Research Article

A Method to Study Group Relationship of College Students Based on Predictive Social Network

Honggao Zhang,[1] Jie Zhang,[2] and Yu Shi [3]

[1]*School of Education, Linyi University, Linyi, Shandong 276000, China*
[2]*School of Pharmacy, Linyi University, Linyi, Shandong 276000, China*
[3]*Hebei Academy of Fine Arts, Shijiazhuang, Hebei 050700, China*

Correspondence should be addressed to Yu Shi; sy@hbafa.edu.cn

With the rise of social network platforms such as WeChat, Weibo, and TikTok, social networks have developed from simple social networks to complex social networks. Researchers have gradually found that the traditional data sampling methods can no longer meet the development needs of the complex social network structure. In order to save network resources, various methods of social relationship prediction have been proposed. In this paper, we propose a BTCS algorithm based on low sampling rate under cognitive model and conduct several sets of comparison experiments under different networks and different sampling rates, and the results show that the BTCS algorithm improves the prediction accuracy and reduces the prediction time under low sampling rate. To address the problems of poor stability and slow prediction speed of random sampling prediction methods, this paper proposes a CCS algorithm in colleges and universities using the characteristics of high awareness among nodes within the same college. It can effectively combine the cognitive characteristics of the nodes with the college attributes and apply them to the relationship prediction to realize the college-oriented relationship prediction. The simulation results show that the CCS algorithm is more stable than other random sampling prediction methods. The results make full use of the cognitive characteristics and college attributes of nodes in social networks; reduce the influence of multiple factors such as response time, data packet loss, and individual behavior on relationship prediction; and improve the efficiency of college student group relationship prediction, which has certain theoretical significance and application prospects.

1. Introduction

With the rapid development of information technology, e-commerce and social networking sites have become inseparable from people's daily life, coupled with the increasing networkedness and various data resources, the network has received wide attention as a new perspective for information analysis and management research. [1–6] Especially since entering the twenty-first century, Internet technology has been rapidly developed and people have rapidly entered the era of online Internet. With the rise of various social networking sites and communication software such as Weibo, WeChat, TikTok, Alipay, and QQ, they provide people with direct and quick platforms for online friendships and online shopping. [7–10] Network group is a collective form by individuals in the network with mobile devices or computers as the communication medium and information as the link. Its purpose is mainly study, interest, communication, or need, while college students' network group is an aggregation formed by college students on the network. The activities of college students on the Internet are not individual activities, but more interactive activities with others. Therefore, the communication behaviors of college students' netizens in the virtual space of the network constitute the group network of college students. The complex and changeable network environment brings abundant resources to college students and satisfies their pursuit of material and spiritual culture. However, the virtual network

world also changes the way of psychological activities of college students, making them show a variety of abnormal thinking or action in the network world. The popularity of these social networks and communication tools has revealed that data processing based on the attributes of the entities themselves ignores the information connections between entities and that such data processing methods cannot meet the needs of big data analysis and research. Therefore, in response to the ever-complex internode relationships, researchers have proposed a novel network, that is, social networks. In fact, social network not only refers to the scope of sociology but also includes various information networks, technology networks, and bio-information networks. [11–14] Of course, the most common social networks in daily life are social networks. For most online social networks, they often have thousands and tens of thousands of nodes, and the relationships between nodes are complex and may change in real time. Therefore, they have the same various properties of complex networks. These social networks with a large number of network nodes and complex relationships between nodes are called complex social networks, which is an important branch of complex networks. In recent years, due to the continuous improvement of the mathematical model of social networks, researchers have been studying complex social networks more and more deeply. [15–18] However, in the analysis of complex social networks, researchers have found that there are always nodes that cannot be directly sampled, and this part of node information plays an extremely important role in the overall study of the network. Therefore, the study of prediction methods for internode relationships in complex social networks is an important element of social network research. [19–21]

Social relationship prediction methods can solve the problem of difficult node information collection in interactive networks. In different social networks, they use different network models, so data collection under different models can have mutual effects. Social relationship prediction can provide a unified mechanism to predict out the relationship between nodes, which can realize the data analysis among interactive networks. College student group relationship prediction methods can achieve accurate prediction of internode relationships at low sampling rates. In the analysis of some specific networks, the sampling rate is often limited due to the influence of node response time and the properties of the nodes themselves. However, lower sampling rates can significantly reduce the accuracy of various analyses of social networks. It then needs to predict the topology of the network at low sampling rates. For the needs of these specific networks, the social relationship prediction method can accurately predict the topology of the overall network based on the information of the sampled nodes under the condition of sampling a small number of nodes, which plays an important role in the analysis of social networks at low sampling rates. The group relationship prediction method for college students has important practical applications. In practical applications, if researchers can make full use of the available data information, build accurate model structures, plan experiments rationally,

and analyze and predict hidden relational information, then relational prediction methods will save time, improve efficiency, and produce accurate and valuable results for people. In social networking sites, friend recommendation is realized based on the information of friends' circle; in shopping sites, interest product recommendation is realized based on the kind of goods users browse; in biological information network, hidden species relationship is predicted based on experimental results, etc. [22–25] As an important tool of social network analysis, college student group relationship prediction plays an important role in the research of all fields of social networks. In recent years, researchers began to pay more and more attention to the study of social relationship prediction and proposed various college student group relationship prediction algorithms; especially, since the twenty-first century, researchers have combined college student group relationship prediction with the theoretical findings of psychology and searched for a new perspective on the integration of network structure prediction and individual perception: the social network cognitive model. [26]

The social network cognitive model is a network structure model that emphasizes the importance of network node perception on network structure prediction. It is an emerging mathematical model of social networks, whose core idea is to analyze the network structure by using the cognitive determination of individual nodes on the relationship between nodes of the whole network. As the field of network perception continues to develop, researchers have found that node perception capabilities can help node relationship prediction. Social relationship perception prediction has also gradually received the attention of social network researchers. Therefore, this paper addresses the research of social relationship prediction methods under the social network perception model with very significant scientific and practical significance.

