IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 13, 2020, accepted December 18, 2020, date of publication December 22, 2020, date of current version January 4, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3046526

Research on the Link Prediction Model of Dynamic Multiplex Social Network Based on Improved Graph Representation Learning

TIANYU XIA, YIJUN GU, AND DECHUN YIN
School of Information Network Security, People's Public Security University of China, Beijing 100038, China

Corresponding author: Yijun Gu (guyijun@ppsuc.edu.cn)

This work was supported by the CCF-NSFOCUS Research Foundation under Grant CCF-NSFOCUS 2020011.

ABSTRACT In the natural and social systems of the real world, various network can be seen everywhere. The world where people live can be seen as a combination of network with different dimensions. Link prediction formalizes the interaction behavior between people. Traditional link prediction methods mainly study the user behavior of static social network. This article studied the dynamic graph representation learning so as to put forward an improved link prediction model in dynamic social network. Besides, the interactions in the real world can be multiple, links at different moments may have different meanings. The proposed model firstly solved the problem of link prediction on multiple kinds of edges. The whole embedding of each node is separated into two parts, basic embedding and edge embedding. Then the proposed model selected time slices for dynamic social network to get the graph embeddings in different snapshots. What's more, the $t + 1$ time step embedding vector was used to validate t time step prediction effect and the proposed model performed better than traditional graph representation learning methods.

INDEX TERMS Link prediction, dynamic social network, graph representation learning.

I. INTRODUCTION

As a complex network data analysis tool, link prediction can be used to process and mine various types of network information, such as assisting scientists in conducting biological protein structure analysis experiments to discover the interactions between different amino acids [1], helping merchants in product recommendation systems to recommend products and services to potential customers [2], assisting data engineers in data processing to retain hidden links and clean up false links [3]. At the same time, link prediction can be used as a criminal investigation tool for network mining and analysis of criminal suspects [4].

Traditional link prediction methods mainly rely on node label information, network topology, machine learning models and maximum likelihood models [5]. Lin *et al.* [6] defined the similarity between nodes based on node attributes, and used classification algorithms for link prediction. Jure *et al.* [7] conducted simulation experiments on a data set containing 1.3 billion undirected edges, and found that people with similar ages, languages, and geographic locations were more likely to be connected. Lü [8] summarized the link prediction based on network topology information. Hasan *et al.* [9] integrated feature classification methods into link prediction, extracted research keywords in the collaboration network of scientists as features, and used common machine learning classification algorithms (Decision Tree, K-Nearest Neighbor Method, Support Vector Machine, etc.) to predict missing links. Yuan *et al.* [10] integrated the probabilistic graph model method into link prediction, extracted the emotional information characteristics of Twitter users, and judged whether two users will form a friend relationship in the future. Kunegis *et al.* [11] integrated the matrix factorization method into link prediction, and combined the algebraic spectrum transformation of the network adjacency matrix to predict the link weight. Clauset *et al.* [12] proposed a method based on maximum likelihood to predict the possible future link relationships in the network. However, these link prediction methods have achieved good experimental results in static network, but they lack flexibility and timeliness when dealing with dynamic network.

In a real world, the network tends to change dynamically. For example, at time step $t = 0$, user A and user B do not know each other, at time step $t = 1$, A and B discuss work

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

at the meeting and establish contact. Moreover, at time step $t = 2$, A and B cooperate closely and publish the research results of the article. For static social network, traditional link prediction methods have achieved satisfied prediction results. However, for dynamic network in which network nodes and links are changing, traditional link prediction methods are not able to capture dynamic characteristics [13]. Therefore, this article introduces an improved link prediction model, which is to take snapshots at equal intervals according to the time sequence, and then conducts graph representation learning based on these network snapshots to obtain the vector representation of nodes at different time steps. Besides, the time snapshot at time step $t + 1$ is used to evaluate the link prediction effect at time t.

Moreover, since social network in reality are mostly multiplex network, links at different moments may have different meanings, such as mobile phone calls with family members, instant chats with friends, email exchanges with colleagues, and face-to-face shopping with businesses. The communication method further refines the tightness of the link. This article proposes the DGATNE model for representation learning of dynamic multiplex network, and compares it with traditional graph embedding algorithms. Both experiments and theories prove that this method supports the modeling and analysis of dynamic multiplex social network, and the model performs better than traditional algorithms in AUC (area under the receiver operating characteristic curve) [14] and F-measure [15] evaluation.

The main contributions of the article include three aspects. Firstly, the model takes into account the dynamic changes of nodes in social network. Secondly, each node may have many different types of relationships, and the model perform embedding representation learning for each node under each edge type, and then use self-attention mechanism to model the relationship between different types of edge embeddings. Thirdly, the model is scalable and can be applied to large graphs.

The following sections are organized as follows. Section II: This section showed the structure of the representation algorithm and introduced the dynamic process of the algorithm. Section III: This section introduced the graph embedding process of dynamic multiplex social network. Section IV: This chapter uses real social network dataset of twitter as a case for evaluation experiment.

