

## Risk Guarantee Prediction in Networked-Loans

Dawei Cheng<sup>1</sup>, Xiaoyang Wang<sup>2</sup>, Ying Zhang<sup>3</sup> and Liqing Zhang<sup>1\*</sup>

<sup>1</sup>MoE Key Lab of Artificial Intelligence, Department of Computer Science and Engineering,  
Shanghai Jiao Tong University, China

<sup>2</sup>Zhejiang Gongshang University, China

<sup>3</sup>University of Technology Sydney, Australia

dawei.cheng@sjtu.edu.cn, xiaoyangw@zjgsu.edu.cn, ying.zhang@uts.edu.au, zhang-lq@cs.sjtu.edu.cn

### Abstract

The guaranteed loan is a debt obligation promise that if one corporation gets trapped in risks, its guarantors will back the loan. When more and more companies involve, they subsequently form complex networks. Detecting and predicting risk guarantee in these networked-loans is important for the loan issuer. Therefore, in this paper, we propose a dynamic graph-based attention neural network for risk guarantee relationship prediction (DGANN). In particular, each guarantee is represented as an edge in dynamic loan networks, while companies are denoted as nodes. We present an attention-based graph neural network to encode the edges that preserve the financial status as well as network structures. The experimental result shows that DGANN could significantly improve the risk prediction accuracy in both the precision and recall compared with state-of-the-art baselines. We also conduct empirical studies to uncover the risk guarantee patterns from the learned attentional network features. The result provides an alternative way for loan risk management, which may inspire more work in the future.

### 1 Introduction

Network-guaranteed loan, a widespread financial phenomenon in East Asia, are attracting increasing attention from the regulators and banks. The existing credit criteria for loans are primarily aimed at major independent players that lag behind the demand from small and medium-sized enterprises (SMEs) [Jian and Xu, 2012]. In order to meet the bank's credit criteria, groups of enterprises back each other to enhance their financial credibility to obtain loans from commercial banks. When more and more enterprises are involved, they form complex directed-network structures [Meng, 2015; Niu *et al.*, 2018]. Thousands of guaranteed-networks with different complexities co-exist and evolve independently for a long time. This requires an adaptive strategy to effectively identify and eliminate any systematic crises [Smith, 2010].

Risk guarantee prediction is the cornerstone of risk management in guaranteed-loans for two reasons: 1) in pre-loan application phase, if we are able to predict the guarantee risk in the loan application, the loan issuer, e.g., banks, could take proper actions in advance, such as requiring the borrower provides more guarantors, reducing the loan grant amount, etc. The risk guarantee detection in the application phase is the first protector of risk management in guaranteed-loans. 2) Post-loan risk management. As the financial and guarantee status of loan network changes, it is vital to predict potential risk guarantee in post-loan risk management. The banks could secure the critical path in advance during risk diffusion across the network if the risk guarantee is predicted so that to avoid systemic loan crises.

Classic loan risk quantitative estimations mainly focus on companies. They aim to infer the default probability by machine learning-based models, such as logistic regression-based credit scorecards [Bravo *et al.*, 2015], decision tree and ensemble learning-based methods [Caire, 2004], neural network and advanced deep neural network [Sigrist and Hirschall, 2019]. Recently, [Niu *et al.*, 2018] point out the potential node and guarantee risk among complex loan networks and design a visual tool to highlight risk patterns. [Cheng *et al.*, 2019a] proposes a graph neural network to predict SMEs' loan risk by learning embeddings from loan networks. Although initial efforts have been made using empirical study [Meng, 2015] and visual analysis to understand the fundamental risks, there is little work of quantitative assessment on risk guarantees in loan networks.

Indeed, the guarantee risk prediction task can be treated as an edge classification task in complex temporal networks. Conventional edge classification task manually generated features before fed into classification methods. Recent developed graph neural networks (GNN) have shown advantages of automatic feature learning in various network-based learning tasks [Battaglia *et al.*, 2018], such as computer version [Niepert *et al.*, 2016], nature language processing [Yao *et al.*, 2019], relational signal process [Ioannidis *et al.*, 2019], chemical science [Coley *et al.*, 2019], recommend system [Wu *et al.*, 2019], etc. In financial literature, [Weber *et al.*, 2019] employ graph convolutional networks on anti-money laundering in bitcoin. [Lv *et al.*, 2019] proposed an auto-encoder based graph neural network to detect online financial frauds. Existing works achieve considerable improvements by learning re-

\*Corresponding Author

lations from networks by GNN. However, guarantee relation in loan networks is dynamically changed, and the hidden pattern exists temporally [Cheng *et al.*, 2019b], which requires the graph neural network could directly extract features from networks and dynamically adjust the importance upon the updates of the networks. Besides, the data sparsity and frequent changes of risk patterns also make the use of dynamic graph information very challenging.