2. BTCS Prediction Algorithm

2.1. Algorithm Improvement. In the prediction of relationships in complex social networks, most researchers find that network data collection is very difficult, especially in interactive networks, which can be affected by multiple factors such as response time, data packet loss, and individual behavior, resulting in unavailability of node information. However, traditional relational prediction methods are highly dependent on the sampling rate and response rate of nodes, and will directly affect the measurement results if the node information is lost or inaccessible. The flow chart of the traditional prediction method is shown in Figure 1. As can be seen from Figure 1, the sampling process of the traditional prediction method requires sampling all nodes in the network to obtain the internode relationship matrix. Then, according to the node relationship matrix, the senders and receivers in the network relationship are accessed and merged two by two. The network relationship exists only when both sender and receiver acknowledge the existence of the network relationship. If one party decides that the relationship does not exist, then the network relationship does not exist. Finally, the network topology matrix is obtained

$P_k > P$

$U_{i,j} \longrightarrow S_{i,j} \longrightarrow P_k$
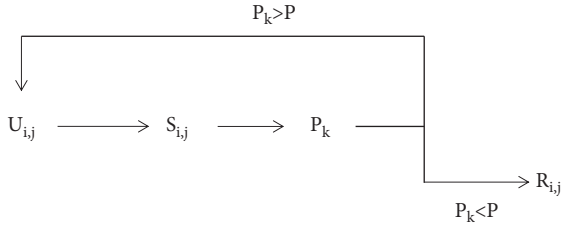
$\longrightarrow R_{i,j}$

$P_k < P$

FIGURE 1: BTCS algorithm flow.

based on the merged results. However, in the traditional relational prediction method, it is affected by response time and sampling error, and data packet loss occurs. Therefore, the error analysis of the traditional method requires multiple repetitions of sampling. In the prediction of high-precision networks, the traditional method needs to complete multiple repetitions of sampling, which is very inefficient.

The BTCS algorithm achieves accurate prediction of social relationships at low sampling rates. The flow chart of the improved BTCS prediction algorithm is shown in Figure 2. Compared with traditional prediction methods, the BTCS algorithm is a randomized, low-sampling-rate relationship prediction method. It only needs to sample a small number of network nodes and obtain the cognitive determination information of the sampled nodes on the relationships among all nodes in the network, and then predict the social relationships among nodes based on the cognitive determination information without sampling all nodes. It adopts a threshold control method to adaptively control the node sampling error and avoid the duplicate sampling of data. Therefore, the BTCS algorithm can reduce the effects of response time, data packet loss, and individual behavior on node sampling and relationship prediction, and improve the efficiency of relationship prediction.

The BTCS prediction algorithm is a cognitive model-based relational prediction method, which is divided into three main steps: sampling, matrix merging, and fault-tolerant control process, as shown in the flow chart of the algorithm. The following section will introduce the specific process of each of the three steps of the BTCS algorithm.

2.2. Sampling Design.

The data sampling process of BTCS prediction algorithm is different from the traditional relationship prediction method. In the traditional relationship prediction method, it needs to sample most of the network nodes and investigate to obtain the social relationships among the sampled members, while in the BTCS algorithm, it just samples a small number of nodes randomly, obtains the cognitive information of the sampled nodes on the social relationships among all network members, and records it in the form of a kind of three-dimensional 0–1 cognitive matrix set.

In the following example of the network (A, B, C, D, and E), Figure 3 represents the five cognitive matrices of members A, B, C, D, and E in the network (A, B, C, D, and E) in order. In each cognitive matrix, it records 20 social relationships among the members. In the sampling, the BTCS algorithm is randomly sampled in the sampling space n. In the network (A, B, C, D, and E), three nodes (A, D, and E) were randomly sampled

when the sampling space n was 3, and the node sampling set m and the cognitive matrix set $R_{i,j,k}$ were obtained, where $m = (1, 4, 5)$.

2.3. Matrix Design.

There are three methods of matrix merging under the cognitive model: the cognitive slice method, the local summary structure method, and consistency structure dispel. Since these three methods merge the known and cognitive relationships of the sampled nodes equally in cognitive matrix merging, a new matrix merging method is proposed. According to the cognitive information of sampled nodes is divided into two categories of self-knowledge relationships and cognitive relationships, this paper divides the merging process into three cases: (1) merging self-knowledge relationships between sampled nodes; (2) merging cognitive relationship between sampled nodes; (3) merging known and cognitive relationship between sampled nodes. Merging rules are set according to these three cases.

We combine the known relationships between the sample nodes. In this paper, we use the merging local aggregation method, as in the following equation:

$$R_{i,j} = \left\{ R_{i,j,i} \cap R_{i,j,j} \right\}, \tag{1}$$

where $R_{i,j,i}$ and $R_{i,j,j}$ represent the known relationships in the cognitive matrix of sampled nodes i and j, respectively, and $R_{i,j}$ denotes the group relationship between node i and j. When both $R_{i,j,i}$ and $R_{i,j,j}$ are 1 and both the sender and receiver of the relationship determine that the relationship exists, then the relationship i to j is judged to exist and $R_{i,j}$ is 1.

