How are You Affected? A Structural Graph Neural Network Model Predicting Individual Social Influence Status

Jiajie Du[1] and Li Pan[1,2](✉)

[1] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
panli@sjtu.edu.cn
[2] National Engineering Laboratory for Information Content Analysis Technology, Shanghai, China

Abstract. Human's daily life is now inextricably bound with online social networks like Weibo and Twitter, where user's opinions and emotions are delivered and get influenced by each other frequently. Therefore, predicting and understanding such phenomenon of social influence, especially on individuals, becomes crucial for applications like recommendation and advertising. Although recent studies achieve some predictions on individual social influence status (active or inactive), the complexity and diversity of the social structures have not been well solved, and such works also lack a deep understanding of the influence mechanism itself like how individuals are affected. Therefore, a structural graph neural network model (SGN) is proposed to learn the diverse relationships and quantify the propagated influence between users. The SGN model is consists of two special representation modules and some global layers. The two representation modules are based on two major social network structures: friendship structure and influence propagation structure. In the modules, the well-developed graph neural networks, GCN (Graph Convolutional Network) and GAT (Graph Attention Networks), are respectively applied to capture spatial correlations. Besides, the global attention mechanism then helps to quantify the relationships between influencers and influencees. Finally, experiments on two real-world social networks show that the proposed SGN not only outperforms other baselines on prediction metrics, but is also able to reveal the intermediate neighbors who affect the target most.

Keywords: Social influence status · Structure · Graph neural network

1 Introduction

Nowadays, social networks have aroused close and profound interactions between users than ever before. Such a phenomenon that a user's emotions or behaviors are affected by others is referred to as social influence. As social influence becomes ubiquitous and widespread among social networks, the related

H. Gao and X. Wang (Eds.): CollaborateCom 2021, LNICST 407, pp. 401–415, 2021.
https://doi.org/10.1007/978-3-030-92638-0_24

researches cover many areas including political election [1], advertising [9], public sentiment [11], academic collaboration [15] and so on. Therefore, a deep understanding towards the underlying dynamics mechanisms of social influence propagation is extremely valuable.

Previously, extensive works about social influence mainly aim to predict some global patterns like cascade size or the influence dynamics of the whole society [10,21,25]. But the differences between individuals and their personal information are ignored before, which turn out to be extremely important then. So recently, there are more user-level influence prediction works, Qiu et al. [24] achieve an automatic end-to-end approach to discover hidden signals and predict user's social influence actions (retweet). Tang et al. [26] take both friendship and temporal dynamics into consideration, and predict several categories of user behaviors jointly. However, these works do not focus a lot on the diversity and complexity of social influence structure itself, thus makes a need for us to analyze the structures and find some classic structural classifications.

Among the user-level prediction tasks of social influence, a core problem is: Given the local structure and some prior characteristic features of a user, how to predict his social influence status, according to [24]. In fact, due to complex user relationships in social networks [5], the cause of the influence is also diverse. As shown in Fig. 1, two typical causes from relationships and structures are friendship and historic influence propagation. Naturally, your friends in the real world also affect you on social platforms, and the netizens whom you interacted with before may also continue influencing you in retweets, comments or likes. Based on this division of the complex social structure, our exact aim is to predict the target u's influence status, which equals to identify whether he will be active about some social contents in the next timestamp.

Besides, the importance of neighbors in social networks is also our concern, as it could enhance our understanding of social influence. Tang et al. [26] used to quantify different importance between friends. But in our problem, the social structures contain not only friends but also some netizens from previous influence propagation records. And we use *neighbors* to refer to all these users who might have a relationship with the target. Then, a method is proposed to quantify the importance of different structures first and all the neighbors next. In this way, some vital neighbors would finally be identified to help us understand who is exactly influencing the targets.

To solve this user-level problem of prediction and identification, also inspired by the rapid development of graph neural networks, we propose an end-to-end individual social influence status prediction framework called SGN (structural graph neural network). Structural features are mined first based on the two structures mentioned above, and the predictions are accomplished next by using the features, finally, quantitative indicators of importance for structures and all the neighbors are calculated, making it reliable to figure out who affects the target's social influence status most. In particular, the proposed SGN contains: a) a friendship interacting module, b) an influence propagating module, c) global attention mechanism, d) outputs. The friendship interacting module captures