II. REALATED WORK

A. CLASSIC LINK PRECTION METHOD IN SOCIAL NETWORK

The development of science and technology has made people's work and life more and more convenient, and the communication between people has become more and more efficient. Correspondingly, social network have become more compact and complex, which inspires people to explore the relationship between nodes. Link prediction methods based on node attributes are generally simple and easy to implement. However, in actual link prediction, the problem of collecting user privacy cannot be solved easily. Therefore, classic link prediction methods are mainly based on the similarity of social network structures. These methods do not need to consider the attribute information of the node itself, but only need to consider the local network structure information where the node is located. The effect of link prediction is determined by the specifically defined similarity index. Suppose a node $v(x)$ in a complex network, and its neighbor set is represented by $\Gamma(x)$, the degree of the node is represented by $k(x)$, and the similarity of node $v(x)$ and node $v(y)$ is represented by S_{xy}.

- Common Neighbors (CN): The main idea is that if two nodes have many directly connected common neighbors, they are considered to be similar.

$$S_{\mathrm{xy}} = |\Gamma(\mathrm{x}) \cap \Gamma(\mathrm{y})| \tag{1}$$

- Salton: The Salton index combines the definition of the cosine formula and standardizes the CN index. The direct common neighbors between two nodes is used as the numerator, and the product of the two nodes' degrees is squared as the denominator.

$$S_{\mathrm{xy}} = \frac{|\Gamma(\mathrm{x}) \cap \Gamma(\mathrm{y})|}{\sqrt{k(x) \times k(y)}} \tag{2}$$

- Jaccard: The Jaccard index is a normalized link prediction method. Take the direct common neighbors between two nodes as the numerator, and the number of all the neighbors of the two nodes as the denominator.

$$S_{\mathrm{xy}} = \frac{|\Gamma(\mathrm{x}) \cap \Gamma(\mathrm{y})|}{|\Gamma(\mathrm{x}) \cup \Gamma(\mathrm{y})|} \tag{3}$$

- Sorenson: The Sorenson index is a research method of ecological statistics. The direct common neighbor of two nodes is used as the numerator, and the sum of the degrees of the two nodes is half as the denominator.

$$S_{\mathrm{xy}} = \frac{2\,|\Gamma(\mathrm{x}) \cap \Gamma(\mathrm{y})|}{k(x) + k(y)} \tag{4}$$

- Hub Promoted Index (HPI): The HPI index believes that nodes which have more degrees in the network are more likely to establish connections with others. The direct common neighbor between two nodes is used as the numerator, and the degree of the two nodes is compared and the minimum value is used as the denominator.

$$S_{\mathrm{xy}} = \frac{|\Gamma(\mathrm{x}) \cap \Gamma(\mathrm{y})|}{\min\{k(x), k(y)\}} \tag{5}$$

- Hub Depressed Index (HDI): The HDI index believes that nodes which have less degrees in the network are more likely to establish connections with others. The direct common neighbor between two nodes is used as the numerator, and the degree of the two nodes is compared and the maximum value is used as the denominator.

$$S_{\mathrm{xy}} = \frac{|\Gamma(\mathrm{x}) \cap \Gamma(\mathrm{y})|}{\max\{k(x), k(y)\}} \tag{6}$$

- Leicht–Holme–Newman Index (LHN-I): The LHN-I index believes that two nodes with more common neighbors are easier to establish social relationships. Take the direct common neighbors between two nodes as the numerator, and the product of the degrees of the two nodes as the denominator.

$$S_{xy} = \frac{|\Gamma(x) \cap \Gamma(y)|}{k(x) \times k(y)} \tag{7}$$

- Preferential Attachment (PA): The PA index considers that the possibility of generating edge connections between nodes which have more degrees in the network is greater.

$$S_{xy} = k(x) \times k(y) \tag{8}$$

- Adamic-Adar (AA): The AA index considers that any node in the network is more likely to be linked to a node with a small peripheral degree.

$$S_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{1}{\lg k(z)} \tag{9}$$

- Resource Allocation (RA): The RA index is considered from the allocation process of resources in the network. And the resources between the two nodes are passed through common neighbors. The similarity between the two nodes depends on the number of resources passed.

$$S_{xy} = \sum_{z \in |\Gamma(x) \cap \Gamma(y)|} \frac{1}{k(z)} \tag{10}$$

B. LINK PRECTION METHOD IN DYNAMIC SOCIAL NETWORK

The real world is mostly a dynamic network, and its edges and nodes are constantly changing over time. For example, the joining of new users and the generation of new friendships in social network will cause new nodes and connections to appear in the network. The time-series information is an important part of the dynamic network, and is the embodiment of the network's evolution mechanism and its dynamics.

Liben-Nowell and Kleinberg [16] divided the dynamic network into two time periods. The edge set in the first time period was used as the training dataset, and the edge set in the second time period was used as the test dataset to verify the prediction effect of the experiment. Sharan and Neville [17] used a weighted static network graph to represent the dynamic graph, and then added the weight of links to the Bayesian classifier for link prediction. Huang and Lin [18] considered a series of time snapshots, and used the structural dependence of internal links and the time dependence of internal links for link prediction. Tylenda *et al.* [19] considered the influence of the current network status and historical network status information, and proved that the method of combining historical timestamp information can improve the link prediction effect. Soares and Prudêncio [20] considered the time series of each pair of unconnected nodes in the network, and then deployed a prediction model on these time series for link prediction.

Taking time-series snapshots is an important processing method for dynamic network. In order to reduce the loss of original network information during division and make the division results easier to handle, it is very important to determine a time snapshot division method with an appropriate size. If the time snapshot is divided too large, the important time-series dynamics and potential interaction structure of the network will be masked or smoothed; but if the divided time snapshot is smaller, the corresponding time series will generate a lot of noise, the average connection in time may be lost.