Therefore, in this paper, we propose an end-to-end dynamic graph-based attention neural network (DGANN) to effectively predict risk guarantees by analyzing the records of networked-loans. We introduce the recurrent graph attention neural network to address the hidden features in complex loan networks. Then, we infer the guarantee risk probability by a prediction network that builds on the top of learned representations in an end-to-end manner. The experiment on real-world dataset show the superior performance of the proposed DGANN, and we explore to uncover the risk guarantee patterns by over six-months empirical studies.

The main contributions of the paper include:

- To the best of our knowledge, we provide the first attempt on predicting risk guarantees in dynamic loan-networks using an end-to-end graph-based temporal attention neural networks.
- We design and implement the dynamic graph-based attention neural network, which enables the model to directly learn temporal features from dynamic network structures in a recurrent graph neural layer. We also propose an attention architecture to automatically learn the importance of each hidden state of graph neural layer and demonstrate the effectiveness of each sub-modules.
- We thoroughly evaluate DGANN in over six months' empirical study. The research is deployed in a real-world risk management system. The result demonstrates that DGANN can effectively predict risk guarantees and discover risk patterns.

## 2 Preliminaries

### 2.1 Business Procedures

In order to apply loans, the borrower need to open an account and provide the bank with basic financial profiles. Banks may not issue the loan immediately because it is tough for SMEs to meet the existing bank's lending criteria, which is intended for big companies. Therefore, a small company must find other companies as an endorsement to guarantee security. The entire process is recorded in the bank's credit management system. As Figure 1 shows, there is often multiple guarantors per loan transaction, and a single guarantor can guarantee multiple loan transactions in the same time. Once the loan has been issued, SMEs usually receive the full loan amount immediately and start repaying it in regular installments until the end of the loan contract.

Thus, monitoring the risk status of guarantee relationship and SMEs are important for banks, as the loan risk may diffuse across the network along with the guarantee relationships. Figure 2a presents an example in over six-month empirical analysis in our collaborated financial institutions.

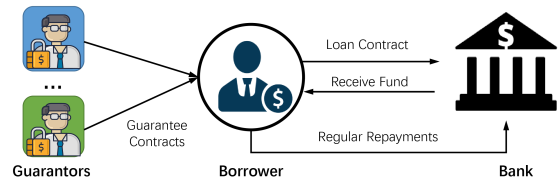


Figure 1: Guarantee loan process. The SME (borrower) that wishes to get a bank loan first sign guarantee loan contracts with guarantors before signing a bank loan contract. The company will repay the loan in a fixed installment.

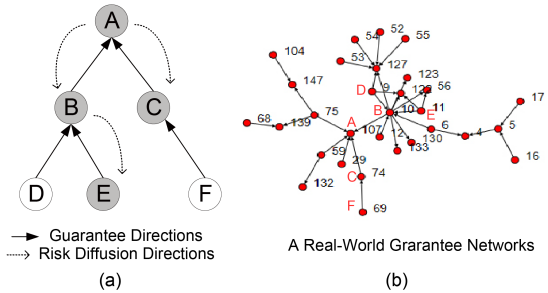


Figure 2: The illustration of risk guarantees. (a) an example of risk diffusion phenomenon observed during empirical studies. (b) presents the global networked view of the involved SMEs described in the left part.

Company A (a paper producer) failed to repay a loan of 0.7 million in December 2015 and subsequently delayed its guarantors (Company B and C, wood, and packaging manufacturers) after four months. Finally, the risk diffused through company B within six months to E (a paper distributor). Figure 2b shows the involved companies in a real-world loan network, which is dynamic and complex. If this situation is not immediately controlled, the risk may continue to spread wider across the network. Therefore, it is urgent for us to develop an adaptive approach to predict and detect risk guarantees effectively.

### 2.2 Problem Definition

Given a set of guaranteed-loan records in time series, the prediction model infers the possibility of risk guarantees. As described above, the probability of whether a guarantee is risk is temporally dynamic and spatially complex depends not only on the financial situation of the company itself but also on the situation of the guarantee and other companies from the temporal network. Therefore, we propose a dynamic graph-based attention neural network to predict the risk guarantees. Here, we formally define the problem of risk guarantee prediction.