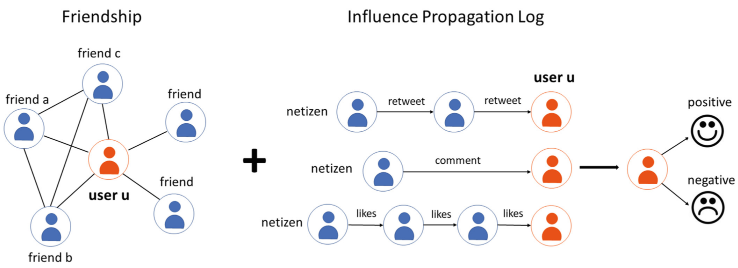


Fig. 1. Two typical relationships and their structures in social network: friendship and historic influence propagation. Friendship refers to the users who are friends in real-world, influence propagation refers to the netizens between whom some interactions have happened online, the interactions include retweets, comments, likes and so on. And the individual social influence status (positive or negative) is always decided by these two structures.

correlations between friends powered by attentioned-GCN, while the influence propagation module chooses multi-head graph attention network (GAT), as GAT is more suitable to the directional and dynamic influence propagation structure [28]. The global attention mechanism is in order to reveal the importance of different structures and aggregate the local attention indicators from the above two modules. Experiments conducted on two real-world social network datasets - Twitter and Weibo demonstrate that the proposed SGN not only outperforms the baselines on individual social influence status prediction, but also helps us identify the neighbors of a target who affect him most. To sum up, our contributions are:

– We present a study about the complexity and diversity of social influence structures and give a typical division: friendship and influence propagation.
– We propose an end-to-end user-level framework SGN, which compute the representation of each structure, predict the individual influence status, and figure out the topk important neighbors.
– Experiments conducted on two real-world datasets indicate that the proposed SGN not only outperform some baselines in prediction task, but also in neighbor importance identification task.

2 Related Work

2.1 Social Influence

Current social influence studies mainly aim to predict global patterns, cascade size is one of the major measurable indicators of social influence [11,34], and influence Maximization is another well-known problem, which David et al. [12] first solve it with famous independent cascade and linear threshold models.

Recently, there have been some efforts to solve these problems using deep learning technologies, e.g., a RNN based cascade prediction model DeepCas [17], a adversarial graph embeddings approach for fair influence maximization [13]. Besides, another major part of studies focuses on the underlying dynamic mechanism of influence propagation, there have been several models trying to explain this mechanism using reinforced [25] or nonhomogeneous [16] poisson processes, neural popularity prediction [2], or heavy tails [20]. However, researches about individual social influence and the explainations are inadequate, although [24] proposed an automatic predictive model, the diversity and homophily of structure raised from [5] is not solved yet.

2.2 Graph Neural Networks

Recently, graph convolutional network (GCN) [14] model and its variants have been widely studied and applied in many areas. GCN simplified from [14] aggregate the node features from their one-hop neighbors, Graph attention network (GAT) [28] introduces attention mechanism and solve the directional graph. GraphSAGE [7] aggregate features in neighborhood with mean/max/LSTM pooling. Graph neural networks are thus widely used in structural representation [30]. Furthermore, combined with LSTM or attention mechanism, these models are now applied in different areas such as disease-gene association identification [8], Traffic Forecasting [32], regional economy prediction [31] and so on.

Based on these contributions of fundamental models, graph neural networks now get evolved in some more specific and complicated areas. The Dynamic Graph Learning is one of the rising patterns [6,18,22], which use dynamic graph solve some time series prediction problems, as traditional GNN models only take static graphs as inputs. So [33] first uses the snapshots of graphs as inputs and demonstrates a ST-GCN model to take time as a dimension in previous CNN layers, and [27] combines the classic neighborhood embedding propagation with innovative self-propagation process in node aggregations, and [6] raised a new sampling method based on time series knowledge. Another interests is in solving some more complicated structures and using a deep model [3,19,23], like [3] introduces a new Generalized PageRank (GPR) GNN architecture to adaptively optimize node feature and topological information extraction and [19] proposes a new supplementary framework (CopulaGNN) which utilizes both representational and correlational information better.

3 Preliminaries

In this section, we define some important notations and problems. First, the *social networks* is a large-scale graph $G_{\text{social}} = (V, E)$ where V represents all the users and $E \subseteq V \times V$ could represent all kinds of the relationships between users. For a specific user u, if there is an undirected edge between him and another user v, they are friends to each other, if there is a directed edge from

u to v, it denotes u may retweet from, comment on or likes v, such directed relationship is defined as *social influence propagation.* For example, if u retweets from v, u also receives social influence from v. Besides, as mentioned in Section I, compared with some global information in the large-scale social network, we mainly focus on individual social influence status.

