanized Computing, 2014, 5(1):21-38

[5] ZHANG F Z, LIU S, LI Z H, et al. Collaborative filtering recommendation algorithm incorporating user's reviews and contextual information[J]. Journal of Chinese Computer Systems,2014, 35(2):228-232.

[6] GUO L, MA J, CHEN Z M, et al. Incorporating item relations for social recommendation[J]. Chinese Journal of Computers, 2014, 37(1):219-228.

[7] TU D D, SHU C C, YU H Y. Using unified probabilistic matrix factorization for contextual advertisement recommendation[J]. Journal of Software, 2013, 24(3):454-464.

[8] MENG X W, LIU S D, ZHANG Y J, et al. Research on social recommender systems[J]. Journal of Software, 2015, 26(6):1356-1372.

[9] FANG H, BAO Y, ZHANG J. Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation[C]//Proceedings of the 28th AAAI Conference on Artificial Intelligence. Menlo Park, CA:AAAI Press, 2014:30-36.

[10] ZHANG L. Cryptanalysis of the public key encryption based on multiple chaotic systems [J]. Chaos, Solitons & Fractals, 2008, 37(3): 669-674.

[11] SALLEH M, IBRAHIM S, ISNIN I F. Image encryption algorithm based on chaotic mapping [J]. Jurnal Teknologi, 2012, 39(1): 1-12.

[12] XUE D, WU L F, LI H B, et al. TraDR:a destination prediction method based on trajectory decomposition and reconstruction in geo-social networks[J]. Computer Science, 2016, 43(3):93-98.

[13] LI Yan, LIU Jiayong. User location prediction model based on author topic model and radiation model. Journal of Computer Applications, 2018, 38(4): 939-944.

[14] LI N, YU Y, ZHOU Z H. Diversity regularized ensemble pruning[C]//Proceedings of the 2012 Joint European Conference on Machine Learning and Knowledge Discovery in Databases, LNCS 7523. Berlin:Springer, 2012:330-345.

[15] PARVIN H, MIRNABIBABOLI M, ALINEJAD-ROKNY H. Proposing a classifier ensemble framework based on classifier selection and decision tree[J]. Engineering Applications of Artificial Intelligence, 2015, 37:34-42.