Critical Firms Prediction for Stemming Contagion Risk in Networked-Loans through Graph-Based Deep Reinforcement Learning

Dawei Cheng[1,4,5], Zhibin Niu[2*], Jianfu Zhang[3,4], Yiyi Zhang[3,4], Changjun Jiang[1,5]

[1] Department of Computer Science and Technology, Tongji University, Shanghai, China
[2] College of Intelligence and Computing, Tianjin University, Tianjin, China
[3] Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
[4] Key Laboratory of Artificial Intelligence, Ministry of Education, Shanghai, China
[5] Shanghai Artificial Intelligence Laboratory, Shanghai, China
{dcheng, cjjiang}@tongji.edu.cn, zniu@tju.edu.cn, {c.sis,yi95yi}@sjtu.edu.cn

Abstract

The networked-loan is major financing support for Micro, Small and Medium-sized Enterprises (MSMEs) in some developing countries. But external shocks may weaken the financial networks' robustness; an accidental default may spread across the network and collapse the whole network. Thus, predicting the critical firms in networked-loans to stem contagion risk and prevent potential systemic financial crises is of crucial significance to the long-term health of inclusive finance and sustainable economic development. Existing approaches in the banking industry dismiss the contagion risk across loan networks and need extensive knowledge with sophisticated financial expertise. Regarding the issues, we propose a novel approach to predict critical firms for stemming contagion risk in the bank industry with deep reinforcement learning integrated with high-order graph message-passing networks. We demonstrate that our approach outperforms the state-of-the-art baselines significantly on the dataset from a large commercial bank. Moreover, we also conducted empirical studies on the real-world loan dataset for risk mitigation. The proposed approach enables financial regulators and risk managers to better track and understands contagion and systemic risk in networked-loans. The superior performance also represents a paradigm shift in addressing the modern challenges in financing support of MSMEs and sustainable economic development.

Introduction

The networked-loans, a kind of understudied financial network, increased their significant role in shaping the systemic risk after the 2008 global crisis. The regulation and risk management attracted massive attention from the academic and banking industry as MSMEs financing support is crucially important for sustainable economic development. Thus, under the economic stimulus program, they are encouraged to guarantee each other to apply for bank loans (Mian and Sufi 2009) and thereby form complex loan networks. This networked-loans differentiates itself from each node representing a borrower company to that in the usual financial networks where each node presents a financial institute. Such

*Corresponding Author.

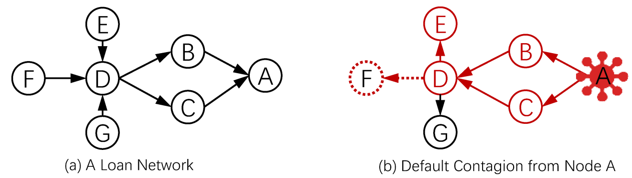


Figure 1: Illustration of risk contagion from Node A.

operation can enhance financial security while introducing more potential risks (Cheng et al. 2018; Wang, Zhang, and Yang 2020). For example, when there is an inevitable negative shock to the nodes, the network structure may magnify the defaults, resulting in large-scale or even systemic risk (Dumitrescu et al. 2022). Only after the capital chain rupture, the financial experts can analyze the cause.

In the banking industry, accidental default is tolerable in most cases; large-scale default or systemic financial crisis is to be firmly prevented. Figure 1 illustrates an example of risk contagion in networked-loans. The default of node (Firm) A, as can be seen, could spread to the whole network. Nodes B and C need to repay the bank's loan from Node A as the result of its default. Therefore, the risk spreads across with the guarantee direction from Firm A to Firm B and C. The banking experts usually rely on the empirical research methodology and have difficulty in foreseeing how to stem the contagion risk effectively (Sui, Li, and Chen 2020; Cheng et al. 2019b). Therefore, in the post-pandemic era, monitoring and mitigating the risk of contagion is more pressing than ever for a healthy economic environment.

Previous work on credit risk modeling utilizes statistics and regression methods as early as the 1950s (Baesens et al. 2003). Then, later in the 1980s, the Basel Committee issued a series of recommendations on banking regulations (Basel Capital Accord I, II, and III) to enhance the understanding of critical regulatory issues and improve macro-prudential oversight (Eubanks 2010; Acemoglu, Ozdaglar, and Tahbaz-Salehi 2015; Georgios 2019). The principles have been widely accepted by banks around the world (Montgomery 2005). Recently, machine learning and deep learning-based approaches have gained success in various sectors (Tu et al. 2020; Zhang et al. 2022), which have also been employed for risk assessment in the current financial situation (Addo, Guegan, and Hassani 2018; Chernyshev 2020) . (Niu et al. 2020; Cheng, Niu, and Zhang 2020) has shown the advantages of utilizing deep models to understand the comprehensive risk in complex financial networks through leveraging graph learning methods to address the challenges for risk rating in networked-loans. The considerable potential of deep graph learning in regulating systemic crises is also proved in (Cheng et al. 2022). However, existing approaches for risk management in networked-loans are targeted at individual or group assessments. They are limited in providing constructive insights to mitigate risk contagion.