**Definition 1** Guarantee. *If a company signs a contract to back another company's loan, we say that the guarantee relationship  $e$  is established.*

*A guarantee contains 1) guarantor  $v_s$ ; 2) borrower  $v_t$ ; 3) amount  $m$ , which indicates the quota of a guarantee; 4) start time  $t_s$ , which is the effective date of the contract; and 5) end time  $t_e$ , which indicates its expiry date.*

**Definition 2** Guarantee Network (or Graph). *A guarantee network (or Graph)  $GN$  is a directed graph  $G = (V, E)$ ,*

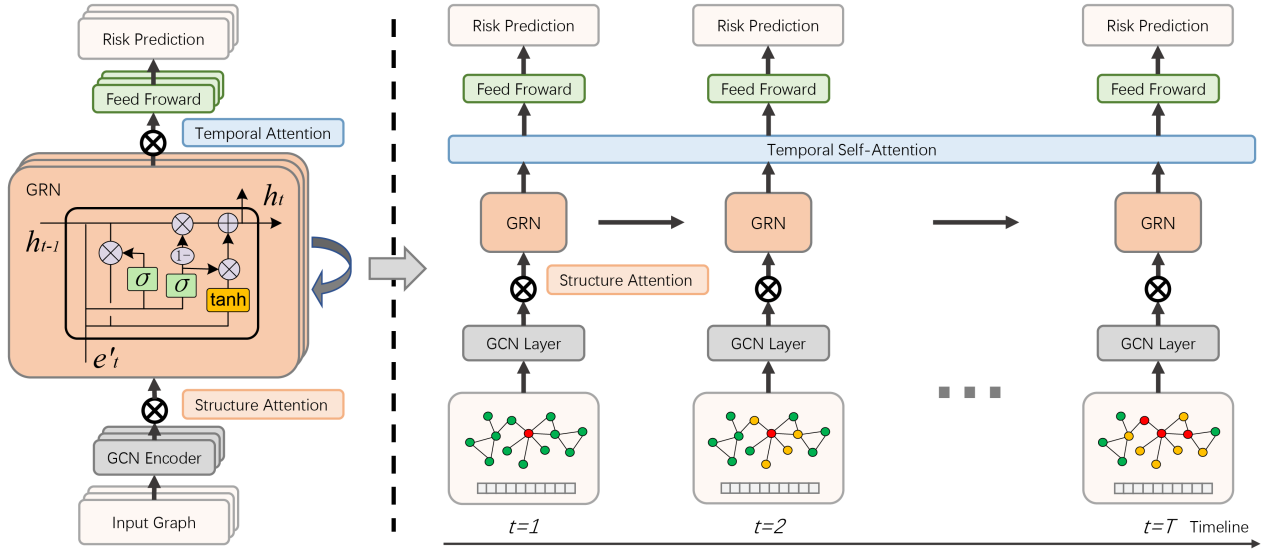


Figure 3: The illustration of the proposed dynamic graph attention neural network (DGANN) model. We first construct temporal graphs from raw loan and guarantee records. A graph convolution network (GCN) with structure attention is then designed on top of graphs to learn high-level guarantee representations. Afterwards, the learned representations are reshaped to vectors and fed into a graph recurrent neural network (GRN) network with temporal attention and prediction layer for risk guarantee prediction. Attentional weights are jointly optimized in an end-to-end mechanism with GCN and GRN modules.

where  $V = \{v_1, v_2, \dots, v_{|V|}\}$  is a set of SMEs, and  $E = \{e_1, e_2, \dots, e_{|E|}\}$  is a set of guarantee relationships (edges).

**Definition 3 Risk Guarantee.** *If a borrower defaults, its guarantor fails to take the legal obligation to repay the guaranteed amounts, we define the guarantee is a risk guarantee.*

We now formalize our default predicting problem as:

*Given a sequential of loan, repayment records, guarantee relationships  $\{e\}$ , and time period  $t_i$ , for every valid guarantee we want to infer the possibility of whether the guarantor will default if its borrower defaults, based on the loan and guarantee status from  $t_1$  to  $t_i$ . The objective is to achieve high accuracy of the risk guarantee prediction, as well as explore the risk patterns of guaranteed loans.*

### 3 The Proposed Approach

In this section, we first introduce the general framework of the dynamic graph attention neural network (DGANN) and then present the input of the model. After that, we present the graph convolution encoder with a structural attention network, graph recurrent network with temporal attention and prediction layers. Lastly, we introduce the optimization strategy of the proposed methods.

#### 3.1 Model Architecture and Inputs

Figure 3 shows the general architecture of our proposed dynamic graph attention neural network for risk guarantee prediction. Generally, the model includes three parts: 1) A graph convolution network (GCN) with structure attention, which takes temporal guarantee networks  $\mathcal{G} = \{G_1, G_2, \dots, G_T\}$  with attributes  $G = (V, E)$  as input, and performs a high order representation of the graphs  $\mathcal{X} = \{x_1, x_2, \dots, x_T\}$  as output. 2) A graph recurrent network (GRN) with temporal

attention, which extracts the edge attributes from outputs of the GCN, and performs a recurrent neural network to learn the temporal patterns from temporal network representations. An attentional network builds an attention mechanism on the output of the sequential hidden layer to learn the importance of temporal information automatically. 3) A prediction layer is introduced to estimate the risk probability of guarantees with the view of global guarantee networks and sequential information.

