



Analysis of Corporate Community of Interest Relationships in Combination with Multiple Network

Yipan Liu , Song Wang^(✉) , Hao Hu , and Shijie Chen 

School of Computer Science and Technology, Southwest University of Science and Technology,
Mianyang 621010, Sichuan, China
wangsong@swust.edu.cn

Abstract. Visualizing the complex relationship among enterprises is ponderable to help enterprises and institutions to find potential risks. Small and medium-size enterprises' (SMEs) loans have higher risk and non-performing rates than other types of enterprises, which are prone to form the complex relationship. Nowadays, the analysis of enterprises' relationships networks mainly focus on the guaranteed relationships among enterprises, but it lacks the holistic analysis of the enterprise community of interest. To address these issues, the concepts of the enterprise community of interest and the investment model withing enterprise community of interest are proposed; The centrality, density, and network diameter algorithms in graph theory are used to evaluate the network of enterprise community of interest; The problem of graph isomorphism are used to query the network of users interested enterprises' relationships; The portrait of enterprise is used to evaluate the enterprise community of interest; In the end, we study the impact of debt relationship among enterprise community of interest. Based on these ideas, to verify the effectiveness of the method, an enterprise relationship network analysis system which included 6745 enterprise nodes and 7435 enterprise relationship data of Shanghai is developed.

Keywords: Enterprise relationship network · Enterprise relationship analysis · Community of interests · Mortgage relations · Enterprise portrait

1 Introduction

According to the data of the fourth Economic Census, the number of people working in SMEs has accounted for 80 percent of the total number of employees. Due to insufficient mortgage guarantees and asymmetric information [1], it is particularly important to analyze the mortgage relationship and enterprise investment. Also, network analysis is widely used in social, Internet [23, 24], transportation, life science [2], and other fields. So it is necessary to study the enterprises' relationships by using network relationships. The phenomenon of enterprises' guarantee networks exists in most countries, which are worth analysing the guarantee relationship to study the influence of risk diffusion [3]. Most SMEs often need "fast, diversified, and flexible" funding and need mortgages to

obtain funds [4]. The short establishment, small size, and lack of detail information of SMEs result in fewer loan opportunities, higher risks, and capital shortfalls [5]. Thus, the SMEs are easy to form a complex enterprise mortgage relationship and enterprise equity pledge relationship. In terms of enterprise risk assessment, research mainly analysis the characteristics of the enterprise, such as the judicial litigation risk of the government, enterprise credit risk [6], and enterprise social news risk reported by some social media. In terms of enterprise relationships, a part of the research only focus on enterprise relationship, such as the guarantee and mortgage relationship [7]. But it ignores the basic attributes of the enterprise, such as the size of assets. The study of enterprise is divided into the study of their own characteristics and the impact of single-enterprise relationships. To better analyze the relationship among enterprises, there are some issues need to be addressed here: firstly, combined with the enterprise characteristics and relationships, the enterprise community of interest characteristics should be analyzed; Secondly, analysis of the influence of debt enterprise to the community of interest is carried out.

2 Related Work

2.1 Enterprise Risk Assessment

In terms of credit, there were many data-driven credit scoring schemes, such as visualization in credit scoring methods based on neural networks [8]. The XGBoost credit rating model were constructed to evaluate the enterprise based on SMOTE series algorithm [9]. Nils [10] used the structural model after Merton (1974) to explore the relationship between a company's stock return and credit risk. Finally he found that the enterprise's stock return increased with the credit default risk premium. Combined with the financial data of bank, Hui [11] conducted regression analysis and Hosmer Lemeshow randomness test to build a logistics risk assessment model. Kang believed that enterprise bond yields reflect the concern about debt deflation and when debt is nominal, unexpectedly low inflation would increase real debt and default risk. Dawei proposed an unbalanced network risk diffusion model to predict the short-term default risk of enterprises and proposed [12] a positive weighted K-nearest neighbor algorithm for the independent case without default contagion to improved prediction accuracy. In the case of credit, commercial banks assess the risk of companies in order to maximize profits [13]. For funds that are likely to default, the impact in the credit business is measured quantitatively to reduce losses [14].