Therefore, in this paper, we present a novel graph-based deep reinforcement learning (DRL) approach to Stem Contagion Risk by Predicting critical Firms (named SCRPF) in networked-loans. In particular, we use the framework of deep reinforcement learning to learn from historical traces and train the policy network to generate critical firm lists. To preserve contagion representations, SCRPF leverages a high-order message passing graph neural network to encode the graph-structured risk diffusion behavior directly on contagion chains, which no longer requires expensive handcraft feature engineering or sophisticated financial knowledge. Then, the DRL agent uses the policy of a trained graph attentional neural network to predict a priority list of critical firms for networked-loans. Finally, we demonstrate remarkable performance improvements contrasted with the baseline models through extensive experiments. By predicting the critical firms more effectively, we enable financial regulatory and loan managers with swifter preventive measures in advance against systemic financial crises. Empirical studies on real-world loan management scenarios are also carried out using the proposed approach. The results show that SCRPF could significantly avoid the potential financial losses for commercial banks with high efficiency. We summarize our paper's main contributions as follows:

1. To the best of our knowledge, this work is the first approach that provides a practical solution by identifying critical firms to stem the contagion risk so that to mitigate systemic financial crises through graph-based deep reinforcement learning.
2. We design and implement the DRL framework, named SCRPF, which naturally represents the contagion process by message-passing graph neural network. The model to learn from networked-loan data directly demonstrates the effectiveness in inferring the multi-level contagion stemming, which is the primary cause of systemic loan risk in the banking industry.
3. We thoroughly evaluate the proposed approach. We compared our approach with existing baselines and reported a state-of-the-art performance. We also conducted empirical studies in real-world risk control applications. The result demonstrates that our method could be invaluable in offering constructive evidence to financial regulators.

Preliminaries

Networked-loan Background

In a number of countries, the networked-loan is a unique secured financing model for MSMEs. In principle, the banks prefer to approve loans to large businesses and unwillingly to MSMEs, who usually lack securities like fixed assets. Helping them to raise money is vital as they can increase the employment rate and economic vitality. Under some simulation programs, the MSMEs are allowed to find guarantors to back their loan applicants. The guarantor takes a legal obligation to repay the loan if its borrower defaults (fails to repay the loan). When more and more firms are engaged, they form complex networks, which bring significant risk management challenges for both the regulatory and banks. A default may spread across the network along the direction of guarantee relationships (contagion risk) and lead to large-scale loan defaults or even systemic financial crises.

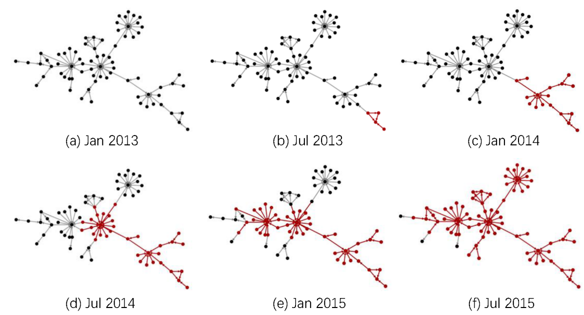


Figure 2: Real-world Systemic Defaults by Risk Contagion.

Various rule-based risk evaluation tools were applied to the borrowers. Such approaches are usually designed for large businesses with the prerequisite that they are independent. Thus, it is inaccurate and even error-prone to assess the small and medium enterprises loan qualifications in the networked-loans as they are interlinked together. They are more secure during the economy's rapid growth but more fragile during business slowing periods.

Risk Contagion in Networked-loans

As illustrated in Figure 1, the default of Node A may diffuse to B and C. Under the same circumstances, the default of Nodes B and C causes the failure of Node D. The same process is followed by E and is leading default for F. If most firms in network-loans default their borrowing, a systemic crisis would outbreak and cause massive financial damage for the lending bank and even the national economy.

Figure 2 shows a real-world case where an accidental corporate default leads to large-scale systemic crises. The red nodes (firms) denote entities that failed to repay loans (i.e., defaulted), and the black nodes are healthy entities. It is clear that in January 2013, all firms were in a healthy situation, which repays loans regularly according to loan contracts. A small number of firms (4 out of 86 firms in the entire network) failed the repayments in July 2013. It caused another 13 firms failures in January 2014. Then, the risk diffused across the loan network like wildfire, which caused 17 more firms to default in July 2014, and another 14 firms in January 2015. Eventually, 83 firms failed to repay their loans in July 2015, accounting for 96.5% of the default ratio in the number of firms and 98.2% failing in the loan amount.

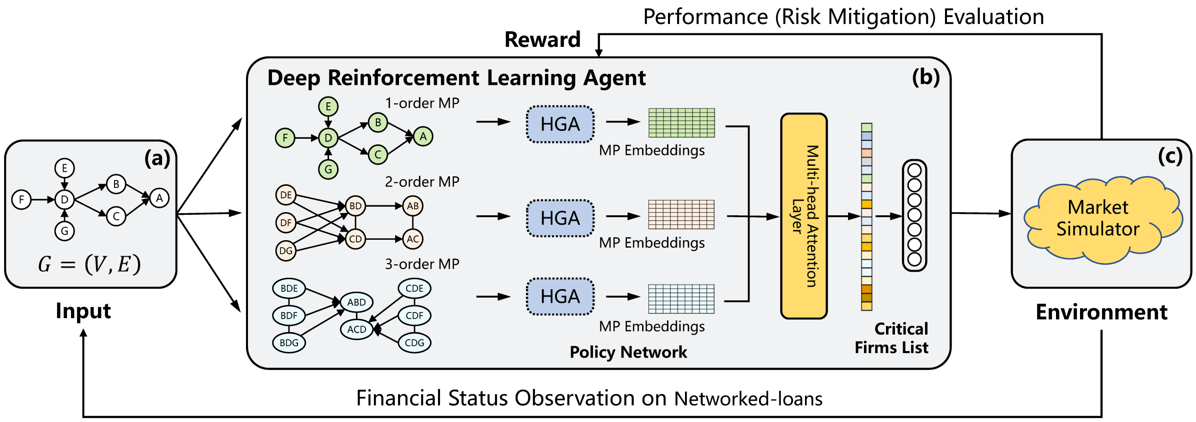


Figure 3: The architecture of our proposed method SCRPF: stemming contagion risk by predicting critical firms in networked-loans. (a) The original input of loan networks; (b) The illustration of deep reinforcement learning agent, which includes contagion chain message passing (MP) process, the policy network with high-order MP graph attentional neural layer and the policy output layer that produce the critical firm's list; (c) display environment of our proposed framework which generates rewards for different policies and simulates the financial status on networked-loans.