In a given timestamp, the input of our proposed model includes the guarantee network  $G = (V, E)$ , where node  $v \in V$  represents the company (SME) and edge  $e \in E$  denotes the guarantee relationship. Our task is to predict risk guarantees  $e$  from given inputs. During feature engineering, we extract attributes of nodes  $v$  from two folds: 1) the SME’s basic profiles, which reflects the capability, condition, and stability of a company. We utilize the enterprise-scale, registered capital, employee numbers and other information as the corporation’s basic profile. 2) Loan behaviors, the SME’s loan behavior feature in the current period. It contains loan times, loan amount, credit history, default times, default amount, etc. Edge attributes  $e$  includes the borrower, the guarantor, guarantee start time, end time, and the guaranteed amount.

#### 3.2 Attention-based Graph Convolution

The structural attention-based graph convolution network (GCN) aims to learn the latent spatial representations of each nodes, which could preserve network information. Given the guarantee networks with nodes features  $V = \{v_1, v_2, \dots, v_{|V|}\}$ , and  $v_i \in \mathbb{R}^F$ , where  $F$  is the dimension of node features, and  $|V|$  is the number of nodes. The attention-based GCN produces hidden representation of nodes  $V' = \{v'_1, v'_2, \dots, v'_{|V|}\}$ ,  $v'_i \in \mathbb{R}^{F'}$ .

In particular, for a given node  $v_i$ , we first conduct convolutional operation on the features matrix, which is constructed by  $v_i$  and its neighborhood nodes  $v_j \in \mathcal{N}_i$ , to learn high-level features.  $\mathcal{N}_i$  denotes the neighborhood collection of node  $v_i$ . We then perform a shared attention mechanism on the nodes to compute the structural attention coefficients between two nodes  $v_i$  and  $v_j$ :

$$\alpha_{i,j} = \frac{\exp(\text{Conv}(W_c v_i, W_c v_j))}{\sum_{k \in \mathcal{N}_i} \exp(\text{Conv}(W_c v_i, W_c v_k))} \quad (1)$$

where  $\text{Conv}$  denotes a convolution layer, and  $W_c \in R^{F' \times F}$  is a weighted matrix on node features which are then fed to convolution networks.  $\alpha_{i,j}$  denotes the attention coefficients between node  $v_i$  and  $v_j$ .

In our implementation, we perform multi-head attention on the output layer of the GCN layer to stabilize the learning process of the shared attention mechanism. Particularly,  $K$  independent attention mechanism executed on the Equation 1, and then we employ averaging function on the output feature representation of  $v_i$ . Therefore, we reach the output feature representation as:

$$v'_i = \sigma \left( \frac{1}{K} \sum_{k=1}^K \sum_{j \in \mathcal{N}_i} \alpha_{i,j}^k \text{Conv}(W_c^k v_j) \right) \quad (2)$$

where  $K$  is the number of attention heads, and  $\sigma$  denotes the sigmoid function,  $\alpha_{i,j}^k$  and  $W_c^k$  are the corresponding attentional coefficient and weight matrix in  $K$ -th independent attention mechanism, respectively. The learned  $V' = \{v'_1, v'_2, \dots, v'_{|V|}\}$  are the input of downstream graph recurrent networks.

### 3.3 GRN and Temporal Attention

For each time stamp, we get the high order node representations from the attention-based GCN layer. As the main task of our work is to predict the high risk guarantees, we then construct edge representations  $e_i^t = \{v_{i,s}^t, v_{i,t}^t, m_i^t\}$  for time stamp  $t$  based on the learned high-order node representations, where  $v_{i,s}^t$  and  $v_{i,t}^t$  denotes the high-order representations of source and target node of edge  $e_i$ , and  $m_i^t$  denotes the guarantee amount. Then, we denote the GCN updated edge feature as  $e_t^t = \{e_{1,t}^t, e_{2,t}^t, \dots, e_{|E|,t}^t\}$ , where  $|E|$  is the number of edges. Thus, we get the input of GRN layer  $e' = \{e'_1, e'_2, \dots, e'_T\}$ ,  $T$  denotes the number of time stamp.

Then, we perform recurrent neural network on the top of updated edge representations. Specifically, we employ gated recurrent unit (GRU) [Chung *et al.*, 2014] with the input sequences  $e' = \{e'_1, e'_2, \dots, e'_T\}$  and produces output sequences  $h = \{h_1, h_2, \dots, h_T\}$ , as shown in Figure 3. The fully gated recursive unit model can be formulated as below:

$$z_t = \sigma(W_z[h_{t-1}, e'_t]) \quad (3)$$

$$r_t = \sigma(W_r[h_{t-1}, e'_t]) \quad (4)$$

$$\tilde{h}_t = \tanh(W_h[r_t * h_{t-1}, e'_t]) \quad (5)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (6)$$

where  $r_t$  and  $z_t$  denote the reset and update gates of the  $t^{\text{th}}$  object respectively.  $\tilde{h}_t$  represents representation of candidate hidden layer.  $W_*$  are the weights dynamically updated during the model training phase.