2.2 Enterprise Relationship Network

By mining the financial characteristic data among enterprises, the network topology layout could be used to represent the mutual relationship between each unit, which was widely used in the financial field [15, 16]. In terms of enterprise relationship, the enterprise loan guarantee relationship was introduced into the network topology layout structure to show the multi-level default risk visualization [17]. Sassa [18] proposed a structure-based suggestion exploration method, which suggested appropriate structures according to user's requests to support the effective exploration of large-scale networks.

Junrui [19] used a knowledge graph construction method for graph neural networks to quantify the correlation among enterprise. In the field of financial relationship data, the relationship among banks was represented through network [25] visualization. For example, text data was collected from financial forums to generate and visualize a jointly mentioned bank network [20]. It also effectively reveals financial crimes such as money laundering and fraud in the financial activity network through up-down interaction [21]. Therefore, the 2-dimensional network [26] topology layout was suitable for discussing specific cases and detailed analysis, but the relationship would be relatively disordered under large-scale data.

The analysis of companies is divided into enterprise characteristic analysis and enterprise relationship analysis. However, the analysis of enterprises characteristics lacks the analysis of the impact among enterprise; In contrast, enterprises relationship analysis is missing enterprise characteristics.

3 Exploration of Enterprise Relationships

The main analysis tasks of this paper are to study enterprise stakeholder networks and enterprise collateral relationships:

- TASK1: Exploring different communities of interest extracted from enterprise relationship network [22], and evaluating enterprise investment relationship networks.
- TASK2: Researching the impact of changes in debt enterprises on the community of interest in combination with enterprise portraits.
- TASK3: Combined with enterprise portrait and enterprise relationship network, we can conduct a custom network similarity search (Fig. 1).

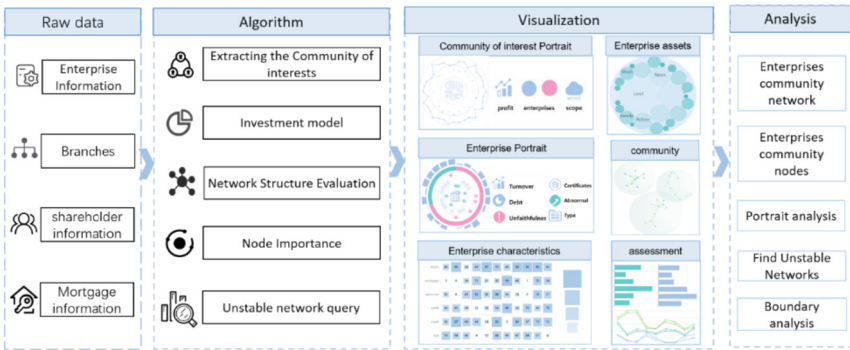


Fig. 1. Overview of the approach

3.1 Analysis the Enterprise Community of Interest

a) Basic concept

Community of interest: The community of interest is a connected graph $G = (V, E)$, the G represents the network topology formed, the V represents set of enterprise nodes, and the E represents the investment relationship of enterprises.

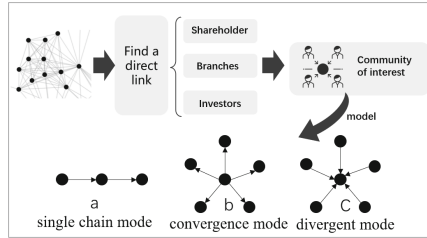


Fig. 2. Mining process of community of interest

Investment model: the investment made by enterprise ‘A’ only to enterprise ‘B’ is called the single chain mode (Fig. 2(a)); Enterprise ‘A’ accepting investment from multiple enterprises is called convergence mode (Fig. 2(b)). Enterprise ‘A’ investment in multiple enterprises is called divergent mode (Fig. 2(c)).

b) Community of interest network characteristics

Network density: As shown in Formula (1), $Density(G)$ means the ratio of already linked vertex pairs to the theoretically possible vertex pairs, $\frac{\sum_{i=1}^n Degree(v_i)}{2}$ represents the number of edges and ‘n’ represents the number of nodes in the graph.

$$Density(G) = \frac{\sum_{i=1}^n Degree(v_i)}{n(n-1)} \quad (1)$$

Average degree: As shown in Formula (2), $\sum_{j=1, j \neq i}^n f(v_i, v_j)$ represents the degree of a node.