Definition 1. ***Social Influence Status*** *of a exact user u at timestamp t is defined as a binary status s_t^u : $\{0, 1\}$, where $s_t^u = 1$ indicates that user u is influenced to be active on timestamp t. Active users may have some actions such as retweet, comment or likes in Twitter. But if u is inactive and has no action, $s_t^u = 0$.*

Problem 1. Given the defination of social influence status, the problem of **predicting individual social influence status** is defined as: for a user u, we aim to compute the possibility of his influence status after a time interval Δt based on his current status s_t^u and his related social network subgraph G^u :

$$P\left(s_{t+\Delta t}^u \mid s_t^u, G^u\right).$$

Due to the diverse and complex relationships in social networks, We divide the structure into two major categories.

Definition 2. ***Friendship Structure Graph*** *G_f^u involves target user and his friends, and the edges indicate friendship or following relationship.*

Definition 3. ***Propagation Structure Graph*** *G_p^u is another major structure, which records the propagated users as nodes and retweeting process as directed edges.*

More concretely, given the user related social graph G_t^u, it is firstly divided into friendship structure graph G_f^u and influence propagation structure graph G_p^u, which denote friends and retweeting relationship respectively. The friendship structure graph is a triplet $G_f^u = \left\{V_f^u, E_f^u, H_f^u\right\}$, where $V_f^u = \{u \cup v \mid d(u, v) \leq 2\}$ consist of target user u and his two-hop friends. E_f^u is a set of edges denoting friendship relationship. $H_f^u \in \mathbb{R}^{n*h}$ is the characteristic feature matrix where each vector of h dimensions represents some basic information of a user. Similarly, in the propagation structure $G_p^u = \left\{V_p^u, E_p^u, H_p^u\right\}$, V_p^u denotes the users in influence propagation routines, the set of directed edges E_p^u represents the propagation paths, e.g., retweet, H_p^u demonstrates characteristic features of involved users in this structure.

Through the friendship and influence propagation structures, **the possibility of prediction** could be modified as:

$$P\left(s_{t+\Delta t}^u \mid s_t^u, G_{f,t}^u, G_{p,t}^u\right) \tag{1}$$

However, only predicting social influence status is inadequate for understanding propagation mechanism, thus we introduce the important intermediate neighbors identity: which is identifying the intermediate users who affect target user most.

Problem 2. **(Important Neighbors Identity)** For a target user u, given his social information G_t^u and all of his neighbors in both structures, compute importance indicators of each person I_u, and identify the top k important neighbors who make most difference on the future influence status of target user u.

4 Our Model: SGN

In this section, we will introduce our proposed model SGN for individual social influence prediction, the overall framework of SGN is shown in Fig. 2.

First in our proposed SGN, as mentioned before, to simplify the diversity and complexity of social network structure, we divide it into two structures: friendship structure and propagation structure, representing two major relationships between users respectively. Then inspired by the well-known graph neural network, two structural modules a) and b) are established taking the graphs of friendship and propagation structure as exact inputs and compute the related representations respectively. In the friendship interacting module, we take the different importance of friends into consideration and aggregate the impacts in the locality of target user u through attentioned-GCN. Meanwhile, the influence propagation module computes the representation by multi-head GAT which is more suitable to the dynamic and directional propagation graphs. Then, global attention c) is attached to learn the different importance of two modules and structures. Finally, the comprehensive attention mechanism through the framework help to predict the individual's social influence status and identify the intermediate users who affect the target most d). The details will be discussed in the following text.

4.1 Friendship Interacting Module

For the friendship interacting module, inspired by the rapidly developed graph convolutional networks where plenty of works have been done in graph representation [14,28,29], we choose the common GCN [14] and an attached neighbor attention mechanism [26] to learn structural information, compute representations, and finally figure out user's dependencies.

As shown in the top-left of Fig. 2, for each input of friendship structure graph G_f^u, GCN represents each node into its embedding vector. Concretely, GCN model is consist of several convolutional layers, and the node embedding matrix $H^{(l)}$ get updated using its structural information on each layer as:

$$
\begin{gathered}
H^{(l+1)} = \mathrm{ReLU}\left(\widetilde{D_f}^{-\frac{1}{2}}\widetilde{A_f}\widetilde{D_f}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \\
\widetilde{A_f} = A_f + I_f
\end{gathered}
\tag{2}
$$

where $W^{(l)}$ is a parameter matrix of each layer, $ReLU$ is a nonlinear activation function. $H^{(l)} \in \mathbb{R}^{n*f^{(l)}}$ is the embedding matrix, where n is the number of nodes, $f^{(l)}$ denotes the dimensions of embedding vector on each layer, specially,