To further capture temporal patterns in a dynamic network, we then perform a temporal self-attention layer. The input of this layer is the outputs of GRU cell denoted as  $h = \{h_1, h_2, \dots, h_T\}$ . The output of temporal attention layer is a new representation of sequences for  $e'$  in each time step, we denoted it as  $u = \{u_1, u_2, u_T\}$ . Then, the temporal self-attention is defined as:

$$u_t = (\|_{j=1}^T \beta_{t,j} W_t h_j)$$

$$\beta_{t,j} = \frac{\exp(\text{NN}(W_n h_t, W_n h_j))}{\sum_{i=1}^T \exp(\text{NN}(W_n h_t, W_n h_i))} \quad (7)$$

where  $\beta_{t,j}$  is the temporal attention weights between time stamp  $t$  and  $j$ . NN denotes a feed forward neural network layer with ReLU as activation function.  $\|$  represents vector concatenation,  $W_*$  are the training parameters in temporal attention network.

The key objective of this self-attention recurrent layer is to capture the temporal variations in the graph edge representations, which includes high-order node feature from structure attention GCN layer, over multiple time steps. Thus, the learning sequence  $u = \{u_1, u_2, u_T\}$  by a pipeline of temporal and structure attention could represent the dependencies between various representations of network structures across different time steps.

### 3.4 Prediction and Optimization

The risk guarantee prediction task takes the high-level representations of edges  $u_t$  in timestamp  $t$ , which is a sequential vector learned by the temporal attention based GRN layer and aims to infer the probability of whether a guarantee is at risk. Thus, the risk guarantee prediction task is a classification task, and the loss function is the likelihood defined as follows:

$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\text{pred}(u_i : \theta)) + \lambda(1 - y_i) \log(1 - \text{pred}(u_i : \theta))] \quad (8)$$

where:  $u_i$  denotes the representation of the  $i$ -th guarantee representations, which is the output of attentional GRN.  $\lambda$  indicates the sample weight according to the biased distribution of fraud and legitimate records.  $y_i$  denotes the label of  $i$ -th records, which is set to 1 if the record indicates risk and 0 otherwise.  $\text{pred}(u_i)$  is the prediction function that maps  $u_i$  to a real-valued score, indicating the probability of whether the current guarantee is risk or not. We implement  $\text{pred}(u_i : \theta)$  with another shallow neural network (two-layer Rectified Linear Units and one-layer Sigmoid).

The proposed dynamic graph attention neural network (DGANN) can be optimized by the standard stochastic gradient descent process. In the implementation, we apply the Adam algorithm as the optimizer. We set the initial learning rate to 0.001, and the batch size to 128 by default.

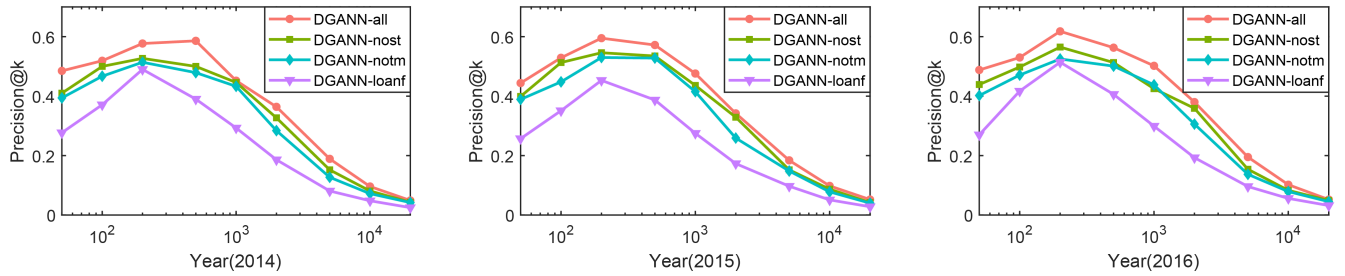


Figure 4: Precision@k results in risk guarantee prediction task, compared by the proposed method and its sub-modules from 2014 to 2016. In all baselines, DGANN-nost shows the most competitive. DGANN-all constantly surpasses all other methods.

