

Event Prediction Based On Large Scale Network Subgraph Convolution

XiaoWei Wu

China People's Public Security University

FanLiang Bu (✉ bufanliang@sina.com)

China People's Public Security University

ZhiWen Hou

China People's Public Security University

Research Article

Keywords: Prediction, Network, Convolution, CSGCN

Posted Date: June 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-622956/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Event Prediction Based on Large Scale Network Subgraph Convolution

XiaoWei Wu^{1,2}, FanLiang Bu^{1,*}, and ZhiWen Hou¹

¹ People's Public Security University of China, College of Information Network Security,
Beijing, 100038, China

² China People's Police University, Overseas Security & Protection School, Langfang,06500,
China

*bufanliang@sina.com

Abstract

Aiming at the problem of event prediction in large-scale event network, a collapse subgraph convolution (CSGCN) algorithm is proposed, which uses event subgraph to predict the subsequent events of event group. CSGCN algorithm collapses the edge induced event subgraph in large-scale event network, removes the irrelevant event nodes from the subgraph, and forms a new event subgraph. GCN algorithm is used to learn the graph embedding representation of the event subgraph, and the subsequent events of the event group are predicted by comparing the similarity between the graph embedding representation of the event group and the subsequent events. Because only some related nodes are processed each time, the application of the model in large-scale data graph is feasible. Through experiments, we explore and verify the effectiveness of extracting features from subgraphs of large-scale graph by using graph convolution training to obtain graph embedding representation. We find that GCN has better event prediction effect than Euclidean distance and co rotation similarity, which further shows that graph convolution algorithm has good performance in the field of graph feature extraction.

Introduction

At present, most of the event prediction methods are based on time event series. Bilod¹ and others use Markov model or RNN model, Mei² and others use LSTM model, Chunyan Sang and others use LSTM and attention mechanism to predict the evolution trend of social network information dissemination. However, in the specific event prediction scenario, the relationship between events is more complex than that of time series, and the length of time window is not consistent, which leads to the effect of dynamic event prediction using time window in many cases is not ideal.

Event chain is used to describe the chain structure formed by the successive triggering of multiple events in a certain time and space. Many scholars also try to use the chain structure to express the temporal and spatial relationship of sequential events and further achieve event prediction. Yuan et al.³ causal chain mining, Li et al.⁴ epidemic model, Kritika Singh⁵ accident path analysis method, etc., mostly use the event pair^{6,7} method to achieve subsequent event prediction, but similar research like Wang et al.⁸ using sequence event chain to obtain high accuracy prediction is rare. At present, event chain and event pair prediction model have a typical feature, that is, the expression ability of the relationship between events is insufficient, and the rich relationship between events has not been fully explored. In 2004, Leveson⁹ pointed out in the research of safety system accident model that the method of analyzing the event process through sequence events ignored the potential relationship between events. In the above research results, this view is also confirmed. People have been exploring methods to fully express the relationship between events and to predict more effective events.

In 2007, Li Ben et al.¹⁰ put forward the concept of link prediction for the first time. Through the analysis of network structure and node information, the link possibility between two unconnected nodes is evaluated to realize the prediction. Traditional link prediction is based on judging the similarity of two nodes or whether there is a link between two event nodes. Common methods include node attribute comparison^{11,12}, network topology comparison^{13,14}, maximum likelihood estimation^{15,16}. With the development of graph neural network, graph embedding representation is more and more used in link prediction^{17,18,19,20} for feature learning. In reference²¹, the covariance matrix adaptive evolution strategy (CMA-ES) is used to optimize the linear combination weight in the similarity comparison of neighborhood and node, and the combination of topology features and node attributes improves the accuracy of link prediction. O'Madadhain et al.²² proposed a local conditional probability model based on network structure and node attributes to predict the time-varying relationship between nodes. At the same time, the research of link prediction promotes people's understanding of network structure characteristics and the exploration of complex network structure evolution law^{23,24}, which provides support for using complex network to express the relationship between events. In most link prediction methods, abnormal events or behaviors are predicted through the prediction of abnormal or missing edges in the network. However, in predictions such as terrorist events, a group of events after occurrence means that other events may occur, and there is no strict timing relationship between the events that have occurred. Traditional methods cannot provide effective solutions for this kind of problems, while link prediction methods based on complex networks provide ideas for solving the problem of event group prediction.