The Proposed Approach

Problem Formulation and Architecture Overview

Given the loan and guarantee relationship records, which could be constructed as loan networks $G = (V, E)$, each node $\{v_i \in V\}$, $V = \{v_1, v_2, \cdots, v_{|V|}\}$, denotes the micro-, small and medium enterprise (MSME), which has a probability $p(v_i)$ of loan defaults. The edges $E = \{e_{ij}\}, (i, j) \in V$, represent the guarantee relationship from node v_i to v_j, and the value of $p(e_{ij})$ denotes the probability of risk contagion to node v_j if v_i default its loan. The symbol $|V|$ denotes the number of nodes, and $|E|$ represents the number of edges in the network. Then, the main purpose of this work is to infer the most critical nodes, which have a higher probability of defaults, and, simultaneously, cause a larger cascade contagion failure across the network. The detailed problem target computation mechanism is reported in the agent's reward design section.

The general architecture of SCRPF is based on the framework of deep reinforcement learning. Figure 3 shows the DRL agent, which is composed of high-order message passing (MP) layer and graph attentional neural network. During the working process, SCRPF takes loan networks as inputs and coverts them into deep graph representations. Then the policy network produces a critical firms list as output and the environment generates the reward for agent training. More specifically, we construct the complex loan networks using raw guarantee records. Our dataset includes the guarantor, borrower, guarantee amount, the contract start time, and end time in the guarantee contract. We create an edge directed from the guarantor to the borrower and consequently construct complex loan networks. Then, we extract contagion chains from each network by a breadth-first search (BFS) algorithm and construct the multi-level orders of message passing networks, which will be introduced in the rest section. The agent of SCRPF learns deep representations from contagion chains by the proposed high-order message passing graph neural network and produces a priority list of critical firms. In each message passing layer, node features are aggregated from their adjacent nodes, which are sources of contagion risks, by graph convolution computation. With the multiple message passing operation, the contagion risk is adequately learned in the output representations.

DRL Training Framework

SCRPF utilizes the interaction tuple $\langle S, A, R\rangle$ to train the policy network of the DRL agent.

- **The State Space:** S describes an agent's perception of the environment, which includes the status and features of networked-loans G.
- **The Action Space:** A denotes the allowed actions an agent can take, which is the critical firm probability list in our task.
- **The Reward:** R calculates the quality of the action, which evaluates the degree of risk mitigation.

In the training process, the agent interacts multiple steps with environment and targets to maximize the cumulated reward $R_T = \sum_{t=0}^{T} r(t)$, where T is the number of interactions. At each step t, the DRL agent gets an observation $s_t \in S$ from the environment and produces an action $a_t \in A$ by the policy network $a_t = \mu_\theta(s_t)$, where θ denotes the parameters of policy network (the high-order MP graph attentional neural network) and is updated by policy gradient method. Particularly, we update θ based on the R_T of each episode and employ a gradient descend optimizer to guide the policy to generate higher reward actions. The parameters are updated by:

$$\theta' = \theta + \eta \sum_{t=1}^{T} \nabla_\theta \log \mu_\theta(s_t, a_t) \left(\sum_{t'=t}^{T} r_{t'} - b_t \right) \quad (1)$$

where the η is the learning rate that controls the update speed of θ during each episode. Following the advances in DRL learning (Sutton et al. 1999; Greensmith, Bartlett, and Baxter 2004; Schulman et al. 2017), we employ $b_t = \mathbb{E}(R_t)$ as a

Algorithm 1: DRL Training Framework of SCRPF

Input: Loan network $G = (V, E)$ and its attributes
Parameter: θ of the policy network
Output: $p(v)$, $p(e)$ and critical firm list.

Orthogonal initialize network parameters;
Pre-train the policy network and output $p(v)$ and $p(e)$;
for epoch = 1 to max_iteration **do**
    Set episode length $l = 32$, $\Delta\theta = 0$;
    Interaction with environment and collect trajectory: $(s_1, a_1, r_1, \cdots, s_l, a_l, r_l) \sim \mu_\theta$;
    for $t = 1$ to l **do**
        Given $p(v)$ and $p(e)$, compute reward by: $R_t = \sum_{t'=t}^{l} r_{t'}$;
        Compute b_t by: $b_t = \mathbb{E}(R_t)$;
        $\Delta\theta = \Delta\theta + \nabla_\theta \log \mu_\theta(s_t, a_t) \left(\sum_{t'=t}^{l} r_{t'} - b_t\right)$
    end for
    $\theta = \theta + \eta\Delta\theta$
end for
Given $s_{\widehat{t}}$, compute action $a_{\widehat{t}} = \mu_\theta(s_{\widehat{t}})$
return $p(v)$, $p(e)$, $a_{\widehat{t}}$

baseline value to reduce the variance of policy gradient. We utilize the recall memory to accelerate the training process and use orthogonal parameter initialization for better gradient optimization. We employ supervised learning by ground-truth loan default labels to pre-train the policy network and generate the probability of nodes (firm default probability) $p(v)$ and edges (default contagion probability) $p(e)$ in advance, which is utilized for reward calculation and will be presented in the later section. The detailed learning process of the DRL agent is reported in Algorithm 1.