## 4 Experiments

### 4.1 Experimental Settings

**Datasets** We evaluate the proposed DGANN on the real-world dataset from a major financial institution in East Asia, from Jan, 01, 2013 to Dec 31, 2016. It includes 112872 nodes (companies), with 124957 edges (guarantees). Given SMEs who default loans, as the guarantors are required to repay it, if the guarantors also delinquent the loan subsequently, we mark the guarantee to 1 as the ground-truth risk guarantee. For the rest part, we set the relationship to 0 as normal guarantees.

**Compared Methods** We compare our proposed DGANN with seven baselines: Graph Factorization (GF) [Ahmed *et al.*, 2013], node2vec [Grover and Leskovec, 2016], GCN [Niepert *et al.*, 2016], GATs [Velickovic *et al.*, 2018], SEAL [Zhang and Chen, 2018], RRN [Palm *et al.*, 2018], GRNN [Ioannidis *et al.*, 2019]. For network representative learning methods, we concatenate the linked node features as guarantee representation and employ the prediction network for comparison. Our model has several variations: DGANN-nost/DGANN-notemp, in which the structural graph attention, temporal recurrent attention are not used. For DGANN-loanf, we only leverage the loan features in the prediction network of DGANN, which is a multi-layer perception with two-layer ReLU and one layer sigmoid.

We set embedding dimension to 128,  $\lambda$  are determined by the data distribution, which is set to 16. The parameters of baseline methods are initialized by their recommended settings. We evaluate the performance of the proposed model by AUC (Area Under the Curve) and Precision@k. Precision@k means the predicted ratio of the true risk guarantees in top k guarantees.

### 4.2 Risk Guarantee Prediction

In this section, we report out the results of risk guarantee prediction, in which the records of the year 2013 are employed as the training data. We then predict the risk guarantees in a recurrent manner for the next three years. Table 1 presents the average AUC of all the baselines.

As we can see, GCN performs better than GF and node2vec, demonstrate the necessity of structural information in risk guarantee prediction. Moreover, the AUC GAT is slightly higher than GCN, which proves the effectiveness of attention mechanism in graph neural networks. Of all

	AUC(2014)	AUC(2015)	AUC(2016)
GF	0.74296	0.74412	0.75512
node2vec	0.78917	0.78832	0.79580
GCN	0.79440	0.80054	0.80718
GATs	0.80175	0.80403	0.81033
SEAL	0.81143	0.81447	0.81957
RRN	0.81743	0.81906	0.82246
GRNN	0.82617	0.82945	0.83202
DGANN-loanf	0.75058	0.75739	0.76570
DGANN-notm	0.80814	0.81190	0.81436
DGANN-nost	0.82668	0.82937	0.83398
<b>DGANN-all</b>	<b>0.84680*</b>	<b>0.84227*</b>	<b>0.85059*</b>

Table 1: Comparison of the risk guarantee prediction accuracy.

baselines, GRNN shows the most competitive because the graph structures are preserved recurrently in temporal aspects. In the variations, DGANN-loanf is not satisfactory, which indicates the essential of extracting temporal structural patterns from loan networks. DGANN-nost is superior than DGANN-notm and GRNN; the importance of temporal attention is strongly proved. It also should be noted that the improvements by temporal recurrent network are more significance than the pure structural ones in graph neural network. DGANN-all surpasses all other methods, including state-of-the-art baselines.

### 4.3 Effects of Attentional Sub-Modules

Figure 4 presents the precision@k results of the proposed DGANN and its sub-modules. As the classic approach, which only employs loan features in risk guarantee prediction tasks, DGANN-loanf achieves sub-optimal performance by employ a multi-layer perception prediction network of our proposed method. By leveraging the graph structures, the DGANN-notm gains significant improvements compared with the method with only loan features, varying from 5-10% improvements, which indicates the importance of preserving network information in risk guarantee prediction. Moreover, the performance of precision@k is improved slightly by DGANN-nost, which demonstrates that temporal patterns play a vital role in the time-sequential prediction task. Both DGANN-nost and DGANN-notm are better than DGANN-



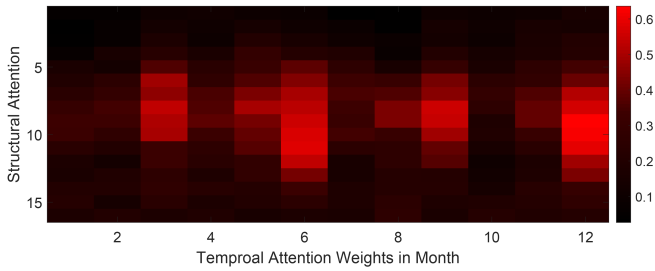


Figure 5: Case study results of the attentional coefficients. Temporally: the last month of each quarter is highlighted. Structurally: edges with seven to twelve adjacent guarantees are highlighted.

loanf, which validate the intuitions of the proposed method.