Cognitive relationship merging between sampled nodes is as follows:

$$R_{i,j} = \begin{cases} 1, \sum_m R_{i,j,m} \geq K, \\ 0, \end{cases} \tag{2}$$

where $\sum_m R_{i,j,m}$ denotes the individual cognitive summary of all sampled nodes for the social relationship between i and j. K is a set threshold valve. When both sender i and receiver j of the relationship are not sampled, the summary value of cognitive information of the relationship is referred to all sampled nodes, and if the summary value is greater than or equal to K, then the relationship $R_{i,j}$ is determined to exist.

The known and perceived relationships between sampled nodes are merged.

$$R_{i,j} = \begin{cases} 1, \ R_{i,j} = 1 \sum_m R_{i,j,m} \geq K, \\ 0, \end{cases} \tag{3}$$

where $R_{i,j} = 1 \& \sum_m R_{i,j,m} \geq K$ denotes that when only one side of both sender i and receiver j of the relationship is sampled, then the known relationship of the sampled node is first considered to be 1. If the known relationship of the sampled node is 1 and the aggregated value is greater than or equal to K, then the relationship $R_{i,j}$ is determined to exist.

The final result θ is obtained by combining the three cases, $\theta = \alpha + \beta + \gamma$.
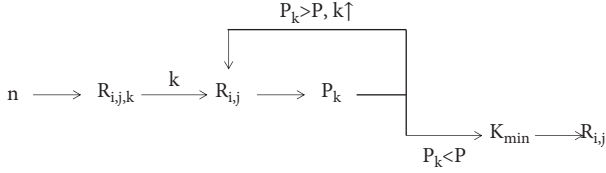
FIGURE 2: Upgraded BTCS algorithm flow.

Taking the network (A, B, C, D, and E) in Figure 3 as an example, the nodes (A, D, and E) are randomly sampled to obtain the sampling set (1, 4, and 5) and the cognitive matrix of A, D, and E. Then, the matrix merging is completed when the threshold value is 2, and the merging process and the merging results are obtained as shown in Figure 4. In Figure 4, α, β, and γ are the matrices obtained by merging in the first, second, and third cases, respectively; θ is the final result to obtain the real network matrix. The social relationship between the sampled nodes (A, D, and E) is obtained in α. For example, when analyzing the social relationship from E to D, by accessing the self-knowledge relationship of the cognitive matrix of D and E, we get $R_{E,D,D}$ and $R_{E,D,E}$ as 1. Then, we can decide that the social relationship from E to D exists, so we get $R_{E,D} =1$ in α.

The social relationships between the unsampled nodes (B and C) are obtained in β. For example, in the analysis of the social relationships from B to C, the cognitive relationships of all sampled nodes are accessed and $\sum_m R_{B,C,m} = 2$, which is equal to the threshold K, so we get $R_{B,C}$ in x as 1. The social relationships between sampled nodes (A, D, and E) and unsampled nodes (B and C) are obtained in γ. For example, when analyzing the social relationship from A to C, the known information of the cognitive matrix of A is accessed first, and Ra is obtained as 1. Then, the cognitive relationships of all sampled nodes are accessed, and $\sum_m R_{A,C,m} = 2$, so $R_{A,C,A}$ is obtained in γ, and $R_{A,C}$ is 1.

The nodal relationship matrix θ is predicted by introducing the K-value merged 3D matrix $R_{i,j,m}$, and different prediction results θ are obtained under different K-values. The random partial sampling method under the social relationship cognitive model is a fast and efficient measurement method, and the prediction results are subject to error, which requires us to control the error within a tolerable range, so the error-tolerant control of the algorithm is important. Therefore, an error-tolerant control process with a threshold value is designed to adaptively regulate the value of K so that the prediction result θ is closest to the true result.

Firstly, we analyze the relationship between the sampling error rate and the threshold value as a function of P_K. These errors are divided into two categories: the first type of relationship error refers to the relationship does not exist in the real network, while the cognitive relationship is judged to exist; the second type of relationship error P_2 is the relationship exists in the real network, while the cognitive relationship is judged not to exist. In the dichotomous threshold algorithm, the threshold K is compared with the network recognition to determine whether the relationship holds, and the recognition is

influenced by P_1, so P_1 is the main factor affecting the threshold reduction result, while P_2 will be gradually reduced in the reduction operation.

In Figure 4, the first type of error is generated by merging the cognitive matrices of sampled nodes A and D, which determine the nonexistence of the relationship A to D, while the cognitive determination exists in the cognitive matrix of E. In complex networks, P_1 increases as the network data increase, and the error P_1 can represent the network error rate P_k. Therefore, we need to analyze the relationship between the error P_1 and the threshold K. Therefore, this section analyzes the effect of the threshold K on θ by comparing the merged matrix θ and the cognitive matrix $R_{i,j,m}$ of the sampled nodes, and obtains the relationship between the value of K and the sampling error rate P_k.

$$P_k = \frac{V}{Q}, \tag{4}$$

where V is the number of errors present in the sampled nodes and Q is the number of possible errors in the sampled nodes.

Setting the fault-tolerance control condition. In this section, based on the maximum tolerable sampling error rate P' (0.1, 0.15, 0.2), an error-tolerant control process is established to find the smallest $K_{\min}$ value such that $P_k < P'$, and the matrix merging result is obtained when the output threshold K is $K_{\min}$, which is the closest to the real network. The prediction result is the closest to the real network.