C. LINK PREDCTION METHOD IN MULTIPLEX SOCIAL NETWORK

In traditional research on complex network, there are no types of edge relationships between nodes. It is believed that as long as there is a connection between nodes, there is an interaction between the two nodes. The network structure formed by this situation is called a single-dimensional complex network. However, in actual situations, in addition to the characteristic that the topology of the real network changes dynamically over time, the connections between its nodes are also complex and multi-dimensional. At this time, it is believed that there may be more than one node between two nodes in the network. By the way, taking the social network as an example, in a small social group, two people may have a relationship between classmates, friends, or colleagues. Every kind of interpersonal relationship corresponds to a dimension, and multiple dimensions are superimposed together to form a multi-dimensional complex network. The analysis of a single dimension may produce wrong analysis results due to the deviation of the observation angle. In actual network, multi-dimensional complex network has more practical significance than single-dimensional complex network. Therefore, it is necessary to introduce more dimensional information in network analysis.

Currently, there are three different methods to deal with the multidimensional structure of complex network. Method 1: Project the different dimensions of the complex network into one dimension to form an aggregated simple network, and calculate the metric on the aggregated network [21], [22]. Method 2: Treat the evaluation indicators independently in each dimension, and then summarize these results in a variety of ways [23], [24]. Method 3: By considering the influence of the interdependence between the dimensions, the measurement calculation in the multi-dimensional network is directly performed [25]–[27].

D. GRPAH REPRESENTATION LEARNING

Graph representation learning, also known as graph embedding, is an important tool for social network analysis and can be applied to link prediction. The inspiration of network representation learning comes from word2vec. Perozzi *et al.* [28] first proposed the DeepWalk algorithm in 2014, which regarded the connection relationship in the graph as a sentence, and introduced the word embedding method into the graph embedding. Tang *et al.* [29] proposed the LINE (Large-scale Information Network Embedding) algorithm in 2015. The LINE algorithm advocated that nodes with the same neighbors, even if they were not directly connected, may have a certain similarity. Wang *et al.* [30] proposed the SDNE (Structural Deep Network Embedding) algorithm in 2016 to capture the high-order linear characteristics of the network structure and use the extracted features for the first-order representation learning reconstruction. Jure *et al.* [31] also proposed the Node2vec algorithm in 2016. Node2vec proposed a random walk strategy and find a better sentence that included node similarity and structural similarity. Ribeiro *et al.* [32] proposed the struc2vec algorithm in 2017, which fully considered the structural similarity. Zhang *et al.* [33] proposed the ProNE algorithm in 2019. The ProNE algorithm decomposed sparse matrix to generate fast graph embedding representation and used spectral propagation as a lifting method for graph embedding.

In simple terms, network representation learning is to express the nodes in the network with a low-dimensional dense vector space through related algorithms (where the dimension of the vector space is much smaller than the total number of nodes), and can maintain the relevant structure of the original network and features. The network representation learning algorithm mainly includes calculation using matrix eigenvectors, calculation using simple neural network, calculation using matrix decomposition and calculation using deep neural network.

In general, network representation learning has the following characteristics: flexible applicability, large-scale scalability and time continuity [34]. Network representation learning can adapt to the evolving real network [35]. And the network embedding algorithm should be able to handle large-scale network in a short time. Besides, the learned vector representation can be represented in space, and the proximity of the spatial distance means that the real network is closely connected.

III. THE LINK PREDICTION MODEL OF DYNAMIC MULTIPLEX SOCIAL NETWORK

A. DEFINITION OF DYNAMIC MULTIPLEX SOCIAL NETWORK

The real world is mostly a dynamic network, in which the edges and nodes are constantly changing over time [36]. For example, the joining of new users and the generation of new friendships in social network will cause new nodes and connections to appear in the network. This time sequence information is an important part of the dynamic network, and is the embodiment of the network's evolution mechanism and its dynamics.

As shown in Figure 1, if we want to predict the possibility of node 1 and node 4 connecting in the future, we can apply the static link prediction method CN indicator on network G, it is found that 1 and 4 have two common neighbors. However, the network G is in a dynamic change process.

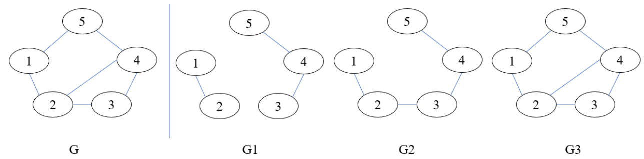

FIGURE 1. Dynamic evolution process of social network.

In the G1 network at time step t_1, node 1 and node 4 have no common neighbors; in the G2 network at time step t_2, node 1 and node 4 have indirectly connected second-order common neighbors; in the G3 network at time step t_3, node 1 and node 4 have one directly connected first-order common neighbor.

Cyber space is an inseparable part of the current modern society, and online social network greatly affect people's study, work and life. Moreover, social network are often multiplex network so there may be multiple links between two users about different types of interactions in the same network [37]. As shown in Figure 2, People on the same social network use different tools to communicate, use WeChat to send and receive general notifications, use email to transfer files, and use twitter to share interesting moments in life. Moreover, the network embedding vector of the same node on different types of edges is not completely independent, but mutual influence [38].

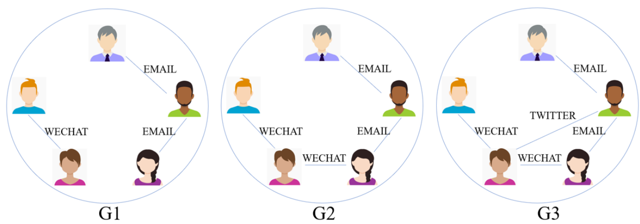

FIGURE 2. Multiplex network at different time steps.