Li²⁸ constructed directed event evolution graph (NEEG) based on script events with event context, and proposed event chain prediction method based on large scale network graph neural network (SGNN) and gating graph neural network (GGNN). Based on graph embedding representation, the interaction between event sequences is modeled and the subsequent events of known event chain are predicted. This method still follows the method of sequence prediction, which requires the event to have obvious temporal characteristics. This paper improves SGNN

algorithm, uses graph convolution neural network algorithm instead of gated graph neural network (GGNN), and proposes collapse subgraph convolution algorithm (CSGCN). Through the association between context event groups, the possible events after a group of events occur are predicted, and the effectiveness of application in large-scale network is verified by experiments.

Theoretical Method

Graph structure has the ability to express complex relationships. Event network graph structure can not only express the relationships between events, but also combine event characteristics and network structure, which helps people to use more accurate and structure sensitive methods²⁵ to study the evolution of event network and predict future events. This paper uses the event network diagram to show the relationship between the events in the text based on the above and below. It takes events as nodes and uses edges to represent the relationship between event nodes. We regard the prediction problem of subsequent target events of event group based on event graph as a binary classification problem. If the causal relationship between event group and target event is known, the sample is marked as positive link, otherwise it is negative link. Extracting useful information from graph structure and node features is very important for event prediction. In contrast, the local relationship coupling of events in event group is closer. Therefore, we extract event subgraphs from large-scale narrative event network, and use graph neural network to obtain graph embedded information representation of event nodes in the graph for event prediction. Through training and learning a large number of sample data, on the basis of graph embedding representation of subgraph structure and event node features, we compare the similarity between event group feature information and target event feature information, train feature information describing the relationship between event group and target event, and further realize event prediction. \in

Event network graph structure and Event prediction description

A graph can be represented as $G = \{E, V\}$. where $V = \{v_1, \dots, v_N\}$ is a collection of N nodes and E is a collection of edges. Graph structure information can also be represented by an adjacency matrix $A \in \mathbb{R}^{N \times N}$. In addition, each node in the graph has a node feature, and we use $X \in \mathbb{R}^{N \times d}$ denotes the node characteristic matrix, where d is the characteristic dimension and N is the number of nodes. Each node in graph G can be seen as a d -dimensional graph signal.

A group of graphs is represented as $\{G_i\}$, and the corresponding tag event is y_i . We combine each graph G_i with its label event y to form an event subgraph, and judge whether the follow-up event of $\{G_i\}$ is y_i through the two classification of the event subgraph, so as to transform the event prediction problem into graph classification problem. In the event graph, we extract a set of subgraphs $\{G_i\}$ composed of K related event $\{v_i^1, v_i^2, v_i^3, \dots, v_i^k\}$. The task of event group prediction is to make graph $\{G_i\}$

(including structural information and node characteristics) is used as input to predict its corresponding event y . The task of event group prediction is to map $\{G_i\}$ as input to predict its corresponding event y_i , using structure information and node characteristics. If features of the $\{G_i\}$ and y_i obtained from the node feature information and the structure information are identical, then y_i is considered to be a subsequent event of $\{v_i^1, v_i^2, v_i^3, \dots, v_i^k\}$, and subsequent event prediction based on event group is implemented.

Event subgraph extraction based on graph collapse

In this paper, we follow the idea of divide and conquer in literature 27 and 28, and only deal with the event related nodes in each training. However, we need to collapse the related nodes of each processing to form the collapse graph of the event subgraph, and on this basis, we use GCN algorithm to obtain the embedded representation of the graph. The experimental results show that since the arrangement order and node features of the event subgraph keep the original features, the graph embedding representation after GCN operation is still meaningful for the whole event graph. The method of extracting event subgraph and collapsing is introduced in this paper.