A error-tolerant control process is designed. Through the above analysis, the relationship P_k between the error rate P and the threshold K is obtained. Therefore, according to the control theory and method, and the fault tolerance range, it can adaptively adjust the threshold K and control the error rate P_k within the fault tolerance rate P'. In words, in BTCS algorithm, fault-tolerant control process is an important part of the algorithm. Too large or too small K value will directly affect the accuracy of prediction results. BTCS algorithm mainly analyzes the direct relationship between K value and network sampling error, and designs the fault-tolerant control process of threshold value, which can adaptively control K value, find the best K value, minimize prediction error, and improve prediction accuracy.

2.4. Mathematical Model. The mathematical model of BTCS algorithm is divided into two parts: matrix dimension reduction model and fault-tolerant control model. In the matrix dimensionality reduction model, it combines the cognitive matrix according to the dimensionality reduction rules under the set K value. In the fault-tolerant control model, it analyzes the sampling error rate of the combined matrix, controls the error rate within the fault-tolerant range, and obtains the best threshold. Therefore, the mathematical model of BTCS algorithm is established step by step according to the matrix dimensionality reduction model and error-tolerant control model.

According to the matrix dimensionality reduction rules designed in the BTCS algorithm description, this subsection

A	A	B	C	D	E
A	0	0	1	0	1
B	0	0	1	0	0
C	1	1	0	0	0
D	0	0	0	0	0
E	1	0	0	0	0

A	A	B	C	D	E
A	0	1	0	0	0
B	1	0	0	0	1
C	0	1	0	0	1
D	0	0	0	0	0
E	0	0	0	0	0

A	A	B	C	D	E
A	0	0	0	0	0
B	1	0	1	0	0
C	1	1	0	0	0
D	0	0	0	0	0
E	0	0	0	0	0

A	A	B	C	D	E
A	0	0	1	0	1
B	0	0	1	1	0
C	1	1	0	0	0
D	0	1	0	0	1
E	1	0	0	1	0

A	A	B	C	D	E
A	0	0	0	1	0
B	0	0	0	0	1
C	0	0	0	0	0
D	1	0	0	0	1
E	0	1	0	1	0

FIGURE 3: 5 node awareness matrix.

A	A	B	C	D	E
A	0	0	1	0	1
B	0	0	1	0	0
C	1	1	0	0	0
D	0	0	0	0	0
E	1	0	0	0	0

D	A	B	C	D	E
A	0	0	1	0	1
B	0	0	1	1	0
C	1	1	0	0	0
D	0	1	0	0	1
E	1	0	0	1	0

E	A	B	C	D	E
A	0	0	0	1	0
B	0	0	0	0	1
C	0	0	0	0	0
D	1	0	0	0	1
E	0	1	0	1	0

α

	A	B	C	D	E
A	0	0	0	0	0
B	0	0	0	0	0
C	0	0	0	0	0
D	0	0	0	0	1
E	0	0	0	1	0

β

	A	B	C	D	E
A	0	0	0	0	0
B	0	0	1	0	0
C	0	1	0	0	0
D	0	0	0	0	0
E	0	0	0	0	0

γ

	A	B	C	D	E
A	0	0	1	0	0
B	0	0	0	0	0
C	1	0	0	0	0
D	0	0	0	0	0
E	0	0	0	0	0

θ

	A	B	C	D	E
A	0	0	1	0	0
B	0	0	1	0	0
C	1	1	0	0	0
D	0	0	0	0	1
E	0	0	0	1	0

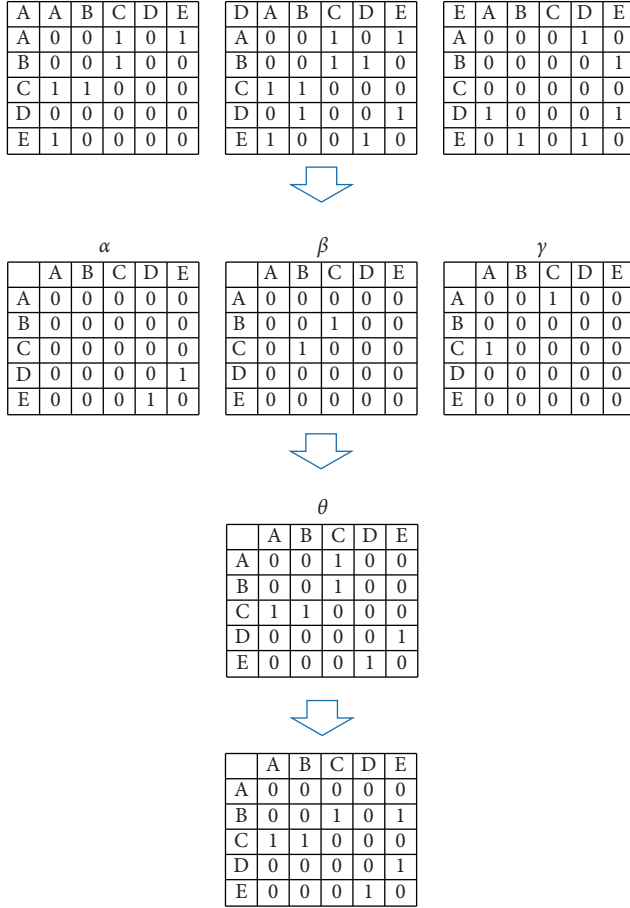|	A	B	C	D	E
A	0	0	0	0	0
B	0	0	1	0	1
C	1	1	0	0	0
D	0	0	0	0	1
E	0	0	0	1	0

FIGURE 4: Merge matrix process.

first obtains the sampling set m, the cognitive matrix Y_i, and the set threshold K. Then, the valid information in the cognitive matrix is extracted in steps, and the mathematical matrix operations are obtained as follows:

(1) separate the cognitive matrix Yi to obtain the known relationship matrix $Y_{Z,i}$, and the cognitive relationship matrix $Y_{R,i}$.