$$Average Degree(G) = \frac{\sum_{i=1}^n \sum_{j=1, j \neq i}^n f(v_i, v_j)}{n} \quad (2)$$

Diameter of the community of interest network: As shown in Formula (3), $MaxP(v_i, v_j)$ represents the maximum distance between two nodes.

$$Diameter(G) = \frac{2Max(MaxP(v_i, v_j), i \neq j)}{n(n-1)} \quad (3)$$

Average path length: As shown in Formula (4), $ShortesP(v_i, v_j)$ represents the shortest distance between two points.

$$AverageP(G) = \frac{2 \sum_{i \geq j} ShortesP(v_i, v_j)}{n(n-1)} \quad (4)$$

Degree centrality: As shown in formula (5), d_{ij} represents the edge among two points.

$$Degree(v_i) = \frac{\sum_{j=1}^n d_{ij} (i \neq j)}{n} \quad (5)$$

Closeness centrality: As shown in Formula (6). The $ShortesP(v_i, v_j)$ indicates the shortest distance among two nodes.

$$CloseCen(v_i) = \frac{1}{\sum_{j=1}^n ShortesP(v_i, v_j)} \quad (6)$$

Between centrality: As shown in formula (7), $sd(j, i, k)$ represents the shortest path that passes through node 'i' from node 'k' to node 'j', and $\sum_{j,k=1}^n s(j, k)$ represents the number of all paths between nodes j and k.

$$BetweenCen(v_i) = \frac{\sum_{j,k=1}^n sd(j, i, k)}{\sum_{j,k=1}^n s(j, k)}, (j \neq k) \quad (7)$$

c) Construction of enterprise interest community and portrait construction

We use the maximum connectivity subgraph algorithm (Eq. (8)) to find the community of interest in the network topology. The G_c represents community of interest; The E_i represents the set of edges in the connected graph network; The f_n represents the set of points in the connected graph; The $GetE$ extract points from edge set.

$$G_c = set \left(\begin{array}{l} \sum_{i=1}^m f_n(E_n) \cap f_n(E_i) = \emptyset \\ E_i = e_i \\ E_n = E - E_i \end{array} E_i = GetE(f_n(E_n) \cap f_n(E_i)) \right) \quad (8)$$

As shown in Fig. 3, the portrait is mainly divided into three layers from inside to outside. The inner word cloud represents the business scope. The outer nodes' position represents

the time when the enterprise joined and the colors represent the different states of the enterprise. The outermost layer is the average profit of the real community of interests.

$$Abnorma(G) = \frac{\sum_i^n f_{abn}(v_i)}{n \times \lambda} \tag{9}$$

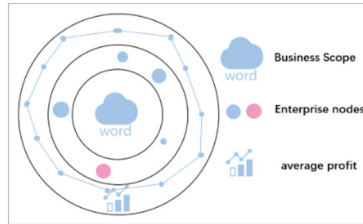


Fig. 3. Portrait of enterprise community of interest

3.2 Analysis of Mortgage Relationship of Enterprise Community of Interest

The Fig. 5 shows the explore the process of enterprise benefit community boundaries. There exists a debtor firm linking the community of interest A and community of interest B. Combined with the analysis of enterprise portrait, the debtor enterprise has great instability and becomes part of community B. Combining the analysis of the community analysis in 3.1, we compare the state of community A and community B. By analyzing the characteristics of the community of interest, we determine whether the original community of interest will change due to debtor.

For the analysis of debt nodes, we analyze them in three aspects. Firstly, we calculate the importance of the evaluation of nodes in the enterprise community of interest (see Sect. 3.1) and obtain the influence of debt nodes on the enterprise community of interest. Combined with enterprise characteristics, we design the enterprise portrait to quickly understand the debt enterprise. We select the indicators that are closely related to enterprise operation and credit, such as profit, credit history, and assets (Fig. 4b). The portraits are introduced to check the operating conditions of debt enterprises over a while (Fig. 4a). The center of the enterprise portrait indicates the industry type to which the enterprise belongs; the outward circular icons represent the enterprise characteristics acquired by the enterprise at different times. In addition, the circular icon represents the proportion of corporate debt and profit. The operating quota of the outermost term is used to check the operating status of the enterprise over a period of time.