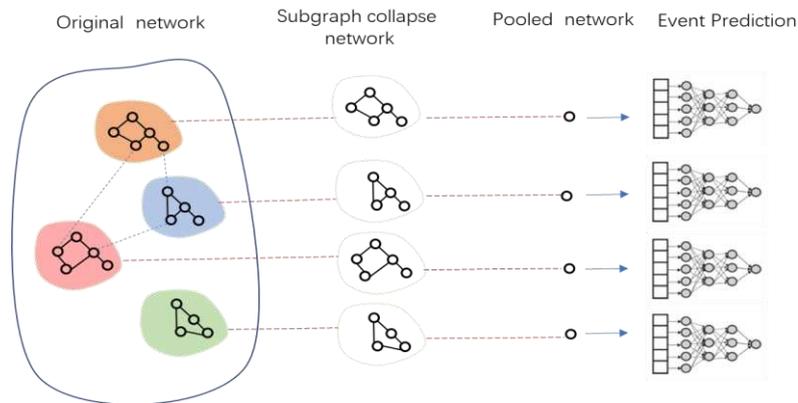


Figure1.Event subgraph extraction based on graph collapse

Definition 1. Graph G refers to a tuple (E, V) consisting of non-empty finite sets V and an unordered pair of sets E of some elements in sets V , record as $G = (E, V)$. V is called the set of vertices of G , and the elements in it are called the vertices of G . E is called the edge set of G , and the elements in it are called the edges of G . If $V = \{V_1, \dots, V_N\}$, the element e in set E corresponds to the unordered pair formed by two elements in set V , record as $e = v_i v_j$ or $e = v_j v_i$.

Definition 2. Let $G = (E, V)$ be a graph if $V' \in V, E' \in E$, and there must be $v_i, v_j \in V'$

for any edge $e \in E'$, $G'=(E', V')$ is also a graph, which is called edge induced subgraph by E' . The subgraph mentioned in this paper refers to edge induced subgraph.

The edge induced subgraphs of event nodes in the event network are divided into event subgraphs, that is, the event network graph is divided into several different event subgraphs. In particular, the subgraph mentioned in this paper refers to the edge induced subgraph. After extracting the event subgraph, the subgraph is separated from the large-scale network structure graph by subgraph collapse. In order to collapse the event subgraph, we first need to extract the event subgraph to form a new graph, and then form a new adjacency matrix according to the arrangement of events in the graph, which is used to represent the network structure of the event subgraph. The selection of scope is the core step of graph collapse process. We divide the event groups and target events in the sample data into an event subgraph, and remove all other nodes in the event network graph that do not belong to the event subgraph. After the event subgraph collapses, we further fuse the attribute information and structure information of the nodes in the graph, and obtain the characteristics of different subgraphs through training. The specific process is shown in Figure (1).

In this paper, we obtain the collapse adjacency matrix A' of edge induced subgraph G' of graph G by Equation(1), where A is the adjacency matrix of graph G . Let $G=(V, E)$ be a graph with n vertices, then the adjacency matrix of G is a square matrix of order n with the following properties:

$$A[i,j]=\begin{cases} 1, & \text{if } (v_i,v_j) \text{ or } \langle(v_i,v_j)\rangle \text{ is an edge of } E(G) \\ 0, & \text{if } (v_i,v_j) \text{ or } \langle(v_i,v_j)\rangle \text{ is not an edge of } E(G) \end{cases} \quad \text{Equation(1)}$$

S is an adjacency matrix composed of nodes in G' ,

$$A'[i,j]=\begin{cases} 1, & \text{if } (v'_i,v'_j) \text{ or } \langle(v'_i,v'_j)\rangle \text{ is an edge of } E'(G') \\ 0, & \text{if } (v'_i,v'_j) \text{ or } \langle(v'_i,v'_j)\rangle \text{ is not an edge of } E'(G') \end{cases} \quad \text{Equation(2)}$$

$$S[i',j']=A[i,j] \text{ if } v'_i \in v_i, v'_j \in v_j \quad \text{Equation(3)}$$

$$A'=S^T A S \in R^{n' \times n'} \quad \text{Equation(4)}$$

$$X'=S^T X \in R^{n' \times d} \quad \text{Equation (5)}$$

The embedding matrix X' of is expressed as Equation (5), where X is the characteristic attribute matrix of G . The matrix representation of event subgraph G' is obtained.

Pooling

It is necessary to train pooling operator to collapse event subgraph. The design of pool operator is based on graph Fourier transform of subgraph $\{G_i\}$. $L=\tilde{D}-\tilde{A}$ represents the Laplacian matrix of subgraph $\{G_i\}$.

The Eigenvector of Laplace matrix L is expressed as u_1, \dots, u_N . according to the literature²⁹, the upper sampling operator c will be used as the basic feature of the subgraph, and the Eigenvector is sampled into the whole graph as $\bar{u}_l = cu_l (l=1 \dots N)$.