Known relationship matrix:

$$Y_{Z,i} = i * Y. \tag{5}$$

Cognitive relationship matrix:

$$Y_{Z,i} = -i * Y_i. \tag{6}$$

Cognitive matrix:

$$Y_i = Y_{z,i} + Y_{R,i}. \tag{7}$$

Sampling node matrix and Z:

$$Z = \sum_{i=m} Y_{Z,i}. \tag{8}$$

Sampling node matrix and R:

$$R = \sum_{i=m} Y_{R,i}. \tag{9}$$

(2) The topological information among the sampled nodes is extracted from the cognitive relationship Z to obtain the matrix α.

Extraction matrix:

$$T = m,$$
$$\alpha = S_1 (T * Z). \tag{10}$$

(3) extract the topological information of the cognitive relationship judgment to obtain $\beta+\gamma$.

2.5. Extraction Matrix

$$W = -m. \tag{11}$$

Cognitive information

$$P_1 = W * Z,$$
$$P_2 = W * R,$$
$$\beta + \gamma = S_k (P_1 + P_2). \tag{12}$$

The three cases are summed to obtain the reduced-dimensional merged matrix θ.

$$\theta = \alpha + \beta + \gamma. \tag{13}$$

2.6. Error-Tolerant Control Mathematical Model Implementation.
According to the error-tolerant control process described by the BTCS algorithm, the sampling error rate P_k of the merged matrix under the set threshold K is analyzed and the mathematical model operation is obtained as follows:

(1) Calculate the number of errors W present in the sampled nodes.

$$W = S_k (T * R). \tag{14}$$

(2) Calculate the number of possible errors Q in the sampling nodes.

$$Q = T - \alpha. \tag{15}$$

(3) Sampling error rate P_k.

$$P_k = \frac{W}{Q}. \tag{16}$$

Finally, the sampling error rate P_k under the control of threshold K is obtained, and the dichotomous iteration method is used to find the optimal threshold value K that satisfies the fault tolerance condition.

3. CCS Relationship Prediction Method within Colleges and Universities

3.1. CCS Algorithm. In the social relationship cognitive model, the overall network random sampling prediction method ignores the different cognitive characteristics among nodes and ignores the high cognitive characteristics of nodes within the same college. Therefore, the stability of this relationship prediction method is poor. To address the shortcomings of the overall network random sampling prediction method, a relationship prediction method with random sampling in colleges and universities is proposed as the CCS prediction algorithm. [27] The CCS algorithm is applicable to networks where all college information is known and the college structures do not overlap. When a node belongs to more than one college at the same time, it is grouped into the college with the higher number of relationships among nodes.

The CCS algorithm takes advantage of the characteristics of high cognitive degree among nodes within the same college; firstly, the CCS algorithm assigns the sampled nodes to each college proportionally; then, the CCS algorithm completes the random sampling within the college according to the sampling space of each community to obtain the cognitive matrix of sampled nodes; finally, for the characteristics of high cognitive degree of nodes within the college, the CCS algorithm designs the matrix merging rules based on the college to get the final prediction results. Compared with other relationship prediction methods, the CCS prediction algorithm effectively combines the cognitive characteristics of nodes and college attributes.

The CCS prediction algorithm is a relationship prediction method in colleges and universities based on a social cognitive model. Figure 5 shows the flow chart of CCS prediction algorithm; in general, it is divided into three main steps: community-based sampling, intracollege node relationship merging, and inter-GA node relationship merging.

The community-based sampling is mainly to reasonably allocate the sampling nodes according to the network community structure and sampling space to obtain the sampling node cognitive matrix. In the community-based cognitive matrix merging, the sampled node cognitive matrix relationships a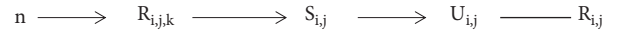re first classified into intracommunity node relationships and intercommunity node relationships; then, the nodes are merged in steps according to the different intracommunity and intercommunity nodes to obtain the prediction results. In this section, four parts of community-based sampling process design, community-based node relationship classification, intracommunity node merging design, and intercommunity node merging design are introduced, respectively.

$$n \longrightarrow R_{i,j,k} \longrightarrow S_{i,j} \longrightarrow U_{i,j} \longrightarrow R_{i,j}$$

FIGURE 5: Flow chart of CCS prediction algorithm.

3.2. Sampling Process Design. In the data sampling process, the CCS prediction algorithm implements a community-based cognitive matrix data sampling, which combines cognitive information with college information. The community-based sampling characteristics of the CCS prediction algorithm are mainly reflected in the following three aspects.

The CCS prediction algorithm is a data sampling process based on a cognitive model. Like other relationship prediction methods under the cognitive model, the data obtained after sampling by the CCS prediction algorithm are also a kind of three-dimensional 0–1 cognitive matrix set $R_{i,j,k}$, where i represents the sender of the social relationship, j represents the receiver, and $R_{i,j,k}$ represents the observer of the social relationship between i and j. If K observes that the relationship from i to j exists, then $R_{i,j,k}$ indicates that it does not exist.