Denote a set of vertices $V = \{v_1, v_2, \cdots, v_m\}$ and the set of undirected edges among these users $E = \{e_{pq}\}$, where each edge e_{pq} represents a type of interaction between v_p and v_q. Consider a series of dynamic multiplex network snapshots $\{G^1, G^2, \cdots, G^t\}$, where $G^t = (V, E^t)$, $E^t = (\cup_{c\in C} E^t_c)$. G^t represents how vertices are connected at time step t, and E^t_c consists of all communication types $c \in C$, and $C > 1$. Notations are summarized in Table 1.

B. THE CONSTRUCTION OF IMPROVED LINK PREDICTION MODEL DGATNE

In this section, we present a model, DGATNE, which is capable of learning representations in dynamic multiplex social network. More specifically, in DGATNE, we divide time steps into snapshots, and at time step t, we split the embedding vector of node v_p into two parts: base embedding and edge embedding. The base embedding is shared between different edge types. And the edge embedding $\boldsymbol{u}^{t,k}_{p,c} \in \mathbf{R}^s$ where k $(1 \leq k \leq K)$ is the dimension of edge embedding. Then, we will introduce the DGATNE model in three parts,

TABLE 1. Description for Notation in Dynamic Multiplex Network.

Notation	Description
G	The input social network
V	The node set of G
E	The edge set of G
N	The neighborhood set of a node on an edge type
n	The number of nodes
m	The number of edge types
c	An edge type
d	The dimension of base/overall embeddings
s	The dimension of edge embeddings
v	A node in the social network
b	The base embedding of a node
u	The edge embedding of a node
a	The self-attention coefficient
x	The node set of G

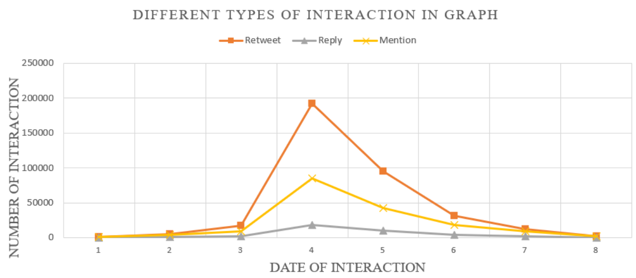

FIGURE 3. The dynamic change in multiplex twitter network.

aggregation [39], concatenation [40] and self-attention mechanism [41].

1) AGGREGATION

As shown in figure3, at time step t, node v_p on edge type c is aggregated from neighbors' edge embedding as following:

$$u_{p,c}^{t,k} = aggregator(\{u_{p,c}^{t,k-1}, \forall v_q \in N_{p,c}\}) \tag{11}$$

In equation (11), $N_{p,c}$ is the neighbors node v_p on edge type c, and the initial edge embedding $u_{p,c}^{t,0}$ is randomly initialized. Apply mean function of GraphSAGE [42] in equation (11), equation (12) is as following:

$$u_{p,c}^{t,k} = \sigma(\mathbf{W}^{t,k} \cdot \text{mean}(\{u_{p,c}^{t,k-1}, \forall v_q \in N_{p,c}\})) \tag{12}$$

2) CONCATENATION

Consider all the embeddings on different types of edges for v_p at time step t, and concatenate the embeddings as U_p^t, equation (13) is as following:

$$U_p^t = (u_{p1}^t, u_{p2}^t, \cdots, u_{pm}^t) \tag{13}$$

3) SELF-ATTENTION MECHANISM

In order to clarify the relationship between the representation of each node under each edge type, self-attention mechanism is introduced, equation (14) is as following:

$$a_{p,c}^t = \text{softmax}((x_c^t)^T \tanh(X_p^t U_p^t))^T \tag{14}$$

In equation (14), x_c and X_c are trainable parameters for edge type c, $a_{p,c}^t$ is the calculated attention weight vector, consider the base embedding and edge embedding of v_p at time step t, the overall embedding of v_p for edge type c is as equation (15) following:

$$v_{p,c}^t = b_p^t + \alpha_c (M_c^t)^T U_p^t a_{p,c}^t \tag{15}$$

In equation (15), b_p^t is the base embedding of v_p at time step t, α_c is a hyper-parameter denoting the importance of edge embeddings and M_c^t is the transformation matrix.

The main flow of DGATNE is as follows:

Algorithm 1 DGATNE

Input: Dynamic network G^1, $G^2, \cdots, G^t$, embedding dimension d, edge type c

Output: Proximity of nodes $v_1, v_2, \cdots, v_m$

```
Initialize all the model parameters
Sample E^t from all existing edges in G^t
Generate random walks on each edge type c
for t ← 1 to T do
    Sample base embeddings b_p^t
    Sample edge embeddings u_p^t on edge type c
    Concatenate all the types of edge embeddings
    Compute the self-attention coefficient
    Calculate the proximity of nodes
    G^t ← G^{t+1}
end
```

C. THE EVALUATION INDEX OF LINK PREDICTION MODEL

Traditional link prediction methods often select test dataset and training dataset randomly, and repeat experiments to evaluate the effect of link prediction. However, the real world is mostly a dynamic network, and the edges and nodes are constantly changing over time [43]. Therefore, the proposed DGATNE model takes snapshots of the dynamic network at equal intervals in time series. Then the proposed model uses real data at time step t as the training dataset and real data at time step $t+1$ as test dataset, and uses AUC and F-measure indicators for evaluation.