DGANN-all consistently surpasses all other methods. The performance gains an average of 5% improvements from the most competitive baselines. The experiment results strongly demonstrate the effectiveness of each sub-module (graph convolution network with structural attention and graph recurrent network with temporal attention).

#### 4.4 Case Studies

In DGANN, to address the complicated structural and temporal patterns in networked-loans, we propose both structural graph attention and recurrent temporal attention to jointly learn the temporal graph patterns in risk guarantee prediction task. In this case study, we explore to visualize the attention weights to validate the effects in learning different importance in temporal networks. For each edge (guarantee), we layout the attention weights by its month and number of adjacent edges in Figure 5 as a matrix format. In the column aspect, the attention weights are highlighted at the end of each quarter, especially in Jun and Dec. In row aspect, an edge with approximately seven to twelve adjacent edges is generally remarked. The result proves the essence of preserving temporal graph information in risk guarantee prediction.

Figure 6 reports the results of empirical studies. In order to validate the effects of the above-learned attention weights, we first report the average risk ratio of guarantees in the historical dataset. As Figure 6a shows, the risk ratio reaches a peak at the end of each quarter. We then investigate the number of guarantee state changes and report out the average curves in Figure 6b. Interestingly, both of them perform periodically every three months, which is consistent with the observation of temporal attention weights. Moreover, we plot the distribution of the risk guarantee ratio according to the number of adjacent edges. The ratio of guarantees with seven to twelve adjacent edges is obviously higher than the rest parts, which also powerfully demonstrates the effects of structural attention.

## 5 Related Work

*Graph Neural Network:* Network science has been considerably improved by the automatic feature learning on graphs by graph neural network [Battaglia *et al.*, 2018], including network embedding [Grover and Leskovec, 2016], graph convolution network [Niepert *et al.*, 2016], graph attention net [Velickovic *et al.*, 2018], etc. Recently, temporal models

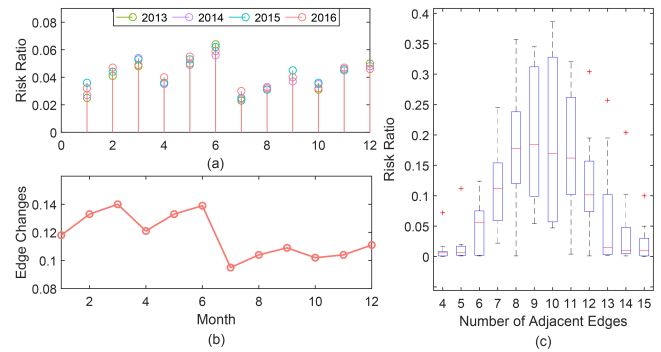


Figure 6: Empirical analysis of risk guarantees.

have been studied to address the learning task in dynamic networks. [Palm *et al.*, 2018] proposed a recurrent relation network on multi-step relational inference. [Ioannidis *et al.*, 2019] studies on dynamic learning multi-relational data by a recurrent graph neural network. Our work develops from a similar intuition but further propose attention mechanism in both temporal and structural aspects, which significantly improves the risk prediction accuracy.

*Risk Control in Loans:* Classic approaches mainly employ data-driven machine learning models on loan risk control, and have been extensively studied [Fitzpatrick and Mues, 2016]. Recently, the advances in deep models and graph learning improve the capability of risk control in loans. For example, [Sigrist and Hirschall, 2019] employs a gradient tree-boosted model for default prediction, which is more accurate than shallow models. [Cheng *et al.*, 2019a] proposed a graph attention neural network to predict delinquent loans by learning from network features. However, there is little work focus on risk guarantee prediction in networked loans. To the best of our knowledge, this is the first study on risk guarantee prediction by preserving both structural and temporal attention in an end-to-end graph neural network.

## 6 Conclusion

In this paper, we propose a dynamic graph-based attention neural network (DGANN) for risk guarantee relationship prediction in networked-loans. In particular, we summarize the model novelty as “structural” and “temporal” availability by presenting 1) graph convolution network with structural attention and 2) graph recurrent layer with temporal attention in an end-to-end framework. This is the first work in which both structural and temporal attention are proposed in a risk guarantee prediction task. During experiments, our models achieve considerable improvements compared with state-of-the-art baselines. Moreover, we conduct empirical studies on the learned attention weights after the model deployed in our collaborated financial institution. The results strongly prove the essential and effectiveness of preserving temporal and structural information in risk guarantee prediction.

## Acknowledgments

The work is supported by the National Key R&D Program of China (2018AAA0100704) and the China Postdoctoral Science Foundation.