The CCS prediction algorithm is a data sampling process based on the node community structure. Based on the network community structure (C1, C2, ... CN), the CCS algorithm will allocate the sampling space to each community proportionally so that the sampling ratio within each community is the same. For communities with more nodes, the CCS algorithm allocates a correspondingly larger number of sampled nodes; while for communities with fewer nodes, the number of samples is correspondingly lowered to ensure an even number of sampled nodes within each community.

The CCS prediction algorithm is a data sampling process based on random sampling within communities. For each community, under the condition that the number of sampled nodes is determined, the CCS algorithm uses a random sampling method in which each node has the same probability of being sampled.

3.3. Stability Analysis. In this subsection, this paper analyzes the stability of the prediction algorithm under random sampling conditions. In order to analyze the stability of the prediction algorithm, multiple sets of experimental predictions were performed in this paper for different node sampling results under the same sampling space using the CCS algorithm, the Central Graph algorithm, and the LAS algorithm. In each set of experiments, 1000 random samples are performed for the same sampling space and 1000 prediction results are obtained, and the confidence interval (CI)
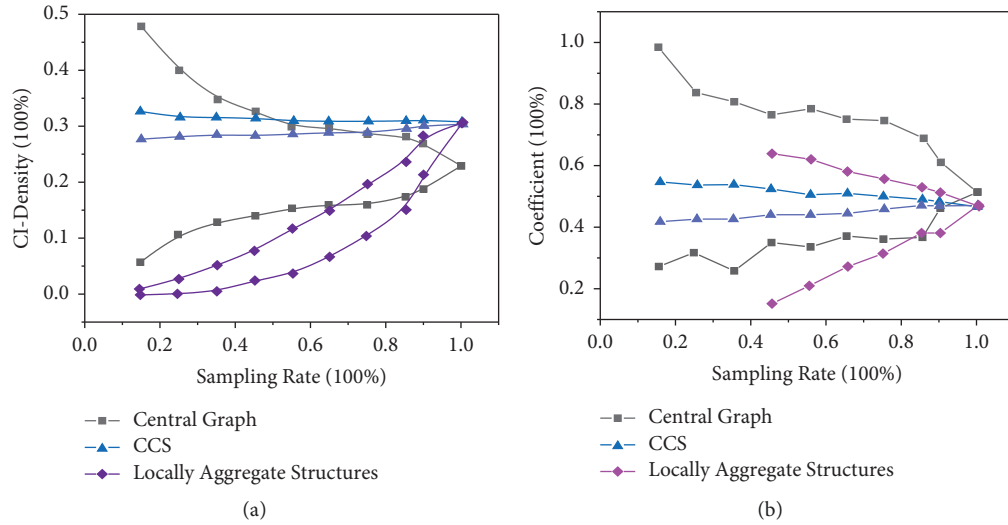
FIGURE 6: Density CI distribution (a) and clustering coefficient CI distribution (b).

numerical processing method is used to analyze the 1000 results.

In the confidence interval processing, this paper first ranks the performance parameters of 1000 prediction network results and then removes the first 2.5% and the last 2.5% to get the distribution curve of the performance parameters, and finally, this paper compares the confidence distribution of the performance parameters of the obtained prediction network results with the performance parameters of the real network results to analyze the stability of the algorithm under the random sampling conditions.

Figures 6(a) and Figure 6(b) show the density CI distribution curves and clustering coefficient CI distribution curves of the CCS algorithm, Central Graph algorithm, and LAS algorithm under different meter sample nodes at the same meter sample rate, respectively; where the horizontal coordinates indicate the sampling rate, the vertical coordinates indicate the network density values and clustering coefficient values of the predicted results, and the dashed lines indicate the performance parameters of the real network. From the analysis in Figures 6(a) and 6(b), it can be obtained that (1) the network density distribution and the clustering coefficient distribution of 1000 measurements at different sampling rates are within the confined curve range. (2) When the sampling space is small, the network density confidence interval distributions and the clustering coefficient confidence interval distributions of the 1000 random sampling prediction results of the CCS algorithm are smaller than those of the LAS algorithm and the Central Graph algorithm, and both are close to the performance parameters of the real network. (3) As the sampling rate increases, the network density confidence interval distribution and the clustering coefficient confidence interval distribution of the prediction results of the three prediction algorithms show convergence and gradually become smaller. However, compared with the other two algorithms, the network density confidence interval distribution and the clustering coefficient confidence interval distribution of the CCS algorithm are closer to the performance parameters of the real

network. (4) When the sampling rate is larger, the CCS prediction algorithm outperforms the LAS algorithm and the Central Graph algorithm, and the CCS prediction algorithm predicts results that are close to the real network density. Therefore, the CCS algorithm is more stable than the other two algorithms, it is less volatile by the different sampling nodes and the stability of the algorithm is higher.

3.4. Prediction Accuracy Analysis.
In terms of accuracy analysis of the prediction results, the mean square error of the three algorithms under different sampling rates is compared. In this paper, the CCS algorithm is compared with the Central Graph algorithm and the LAS (locally aggregates structures) algorithm.

