

Corporate Credit Rating Evaluation Based on Community Detection of Label Propagation

Lihong Guo and Xiaoguang Luo

Harbin University of Science and Technology, Harbin 150080, China
lihong_guo@sina.com, luoxg@126.com

Abstract

Current corporate credit rating methods tend to focus on rating single corporations and usually fail to assess corporate groups from different industries or geographical areas. This paper presents a method of evaluating corporate credit ratings using the community detection of label propagation (CCRELP). First, the study used a label propagation algorithm to group corporations with similar characteristics into communities based on the connections between the companies, their industries, and their geographical locations. Then, the study performed an evaluation using various indicators and weights to acquire credit ratings for companies from different communities. Experiments using real-world datasets showed that, when using various indicators to rate corporate credit, CCRELP could detect differences among the results for companies from different communities, thus indicating the possibility of developing a more reasonable evaluation system for companies with similar characteristics. The results also proved that the community-based credit rating system is an effective and feasible means of corporate credit rating.

Keywords: *Corporate credit rating, evaluation indicator, label propagation, community detection*

1. Introduction

Credit relationships have received growing attention amid the intensifying social and economic development. As a fundamental social factor, credit relationships form the basis of not only the existence and development of the economic society but also the entire market economy. In a way, the development of credit relationships can promote the progress of society itself. Since the market economy is, in fact, a credit economy, a good social credit system serves as a guarantee of an orderly market economy, an important condition for the effective prevention of financial risks, as well as a foundation for the normal operation of a modern economy [1]. Hence, increasing credit awareness, strengthening credit management, and improving credit rating systems are very important in promoting the development of the economic society.

Since credit ratings are a major measure of credit management, their importance has become increasingly prominent. Credit ratings are normally conducted by professional agencies or departments, who, by applying certain methods and procedures and after thorough examination, investigation, research, and analyses of the target business, provide a comprehensive assessment of its financial capacity and credibility. The essence of a credit rating is to evaluate the safety and reliability of a company's credit behavior and reveal its risks.

However, the current rating system targets single corporations and fails to comprehensively assess and analyze companies from different fields and geographical areas. As a result, horizontal comparisons of credit ratings among companies in the same industries and within the same areas are few. Due to the industrial and geographical differences among businesses, banks and related administrative departments tend to give

more support to companies located in economically developed areas in industries with better prospects. Consequently, an asymmetry appears in the development of enterprises from different areas and industries. In addition, the current rating system cannot identify the determining indicators of corporate credit in a given industry, causing inaccuracy in many corporate credit rating methods.

The community detection methods used in social network theory can help resolve these problems because, in a social network, network connections are usually dense within communities but sparser between them [2], and elements within the same community tend to have more similar characteristics. The connections among today's corporations usually exhibit the characteristics of a network [3]. Breaking this network down into different communities using community detection methods reveals that companies within the same communities are often the most similar to one another. Conducting corporate credit ratings on this basis can not only provide comprehensive assessments and analyses of enterprises within the same communities but also uncover the distinct impact each corporate credit rating evaluation indicator has on companies from different communities.

Following the above analysis, this paper proposes an evaluation method, the Corporate Credit Rating Evaluation, based on the community detection of label propagation (CCRELP) process. This method not only considers the density of the links among businesses but also divides the companies into several communities based on their industries and geographical locations and then calculates the group results for the evaluation indicator of each corporate community. The advantage of the CCRELP's classification of companies into different communities is that the connections among the companies within each community are denser. In addition, since the grouped companies belong to the same industry and are located in the same area, they tend to share more similarities, making the credit rating of the corporate groups more meaningful and providing a more accurate analysis of the importance of each evaluation indicator.

2. Literature Review

2.1. Corporate Credit Rating

Using 33 bankrupt US firms as a reference and comparing them to 33 non-bankrupt firms via 22 financial indicators, Altman (1968) proposed a linear model, the Z-score model, to assess business performance [4]. Altman *et. al.*, (1977) later built the Zeta model by adding two variables to the original Z-score model, marking the beginning of an era of using mathematical models for corporate credit rating [5].