The pool operator is obtained by using the upper sampling Eigenvector, according $X_l = \bar{u}_l^T X$. The $X_l \in \mathbb{R}^{N \times d}$ is a pooling operator represented by the Eigenvector u_1, \dots, u_N of Laplace matrix L and the sampling operator. For the convenience of calculation, the first H Eigenvectors can also be used for $X_p = [X_0, \dots, X_H]$.

Embedding representation of event subgraphs

Convolution network on graph is divided into spectral domain convolution and spatial domain convolution. The graph convolution algorithm GCN²⁶ mentioned by Thomas Kipf is spectral convolution, which transforms the convolution core and the graph signal to the Fourier domain to extract the characteristics of the graph signal. After sampling a fixed number of domain nodes, the spatial convolution network obtains the embedded representation of the network by inner product of neighborhood points and convolution kernel parameters. In comparison, spectral convolution algorithm has higher accuracy, but the scalability of network nodes is not strong, so it is not suitable for large-scale graph classification and other functions training. On the contrary, spatial convolution algorithm has strong expansibility, but because it is based on sampling technology, the accuracy of its model is slightly inferior to spectral convolution. In this paper, we apply graph convolution network to event subgraph. The network scale is small, so spectral domain convolution algorithm GCN is used to realize information embedding and message propagation of graph structure nodes, so as to obtain the best effect. SGCN algorithm uses Equation(6) to realize the spectral domain convolution process of event subgraph.

$$H^{(K)} = M(A, H^{(K-1)}; W^{(K)}) = \text{ReLU} \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(K-1)} W^{(K-1)} \right) \quad \text{equation (6)}$$

Where M is the message function, this paper uses the frequency domain graph convolution algorithm to extract the typical feature structure, and $W^{(K)} \in \mathbb{R}^{d \times d}$ is the weight matrix that needs to be learned and trained. \tilde{A} is the sum of adjacency matrix of G and identity matrix, \tilde{D} is the degree matrix of degree \tilde{A} .

$$\tilde{A} = A + I, \tilde{D} = \sum_j \tilde{A}_{ij} \quad \text{equation (7)}$$

Event prediction based on event subgraph convolution

After the embedded representation of the graph is obtained by CSGCN algorithm, the graph is pooled into a node, and the Eigenvalues of each node in the collapse subgraph are averaged as the Eigenvalues of the pooled node. The similarity between the pooled node and the subsequent event node is compared, so as to realize the event prediction. The method of embedding representation of nodes and averaging Eigenvalues cannot further reflect the structure information of subgraphs. We try to integrate and extract the information in each subgraph. EigenPooling³⁰ uses graph Fourier transform to integrate and output the graph structure information and attribute information. In this paper, after convoluting the subgraph to obtain the graph embedded representation, we further apply EigenPooling algorithm to enhance the fusion of graph structure and feature information, but the effect is not obvious. At the same time, we also try to use the attention mechanism to obtain the complex relationship between events in the event group, and achieve good results.

Experiments and results

We tested our idea in the experiment, through the operation of graph convolution after the event subgraph collapses to realize the prediction of subsequent events. Firstly, the undirected event evolution graph (UDEEG) is constructed to represent the event collapse subgraph. Its nodes are events, and its edges represent the temporal or causal relationship between events. In the experiment, we use the relationship between known event groups and subsequent events in the sample data to form undirected UDEEG.

Data sets

The data set for this study is from the billion word corpus of the New York Times. After POS tagging and dependency parsing, phrase structure parsing and coreference resolution, independent context event groups and candidate events are extracted. Based on the data set used in reference 1, this paper improves it. Each group of context events and the correct events in its five candidate follow-up events are taken as the positive samples for training, and an error event is randomly selected to form a negative sample. The ratio of positive and negative samples in the test set is 1:1, forming the data set used in this experiment. Table 1 shows the detailed amount of sample data in the dataset.

Experimental setting

In the experiment, the number of event groups processed in a batch is Batchsize. According to Equations 6, 7 and 11, the collapse subgraph convolution is performed

for each event group in the data, and the convolution level is $K = 2$.