Contagion Chain Message Passing

In the proposed DRL framework, we employ the basic profile as state features, which does not require manual feature engineering. In particular, node's (MSMEs) attributes, denoted as $rep_{\widehat{v}}$, include assets, liabilities, registered capital, number of employees, etc. The edge's (guarantee) attributes, denoted as $rep_{\widehat{e}}$, include guarantee amount, loan amount, loan interest, etc. It should be noted that the guarantee contracts only take effect during the contract term, which will impact the edge and the structure of the network – making the guarantee loan networks temporal. Thus, we dynamically update the network and the contagion chains in training and predicting phase.

As the relationship is temporal, a temporary edge e will be inserted into the loan network G when the guarantee is issued and removed when the contract expires. Correspondingly, the contagion chains are also updated $C = \{c_1, c_2, \cdots, c_{|C|}\}$, in which node v_i's contagion chain c_i^n is the node's n-order risk diffusion scope. The edges $\mathrm{Diff}(v_i, G)$ that directed from v_i across the network G are denoted as $c_i^n = \{(v, e)|v \in \mathrm{Diff}^n(v_i, G) \cup e \in \mathrm{Diff}^n(v_i, G)\}$. To generate the contagion chains of each node, we utilize a directed breadth-first search (BFS) algorithm in practice. We then aggregate n adjacent nodes along contagion chains and form n^{th}-order message-passing networks, as illustrated in Figure 3b. For example, in 2-order MP, we combine every two node diffusions and form nodes (AB) and (BD) from contagion chain ABD. Meanwhile, we new directed edges for diffused relations, such as (AB) to (BD), while constructing undirected edges for paralleling collections, such as (AB) and (AC). The feature of collected nodes and edges in the MP layer is the combination of individual MSMEs within the collection. So long as we initialize the network and contagion chains, it is quite efficient to update them. Moreover, our proposed method only leverages the original attributes for the training without requiring heavy feature engineering. Thus, with the high computational efficiency, the MP layer meets the low-legacy demands in the financial industry.

Graph-based Policy Network

As shown in Figure 3b, the policy network of the DRL agent takes the MP networks as inputs and produces MP embeddings for the downstream multi-head attention layer. As defaults diffuse across contagion chains in a cascade way (see Figure 1), we develop a high-order MP network and HGA layers to learn the cascade contagion process. In particular, for each node collection $\widehat{v}$ in n-order contagion chain c^n, we introduce $\mathcal{N}_{\widehat{v}}$ to denote neighbors of node collection $\widehat{v}$. Then, the n^{th} order message passing layer is denoted as:

$$h_{i,j}^n = \mathrm{LeakyReLU}\left(W_n rep_{\widehat{v}_i}^n \parallel W_n rep_{\widehat{v}_j}^n\right) \quad (2)$$

$$\alpha_{ij}^n = \frac{\exp\left(h_{i,j}^n\right)}{\sum_{k \in \mathcal{N}_{\widehat{v}}} \exp\left(h_{i,k}^n\right)} \quad (3)$$

where $\mathrm{LeakyReLU}(\cdot)$ denotes Leaky Rectified Linear Unit serving as a nonlinear transformation. $rep_{\widehat{v}_i}^n$ means the original aggregated attributes of collection $\widehat{v}_i$ in n^{th}-order MP layer. W_n represents the learned weighted matrix and α_{ij}^n denotes each HGA layer's attentional weights. Then, we reach the representation of outputs of each HGA layer as:

$$rep_i^n = \sum_{j \in \mathcal{N}_{\widehat{v}}} \alpha_{ij}^n rep_{\widehat{v}_j}^n \quad (4)$$

where the α_{ij}^n denotes the trainable attentional weights of HGA layers. rep_i^n means the output of each HGA layer. We conduct the raw node attribute as the initial node feature and the edge attribute as the edge feature in the implementation. Compared with the graph neural network that mainly focuses on network structure learning, our proposed message-passing layer learns the high-order representation from both the structure and the feature transformation of nodes and edges, as well as contagion chains. Therefore, our proposed feature representative learning directly takes attributes of loan behavior as input and no longer requires expensive feature engineering in this process.

Then, the multi-head attentional prediction network takes the HGA layers' output as input. As different order of MP contributes differently to the risk contagion, to obtain sufficient expressive power to transform the representations into a high-level feature, we perform a shared linear transformation, parameterized with W_α, to every order of MP. We then perform a self-attention on the MP embeddings, which computes attention coefficient α_{ij} that indicates the importance of rep_v^i's features to rep_v^j. For the brevity of notations, we utilize rep^i denotes the combination of $\{rep_1^i, rep_2^i, \cdots, rep_{|v|}^i\}$.

$$\alpha_{ij} = \frac{\exp\left(\sigma\left(W_\alpha\left[rep^i \parallel rep^j\right]\right)\right)}{\sum_{k=1}^{|n|}\exp\left(\sigma\left(W_\alpha\left[rep^i \parallel rep^k\right]\right)\right)} \tag{5}$$

where σ denotes a nonlinear transformation like sigmoid. W_α represents the learned weighted matrix and $\parallel$ denotes the concatenation operation. $|n|$ denotes the number of orders in the MP layer, which is a hyper-parameter.