AUC (Area Under Curve) [14] is one of the main offline evaluation indicators used by the two-class model. In this article, AUC refers to the probability that an edge selected in the test dataset gets a higher score than the edge selected randomly in the non-existent edge dataset. The AUC value can be calculated as follows: each time an edge is randomly selected from the test set and the set without edges for comparison, if the score value of the selected edge of the test set is greater than the score value of the edge in the set without edges, we get 1 point; if the two points are equal, we get 0.5 point. Repeat the independent evaluation experiments n times. Assuming that there are n' times where the test dataset selects a large side score, and there are n'' times where the two scores are equal, the AUC calculation is as following:

$$AUC = \frac{n' + 0.5n''}{n} \tag{16}$$

F-measure [15] considers both Precision [44] and Recall [45]. Precision refers to the difference between the average value and the known true value in each independent experiment. Recall is a measure of coverage. Based on the previous precision and recall, the F-measure is calculated as following:

$$F\text{-}measure = \frac{2 \times Recall \times Precision}{n} \tag{17}$$

IV. DATASET

In this section, we employ the Higgs twitter dataset [46] to evaluate the effectiveness of the proposed model, DGATNE. The Higgs dataset has been built after monitoring the spreading processes on Twitter before, during and after the announcement of the discovery of a new particle with the features of the elusive Higgs boson on 4th July 2012. The Higgs twitter dataset is a dynamic multiplex network, which consists of more than 14 million interactions between 456,626 users over 8 days. Besides, interactions can be retweet, reply and mention relationship. As shown in Table 2, retweet network, reply network and mention network are included in this dataset.

TABLE 2. Statistics of Datasets.

	Nodes	Edges	Days
Social	456626	1485584	8
Retweet	256491	328132	8
Reply	38918	32523	8
Mention	116408	150818	8

In order to better show the relationship between retweet, reply and mention network dynamically, figure 3 is shown.

As figure 3 is shown, we can summarize the events before and after the discovery of the boson, dividing them into 4 different periods:

- Period I: In the first two days, there were some rumors about the discovery of a Higgs-like boson at Tevatron;
- Period II: In the second two days, scientists from CDF and Dzero experiments, based at Tevatron, found that the Higgs particle should have a mass between 115 and 135 GeV/c^2. And there were many rumors about the Higgs boson discovery at Large Hadron Collider;
- Period III: In the third two days, popular media covered the event, and people are very interested in the Higgs boson, which can explain the mystery of the mass of matter.
- Period IV: In the fourth two days, the number of people paying attention to the Higgs boson event was gradually decreasing, and the popularity of the event decreased.

V. EXPERIMENTATION

A. BASELINE METHODS

The link prediction method based on common neighbors stores the relationship between users in the adjacency matrix for calculation. This method can handle smaller social network data sets, but when the user node counts at 100,000, the calculation space will reach Ten billion. The link prediction method based on network representation learning stores the relationship between nodes in the form of vectors, so it can handle large-scale complex network dataset. In this section, we compare the proposed model with classical link prediction methods based on graph representation learning.

- DeepWalk: Inspired by word2vec [47] in NLP(Nature Language Processing) [48], the DeepWalk method builds a series of walks through random walks to capture the network topology. Besides, the DeepWalk method proposes a statistical result of similar node co-occurrence frequency and vocabulary co-occurrence frequency [49]. In other words, the sequence of wandering can be analogous to the sentences in the corpus, and the nodes in the sequence can be analogous to the words in the sentence, and the co-occurrence of vocabulary is similar to the co-occurrence of nodes in the sequence of walking.
- LINE: By designing an objective function, the LINE method retains local and global structural information. Considering both first-order similarity and second-order similarity, LINE adopts the edge-sampling method, that is, the edge weight is regarded as the probability of the appearance of the edge, and the sampled edge is regarded as a binary edge through weight sampling to update the model, so as to ensure that the objective function is unchanged and the weight coefficient does not affect the gradient.
- Node2vec: Compared with DeepWalk and LINE methods that both focus on node similarity, Node2vec advocates that if similar structures exist in neighbors, embedding should also be similar. In other words, assuming that there are two nodes in different communities, if the two nodes have the same structure and play similar roles in their respective communities, their embeddings should be similar. The wandering strategy of node2vec uses two parameters to control the BFS (Breath First Search) [50] and DFS (Depth First Search) [51], and dig out the characteristics of the node in the network.
- ProNE: ProNE further integrates high-level image information into the image embedding based on the spectrum propagation strategy. Firstly, the definition of graph embedding representation is reduced to a sparse matrix factorization definition. Then, the high-order Cheeger's inequality [52] in Riemannian Geometry is used to modulate the spectral space of the graph, and the embedding representation learned in the first step is spread on the adjusted graph, so as to make the localized smoothing information [53] and the global clustering information [54] integrated into the graph representation learning.