Super parameters are optimized on the training set, especially CSGCN is actually two classification problem, so cross entropy loss function is used as the objective function.

$$H(x, W) = -\sum_{i=1}^n P(x_i) \log(q(x_i)) + \frac{\lambda}{2} \|W\|^2 \quad (8)$$

Where $P(x_i)$ is the true value and $q(x_i)$ is the predicted value. W is the model parameter set. λ Is the parameter of L2 regularization, which is set to 0.00001. The learning rate is 0.0001, the batch size is 4000, and the convolution level K is 2. In this paper, the initial node feature of the node is represented by the embedding result of event predicate-GR trained by DeepWalk algorithm in reference 28, and the embedding dimension d is 128. The model parameters are optimized by RMSprop algorithm. The ratio of training set to test set is 9:1.

Comparison of experimental results and baselines

In previous studies, the prediction of subsequent target events of narrative text event sequence is realized by selecting one of several candidate events. For example, Li²⁸ divided the five candidate events into five categories, and set their labels to 0, 1, 2, 3 and 4 respectively according to the order of candidate events. In the event prediction process, one of the five candidate events was selected. In this case, the label of candidate events does not have practical significance, and the candidate events belonging to the same category do not have the same characteristics. The neural network algorithm is mainly used for feature learning, which is one of the important reasons for the poor effect in narrative text event prediction.

In this paper, the candidate events are divided into positive and negative samples, and the classification has practical significance. Although facing the complex event group prediction, the experiment still achieves good prediction results. The learning curve of the optimized CSGCN that can express the change of accuracy with the training period is shown in Figure (2), and the accuracy of the optimized algorithm reaches 62.56%.

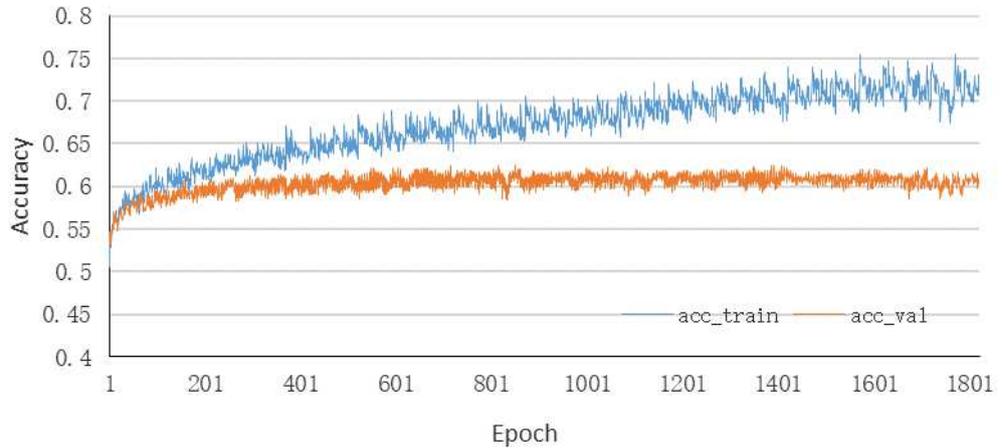


Figure 2. learning curve of optimized CSGCN (accuracy changes with training cycle)

In the experiment, we optimize the experiment on the basis of CSGCN algorithm. Firstly, the graph embedding obtained from the convolution of subgraphs is further combined with the graph structure information, and the Eigenvector of the adjacency matrix of subgraphs is combined with the graph embedding representation (CSGCN + Eigen). The interactive utilization of local structure information and node information is attempted, and the improvement of 1.33% is obtained. We use the attention mechanism (CSGCN + Attention) to learn the event group relationship, and get a better learning effect, which is 3.32% higher than CSGC algorithm. The comparison of accuracy between our algorithm and baseline algorithms is shown in Table 2.

Analysis of experimental

In CSGCN algorithm, it is necessary to compare the similarity between the feature representation of event group and target event. In this paper, we try a variety of methods in event representation, event similarity comparison and embedded representation of event group in the convolution layer, and analyze these methods through experiments.

- Analysis of convolution layer event representation. This paper attempts to use the following methods and makes a comparison. In the second convolution output of two-level graph convolution, the Eigenvalue output method is used, and no obvious effect is achieved. If the dimension of the second layer convolution is set to 128 or 12, the effect is not obvious. Finally, the output dimension of the second level graph convolution is set to 2. After similarity and comparison, good results are obtained.
- Analysis of similarity comparison method. We use the similarity of event group and target event to get the probability that the sample belongs to positive sample. When the convolution output dimension is 128 or 12, this paper uses the methods of cosine similarity and vector subtraction to train, and the training accuracy is not significantly improved, which leads to a higher recall rate in the training process. When the convolution output dimension is 2, the vector subtraction method is used for similarity comparison, the training accuracy is improved greatly, and the training recall rate is relatively low. The specific learning results are shown in Figure (3). The experimental results show that the convolution dimension reduction representation of graph is more meaningful for event similarity calculation, and the convolution training of graph can better reflect the complex relationship between events.