Finally, the learned attention coefficients are employed to compute the linear combination of the features corresponding to the target output. Thus, we get the output of graph attention embeddings as follows:

$$\begin{aligned} \widehat{rep^i} &= \tanh\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{|n|}\alpha_{ij}W_\alpha rep^j\right) \\ \widehat{rep} &= \parallel_{i=1}^{|n|}\widehat{rep^i} \end{aligned} \tag{6}$$

where K is the number of multi-head attention, which means that K independent attention mechanism is deployed on the final layer of the policy network, followed by the linear combination of the output features. They are then aggregated as the output feature representation for optimization.

Model Pre-train and Reward Function

As described above, critical nodes often have a higher default probability. We pre-train the policy network by ground-truth node default label after obtaining the graph attention embeddings form of representation $\widehat{rep}$. The loss function of the pre-training process is defined as follows:

$$\begin{aligned} \mathcal{L}(\theta) = -\frac{1}{N}\sum_{v=1}^{|v|}&\left[y_v \log\left(\mathrm{NN}_v(\widehat{rep}_v;\theta_v)\right)\right. \\ &\left.+(1-y_v)\log\left(1-\mathrm{NN}_v(\widehat{rep}_v;\theta_v)\right)\right] \end{aligned} \tag{7}$$

where $\widehat{rep}_v$ denotes the feature of the v^{th} nodes which is the output of the attentional prediction network. y_v denotes the ground-truth label of the v^{th} sample, which is set to 1 if the firm is default and 0 otherwise. $\mathrm{NN_v}$ is a shallow neural network with one-layer ReLU (Rectified Linear Units) (Nair and Hinton 2010) and one layer of sigmoid. Similarly, we employ $\mathrm{NN_e}$ as the edge probability prediction network, which is trained by the ground-truth contagion label. Finally, we introduce $\mathrm{NN_a}$ as the action output layer that produces the probability of critical nodes for policy gradient optimization, which is trained independently and shares the same structure with $\mathrm{NN_v}$.

$$\begin{aligned} &a = \mathrm{NN}_a(\widehat{rep}_v), \;\; p(v_i) = \mathrm{NN}_v(\widehat{rep}_{v_i}) \\ &p(e_{ij}) = \mathrm{NN}_e(\widehat{rep}_{v_i}, \widehat{rep}_{v_j}, rep_{\widehat{e}_{ij}}) \end{aligned} \tag{8}$$

So far, given the probability of $p(v)$, $p(e)$ and action a, the DRL framework evaluates the policy by calculating the value of risk mitigation as rewards. For each episode $(s_1, a_1, r_1, \cdots, s_l, a_l, r_l)$, the reward is computed as:

$$r = \sum_{i=1}^{l} a_{v_i}\left(1-(1-p(v_i))\prod_{v_j\in\mathcal{N}(v_i)}(1-p(e_{ji})p(v_i))\right) \tag{9}$$

Finally, the proposed approach can be learned during the optimization process through standard SGD (Stochastic Gradient Descent) based algorithms (Zinkevich et al. 2010). We employ the Adam method (Dozat 2016) with momentum schedule as the optimizer.

Experiments

Experimental Settings

Datasets. We collect the dataset from a major commercial bank with loan records from January 2004 to December 2015, containing monthly information about all loan repayments of MSMEs and the guarantee status. There are around 0.83 million guarantee relationships, including approximately 0.57 million MSMEs. The data contains loan-level information for each participant (both borrower and guarantor, the amount of each loan, and period of validity) as well as firm-level fundamentals (e.g., assets, liabilities, registered capital) of the MSME with the total amount of networked-loans over 3 trillion US dollars. Due to the non-disclosure agreement (NDA), we can not introduce firms' privacy information in the dataset.

Baseline Methods and Setting. To emphasize the effectiveness of our proposed methods, we utilize three kinds of methods as baselines: 1) the widely used approaches in the financial industry: ScoreCard (Thomas, Crook, and Edelman 2017), GBDT (Ke et al. 2017), xgboost (Chen and Guestrin 2016); 2) Graph-based methods: GNN (Battaglia et al. 2018), GAT (Velickovic et al. 2018), HGAR (Cheng et al. 2019a), SuperGAT (Kim and Oh 2020), SA-GAT (Gao et al. 2021); 3) DRL methods: PPO+FC (Schulman et al. 2017),PPO+GAT, using full-connected (FC) layers and GAT as policy network respectively, DeepWeave(Sun et al. 2021). Our model has two variations: SCRPF-noRL, in which we only employ the supervised pre-train model. SCRPF-noPRT, in which we remove the pre-train process on the policy network. SCRPF-ALL denotes the full model with all components proposed in this paper. We set the attention head number K to 4, the hidden size of NN_v to 128, the episode length to 32, the learning rate to 0.001 and the batch size to 200.

Evaluation Metrics. The performance of our approach is evaluated by classification metrics, including F-Score and AUC (area under the ROC curve). We set ground-truth labels of each firm by the spread scope of real default contagion. In particular, for each firm in contagion chains, if we observe over 40% of MSMEs within the chain defaults spread by the firm's failure, we label it as a critical firm. We employ the data of the year 2004 as the training set and slide the time window by month for validation.