B. THE PROPOSED DGATNE MODEL

As shown in Figure 4, we propose a link prediction model in dynamic multiplex social network, and the model is meaningful to solve practical problems in the real world. As traditional link prediction methods mainly regard the network as a static network and ignore the time information and evolution

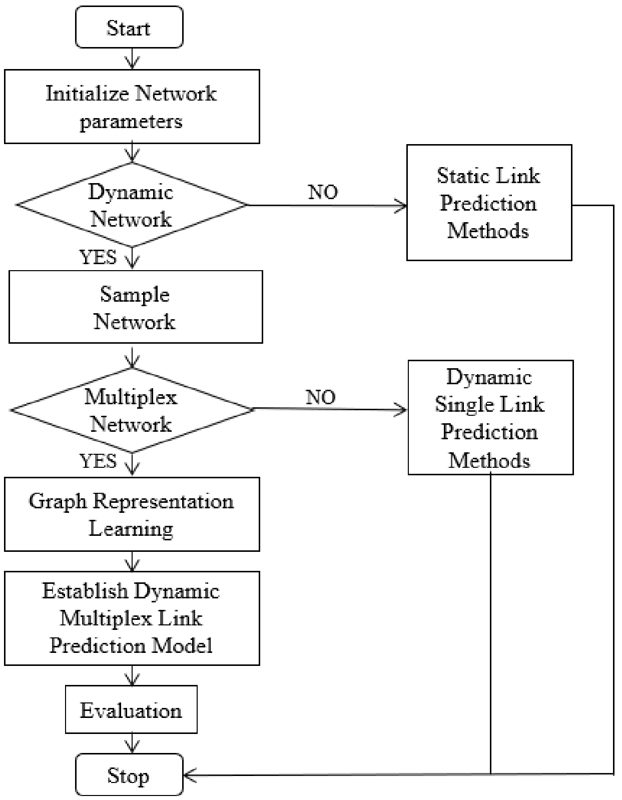

FIGURE 4. The flowchart of dynamic multiplex link prediction model.

process of the network. The proposed DGATNE model takes user's historical interaction information and network structure information into account. Besides, considering the links between people in the real world are usually multi-dimensional, the proposed model combines self-attention mechanism with the improved graph representation learning algorithm so as to achieve more accurate link prediction results.

The proposed DGATNE model considers the influence of network structure on network evolution from both the micro and macro perspectives, and the effectiveness of the proposed model is proved through experimentations. From a macro perspective, we find that nodes with similar behaviors in history would behave similarly in the future. That is to say, an improved graph representation learning method can be used to encode the historical behavior of nodes, so as to achieve link prediction in dynamic multiplex social network. From a microscopic point of view, the dynamic changes of the network structure can be represented by dynamic triads, and the generation, continuation and disappearance of links are affected by the influence of surrounding nodes.

Different from the traditional link prediction experiments, which randomly divides the training dataset and test dataset, this article uses the time snapshot at time step $t + 1$ to verify the link prediction effect at time step t.

Firstly, according to the spread process of the event period, select two days as the event slice, and the experimental results are shown in Table 3.

TABLE 3. Model Performance in Event Snapshots.

	AUC	F-Measure
DeepWalk	0.8055	0.6888
LINE	0.7679	0.6623
Node2vec	0.8203	0.6989
ProNE	0.7318	0.6708
DGATNE	0.8233	0.7010

According to the results that are shown in Table 3, the proposed DGATNE model performs better in AUC and F-Measure evaluation experiments than the baseline methods. Moreover, DGATNE and Node2vec methods perform better than LINE and ProNE methods in AUC evaluation experiments. And DGATNE model is slightly better than DeepWalk and Node2vec methods in F-Measure evaluation experiments.

Secondly, according to people's work behavior every day, one day is selected as the time slice. The experimental results are shown in Table 4.

TABLE 4. Model Performance in Time Snapshots.

	AUC	F-Measure
DeepWalk	0.7779	0.6640
LINE	0.7351	0.6678
Node2vec	0.7817	0.6734
ProNE	0.7408	0.6613
DGATNE	0.8023	0.6777

According to the results that are shown in Table 4, the proposed DGATNE model performs better in AUC and F-Measure evaluation experiments than the baseline methods. Moreover, DGATNE and Node2vec methods perform better than LINE and ProNE methods in AUC evaluation experiments. And DGATNE model is slightly better than DeepWalk and Node2vec methods in F-Measure evaluation experiments.

VI. CONCLUSION

In the beginning, this article analyzed the existing link prediction models and pointed out their two shortcomings. Firstly, the traditional link prediction models mainly focused on static social network, ignoring the dynamic changes in social network. Secondly, the interactions in the real world could be multiple, links at different moments may have different meanings, such as mobile phone calls with family members, instant chat with friends, email exchanges with colleagues, and face-to-face shopping with businesses. Considering the above problems, this study firstly used DGATNE model combined with self-attention mechanism in constructing the dynamic multiplex network. And, the traditional link prediction methods were compared with DGATNE model in this article. On the one hand, the model realized the research of link prediction model of dynamic multiplex social network. On the other hand, the model performed better than the traditional graph representation learning methods in large-scale real social network dataset. In the next study of link prediction, we will combine node attribute information and link generation algorithms to further characterize the dynamic network from wider dimensions to achieve better link prediction results.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their suggestions which helped in improving the quality of the article.

REFERENCES

[1] H. Mamitsuka, "Mining from protein-protein interactions," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 2, no. 5, pp. 400–410, Sep. 2012.

[2] L. Zhang, J. Li, Q. Zhang, F. Meng, and W. Teng, "Domain knowledge-based link prediction in customer-product bipartite graph for product recommendation," *Int. J. Inf. Technol. Decis. Making*, vol. 18, no. 1, pp. 311–338, Jan. 2019.