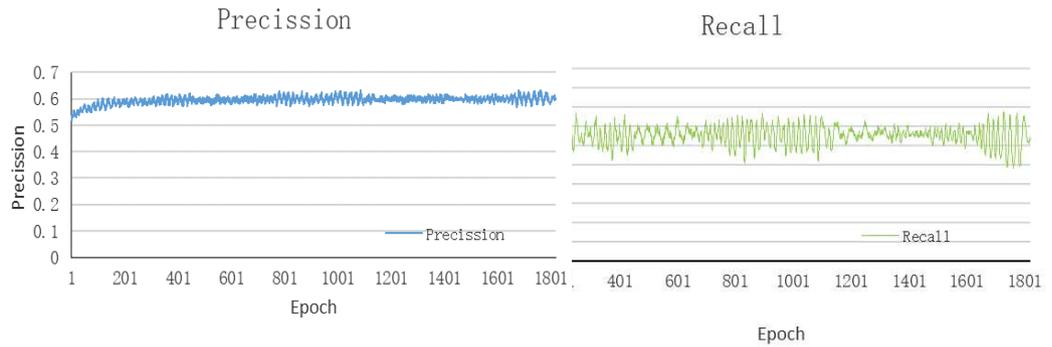


Figure 3-a. Experimental results of vector subtraction after dimensionality reduction using GCN

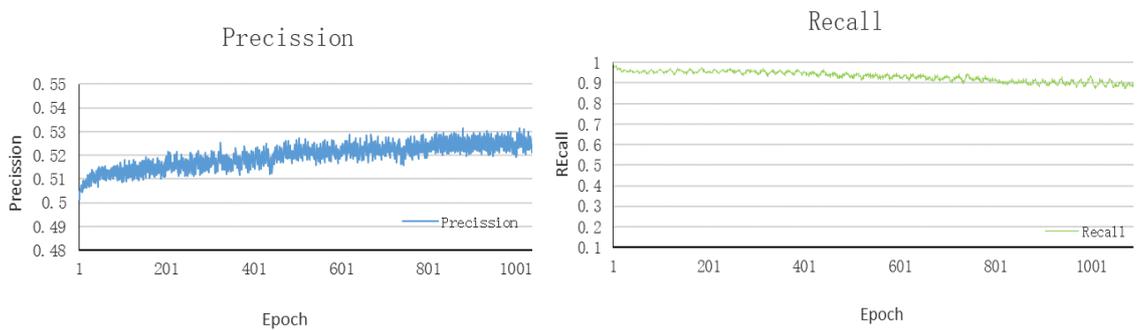


Figure 3-B. Experimental results of GCN output of 12 and comparison using cosine similarity

Figure 3. Comparison of accuracy and recall using different similarity comparison methods

- Analysis of graph embedding representation of event group. The events in the event group have different values. The attention mechanism is introduced to train the weight distribution of the events in the event group. The graph embedding representation of the event group is obtained by calculating the dot product of the weight coefficient and the graph embedding representation of the events. With the introduction of attention mechanism, the efficiency of the algorithm has been effectively improved, but it has caused serious over fitting. This phenomenon shows that it is of great significance to identify the important events in the group events. Although the algorithm is improved by adding parameters, the phenomenon of over fitting shows that the important events in the event group can get better recognition effect through training, but the universality of its law is not enough. The phenomenon first shows that it is of great significance to identify the important events in group events. However, the over fitting phenomenon after increasing the parameters to improve the accuracy of the algorithm shows that the important events in the event group can get better recognition effect through training, but the universality of the rule is not enough.

Conclusion

CSGCN algorithm only disposes some related nodes at a time, which makes the application of the model in large-scale data graph feasible. Through experiments, we explore and verify the

effectiveness of extracting features from subgraphs to obtain graph embedding representation for event prediction.