Methods	F-Score	AUC	N Features
ScoreCard	0.62351	0.79877	∼250
GBDT	0.65972	0.82247	∼250
xgboost	0.66037	0.83620	∼250
GNN	0.68937	0.85336	17
GAT	0.70062	0.87025	17
HGAR	0.71351	0.88213	17
SuperGAT	0.71566	0.88319	17
SA-GAT	0.72891	0.88720	17
PPO+FC	0.68571	0.84951	∼250
DeepWeave	0.73117	0.89149	17
PPO+GAT	0.74825	0.89922	17
SCRPF-noRL	0.73283	0.89031	17
SCRPF-noPRT	0.75006	0.89986	17
SCRPF-ALL	**0.77493****	**0.92830****	17

Table 1: The results of critical firm prediction.

Critical Firm Prediction

This section evaluates the critical firm prediction accuracy in the networked-loans, which is the main task of this paper. Table 1 shows the F-Score, AUC, and number of features (N Features) of different baselines. The first eleven rows of Table 1 report the result of baseline methods.

As we can see, xgboost performs better than the widely-used ScoreCard method, indicating that the critical firm prediction problem in complex loan networks is over complicated for a shallow model to handle. By leveraging the deep learning approach directed on the graph structure, SuperGAT and SA-GAT improve both the F-Score and AUC. Meanwhile, PPO+FC outperforms xgboost by utilizing a DRL mechanism with the same number of features. PPO+GAT is considerably better than SA-GAT, demonstrating the effectiveness of the DRL framework in this task. SCRPF-noRL outperforms all GNN-based baselines, which shows the essence of our designed high-order MP graph networks. SCRPF-noPRT is better than PPO+GAT, proving the importance of high-order MP in the DRL-based critical prediction task. The proposed SCRPF significantly outperforms all baselines in both F-Score and AUC, demonstrating the effectiveness of all sub-components of the proposed approach. Our method's significant advantages can also be proved by the observation that SCRPF-ALL achieves remarkable improvements by only 17 original loan attributes, compared with 250 handcraft features of classic methods.

Stemming Contagion Risk Experiment

In this experiment, we simulate the financial industry's risk mitigation process by manually securing the top 1% of critical firms from being defaulted by contagion risk. In real-world scenarios, commercial banks or other financial institutions usually refinance these critical firms to prevent cascade defaults and systemic crises.

Figure 4 shows the averaged contagion risk reduction by mitigating risk from the predicted critical firms. For brevity, we select the best competitive baselines in each category. As we can see, xgboost and SA-SAT are not satisfactory, while DeepWave and PPO+GAT considerably outperform them by averaging 4% improvements. Our proposed method is significantly and constantly better than all baselines across the ten years. In 2008 and 2013, stemming contagion risk's capability improved near 7% compared with the best baseline. We also observe that the improvements vary in a different years. For example, the gap enlarges with the outbreak of the global financial crisis in the year 2008. Over 30% of MSMEs face a financial dilemma (Wang, Zhang, and Yang 2020) and the government implements an economic stimulus program so that selected companies are refinanced in this process. The superior performance of our method could significantly reduce the systemic risk in the networked-loans. With the same number of firms being secured, our method reduces the risk by over 35%, which is almost twice the most popular xgboost method.

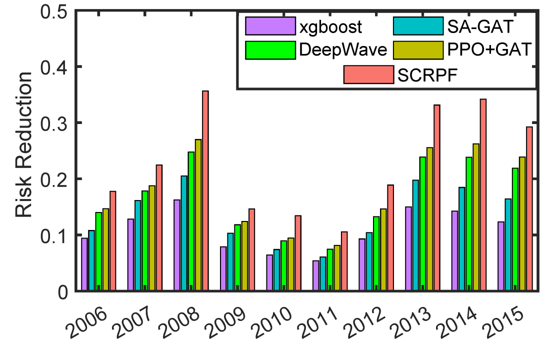


Figure 4: The contagion risk mitigation by securing the top 1% of critical firms predicted by our SCRPF and baselines. The x-axis denotes the averaged values in ten years back test and y-axis mean the ratio of contagion risk reduction.

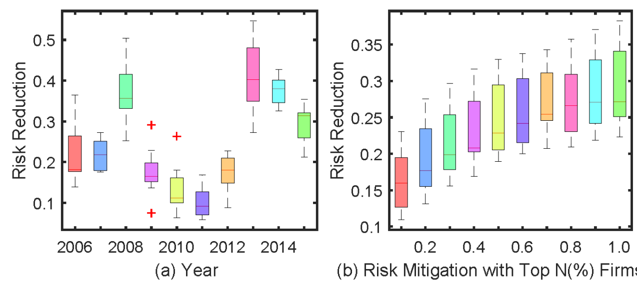
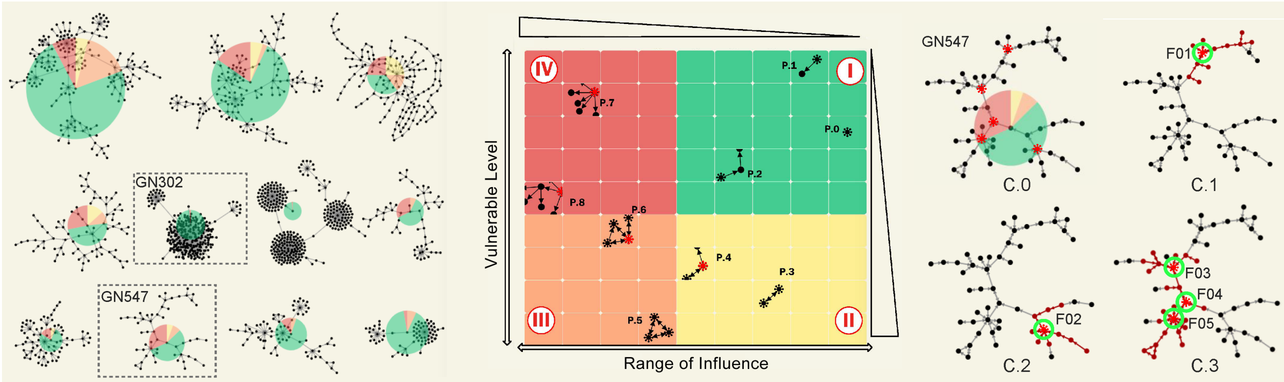