3.3 Explore the Similar Relationship to Avoid Risk in Time

Combined with the topology structure and abnormal characteristics of the enterprise’s community of interests, the enterprise’s community of interests is evaluated and

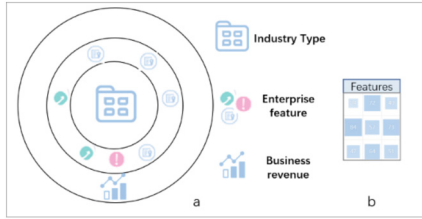


Fig. 4. The portrait of enterprise

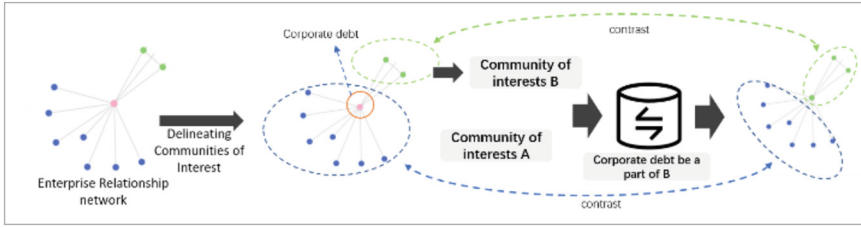


Fig. 5. Research on mortgage boundary

described from the aspects of the overall network structure, node importance, and comprehensive portrait. If the important nodes or an important paths section are the abnormal enterprises, they will have a significant impact on the enterprise relationship network and easily cause changes in the enterprise relationship network.

To discover the hidden risk in time, it is necessary to analyze the enterprise relationship and search for similar relationship, which can provide early warning and avoid risks in time.

We use the isomorphism problem in graph theory to solve similar network problems. For two enterprise communities of interest $G(V, E)$ and $G'(V', E')$. If there is a bijection $f : V \rightarrow V'$, such that for any two points u, v in V , the edges $f(u)f(v) \in E'$ and $uv \in E$, we say that $G(V, E)$ and $G'(V', E')$ are isomorphic.

Firstly, the user choose the model network which has the abnormal enterprise node. The next step we choose a part of network to explore. Then we extract the subnetwork from the target network which has the same number of nodes of model network. Finally, we use the isomorphism algorithm to determine whether the structure is the same.

Algorithm Search similar Community of interests

Input:

Model: Community of interest $G(V, E)$

To be matched: $G'(V', E')$

Output: search for the network that is similar to the network $G(V, E)$

- 1 Initialize parameters $mList = []$, $tList = []$
- 2 $m = CountNodes(G)$, $n = CountNodes(G')$
- 3 **if** $m \geq n$ or $m < 1$ **then**
- 4 **return** false, the selected network does not comply with the rules
- 5 **else**
- 6 **for** ($i = 1$; $i \leq c_i$; $i++$) **do**
- 7 $temGraph = f_{select}(G, m)$
- 8 **if** $f_{connect}(temGraph) == true$
- 9 $mList.append(temGraph)$
- 10 **for** ($i = 1$; $i \leq len(mList)$; $i++$) **do**
- 11 $temGraph = mList[i]$
- 12 **if** $isomorphic(temGraph, G) == true$
- 13 $tList.append(temGraph)$

final:

return $tList$ (the similar network)



Fig. 6. System interface of enterprise relationship

4 The Influence of Debt Enterprises on Interest Communities

4.1 Explore the Similar Relationship to Avoid Risk in Time

We developed an enterprise relationship analysis system using real data from Shanghai (Fig. 6). Large spherical nodes in Fig. 6a represent the community of interest. Figure 6b shows the enterprise community, 6c presents its basic characteristics, 6d illustrates node importance, and 6e displays the evaluation network. Figure 6f details enterprise assets, 6g depicts custom matching, 6h shows enterprise portraits, and 6i outlines basic enterprise characteristics.

4.2 Analysis of the Community of Interest

a) Analysis of investment model of a community of interest

In Fig. 7. We conducted a study as shown in Table 1.