[3] I. Ahmad, M. U. Akhtar, S. Noor, and A. Shahnaz, "Missing link prediction using common neighbor and centrality based parameterized algorithm," *Sci. Rep.*, vol. 10, no. 1, pp. 167–256, Dec. 2020.

[4] F. Calderoni, F. Calderoni, N. Parolini, M. Verani, and C. Piccardi, "Robust link prediction in criminal networks: A case study of the Sicilian Mafia," *Expert Syst. Appl.*, vol. 161, no. 15, p. 11366, 2020.

[5] T. Zhou, L. Lü, and Y.-C. Zhang, "Predicting missing links via local information," *Eur. Phys. J. B*, vol. 71, no. 4, pp. 623–630, Oct. 2009.

[6] L. Cazzanti and M. Gupta, "Information-theoretic and set-theoretic similarity," in *Proc. IEEE Int. Symp. Inf. Theory*, San Francisco, CA, USA, Jul. 2006, pp. 296–304.

[7] J. Leskovec and E. Horvitz, "Planetary-scale views on a large instant-messaging network," in *Proc. 17th Int. Conf. World Wide Web (WWW)*, Beijing, China, 2008, pp. 915–924.

[8] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Phys. A, Stat. Mech. Appl.*, vol. 390, no. 6, pp. 1150–1170, 2011.

[9] M. Hasan, V. Chaoji, S. Salem, and M. Zaki, "Link prediction using supervised learning," in *Proc. ICDM*, Bethesda, MD, USA, 2006, pp. 798–805.

[10] G. Yuan, P. K. Murukannaiah, Z. Zhang, and M. P. Singh, "Exploiting sentiment homophily for link prediction," in *Proc. 8th ACM Conf. Recommender Syst. (RecSys)*, Silicon Valley, CA, USA, 2014, pp. 17–24.

[11] J. Kunegis and A. Lommatzsch, "Learning spectral graph transformations for link prediction," in *Proc. 26th Annu. Int. Conf. Mach. Learn. (ICML)*, Montreal, QC, Canada, 2009, pp. 561–568.

[12] A. Clauset, C. Moore, and M. E. J. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, no. 7191, pp. 98–101, May 2008.

[13] W. Li, Y. Gu, D. Yin, T. Xia, and J. Wang, "Research on the community number evolution model of public opinion based on stochastic competitive learning," *IEEE Access*, vol. 8, pp. 46267–46277, 2020.

[14] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, p. 29, 1982.

[15] G. Hripcsak, "Agreement, the F-measure, and reliability in information retrieval," *J. Amer. Med. Inform. Assoc.*, vol. 12, no. 3, pp. 296–298, Jan. 2005.

[16] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 58, no. 7, pp. 1019–1031, 2007.

[17] U. Sharan and J. Neville, "Temporal-relational classifiers for prediction in evolving domains," in *Proc. 8th IEEE Int. Conf. Data Mining*, Pisa, Italy, Dec. 2008, pp. 540–549.

[18] Z. Huang and D. K. J. Lin, "The time-series link prediction problem with applications in communication surveillance," *Informs J. Comput.*, vol. 21, no. 2, pp. 286–303, May 2009.

[19] D. M. Dunlavy, T. G. Kolda, and A. Evrim, "Temporal link prediction using matrix and tensor factorizations," in *Proc. ACM Trans. Knowl. Discovery Data (TKDD)*, New York, NA, USA, 2011, pp. 1–27.

[20] P. R. da Silva Soares and R. B. C. Prudencio, "Time series based link prediction," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Brisbane, QLD, Australia, Jun. 2012, pp. 1–7.

[21] F. Battiston, V. Nicosia, and V. Latora, "Structural measures for multiplex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 89, no. 3, Mar. 2014, Art. no. 032804.

[22] M. De Domenico, A. Solé-Ribalta, S. Gómez, and A. Arenas, "Navigability of interconnected networks under random failures," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 23, pp. 8351–8356, 2014.

[23] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer network," *J. Complex Netw.*, vol. 2, no. 3, pp. 203–271, 2014.

[24] T. G. Kolda, B. W. Bader, and J. P. Kenny, "Higher-order Web link analysis using multilinear algebra," in *Proc. 5th IEEE Int. Conf. Data Mining (ICDM)*, Houston, TX, USA, Nov. 2005, pp. 242–249.

[25] J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L. M. Floría, "Evolution of cooperation in multiplex networks," *Sci. Rep.*, vol. 2, p. 620, Aug. 2012.

[26] A. Halu, R. J. Mondragón, P. Panzarasa, and G. Bianconi, "Multiplex PageRank," *PLoS ONE*, vol. 8, no. 10, Oct. 2013, Art. no. e78293.

[27] T. Chakraborty and R. Narayanam, "Cross-layer betweenness centrality in multiplex networks with applications," in *Proc. IEEE 32nd Int. Conf. Data Eng. (ICDE)*, Helsinki, Finland, May 2016, pp. 397–408.

[28] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, New York, NY, USA, 2014, pp. 701–710.

[29] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "LINE: Large-scale information network embedding," in *Proc. 24th Int. Conf. World Wide Web*, Florence, Italy, May 2015, pp. 1067–1077.

[30] D. Wang, P. Cui, and W. Zhu, "Structural deep network embedding," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 1225–1234.

[31] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. 22nd ACM Int. Conf. Knowl. Discovery Data Mining (SIGKDD)*, San Francisco, CA, USA, Aug. 2016, pp. 855–864.