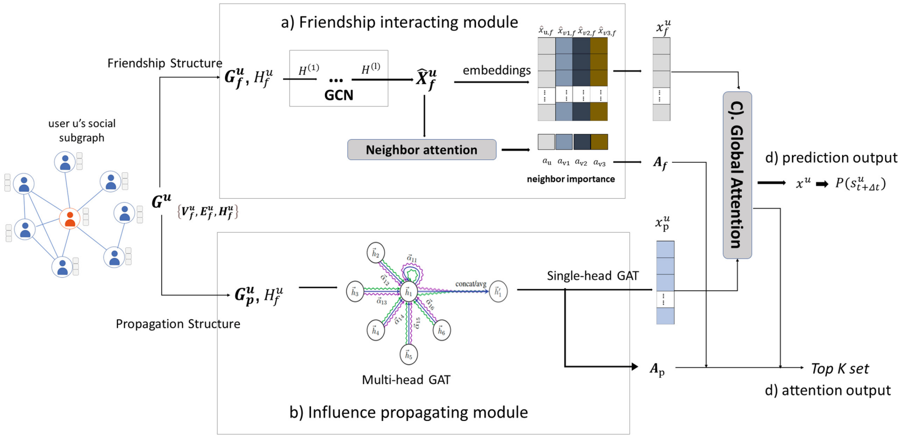


Fig. 2. The framework of Structural Graph Neural Network Model (SGN): each user's social subgraph G^u is framework input, and the exact inputs of each module are: for example, friendship subgraph G^u_f and the feature matrix H^u_f. The final outputs are the prediction possibility of status $P(s^u_{t+\Delta t})$ and the topk important neighbors set. For details, the SGN framework contains: a) a friendship interacting module which captures friendship structural embedding matrix $\hat{X}^u_f$ and output the embedding $\hat{x}_{u,f}$ of target u from GCN layers. Besides, neighbor attentions A^f are computed to demonstrate the importance of friends. b) an influence propagating module which uses multi-head GAT to get the attention matrix A^p, indicating how much different neighbors influence the target user u, the module also outputs an embedding $\hat{x}_{u,p}$. c) the global attention mechanism indicates the importance of two structures and computes the comprehensive embedding x^u. d) outputs.

$H^{(0)}$ is the characteristic feature matrix H^u_f of friendship structure G^u_f. Besides, $\widetilde{A_f}$ is the Laplaician matrix of graph G^u_f, where A_f and I_f denotes adjacency and degree matrix respectively. The final embedding matrix is $\hat{X}^u_f$ which contains the representations of each user in friendship structure $(\hat{x}_{u,f}, \hat{x}_{v_1,f}, ..., \hat{x}_{v_{n-1},f})$.

However, an obvious disadvantage of common GCN is that the information aggregation during convolution is equal and the differences between friends are ignored. But according to [26], 80% of users tend to contact with 20% of their friends frequently, which makes a need to quantify the exact impact from different friends to the target user on his individual influence status. Thus, we introduce the neighbor attentions mechanism, which, for each pair of friends, an attention factor is computed to indicate the clossness and importance of the friendship. For each friend $v \in N(u)$, the relative attention factor a_v is:

$$a_v = \frac{\exp\left(\phi\left(W\hat{x}_{v,f}\right)\right)}{\sum_{v\in N(u)}\left(\exp\left(\phi\left(W\hat{x}_{v,f}\right)\right)\right)} \tag{3}$$

where $\phi()$ denotes an activation function, W is the attention parameter matrix, $\hat{x}_{v,f}$ refers to the embedding vectors gained from the previous GCN. Through this neighbor attention mechanism, an averaged weighted embedding vector of target based on structural information could be denoted as:

$$x_f^u = \hat{x}_{u,f} \oplus \Sigma_{v \in N(u)} a_v * \hat{x}_{v,f} \tag{4}$$

where $\oplus$ indicates an aggregate operation, e.g., concatenation. Meanwhile, A set of the quantitative importance of target user's friends A_f could be collected.