Figure 5: Detailed experimental results of stemming the contagion risk. (a) The top 1% of critical firms are secured to simulate the ratio of risk shrinkage. (b) We increase the proportion of secured firms from 0.1% to 1% and observe the ratio of contagion risk shrinkage.

In the risk mitigation simulation, we backtest the ratio of risk shrinkage ten times and report the result in Figure 5. Figure 5a presents the boxplot of reduction by mitigating detected top 1% of critical firms. As we can see, the performance reached a peak in the years 2008 and 2013. We then employ empirical studies with financial experts and observe that the superior performance is in accordance with the systemic risk ratio each year. In 2008, the global financial crisis outbroke and selected MSMEs were refinanced to decrease from 2009 to 2012, which are required to repay in five years, so default risks rose significantly in 2013. To our best knowledge, SCRPF is especially optimal in these high-



(a) Empirical Study of Contagion Risk in Networked-loans (b) Critical Firms in Contagion Effect Matrix View (c) Critical Firms Predicted in GN547 Network

Figure 6: Empirical study on critical firms of contagion chains in networked-loans. (a) displays the risk view on loan networks. (b) present critical firms located in the contagion effect matrix view. (c) The predicted critical firms in a typical GN547 network.

risk periods, stemming a risk of about 35% with mitigating only 1% of firms. We then increase the proportion of secured firms from 0.1% to 1% with the step size set to 0.1% and report the risk reduction in Figure 5b. It is clear that with more firms secured, the higher value of risk reduction. Nevertheless, the increase is becoming more gentle from 0.5% to 1%, demonstrating the superior performance of our methods. In practice, commercial banks could employ the elbow point method to choose the best percentage to mitigate the risk of MSMEs, as well as promote economic development and secure job opportunities.

Case Studies

We integrate SCRPF in our collaborated bank's loan management system and leverage empirical study on real-world risk control scenarios. We report the predicted critical firms in case studies. Figure 6b displays typical contagion patterns and their locations in the risk effect matrix—the higher the quadrant, the more risk of contagion. We colored the quadrant according to the risk level with Q.I green (lowest), Q.II yellow, Q.III orange, and Q.IV red (highest). We mark predicted critical firms in each frequent contagion path by a virus-shaped icon (). As we can see, 68.1% of critical firms are located in most risk Q.IV and 84.6% of them in the risk quadrant (Q.III and Q.IV). Besides, contagion chains with predicted critical firms account for over 87.5% of all patterns. Then, we visualize the contagion effect matrix in loan networks as illustrated in Figure 6a. We mark the total amount of default at risk as the circle size and divide their contagion risk quadrants in a pie chart. For example, the default at risk of loan network coded as GN302 is very close to GN 547. Thus, existing loan management methods are likely to treat them as the same level of risk. However, in the empirical study, they are significantly different in terms of risk components. GN302 is more healthy than GN547 as most of its risk consists of small-scale contagions. However, the situation in GN547 is different; near 1/3 of defaults is Q.IV (highest). In other words, these kinds of risks have a higher probability of triggering contagion to healthy firms and would cause systemic defaults.

Then, we leverage case studies on GN547 and display their details in Figure 6c. The general amount of risk at defaults and each component of four quadrants are shown in C.0. There are three leading contagion chains in GN547, and we report them in C.1, C.2, and C.3 of Figure 6c. Our proposed method reports F01 as a critical firm in C.1, F02 in C.2, F03, F04, and F05 in C.3. It is clear that if F01 defaults, the risk would diffuse along edges and cause four firms at risk in first-level contagion, seven firms in second-level contagion, and nine firms at risk with third-level contagion. Our method successfully detects the most critical firm in contagion chain C.1. A similar phenomenon is observed in chain C.2 that F02 is the critical firm. Interestingly, our collaborated financial experts find that F04 is among predicted critical firms in chain C.3, which is not observed by domain knowledge. Only three nodes are connected to F04, while more than four are connected to the most critical firms. After in-depth analysis, we find that even though only three nodes are in the first contagion of F04, 12 nodes are involved in its second level contagion. Moreover, 17 firms would be at risk by the third-level contagion if F04 defaults. Our proposed method successfully predicts all critical firms in the case study. It provides meaningful insight into risk management in complex loan networks. These empirical studies and case analyses demonstrate our method's effectiveness in predicting critical firms to stem contagion risk in networked-loans.

Conclusion

This paper proposes a novel graph-based deep reinforcement learning method for critical firm prediction to stem contagion risk in networked-loans. Our method leverages high-order message-passing graph layers as the policy network to directly learn representations from risk contagion over complex financial networks. We thoroughly evaluate our method compared with popular benchmarks on the historical datasets and achieve superior performance. The ability of the proposed method in predicting critical firms for stemming contagion risk is demonstrated. The empirical study proves that our work could support financial authorities and banks for more prompting prevention measures. The idea of modeling risk based on contagion can be applied widely in financing MSMEs for sustainable economic development.