The following conclusions are verified by this experiment:

- Using graph convolution training on subgraphs of large-scale graph can obtain more effective graph embedding representation.
- Using event group based on event subgraph to predict subsequent events can obtain a relatively high accuracy.

References

1. Bilo, M., Charpentier, B., & Günnemann, S. Uncertainty on Asynchronous Time Event Prediction, DOI: [https://arxiv.org/abs/1911.05503v2\(2019\)](https://arxiv.org/abs/1911.05503v2(2019)).
2. Yuting, Mei., & Fa, Li. Predictability Comparison of Three Kinds of Robbery Crime Events Using LSTM. Proceedings of The 2nd International Conference on Data Storage and Data Engineering (DSDE 2019). 48-52(2019).
3. Yuan, C. F., Wang, H., & Chen, Y. The hierarchical risk source identification method connected with event causal chain in the emergency process of fire accident of petroleum storage and transportation. Applied Mechanics and Materials. 501-504, 2411-2414, DOI: [http://doi.org/10.4028/www.scientific.net/AMM.501-504.2411\(2014\)](http://doi.org/10.4028/www.scientific.net/AMM.501-504.2411(2014)).
4. Li, S., Liu, Z., & Li, Y. Temporal and spatial evolution of online public sentiment on emergencies. Information Processing & Management. 57(2), 102177(2020).
5. Singh, K., Maiti, J., & Dhalmahapatra, K. Chain of events model for safety management: data analytics approach. Safety Science. 118, 568-582(2019).
6. Nathanael, C., & Daniel, J. Unsupervised learning of narrative event chains. In ACL. 94305, 789-797(2008).
7. Mark, G. W., & Stephen, C. What happens next? event prediction using a compositional neural network model. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16). 2727-2733(2016).
8. Zhang, Y., & Chang, C. Integrating order information and event relation for script event prediction. In EMNLP. 57-67(2017).
9. Leveson, N. A new accident model for engineering safer systems. Safety Science. 42(4), 237-270 (2004).
10. LibenNowell, D., & Kleinberg, J. The link-prediction problem for social networks. Journal of The American Society For Information Science and Technology. 58(7), 1019-1031(2007).
11. Popescul, A., & Ungar, L. Statistical Relational Learning for Link prediction IJCAI workshop on learning statistical models from relational data(2003).
12. Xin, L., & Chen, H. Recommendation as link prediction in bipartite graphs: a graph kernel-based machine learning approach. Decision Support Systems. 54(2), 880-890 (2013).
13. Zhou, T., Lü L., & Zhang, Y. C. Predicting missing links via local information. European Physical Journal B. 71(4), 623-630, DOI: [http://doi.org/10.1140/epjb/e2009-00335-8\(2009\)](http://doi.org/10.1140/epjb/e2009-00335-8(2009)).