## References

- [Ahmed *et al.*, 2013] Amr Ahmed, Nino Shervashidze, Shравan Narayanamurthy, Vanja Josifovski, and Alexander J Smola. Distributed large-scale natural graph factorization. In *Proceedings of the 22nd international conference on World Wide Web*, pages 37–48. ACM, 2013.
- [Battaglia *et al.*, 2018] Peter Battaglia, Jessica Blake Chandler Hamrick, Victor Bapst, Alvaro Sanchez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. 2018.
- [Bravo *et al.*, 2015] Cristián Bravo, Lyn C Thomas, and Richard Weber. Improving credit scoring by differentiating defaulter behaviour. *Journal of the Operational Research Society*, 66(5):771–781, 2015.
- [Caire, 2004] Dean Caire. Building credit scorecards for small business lending in developing markets. *Bannock Consulting*, 2004.
- [Cheng *et al.*, 2019a] Dawei Cheng, Yi Tu, Zhenwei Ma, Zhibin Niu, and Liqing Zhang. Risk assessment for networked-guarantee loans using high-order graph attention representation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 5822–5828. AAAI Press, 2019.
- [Cheng *et al.*, 2019b] Dawei Cheng, Yiyi Zhang, Fangzhou Yang, Yi Tu, Zhibin Niu, and Liqing Zhang. A dynamic default prediction framework for networked-guarantee loans. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 2547–2555. ACM, 2019.
- [Chung *et al.*, 2014] Junyoung Chung, Caglar Gulcehre, Kyunghyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning, December 2014*, 2014.
- [Coley *et al.*, 2019] Connor W Coley, Wengong Jin, Luke Rogers, Timothy F Jamison, Tommi S Jaakkola, William H Green, Regina Barzilay, and Klavs F Jensen. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical science*, 10(2):370–377, 2019.
- [Fitzpatrick and Mues, 2016] Trevor Fitzpatrick and Christophe Mues. An empirical comparison of classification algorithms for mortgage default prediction: evidence from a distressed mortgage market. *European Journal of Operational Research*, 249(2):427–439, 2016.
- [Grover and Leskovec, 2016] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864. ACM, 2016.
- [Ioannidis *et al.*, 2019] Vassilis N Ioannidis, Antonio G Marques, and Georgios B Giannakis. A recurrent graph neural network for multi-relational data. In *ICASSP 2019*, pages 8157–8161. IEEE, 2019.
- [Jian and Xu, 2012] Ming Jian and Ming Xu. Determinants of the guarantee circles: The case of chinese listed firms. *Pacific-Basin Finance Journal*, 20(1):78–100, 2012.
- [Lv *et al.*, 2019] Le Lv, Jianbo Cheng, Nanbo Peng, Min Fan, Dongbin Zhao, and Jianhong Zhang. Auto-encoder based graph convolutional networks for online financial anti-fraud. In *CIFER*, pages 1–6. IEEE, 2019.
- [Meng, 2015] Xinhai Liu Xiangfeng Meng. Credit risk evaluation for loan guarantee chain in china. 2015.
- [Niepert *et al.*, 2016] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In *International conference on machine learning*, pages 2014–2023, 2016.
- [Niu *et al.*, 2018] Zhibin Niu, Dawei Cheng, Liqing Zhang, and Jiawan Zhang. Visual analytics for networked-guarantee loans risk management. In *2018 IEEE Pacific Visualization Symposium (PacificVis)*, pages 160–169. IEEE, 2018.
- [Palm *et al.*, 2018] Rasmus Palm, Ulrich Paquet, and Ole Winther. Recurrent relational networks. In *Advances in Neural Information Processing Systems*, pages 3368–3378, 2018.
- [Sigrist and Hirsenschall, 2019] Fabio Sigrist and Christoph Hirsenschall. Grabit: Gradient tree-boosted tobit models for default prediction. *Journal of Banking & Finance*, 102:177–192, 2019.
- [Smith, 2010] Stan W Smith. An experiment in bibliographic mark-up: Parsing metadata for xml export. In *Proceedings of the 3rd. annual workshop on Librarians and Computers (LAC’10), Reginald N. Smythe and Alexander Noble (Eds.)*, volume 3, pages 422–431, 2010.
- [Velickovic *et al.*, 2018] Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. In *ICLR*, 2018.
- [Weber *et al.*, 2019] Mark Weber, Giacomo Domeniconi, Jie Chen, Daniel Karl I Weidele, Claudio Bellei, Tom Robinson, and Charles E Leiserson. Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. *arXiv preprint arXiv:1908.02591*, 2019.
- [Wu *et al.*, 2019] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. Session-based recommendation with graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 346–353, 2019.
- [Yao *et al.*, 2019] Liang Yao, Chengsheng Mao, and Yuan Luo. Graph convolutional networks for text classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 7370–7377, 2019.
- [Zhang and Chen, 2018] Muhan Zhang and Yixin Chen. Link prediction based on graph neural networks. In *Advances in Neural Information Processing Systems*, pages 5165–5175, 2018.