4.2 Influence Propagating Module

As mentioned in Section I, influence propagation routine (mainly refers to the retweeting routines) is another crucial component of structural information. But the common GCN are not suitable enough as propagation data is often directed, dynamic and random. Besides, in this structure, intermediate users from higher-hop (not only one-hop neighbors) could have impact on the individual influence status of the target user. Therefore, we apply the state-of-the-art algorithm GAT on the propagation graph G_p^u. The details are shown in the bottom-left of Fig. 2. In fact, a main idea of GAT is to compute an attention coefficient between all the linked pair of nodes, thus for each node v_i, his attention $a_{i,j}$ to his propagation target v_j is:

$$a_{i,j} = \frac{\exp\left(\text{LeakyReLU}\left(a^T\left[Wh_i \| Wh_j\right]\right)\right)}{\sum_{v_j \in N(v_i)} \exp\left(\text{LeakyReLU}\left(a^T\left[Wh_i \| Wh_j\right]\right)\right)} \tag{5}$$

where $\|$ denotes concatenating operation between vectors, a^T and W are parameter vector and matrix, LeakyReLU is the activation function referred to [28]. Then, based on this attention martix $A_p = [a_{i,j}]_{n*n}$ and input of the characteristic feature matrix $H \in \mathbb{R}^{n*f}$ of propagation structure, the embedding matrix $\widehat{H} \in \mathbb{R}^{n*f'}$ should be:

$$\widehat{H} = \text{ReLU}\left(A_p H W'\right) \tag{6}$$

where $W' \in \mathbb{R}^{f*f'}$ is the embedding parameter matrix. To strengthen this process, we actually use the suggested multi-head GAT [28] which makes a parallel computation on K categories of attentions. For the set of parameters matrix $W' = \{W_1', W_2', \dots W_K'\}$ and attention matrix $A_p = \{A_1, A_2, \dots, A_K\}$, the output embedding matrix $\widehat{H}$ is concatenated as:

$$\widehat{H} = \text{ReLU}\left(A_1 H W_1' \| A_2 H W_2' \| \dots \| A_K H W_K'\right) \tag{7}$$

In addition, as shown in bottom-right of Fig. 2, compared with the friendship interacting module, we further introduce another single-head GAT (which has only one kind of attention) as:

$$H = \text{ReLU}\left(A_{\text{out}}\, \widehat{H} W_{\text{out}}'\right) \tag{8}$$

to obtain a same-dimentional embedding vector of target user x_p^u and the attention matrix A_{out}. As the attention matrix A_{out} only compute impact between direct neighbors, we use a breadth first search (bfs) algorithm to compute a similar set of quantitative importance A_p, which indicates the importance of all the nodes in propagation structure towards target user u.

4.3 Global Attention and Output

Given the embedding vector from friendship structure x_f^u and from propagation structure x_p^u. A global attention mechanism is introduced to differ the importance of the two structures. Take the attention coefficient of friendship structure a_f as an example:

$$a_f = \frac{\exp\left(Wx_f^u\right)}{\exp\left(Wx_p^u\right) + \exp\left(Wx_f^u\right)} \tag{9}$$

and the structural representation x^u is:

$$x^u = a_f \cdot x_f^u + a_p \cdot x_p^u \tag{10}$$

Then, after a neural layer and softmax function, the model will finally output a two-dimentioned vector represent whether the user's individual social influence is active or not. The negative log-likelihood loss function will be optimized next.

$$\text{Loss} = -\sum \log\left(P\left(s_{t+\Delta t}^u \mid s_t^u, G_{f,t}^u, G_{p,t}^u\right)\right) \tag{11}$$

Besides, in order to identify the intermediate users who have most impact on target individual influence status, plenty of related attentions are learned in either module or globally. Based on these quantitative users attention set A_f, A_g and the relative global attention coefficient a_f and a_g, we could first obtain the final set of impact factors from target's friends in friendship structure I_f, and the set of impact factors from all the users who participate the influence propagation in propagation structure I_p. Then through concatenating and sort operation, we can final figure out the Top k important intermediate users who affect target user's individual social influence most as shown below:

$$\text{Top}\, K\left(\text{sorted}\left(\{I_f \mid I_f = a_f * A_f\} \oplus \{I_p \mid I_p = a_p * A_p\}\right)\right) \tag{12}$$

5 Experiment

5.1 Experiment Setings

Dataset. We use two real-world social network datasets to evaluate the proposed SGN framework on user-level social status prediction – Twitter and Weibo.

- **Twitter** [4] The Twitter dataset records 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 friendship and propagation structure refer to its follower network and retweet network respectively.
- **Weibo** Weibo is a famous large-scale Chinese microblogging network. The dataset here is from [34] and contains the directed following relationship and their retweets behaviour between 1,776,950 users from 28th July 2012 to 29th October 2012.