[32] L. F. R. Ribeiro, P. H. P. Saverese, and D. R. Figueiredo, "struc2vec: Learning node representations from structural identity," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Halifax, NS, Canada, Aug. 2017, pp. 385–394.

[33] J. Zhang, Y. Dong, Y. Wang, J. Tang, and M. Ding, "ProNE: Fast and Scalable Network Representation Learning," in Proc. 28th Int. Joint Conf. Artif. Intell., Macao, China, Jul. 2019, pp. 4278–4284.

[34] C. Wu, Y. Zhou, L. Tan, and C. Teng, "Link prediction based on graph embedding method in unweighted networks," in *Proc. 39th Chin. Control Conf. (CCC)*, Jul. 2020, pp. 736–741.

[35] D. Zhu, P. Cui, Z. Zhang, J. Pei, and W. Zhu, "High-order proximity preserved embedding for dynamic networks," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 11, pp. 2134–2144, Nov. 2018.

[36] L. Zhu, D. Guo, J. Yin, G. Ver Steeg, and A. Galstyan, "Scalable temporal latent space inference for link prediction in dynamic social networks," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 10, pp. 2765–2777, Oct. 2016.

[37] Y. Cen, X. Zou, J. Zhang, H. Yang, J. Zhou, and J. Tang, "Representation learning for attributed multiplex heterogeneous network," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, Jul. 2019, pp. 1358–1368.

[38] J. Li, L. Wu, R. Guo, C. Liu, and H. Liu, "Multi-level network embedding with boosted low-rank matrix approximation," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Vancouver, BC, Canada, Aug. 2019, pp. 49–56.

[39] H. R. Singh, "A model of computing trust in Web based social network using new aggregation and concatenation operators," *Int. J. Comput. Sci. Netw.*, vol. 2, no. 4, pp. 70–76, 2013.

[40] V. W. Y. Liao, A. S. Kusay, T. Balle, and P. K. Ahring, "Heterologous expression of concatenated nicotinic ACh receptors: Pros and cons of subunit concatenation and recommendations for construct designs," *Brit. J. Pharmacol.*, vol. 177, no. 18, pp. 4275–4295, Sep. 2020.

[41] M. Li, W. Hsu, X. Xie, J. Cong, and W. Gao, "SACNN: Self-attention convolutional neural network for low-dose CT denoising with self-supervised perceptual loss network," *IEEE Trans. Med. Imag.*, vol. 39, no. 7, pp. 2289–2301, Jul. 2020.

[42] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. NIPS*, Red Hook, NY, USA, 2017, pp. 1025–1035.

[43] Z. Zhang, J. Wen, L. Sun, Q. Deng, S. Su, and P. Yao, "Efficient incremental dynamic link prediction algorithms in social network," *Knowl.-Based Syst.*, vol. 132, pp. 226–235, Sep. 2017.

[44] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proc. 23rd Int. Conf. Mach. Learn. (ICML)*, Pittsburgh, PA, USA, 2006, pp. 233–240.

[45] C. De Bacco, E. A. Power, D. B. Larremore, and C. Moore, "Community detection, link prediction, and layer interdependence in multilayer networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 95, no. 4, Apr. 2017, Art. no. 042317.

[46] M. De Domenico, A. Lima, P. Mougel, and M. Musolesi, "The anatomy of a scientific rumor," *Sci. Rep.*, vol. 3, no. 1, p. 2980, Dec. 2013.

[47] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. Adv. Neural Inf. Process. Syst.*, Red Hook, NY, USA, 2013, pp. 3111–3119.

[48] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Doha, Qatar, 2014, pp. 1532–1543.

[49] K. Schouten, O. van der Weijde, F. Frasincar, and R. Dekker, "Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1263–1275, Apr. 2018.

[50] N. Banerjee, S. Chakraborty, V. Raman, and S. R. Satti, "Space efficient linear time algorithms for BFS, DFS and applications," *Theory Comput. Syst.*, vol. 62, no. 8, pp. 1736–1762, Nov. 2018.

[51] L. Cui, G. Li, Q. Lin, Z. Du, W. Gao, J. Chen, and N. Lu, "A novel artificial bee colony algorithm with depth-first search framework and elite-guided search equation," *Inf. Sci.*, vols. 367–368, pp. 1012–1044, Nov. 2016.

[52] L. Massoulie and R. Varloot, "Rapid mixing of dynamic graphs with local evolution rules," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1344–1354, Jul. 2020.

[53] X.-C. Liu, X.-Z. Zhu, H. Tian, Z.-P. Zhang, and W. Wang, "Identifying localized influential spreaders of information spreading," *Phys. A, Stat. Mech. Appl.*, vol. 519, pp. 92–97, Apr. 2019.

[54] P. Zhao and L. Lai, "Analysis of KNN information estimators for smooth distributions," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3798–3826, Jun. 2020.

TIANYU XIA received the B.S. degree from the People's Public Security University of China, in 2019, where he is currently pursuing the master's degree with the School of Information Network Security. His research interests include complex network analysis and data mining.

YIJUN GU received the Ph.D. degree from the School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China. He is currently a Professor with the School of Information Network Security, People's Public Security University of China. His research interests include big data analysis, data mining, cyber security, and social network analysis.

DECHUN YIN received the Ph.D. degree from the School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China. He is currently an Associate Professor with the School of Information Network Security, People's Public Security University of China. His research interests include natural language processing, big data analysis, data mining, and social network analysis.