Miller and Rojas used the Small Business Credit Scoring (SBCS) model in an empirical analysis of small and medium-sized enterprises (SMEs) in Brazil, Mexico, and Colombia [6], suggesting that the SBCS model could help simplify the estimation of default probabilities for risk classes of borrowers with statistical models during the credit origination process. In this study, we used a correlation between corporate credit and the personal financial status of its owner.

Angelini *et. al.*, (2008) applied two neural network systems developed for the credit assessments of Italian companies. The results showed that, as long as the data were pre-processed carefully, neural networks could effectively assess a borrower's default tendency. Their study exemplified the use of artificial intelligence technologies in the corporate credit rating process [7].

Gumparathi *et al.*, (2011) designed a corporate credit rating model for state-owned banks in India. They argued that improvements should be made to the SME credit evaluation model proposed by the new Basel Accords (Basel II), especially as most of the existing systems lacked appropriate weights. Based on a survey conducted among 30 credit-rating managers from 15 banks, they established a weighted average model for predicting credit worthiness and assigning reasonable weights to relevant indicators [8].

Vieira *et. al.*, (2012) suggested the Isomap algorithm for the credit assessment of SMEs. This algorithm produced a smooth distribution of ratings by projecting firms' balance sheet data into a one-dimensional component; several French SMEs were selected as samples to test the algorithm [9].

Although the abovementioned methods achieved good results in rating corporate credit, they could be used to evaluate the credit of only a single enterprise and could not collectively assess companies from a given industry or area. They thus failed to provide banks and relevant authorities with a fully accurate analysis of horizontal data comparisons, making the resultant data incomplete. This paper therefore proposes a corporate group credit rating system based on the detection of network communities.

2.2. Community Detection

Many community detection algorithms have been proposed in recent years. For example, the Girvan–Newman (GN) algorithm, proposed by Girvan and Newman (2001) [10], posits that the betweenness centrality of the edges that connect communities should be higher than the betweenness centrality of the edges within a community—the betweenness centrality of a given node being defined as the number of shortest paths between pairs of the other nodes running through the given node. By interactively removing the edges with the highest betweenness, the GN algorithm revealed a hierarchical clustering tree of the community structure. To obtain the optimal community structure of the clustering tree, Newman (2004) defined a modularity measure (Q) [2]. Thereafter, the optimization algorithm based on the modularity measure became one of the mainstream algorithms used in community detection-related studies. The modularity measure Q can be defined as $Q = \sum_i e_{ii} - a_i^2$, where e_{ii} indicates the proportion of edges that connect two nodes of the same community i out of the total edges in the network, while a_i indicates the proportion of edges with at least one connected node in community i , out of the total edges in the network.

Based on the idea of modularity, Newman further proposed a fast algorithm for community detection known as “Fast Newman” (FN) [11]. This algorithm started by considering each node in the network as an independent community and then joined the communities that shared the same edge into pairs. Next, the increment in Q after the pairing was calculated according to $\Delta Q = 2(e_{ij} - a_i a_j)$ wherein, if there was an edge connecting node i and j , then $e_{ij} = 1/2m$; otherwise $e_{ij} = 0$, $a_i = k_i/2m$, k_i is the degree of node i . The objective of joining communities was to choose the step with the maximum increase or the minimum decrease in Q . This action was repeated until the entire network was joined into one single community, at which point a *dendrogram* of the community structure could be obtained, and the optimal community structure could be acquired by cutting through this dendrogram at the places with the maximal value of Q .

Raghavan *et. al.*, (2007) were the first to propose the term “propagation algorithm” (LPA) [12]. The basic goal of this algorithm was to, first, initialize each node in the network with a unique label denoting its community and then, at each iteration, replace this label with the one most commonly shared by its neighboring nodes; when the community label of each node in the network was changed into the label shared by the maximum number of neighbors, the algorithm stopped running. Finally, the nodes with the same community labels were grouped into communities.