Acknowledgments

This work was supported by the National Science Fundation of China (62102287), the foundation of Key Laboratory of Artificial Intelligence, Ministry of Education, P.R. China and the Shanghai Science and Technology Innovation Action Plan Project (22511100700).

References

Acemoglu, D.; Ozdaglar, A.; and Tahbaz-Salehi, A. 2015. Systemic risk and stability in financial networks. *American Economic Review*, 105(2): 564–608.

Addo, P. M.; Guegan, D.; and Hassani, B. 2018. Credit risk analysis using machine and deep learning models. *Risks*, 6(2): 38.

Baesens, B.; Setiono, R.; Mues, C.; and Vanthienen, J. 2003. Using neural network rule extraction and decision tables for credit-risk evaluation. *Management science*, 49(3): 312–329.

Battaglia, P. W.; Hamrick, J. B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; et al. 2018. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*.

Chen, T.; and Guestrin, C. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Cheng, D.; Niu, Z.; Li, J.; and Jiang, C. 2022. Regulating systemic crises: Stemming the contagion risk in networked-loans through deep graph learning. *IEEE Transactions on Knowledge and Data Engineering*.

Cheng, D.; Niu, Z.; Tu, Y.; and Zhang, L. 2018. Prediction defaults for networked-guarantee loans. In *2018 24th International Conference on Pattern Recognition (ICPR)*, 361–366. IEEE.

Cheng, D.; Niu, Z.; and Zhang, L. 2020. Delinquent events prediction in temporal networked-guarantee loans. *IEEE Transactions on Neural Networks and Learning Systems*.

Cheng, D.; Tu, Y.; Ma, Z.; Niu, Z.; and Zhang, L. 2019a. Risk assessment for networked-guarantee loans using high-order graph attention representation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 5822–5828. AAAI Press.

Cheng, D.; Zhang, Y.; Yang, F.; Tu, Y.; Niu, Z.; and Zhang, L. 2019b. A dynamic default prediction framework for networked-guarantee loans. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2547–2555.

Chernyshev, A. 2020. Standard Risk management: principles and guidelines. https://www.iso.org/iso-31000-risk-management.html. Accessed 13-December-2020.

Dozat, T. 2016. Incorporating nesterov momentum into adam. *ICLR workshop*.

Dumitrescu, E.; Hué, S.; Hurlin, C.; and Tokpavi, S. 2022. Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, 297(3): 1178–1192.

Eubanks, W. W. 2010. *Status of the Basel III Capital Adequacy Accord*. DIANE Publishing.

Gao, J.; Gao, J.; Ying, X.; Lu, M.; and Wang, J. 2021. Higher-order interaction goes neural: A substructure assembling graph attention network for graph classification. *IEEE Transactions on Knowledge and Data Engineering*.

Georgios, K. 2019. Credit risk evaluation and rating for SMES using statistical approaches: the case of European SMES manufacturing sector. *Journal of Applied Finance and Banking*, 9(5): 59–83.

Greensmith, E.; Bartlett, P. L.; and Baxter, J. 2004. Variance Reduction Techniques for Gradient Estimates in Reinforcement Learning. *Journal of Machine Learning Research*, 5(9).

Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; and Liu, T.-Y. 2017. Lightgbm: A highly efficient gradient boosting decision tree. In *NeurIPS*, 3146–3154.

Kim, D.; and Oh, A. 2020. How to find your friendly neighborhood: Graph attention design with self-supervision. In *International Conference on Learning Representations*.

Mian, A.; and Sufi, A. 2009. The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *The Quarterly Journal of Economics*, 124(4): 1449–1496.

Montgomery, H. 2005. The effect of the Basel Accord on bank portfolios in Japan. *Journal of the Japanese and international economies*, 19(1): 24–36.

Nair, V.; and Hinton, G. E. 2010. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning*, 807–814.

Niu, Z.; Li, R.; Wu, J.; Cheng, D.; and Zhang, J. 2020. iConViz: Interactive Visual Exploration of the Default Contagion Risk of Networked Guarantee Loans. *2020 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 1–11.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Sui, X.; Li, L.; and Chen, X. 2020. Risk contagion caused by interactions between credit and guarantee networks. *Physica A: Statistical Mechanics and its Applications*, 539: 122867.

Sun, P.; Guo, Z.; Wang, J.; Li, J.; Lan, J.; and Hu, Y. 2021. Deepweave: Accelerating job completion time with deep reinforcement learning-based coflow scheduling. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 3314–3320.

Sutton, R. S.; McAllester, D.; Singh, S.; and Mansour, Y. 1999. Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems*, 12.

Thomas, L.; Crook, J.; and Edelman, D. 2017. *Credit scoring and its applications*. SIAM.

Tu, Y.; Niu, L.; Chen, J.; Cheng, D.; and Zhang, L. 2020. Learning from web data with self-organizing memory module. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12846–12855.

Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; and Bengio, Y. 2018. Graph attention networks. In *ICLR*.

Wang, Y.; Zhang, Q.; and Yang, X. 2020. Evolution of the Chinese guarantee network under financial crisis and stimulus program. *Nature Communications*, 11(1): 1–11.

Zhang, Z.; Zhang, A.; Li, M.; and Smola, A. 2022. Automatic Chain of Thought Prompting in Large Language Models. *arXiv preprint arXiv:2210.03493*.

Zinkevich, M. A.; Weimer, M.; Smola, A.; and Li, L. 2010. Parallelized stochastic gradient descent. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems-Volume 2*, 2595–2603.