To analyze the prediction accuracy of the CCS algorithm in communities, multiple sets of prediction experiments were conducted in the cognitive social network data package using the CCS algorithm with the Central Graph algorithm and the LAS algorithm, respectively, and network density MSE curves and average clustering coefficient MSE curves were obtained, as shown in Figures 7(a) and 7(b), where the horizontal coordinates denote the sampling rate, and the vertical coordinates indicate the MSE values of network density and clustering coefficients. From Figures 7(a) and 7(b), we can get that (1) the MSE values of CCS prediction algorithm fluctuate less than Central Graph algorithm and LAS algorithm at different sampling rates, and the prediction results of both CCS algorithms are very close to the real network. (2) When the sampling rate is small, the MSE values of Central Graph algorithm and LAS algorithm are larger, while the MSE values of CCS algorithm are smaller, and the prediction accuracy of CCS algorithm is still very high. (3) When the sampling rate is larger, the error value of the LAS algorithm gradually decreases and approaches the real network results, while the CCS algorithm still has a certain smaller error when the sampling rate is 1. Because in the sampling process of the CCS algorithm, it makes full use of the network community information and subcommunity
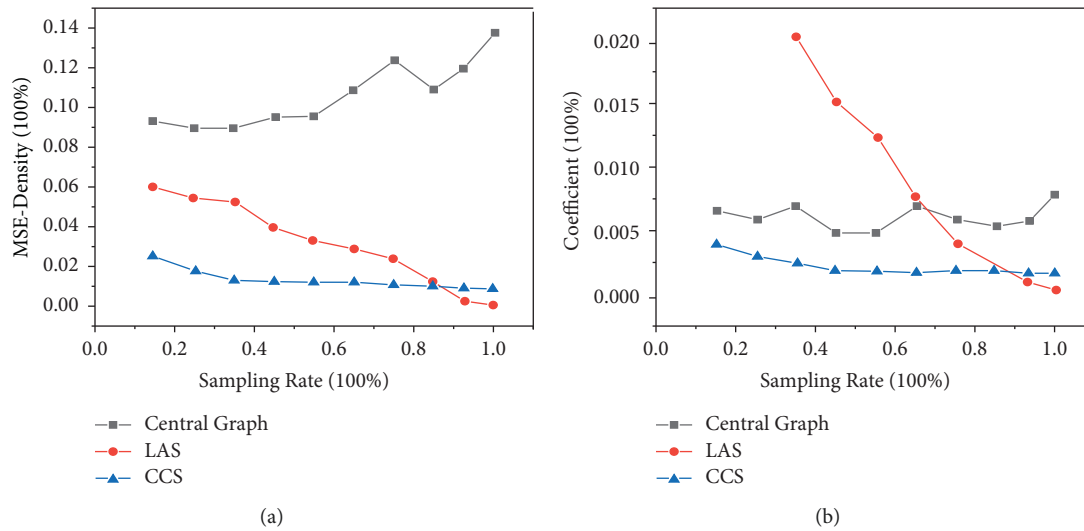
FIGURE 7: MSE of network density at different sampling rates (a) and MSE of clustering coefficients at different sampling rates (b).

random sampling to ensure the balance of the overall network cognition of the sampled nodes, it can always maintain a high prediction accuracy. When the sampling rate is 1, the community merging rule of CCS algorithm relies on the overall cognition of the community to complete the relationship determination instead of directly accessing the relationship sender and receiver, and thus, there is a small error even if all nodes are sampled.

The CCS prediction algorithm in the community is more stable, more accurate, and faster than other relationship prediction algorithms under cognitive models.

4. Conclusion

This paper addresses the relationship prediction method under the social network cognitive model. Based on the relationship prediction method under the cognitive model, a relationship prediction method under the cognitive model of college student group is proposed, that is, BTCS prediction algorithm. The cognitive nature of the nodes in the network is exploited, and the cognitive model of social relationships is combined with the relationship prediction technique to design a new method of cognitive matrix merging under the cognitive model of university student groups, which achieves the relationship prediction of complex social networks with low sampling rate and reduces the sampling time; secondly, two methods of threshold control and dichotomous lookup are introduced, and the interrelationship between threshold and network topological relationship error is analyzed, and the dichotomous lookup is used to optimize the measurement process. The method of dichotomous lookup is used to optimize the measurement process and improve the efficiency of threshold control. Based on the characteristics of high node awareness within the same college, a method for predicting college student group relationships in colleges and universities is proposed, that is, the CCS prediction algorithm. In social networks, network nodes not only have cognitive properties but also have social attributes. Due to the different social attributes among

nodes, the closeness and cognitive degree among nodes are different. The group relationship prediction method for college students with random sampling under the cognitive model exploits the node cognitive characteristics to reduce the effects of response time, data packet loss, and individual behavior. Experimental results show that the CCS prediction algorithm in colleges and universities is more stable, more accurate, and faster than other group relationship prediction algorithms for college students under cognitive models.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] G. H.-L. Cheng, P. Sung, A. Chan, S. Ma, and R. Malhotra, "Transitions between social network profiles and their relation with all-cause mortality among older adults," *Social Science & Medicine*, vol. 292, Article ID 114617, 2022.

[2] T. Hu, X. Zheng, X. Zhang, and D. D. Zeng, "ϵ-k anonymization and adversarial training of graph neural networks for privacy preservation in social networks," *Electronic Commerce Research and Applications*, vol. 50, Article ID 101105, 2021.

[3] Y.-J. Zhang, X.-l. Cai, H.-X. Hu et al., "Social brain network predicts real-world social network in individuals with social anhedonia," *Psychiatry Research: Neuroimaging*, vol. 317, Article ID 111390, 2021.

[4] J. E. Losee, G. D. Webster, and C. McCarty, "Social network connections and increased preparation intentions for a

disaster," *Journal of Environmental Psychology*, vol. 79, Article ID 101726, 2022.