Rosvall and Bergstrom (2008) proposed an Infomap algorithm based on minimum description length [13] using the dissemination and diffusion of information to identify the community structure within a network. The key idea of the minimum description length was that any regularity in a given network could be used to effectively compress the data length. If the random walk path was encoded and the community structure was considered

as a regularity—and thus compressed—the community structures that corresponded to the minimum description length of the path became the effective community structures. The Infomap algorithm assumed that a random walker who has stepped into a community tends to spend longer periods of time within the community looking for the desired code length of the random walk path, of which the shorter the code length, the better the community’s identifiable structure.

Palla *et. al.*, (2005) formulated a critical path method (CPM) algorithm based on overlapping communities [14], suggesting that edges within communities were likely to form large complete subgraphs, while edges between communities would almost never do so. The CPM algorithm could identify the community structures within a network by uncovering the overlapping communities.

In this paper, we used community detection technology in network theories to divide companies into communities based on the degree of their connections, their industries, and geographical areas; as companies within communities share the most characteristics, we then conducted corporate group credit ratings related to their community classification.

3. The CCRELP

The CCRELP proposed by this study was applied as follows. First, an analytic hierarchy process (AHP) [15-16] was used to determine the weights of the indicators required for community detection (such as connections between enterprises, the corresponding industry, and geographical location). Then, label propagation was conducted for community detection after the degree of the similarity between each enterprise and its community was assessed. Finally, based on the corresponding indicators and their weights, as well as the actual company situation, scores were determined for each company, per indicator, which were then multiplied by the indicator weights to calculate the corporate group credit rating.

3.1. Determining the Indicator Weight for Community Detection

The indicators used in the CCRELP to detect communities comprise the connections between enterprises, the firms’ industry, and their location. Since the three indicators played different roles in detecting the communities within a corporate network, weights were assigned to each indicator using AHP.

As the businesses within corporate relations networks that have closer connections tend to be more similar, the connections between enterprises were more important during community structure detection than their industry or area. The AHP matrix used in this study is presented in Table 1.

Table 1. Comparison Matrix of the Three Indicators for Community Detection

C	C1	C2	C3
C1	1	2	2
C2	½	1	1
C3	½	1	1

Remarks: C1- Connections Between Enterprises, C2- Industry, C3-Area

The Sum-Product Method was used to calculate the weights of the three indicators as $W = (W_{c1}, W_{c2}, W_{c3})^T = (0.5, 0.25, 0.25)^T$. The greatest characteristic root λ_{max} was 3, with

its corresponding consistency index $CI = 0$. The correlation coefficient was $CR = CI/RI$, $CR = 0 < 0.1$. This also indicated that the matrix satisfied the consistency requirements.

Therefore, the weights assigned to the connections between enterprises, industry, and area in the community detection process were 0.5, 0.25, and 0.25 respectively.

3.2. Community Detection in Terms of Corporate Relations Networks

Since this study aimed to rate the group credit of close-knit corporate communities to perform horizontal comparisons of group credit between communities, we proposed a community detection method based on corporate relations networks for the enterprise identification methods—the Community Detection of Label Propagation in terms of Corporate Relations Networks (CDLP). The relevant terms are defined below.

Definition 1 (Community) [17] assumes that i is a node, k_i is the degree of node i , as for the adjacency matrix A_{ij} , $k_i = \sum_j A_{ij}$. G represents the entire corporate relations

network, whilst C represents a subgraph of G that includes node i —in other words, $C \subseteq G$. Thus, the degree of node i can be split into two contributions—the degree calculated based on the edges connecting node i to other nodes within the community C $k_i^{in}(C)$ and the degree calculated based on the edges connecting node i to nodes outside of community C $k_i^{out}(C)$, or

$$k_i(C) = k_i^{in}(C) + k_i^{out}(C) \quad (1)$$

giving us a strong community structure and a weak community structure. The community is considered strong if

$$k_i^{in}(C) > k_i^{out}(C), \forall i \in C \quad (2)$$

and weak if

$$\sum_{i \in C} k_i^{in}(C) > \sum_{i \in C} k_i^{out}(C) \quad (3)$$

The CDLP is intended to reveal the pattern in a corporate relations network where the connections were strong within communities composed of companies from the same industry and area and were weaker between communities.