Table 1. Overall index

project	a	b	c
Number of nodes	10	10	10
Number of edges	9	9	9
The density of the figure	0.20	0.20	0.20
The average degree	1.8	1.8	1.8
The network diameter	0.067	0.089	0.20
The average path length	1.8	3.11	3.64

In Table 1, when the number of nodes is the same in the three models, edges, graph density, and average degree also match. However, network diameter and average path length differ: Fig. 7a’s star mode is the smallest, 7c’s single-chain mode is the largest, and 7b’s hybrid mode falls in between. The star model has a small diameter and segmented paths due to its centralization, while the single-chain pattern has larger diameters and longer paths, indicating a looser structure. The mixed mode sits between the star and single modes.

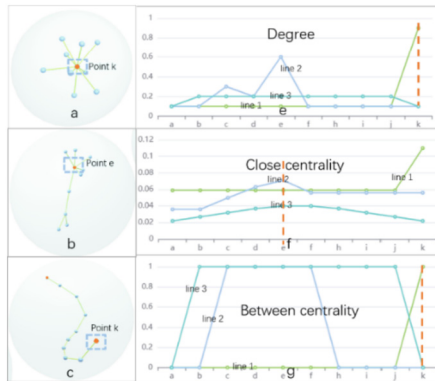


Fig. 7. Features of enterprise nodes

Then we analyze the node centrality of the enterprise community of interest and obtain the results shown in Fig. 7. Figure 7 e represent the degree centrality of each enterprise node in the enterprise interest community. Figure 7 f represent the closeness centrality of each enterprise node. Figure 7 g represent the between centrality of the

enterprise node. Line 1 represent the star-shaped community of interest in Fig. 7 e. Line 2 represents the mixed community of interest in Fig. 7 f. Line 3 represent the single-chain community of interest in Fig. 7 g.

In Fig. 7, Line 1 remains straight until point k, which acts as a central hub connecting multiple enterprises in a star-like pattern. Line segment 2, particularly node e, maintains high centrality within a mixed-mode community. Line 3 is symmetric across all nodes. Higher closeness centrality of a middle enterprise node in Fig. 7b significantly impacts the entire enterprise community. The star model consistently shows high degree and betweenness centrality, indicating close interconnections among enterprises. In contrast, the single-chain model's high closeness centrality highlights the community's vulnerability to disruptions from individual path interruptions.

b) Community of interest portrait and matching query

We created a community portrait in Fig. 8c, with dots representing different companies. In Fig. 8b, the horizontal axis denotes mean path, and the vertical axis represents anomaly index. Figure 8f reveals an enterprise with low profit but high debt. When we identify an anomalous community in the network, users can define a matching corporate community (Fig. 8d) and search for similar structures across all enterprises (Fig. 8e), with results highlighted in red.

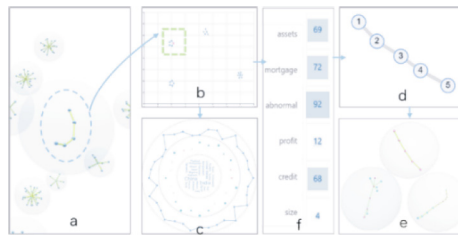


Fig. 8. Enterprise node characteristics

Degree and closeness centrality measure community structure, while betweenness centrality identifies key bridge nodes. Key nodes are vital in complex interest communities, as abnormalities in non-endpoint enterprises can harm the entire community. Simplified network topology reduces disruptions, benefiting the enterprise network. In diverse interest communities, decompose into basic models, like star and betweenness centrality, to prevent disruptions and fragmentation, which investors dislike.

4.3 Analysis of the Mortgage Community of Interest Relationship

In Fig. 9f, enterprise has substantial land resources. Despite past abnormal information, its centrality in Fig. 9b (9g, 9h, 9i) suggests a non-critical position within the community. The profile shows profitable performance surpassing debt, making it less critical for the entire community in case of collateral risks.

But in Fig. 9b community of interest, firm node d has this severe business operating condition. In Fig. 9b community of interest, enterprise node 'd' (Fig. 9e) has a serious

business operating condition; Observing and analyzing the centrality of each node of the enterprise community of interest b, we learn that the enterprise node ‘d’ is at a high index of degree centrality, closeness centrality and between central.