[5] L. Chen, J. Chen, and C. Xia, "Social network behavior and public opinion manipulation," *Journal of Information Security and Applications*, vol. 64, Article ID 103060, 2022.

[6] D. Jeong, J. Kim, D. Choi, and E. Park, "Social networking services as new venue for public perceptions of energy issues: the case of Paris agreement," *Energy Strategy Reviews*, vol. 39, Article ID 100758, 2022.

[7] H. Suk and R. Grace, "Influence of social networks and opportunities for social support on evacuation destination decision-making," *Safety Science*, vol. 147, Article ID 105564, 2022.

[8] J. Gao, H. Hu, and H. He, "Household indebtedness and depressive symptoms among older adults in China: the moderating role of social network and anticipated support," *Journal of Affective Disorders*, vol. 298, pp. 173–181, 2022.

[9] J. Jabari Lotf, M. Abdollahi Azgomi, and M. R. Ebrahimi Dishabi, "An improved influence maximization method for social networks based on genetic algorithm," *Physica A: Statistical Mechanics and Its Applications*, vol. 586, Article ID 126480, 2022.

[10] J. Pulgar, "Classroom creativity and students' social networks: theoretical and practical implications," *Thinking Skills and Creativity*, vol. 42, Article ID 100942, 2021.

[11] C. K. Tanui, S. Karanth, P. M. K. Njage, J. Meng, and A. K. Pradhan, "Machine learning-based predictive modeling to identify genotypic traits associated with *Salmonella enterica* disease endpoints in isolates from ground chicken," *Lebensmittel-Wissenschaft & Technologie*, vol. 154, Article ID 112701, 2022.

[12] M. He, C. Li, Y. Kang, Y. Zuo, L. Duo, and W. Tang, "Clinical predictive model for the 1-year remission probability of IgA vasculitis nephritis," *International Immunopharmacology*, vol. 101, Article ID 108341, 2021.

[13] M. J. Wessel, P. Egger, and F. C. Hummel, "Predictive models for response to non-invasive brain stimulation in stroke: a critical review of opportunities and pitfalls," *Brain Stimulation*, vol. 14, no. 6, pp. 1456–1466, 2021.

[14] L. Wisnieski, D. E. Amrine, and D. G. Renter, "Predictive modeling of bovine respiratory disease outcomes in feedlot cattle: a narrative review," *Livestock Science*, vol. 251, Article ID 104666, 2021.

[15] J. Chu, Y. Wang, X. Liu, and Y. Liu, "Social network community analysis based large-scale group decision making approach with incomplete fuzzy preference relations," *Information Fusion*, vol. 60, pp. 98–120, 2020.

[16] Y. Muto, N. Awano, M. Inomata et al., "Predictive model for the development of critical coronavirus disease 2019 and its risk factors among patients in Japan," *Respiratory Investigation*, vol. 59, no. 6, pp. 804–809, 2021.

[17] H. Hesabi, M. Nourelfath, and A. Hajji, "A deep learning predictive model for selective maintenance optimization," *Reliability Engineering & System Safety*, vol. 219, Article ID 108191, 2022.

[18] R. S. Chavez, D. T. Tovar, M. S. Stendel, and T. D. Guthrie, "Generalizing effects of frontostriatal structural connectivity on self-esteem using predictive modeling," *Cortex*, vol. 146, pp. 66–73, 2022.

[19] R. Sandeepkumar, S. Rajendran, R. Mohan, and A. Pascoal, "A unified ship manoeuvring model with a nonlinear model predictive controller for path following in regular waves," *Ocean Engineering*, vol. 243, Article ID 110165, 2022.

[20] J. Casanelles-Abella, Y. Chauvier, F. Zellweger et al., "Applying predictive models to study the ecological properties of urban ecosystems: a case study in Zürich, Switzerland," *Landscape and Urban Planning*, vol. 214, Article ID 104137, 2021.

[21] X. Xu, Q. Zhang, and X. Chen, "Consensus-based non-cooperative behaviors management in large-group emergency decision-making considering experts' trust relations and preference risks," *Knowledge-Based Systems*, vol. 190, Article ID 105108, 2020.

[22] Y. J. Lee and E. M. Anderman, "Profiles of perfectionism and their relations to educational outcomes in college students: the moderating role of achievement goals," *Learning and Individual Differences*, vol. 77, Article ID 101813, 2020.

[23] D. Hadar-Shoval, M. Alon-Tirosh, and H. Morag, "Social relations between students from two groups in conflict: differences in stereotypes and perceived social distance between Jewish and Arab nursing students," *Nurse Education Today*, vol. 78, pp. 5–9, 2019.

[24] J. C. Rodil, A. Meca, K. K. Allison et al., "Measurement invariance testing for the United States identity scale (USIS) across non-hispanic black and white college students," *International Journal of Intercultural Relations*, vol. 86, pp. 134–144, 2022.

[25] J. Zhou, H. Scott, E. S. Huebner, and L. Tian, "Co-developmental trajectories of psychological need satisfactions at school: relations to mental health and academic functioning in Chinese elementary school students," *Learning and Instruction*, vol. 74, Article ID 101465, 2021.

[26] C. S. Song, C. Xu, E. A. Maloney et al., "Longitudinal relations between young students' feelings about mathematics and arithmetic performance," *Cognitive Development*, vol. 59, Article ID 101078, 2021.

[27] S. Du, Z. Chen, H. Wu, Y. Tang, and Y. Q. Li, "Image recommendation algorithm combined with deep neural network designed for social networks," *Complexity*, vol. 9, Article ID 5196190, 2021.