Definition 2 (Label) assumes that node i belongs to community C ; then, the label for community C is $L(C)$. If $L(i) = L(C)$, we say that $L(i)$ is the community label of node i .

Based on the weights of the three indicators calculated from the AHP in the previous section, the definition of the community similarity proportion is given below.

Definition 3 (Community Similarity Proportion) assumes that C is a community and i is a node in the corporate relations network. The number of node i 's neighboring nodes labeled as $L(C)$ is N_i^C ; the number of its neighboring nodes labeled as $L(C)$ and belonging to the same industry as node i is I_i^C ; the number of its neighboring nodes labeled as $L(C)$ and belonging to the same area as node i is A_i^C . Thus, the community similarity proportion of node i to community C can be calculated as

$$F_i^C = 0.5 * \frac{N_i^C}{adj(i)} + 0.25 * \frac{I_i^C}{adj(i)} + 0.25 * \frac{A_i^C}{adj(i)} \quad (4)$$

wherein $adj(i)$ is the number of nodes adjacent to node i .

Based on the above definitions, the CDLP method proceeds as follows:

(1) Each node i in the corporate relations network is initialized with a unique label, and its industry and area are determined. Each node represents a company in the corporate relations network;

(2) An asynchronous update method is used to conduct label propagation on the label $L(i)$ of each node i in the network based on the formula below;

$$L(i) = \arg \max_L \sum_{j \in \text{adj}(i)} \delta(L(j), L) \quad (5)$$

(3) With the asynchronous updates, the label $L(i)$ of each node i is updated with the label of the adjacent community with the highest community similarity proportion with node i ;

(4) Step (3) is repeated until no more updates occur to any node in the network; the nodes with the same labels are then defined as a community.

These steps can divide companies within a corporate relations network into different communities. Step (1) initializes a unique label for each company in the network (*i.e.*, this step is the initialization of each company as a community); Step (2) conducts a label update operation to generate a primeval community structure with a certain degree of accuracy based on the density of the connections between companies; Step (3) refers to this primeval structure to assign companies to communities according to the firms' community similarity proportion with adjacent firms; Step (4) sets the conditions under which the operation should cease repeating, when companies with the same label are grouped into communities. On this basis, corporate group credit rating evaluation and analysis can be conducted for each community.

3.3. Corporate Group Credit Rating Evaluation

After the communities in the corporate relations network are detected, each company is assigned to a community composed of the enterprises from the same industry and area that are highly connected to it, simplifying the corporate group credit rating process.

Table 2. Corporate Credit Rating and Score System

Rating	Rating Description	Credit Score
A+	High grade	100
A	Upper medium grade	90
B	Average grade	80
C	Lower medium grade	70
D	Low grade	60

According to the industry and area of each community calculated in the previous steps, we assigned a different weight W_i to the indicators used in our evaluation system. Moreover, through an expert opinion survey, indicator scores were given to each company according to their actual operation situation S_i . Next, we used the formula

$$T = \sum_{i=1}^l (W_i \times S_i),$$

where l is the number of indicators included in the calculation, to return

the value of the corporate credit of each company. Finally, according to the corporate credit values, we rated companies as A+, A, B, C, or D. Details of the system are provided in Table 2. Thus, an integrated analysis of corporate group credit can be conducted to

provide banks and relevant authorities with more comprehensive data. In addition, further analysis can be conducted to evaluate the importance of each individual indicator.

4. Empirical Analysis

To verify the effectiveness and feasibility of the CCRELP, we applied it to the Karate dataset [18] and the dataset of 31 IT companies from Harbin, Heilongjiang Province, China. The datasets were divided into two experimental groups. In the Karate dataset, the accuracy of the CDLP's community detection was examined; in the second group, with the 31 Chinese IT companies, the indicator weights' impact on the credit rating was analyzed. The CDLP programming was conducted with Matlab 7.13.