14. Thi, D. B., Ichise, R., & Le, B. Link Prediction in Social Networks Based on Local Weighted Paths. *Future Data and Security Engineering*. 21(19), 151-163(2014).
15. Simone, D., Josephine, M. T., Claudio, D., & Carlo, V. C. Common neighbours and the local-community-paradigm for topological link prediction in bipartite networks. *New Journal of Physics*. 17(11), 113037(2015).
16. Desislava, H., Anastasios, N., Chloë. B., & Cecilia, M. A multilayer approach to multiplexity and link prediction in online geo-social networks. *EPJ Data Science*. 5(1), DOI: [http://dx.doi.org/10.1140/epjds/s13688-016-0087-z\(2016\)](http://dx.doi.org/10.1140/epjds/s13688-016-0087-z(2016)).
17. Jiang, J., Liu, L. P., & Hassoun, S. Learning graph representations of biochemical networks and its application to enzymatic link prediction. *Bioinformatics(Oxford,England)*. 37(6), DOI: [http://doi.org/10.1093/BIOINFORMATICS/BTAA881\(2021\)](http://doi.org/10.1093/BIOINFORMATICS/BTAA881(2021)).
18. Zhou, Y., Wu, C., & Tan, L. Biased random walk with restart for link prediction with graph embedding method. *Physica A: Statistical Mechanics and its Applications*. 570, DOI: [http://doi.org/10.1016/J.PHYSA.2021.125783\(2021\)](http://doi.org/10.1016/J.PHYSA.2021.125783(2021)).
19. Zhang, Q., Wang, R., Yang, J., & Xue, L. Knowledge graph embedding by translating in time domain space for link prediction. *Knowledge-Based Systems*. 212, DOI: [http://doi.org/10.1016/j.knosys.2020.106564\(2021\)](http://doi.org/10.1016/j.knosys.2020.106564(2021)).
20. Wang, K., Liu, Y., Xu, X., & Sheng, Q. Z. Enhancing knowledge graph embedding by composite neighbors for link prediction. *Computing(prepublish)*.102, DOI: [http://doi.org/10.1007/s00607-020-00842-5\(2020\)](http://doi.org/10.1007/s00607-020-00842-5(2020)).
21. Catherine, A. B., Morgan, R. F., Christopher, M. D., & Peter S. D. An evolutionary algorithm approach to link prediction in dynamic social networks. *Journal of Computational Science*. 5(5), DOI: [http://doi.org/:10.1016/j.jocs.2014.01.003.\(2014\)](http://doi.org/:10.1016/j.jocs.2014.01.003.(2014)).
22. O'Madadhain, J., Hutchins, J., & Smyth, P. Prediction and ranking algorithms for event-based network data[J]. *ACM SIGKDD Explorations Newsletter*. 7(2), 23-30, DOI: [http://doi.org/10.1145/1117454.1117458\(2005\)](http://doi.org/10.1145/1117454.1117458(2005)).
23. Bai, S., Zhang, Y., Li, L., Shan, N., & Chen, X. Effective link prediction in multiplex networks: A Topsis method - ScienceDirect. *Expert Systems with Applications*.177, DOI: [http://doi.org/10.1016/J.ESWA.2021.114973\(2021\)](http://doi.org/10.1016/J.ESWA.2021.114973(2021)).
24. Zhang, Q. M. ,Lü, L., Wang, W. Q. , Yu, X., & Tao, Z. Potential theory for directed networks. *Plos One*, 8(2), DOI: [http://Doi.org/10.1371/journal.pone.0055437\(2013\)](http://Doi.org/10.1371/journal.pone.0055437(2013)).
25. Juszczyszyn, K., Musial, K., & Budka, M. Link Prediction Based on Subgraph Evolution in Dynamic Social Networks. 3rd IEEE International Conference on Social Computing. IEEE(2011).
26. Kipf, T. N., & Welling, M. Semi-Supervised Classification with Graph Convolutional Networks(2016).
27. Li, Y., Tarlow, D., & Brockschmidt, M. Gated Graph Sequence Neural Networks. *Computer Science*(2015).
28. Li, Z., Ding, X., & Liu, T. Constructing Narrative Event Evolutionary Graph for Script Event Prediction. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial*

29. Ma, Y., Wang, S., Aggarwal, C., & Tang, J. Graph Convolutional Networks with EigenPooling. Knowledge Discovery & Data Mining, 723-731(2019).

Figure Legends

Figure1-Event subgraph extraction based on graph collapse. Describes the process of extracting and collapsing event subgraphs.

Figure2- learning curve of optimized CSGCN (accuracy changes with training cycle). Represents the situation that the learning accuracy of the CSGCN optimization algorithm changes with the training cycle. The blue curve  represents the training accuracy. The orange curve  represents the validation accuracy.

Figure3-Comparison of accuracy and recall using different similarity comparison methods. The figure shows the comparison of training accuracy under different methods. The blue curve  represents the training accuracy, and the green curve  represents the recall rate.

Tables

Table1. Statistics of our datasets

Algorithm	Training	Development	TEST
Documents	830,643	103,583	103,805
NEEG	5,997,385	-	-
Subgraph of CSGCN	28000,	20,000	20,000

Table 2. algorithm accuracy comparison

Methods	Accuracy
Word2vec [Mikolov et al., 2013]	-
DeepWalk [Perozzi et al., 2014]	-
EventComp [Granroth-Wilding and Clark, 2016]	-
PairLSTM [Wang et al., 2017]	-
SGNN	52.45
SGNN+ PairLSTM	52.17
SGNN+ EventComp	54.15

SGNN+ EventComp+ PairLstm	54.93
CSGCN	59.24
CSGCN+ Eigen	60.57
CSGCN+ Attention	62.56

Competing

interests (mandatory)

The authors declare no competing interests.

Author contributions

F. B. and X. W. conceived the experiments. X.W. conducted the experiments, X.W. and Z. H. performed statistical analysis and figure generation. All authors reviewed the manuscript.