Data Preprocess: Following the practice in existing work [4,24], The users who are influenced in an extended period T_1 and be active in the next rather short T_2 are considered as positive samples, while the inactive ones are negative. Meanwhile, because of the imbalance in users' neighbors, we rebuilt the friendship and propagation structure with RWR(random walk with restart)

Baselines. We validate the effectiveness of SGN by comparison with some basic and state-of-the-art baselines as follow:

- **Logistic Regression (LR)** We use LR as a classic classification model. Some of the user's characteristics features are used as the input of the model which includes: Coreness, Pagerank, Hub score and authority score, Eigenvector Centrality, Clustering Coefficient, The number/ratio of active neighbors and Density of subnetwork induced by active neighbors.
- **Support Vector Machine (SVM)** SVM with linear kernel is another classic supervised classification model. And the same features as LR are the input of SVM.
- **Deepinf** [24] Deepinf is now the state-of-the-art model on user-level social influence prediction problems. Firstly, It maps each user to her D-dimensional representation through Deepwalk, then concatenates the representation and the characteristics features as an input of GCN/GAT layers, finally predicts the user's social influence status.
- **FATE** [26] FATE can predict different categories of behaviour of a user by modeling the friendship and user actions through an attentioned GCN and temporal dynamic through tLSTM. Here, we only predict one behaviour(retweet) and set the timestamp of tLSTM as 1.

Parameter Setting. First, the restart probability in random walk is set as 0.8, and the size of either structure graph is 50. Next, in SGN Framework, we trained two GCN layers with 32/8 hidden units, and two multi-head GAT layers consist of 8/1 attention heads and 16/16 hidden units. For the global attention mechanism, the hidden parameter dimension is 16 and then output 2 units for binary prediction. For detail, we adopted elu as the nonlinearity function and trained adam optimizer with learning rate 0.005, weight decay 1e-4, and dropout rate 0.2. In addition, We use 80%, 10%, 10% of the dataset for training, validation and test, respectively.

Table 1. Prediction performance between baselines on Twitter and Weibo (%)

Dataset	Module	Accuracy	Precision	Recall	F1-score
Twitter	LR	77.33	72.92	53.41	61.66
	SVM	78.15	76.48	54.49	63.64
	DeepInf	84.26	85.00	63.93	72.97
	FATE	83.95	85.30	63.27	72.67
	SGN	**85.78**	**89.36**	**65.90**	**75.86**
Weibo	LR	69.23	62.92	42.25	50.55
	SVM	71.34	55.24	44.14	49.07
	DeepInf	73.33	85.40	48.12	61.55
	FATE	73.73	88.68	48.61	62.80
	SGN	**74.86**	**94.84**	**49.85**	**65.35**

5.2 Experiment Result of Prediction Performances

The prediction performances between all the baselines on Twitter and Weibo datasets are compared in Table 1. The evaluation metrics are accuracy, precision, recall and F1-score. We have the following observations:

Our proposed model SGN achieves the best performance on all datasets and different evaluation metrics. Especially when compared with LR and SVM, SGN (Twitter) achieves about 10% improvement on AUC and 20% improvement on F1-score.

And it's obvious that graph neural networks related models Deepinf, FATE and SGN achieve significantly better performance compared with 2 classic ML models. It indicates that the structural features learned from GNN can be beneficial to user-level social influence status prediction. In addition, the better performances of SGN among GNN models also indicate the effectiveness of our division of structural graph from both friendship and influence propagation.

What's more, we can witness the significant improvement of GNN related models are mainly from Precision, which indicates that the structural features improve the classification ability especially on positive samples thus can identify users who will be more likely affected active in the future. It reveals the promoting effects of structural information and implies the existence of some important users either from friendship structure or propagation structure who effect the target user to be active. By contrast, the relatively poor correlation between negative samples and structural information may demonstrate the rather more complicated reasons for the formulation of negative attitudes. In fact, negative attitudes should not only come from the neighborhoods, but also be affected by various objective factors.

5.3 Attention Analysis

Global Attention. We compute the global attention of friendship and propagation structure in order to investigate their effectiveness and different importance. The Fig. 3 shows the attention values of propagation structure are higher than

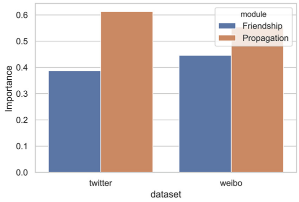


Fig. 3. The global attention importance between Friendship interacting module and Influence propagating module