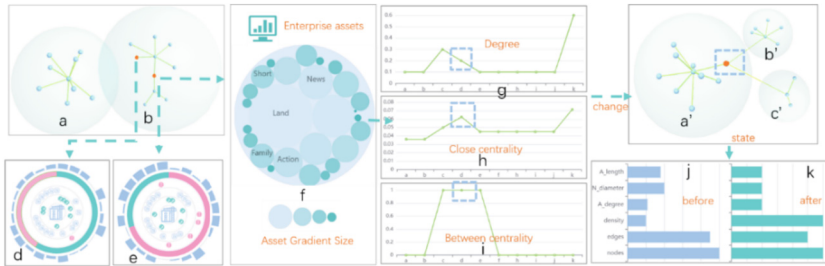


Fig. 9. Boundary analysis

This means that the enterprise node is in a key position and at the center of the entire community of interest; Suppose a collateral event occurs at this node, causing a split in the enterprise community of interest (Fig. 9a’, b’, c’), which could easily lead to a series of incalculable reactions; Analyzing the community of interest before and after the change b (Fig. 9j, k), it is found that the graph density possesses a large change. Therefore, we suggest focusing on the enterprises with a larger degree.

5 Evaluation

We invited 15 volunteers to evaluate the analysis model of this paper, including 5 undergraduate students, 9 graduate students, and 1 visualization professional in the field of finance. The evaluation method adopts a five-point rating scale, which is divided into 1 and 5 points, representing “very dissatisfied”, “dissatisfied”, “fair”, “satisfied”, and “very satisfied” respectively (Table 2).

Table 2. Questionnaire evaluation.

Question	description
Q1	Whether the structure of the whole community of interest can be observed
Q2	Be able to effectively present and analyze large-scale corporate relationships
Q3	Whether it can effectively evaluate the network structure and node importance of the enterprise community of interests
Q4	Whether network matching can be done effectively
Q5	Be able to effectively demonstrate that the mortgage relationship arises
Q6	Whether it is easy to analyze the impact of debt on enterprises

Figure 10 displays user evaluation results. The paper excels in large-scale data presentation analysis (Q2). It effectively observes community structure (Q1) and analyzes

network nodes' importance (Q3). The method enables efficient querying of unstable investment networks (Q4), enhances the display of collateral relationships between companies (Q5), and effectively analyzes the impact of mortgage relationships on interest communities (Q6).

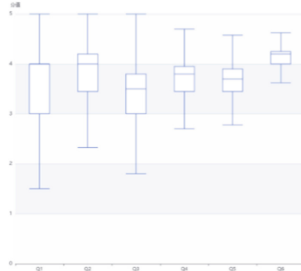


Fig. 10. Evaluation results

6 Conclusion

This paper introduces the concept of interest communities, analyzes enterprise relationships, and evaluates them through network analysis. It also examines the impact of collateral relationships among these communities. User evaluations validate the model, revealing that diverse perspectives enhance stability assessment within and among enterprise communities. Areas for improvement include considering additional characteristics for individual enterprise evaluation and exploring the impact of debt firms on critical path propagation.

Acknowledgement. This work was supported by Natural Science Foundation of Sichuan Province (Grant No. 2022NSFSC0961) the Ph.D. Research Foundation of Southwest University of Science and Technology (Grant No. 19zx7144) the Special Research Foundation of China (Mianyang) Science and Technology City Network Emergency Management Research Center (Grant No. WLYJGL2023ZD04).

References

1. Stroebel, J.: Asymmetric information about collateral values. *J. Financ.* **71**(3), 1071–1112 (2016)
2. Silva, J.S., Saraiva, A.M.: A methodology for applying social network analysis metrics to biological interaction networks. In: *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 1300–1307 (2015)
3. Niu, Z., Li, R., Wu, J., et al.: iconviz: Interactive visual exploration of the default contagion risk of net-worked-guarantee loans. In: *2020 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, pp. 84–94 (2020). Author, F.: Contribution title. In: *9th International Proceedings on Proceedings*, pp. 1–2. Publisher, Location (2010)