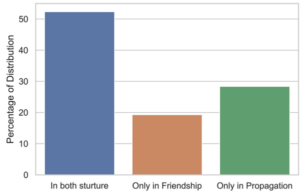
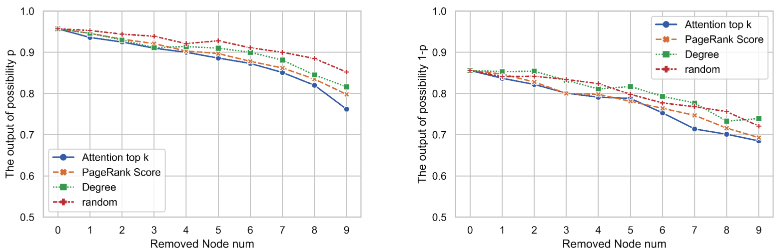


Fig. 4. Distribution of Top 10 important intermediate neighbors who affect target user most



(a) positive sample (b) negative sample

Fig. 5. The prediction possibility p after removing some nodes from original graph according to their importance computed by 4 methods

friendship after 20 training epochs. This implies that information in propagation structure should be more important, and incorporation with Table 1, positive users are affected by those from their influence propagation routines more.

User Attention. For each user, Their top 10 most important users are computed from the learned user attention mechanism in SGN framework. According to their belonging structure: friendship only, propagation only, or both, the distribution was shown in Fig. 4. The most important users are mainly from both structures, and again, propagation structure turns to be more crucial.

Further, in order to quantify the effectiveness of the attention mechanism and the top 10 most important users, we chose a positive and negative sample and output their predicted possibilities after removing part of his neighbors according to the top 10 list. By contrast, other three groups of nodes are established based on random, degree and PageRank score. For each group, neighbor nodes would be successively removed from the original graph according to its rank, then the model will output the new possibilities of the target user's status. As shown in Fig. 5, rather in positive or negative sample, prediction possibility p declines more when removing the users in our attention top10 list, which indicates the significant impact in target user's neighborhood structure. For details in the positive sample as an example, first, the head of each group which is removed in prior refers to important intermediate users selected by each method, thus the differences between random and others prove the effectiveness of this strategy; besides, our SGN attention even outperforms degree and pagerank which are commonly used in nodes importance ranking, this indicates the effectiveness of our neighbor attentions. Meanwhile, the decline range and trend of possibility p between two samples also imply that positive users suffer from their important intermediate neighbors more.

6 Conclusion

In this paper, we study the problem of individual social influence status prediction on diverse structures, and first introduce a division of the structural information which contains friendship and propagation structure respectively. Next, we propose an end-to-end prediction model SGN to learn representations from the two structures through attentioned-GCN and multi-head GAT. Besides, global and local attention mechanisms help to complete the predictions and to identify the most important intermediate neighbors who affect the target most. Experiments are conducted based on two real-world large-scale datasets, and the proposed SGN is compared with 4 baselines. The result shows SGN outperforms both the classic ml models and some current GNN models. Meanwhile, we validate the importance of the topk neighbors set by comparing our neighbor attention mechanism in SGN with some classic node ranking algorithms like degree and pagerank. What's more, those intermediate neighbors seem to be more important in positive samples, which means they are more likely to make target users active. For some future work, we would like to come up with some innovate graph encoders to substitute our model stack here, and also be able to get better representations of the complex social interaction graphs Besides, we'd also like to imply time as another important dimension in social networks and also the social influence prediction problems.

Acknowledgment. This work is supported by National Key Research and Development Plan in China (2018YFC0830500), and by National Natural Science Foundation of China (U1636105).

References

1. Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election. Social Science Electronic Publishing (2017)
2. Chen, G., Kong, Q., Xu, N., Mao, W.: NPP: a neural popularity prediction model for social media content. Neurocomputing **333**, 221–230 (2019)
3. Chien, E., Peng, J., Li, P., Milenkovic, O.: Adaptive universal generalized pagerank graph neural network. In: International Conference on Learning Representations (2021). https://openreview.net/forum?id=n6jl7fLxrP

4. De Domenico, M., Lima, A., Mougel, P., Musolesi, M.: The anatomy of a scientific rumor. Sci. Rep. **3**, 2980 (2013)
5. Dong, Y., Johnson, R.A., Xu, J., Chawla, N.V.: Structural diversity and homophily: a study across more than one hundred big networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 807–816 (2017)
6. Goyal, P., Chhetri, S.R., Canedo, A.: dyngraph2vec: capturing network dynamics using dynamic graph representation learning. Knowl.-Based Syst. **187**, 104816 (2020)
7. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: Advances in Neural Information Processing Systems, pp. 1024–1034 (2017)
8. Han, P., et al.: GCN-MF: disease-gene association identification by graph convolutional networks and matrix factorization. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 705–713 (2019)
9. Hennig-Thurau, T., Wiertz, C., Feldhaus, F.: Does twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. J. Acad. Mark. Sci. **43**(3), 375–394 (2015)
10. Hoang, T.B.N., Mothe, J.: Predicting information diffusion on twitter - analysis of predictive features. J. Comput. Sci. **28**(Sep), 257–264 (2018)
11. Jie, T., Sun, J., Chi, W., Zi, Y.: Social influence analysis in large-scale networks. In: ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2009)
12. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. In: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 137–146 (2003)
13. Khajehnejad, M., Rezaei, A.A., Babaei, M., Hoffmann, J., Jalili, M., Weller, A.: Adversarial graph embeddings for fair influence maximization over social networks. arXiv preprint arXiv:2005.04074 (2020)
14. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: ICLR (2017)
15. Kong, X., Shi, Y., Yu, S., Liu, J., Xia, F.: Academic social networks: modeling, analysis, mining and applications. J. Netw. Comput. Appl. **132**, 86–103 (2019)
16. Lee, C., Wilkinson, D.J.: A hierarchical model of nonhomogeneous poisson processes for twitter retweets. J. Am. Stat. Assoc. **115**(529), 1–15 (2020)
17. Li, C., Ma, J., Guo, X., Mei, Q.: Deepcas: an end-to-end predictor of information cascades. In: Proceedings of the 26th International Conference on World Wide Web, pp. 577–586 (2017)
18. Lu, Y., Wang, X., Shi, C., Yu, P.S., Ye, Y.: Temporal network embedding with micro-and macro-dynamics. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pp. 469–478 (2019)
19. Ma, J., Chang, B., Zhang, X., Mei, Q.: CopulaGNN: towards integrating representational and correlational roles of graphs in graph neural networks. In: International Conference on Learning Representations (2021). https://openreview.net/forum?id=XI-OJ5yyse
20. Mathews, P., Mitchell, L., Nguyen, G., Bean, N.: The nature and origin of heavy tails in retweet activity. In: Proceedings of the 26th International Conference on World Wide Web Companion, pp. 1493–1498 (2017)

21. Matsubara, Y., Sakurai, Y., Prakash, B.A., Li, L., Faloutsos, C.: Rise and fall patterns of information diffusion: model and implications. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2012, pp. 6–14. Association for Computing Machinery, New York (2012). https://doi.org/10.1145/2339530.2339537
22. Pareja, A., et al.: Evolvegcn: evolving graph convolutional networks for dynamic graphs. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 5363–5370 (2020)
23. Puny, O., Ben-Hamu, H., Lipman, Y.: Global attention improves graph networks generalization (2021). https://openreview.net/forum?id=H-BVtEaipej
24. Qiu, J., Tang, J., Ma, H., Dong, Y., Wang, K., Tang, J.: DeepInf: social influence prediction with deep learning. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining, KDD 2018, pp. 2110–2119. Association for Computing Machinery, New York (2018). https://doi.org/10.1145/3219819.3220077
25. Shen, H., Wang, D., Song, C., Barabási, A.L.: Modeling and predicting popularity dynamics via reinforced poisson processes. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 28 (2014)
26. Tang, X., Liu, Y., Shah, N., Shi, X., Mitra, P., Wang, S.: Knowing your fate: Friendship, action and temporal explanations for user engagement prediction on social apps. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2020, pp. 2269–2279. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3394486.3403276
27. Trivedi, R., Farajtabar, M., Biswal, P., Zha, H.: DyRep: learning representations over dynamic graphs. In: International Conference on Learning Representations (2019)
28. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)
29. Wang, X., Zhu, M., Bo, D., Cui, P., Shi, C., Pei, J.: AM-GCN: adaptive multi-channel graph convolutional networks. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2020, pp. 1243–1253. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3394486.3403177
30. Xiong, H., Yan, J.: BTWalk: branching tree random walk for multi-order structured network embedding. IEEE Trans. Knowl. Data Eng. 1 (2020). https://doi.org/10.1109/TKDE.2020.3029061
31. Xu, F., Li, Y., Xu, S.: Attentional multi-graph convolutional network for regional economy prediction with open migration data. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2225–2233 (2020)
32. Yu, B., Yin, H., Zhu, Z.: Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI) (2018)
33. Yu, B., Yin, H., Zhu, Z.: Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp. 3634–3640 (2018)
34. Zhang, J., Liu, B., Tang, J., Chen, T., Li, J.: Social influence locality for modeling retweeting behaviors. In: IJCAI, vol. 13, pp. 2761–2767 (2013)