4. Deyoung, R., Gron, A., Torna, G., et al.: Risk overhang and loan portfolio decisions: small business loan supply before and during the financial crisis. *J. Financ.* **70**(6), 2451–2488 (2015)
5. Garvin, W.J.: The small business capital gap: the special case of minority enterprise. *J. Financ.* **26**(2), 445–457 (1971)
6. Ng, C.K., Smith, J.K., Smith, R.L.: Evidence on the determinants of credit terms used in interfirm trade. *J. Financ.* **54**(3), 1109–1129 (1999)
7. Niu, Z., Cheng, D., Zhang, L., et al.: Visual analytics for networked-guarantee loans risk management. In: 2018 IEEE Pacific Visualization Symposium (PacificVis). IEEE, pp. 160–169 (2018)
8. Sarlin, P., Marghescu, D.: Visual predictions of currency crises using self-organizing maps. *Intell. Syst. Account. Financ. Manag.* **18**(1), 15–38 (2011)
9. Shen, F., Zhao, X., Kou, G., et al.: A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique. *Appl. Soft Comput.* **98**, 106852 (2021)
10. Friewald, N., Wagner, C., Zechner, J.: The cross-section of credit risk premia and equity returns. *J. Financ.* **69**(6), 2419–2469 (2014)
11. Sun, H., Guo, M.: Credit risk assessment model of small and medium-sized enterprise based on logistic regression. In: 2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE (2015)
12. Cheng, D., Niu, Z., Zhang, L.: Risk research of large-scale unbalanced guarantee network loan. *Chin. J. Comput.* **43**(04), 668–682 (2020)
13. Bian, W.-L.: Research on credit risk management of commercial banks in China. *New Econ.* **2**(2), 50–51 (2016)
14. Nijskens, R., Wagner, W.: Credit risk transfer activities and systemic risk: how banks became less risky individually but posed greater risks to the financial system at the same time. *J. Bank Financ.* **35**(6), 1391–1398 (2011)
15. Flood, M.D., Lemieux, V.L., Varga, M., Wong, B.W.: The application of visual analytics to financial stability monitoring. *J. Financ. Stab.* **27**, 180–197 (2016)
16. Sarlin, P.: Macroprudential oversight, risk communication and visualization. *J. Financ. Stab.* **27**, 160–179 (2016)
17. Cheng, D., Niu, Z., Liu, X., Zhang, L.: Risk assessment of contagion paths in complex guarantee networks. *Sci. China: Inform. Sci.* **51**(07), 1068–1083 (2021)
18. Chen, W., Guo, F., Han, D., et al.: Structure-based suggestive exploration: a new approach for effective exploration of large networks. *IEEE Trans. Visual Comput. Graphics* **25**(1), 555–565 (2018)
19. Zhang, J., Song, Z.: Research on knowledge graph for quantification of relationship between enterprises and recognition. In: 2019 4th IEEE International Conference on Cybernetics (Cybconf). IEEE (2019)
20. Rönqvist, S., Sarlin, P.: Bank networks from text: interrelations, centrality and determinants. *Quant. Financ.* **15**(10), 1619–1635 (2015)
21. Didimo, W., Liotta, G., Montecchiani, F., Palladino, P.: An advanced network visualization system for financial crime detection. In: Visualization Symposium (PacificVis), 2011 IEEE Pacific, pp. 203–210. IEEE (2011)
22. Suh, A., Hajij, M., Wang, B., et al.: Persistent homology guided force-directed graph layouts. *IEEE Trans. Visual Comput. Graphics* **26**(1), 697–707 (2019)
23. Goswami, S.: Enhancement of the effective bandwidth available to a client host in a network by using multiple network devices concurrently. In: International Conference on Information and Communication Technology for Competitive Strategies. Springer Nature, Singapore (2022). https://doi.org/10.1007/978-981-19-9638-2_13
24. Yadav, S., Raj, R.: Power efficient network selector placement in control plane of multiple networks-on-chip. *J. Supercomput.* **78**(5), 6664–6695 (2022)

25. Bharatula, S., Murthy, B.S.: A novel spectrum sensing technique for multiple network scenario. In: Proceedings of International Conference on Communication and Artificial Intelligence: ICCAI 2021. Springer Nature, Singapore (2022). https://doi.org/10.1007/978-981-19-0976-4_9
26. Xing, G., et al.: Improving UWB ranging accuracy via multiple network model with second order motion prediction. Cluster Comput. 1–12 (2023)