Experiment 1: Analysis on the Karate dataset

Karate was the Karate Club Network at an American university. The network was constructed based on a two-year observation and the analysis of an American university karate club. The club had 34 members, all of whom were considered network nodes. The friendship between each set of two members was considered the edge connecting these two nodes. Due to a dispute, the club was eventually divided into two new clubs, one with the principal as the center (node 34) and the other with the coach as the center (node 1). Since the dataset of this network was established with a real community structure, it could be used as a reference to verify the accuracy of the CDLP in community detection. Hence, we used this dataset to measure the CDLP's accuracy in detecting communities in a network. In the Karate dataset, we assigned unified industry and area information to the nodes that were in the same community as node 1; this was done for node 34 as well.

After using the CDLP to detect communities in this network, we acquired the community structure shown in Figure 1. As can be seen from the Karate network graph, the community structure detected through our proposed method is identical to the real structure, indicating that the proposed CDLP can be used to accurately identify community structures and provides the basis for the accurate analysis of corporate group credit.

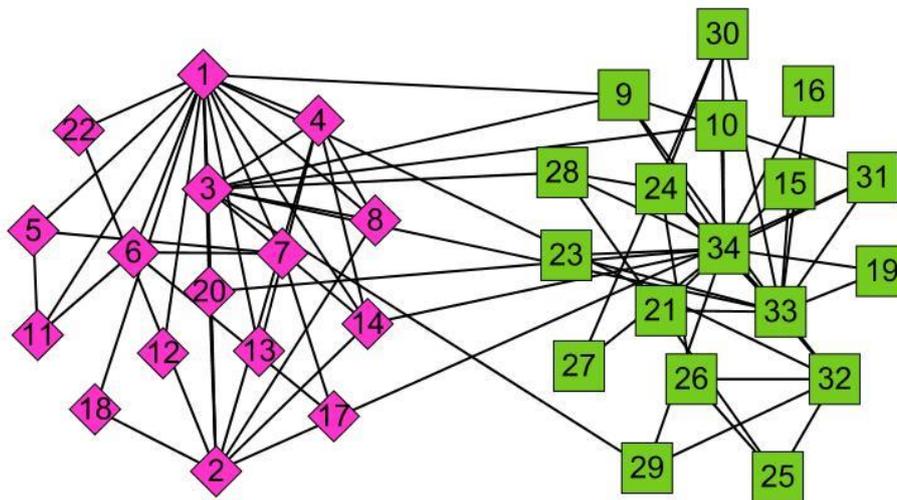


Figure 1. Using CDLP to Detect the Community in the Karate Network

Table 3. Corporate Credit Rating Evaluation Indicators

Level 1 Indicator	Level 2 Indicator	Indicator Properties	A+	A	B	C	D
Company's Basic Qualities	Quality of the Managers	Qualitative	10	8	6	4	2
	Quality of the Employees	Qualitative	9	7	5	3	1
	Quality of Technology	Qualitative	12	10	8	3	1
	Company Size and System	Qualitative	9	7	5	3	1
Company's Ability in Technology Innovation	Number of Patents and Proprietary Technology	Quantitative	30	20	15	10	3
	Investment Ratio in R&D (%)	Quantitative	5	4	3	2	1
	New Product Development Capability	Qualitative	15	11	10	6	4
Company's Financial Situation	Advance Level of Technology and Equipment	Qualitative	10	6	4	2	0
	Assets & Liabilities (%)	Quantitative	40	60	80	100	120
	Liquidity Ratio (times)	Quantitative	1.5	1.1	0.7	0.4	0.1
	Quick Ratio (times)	Quantitative	2.5	2	1.5	1	0.5
	Return On Total Assets Ratio (%)	Quantitative	8	4	2	1	-1
	Return On Equity Ratio (%)	Quantitative	15	8	5	2	-2
	Total Assets Turnover Ratio (%)	Quantitative	2	1.6	1	0.5	0
Company's Operational Management Ability	Accounts Receivable Turnover Ratio (%)	Quantitative	13	10	7	3	1
	Sales Contract Compliance Rate (%)	Quantitative	90	80	70	60	50
	Net Profit Growth (%)	Quantitative	25	18	10	3	-2
Business Growth and Development Capacity	Sales Growth Rate (%)	Quantitative	24	15	10	1	-8
	Growth Rate of Technology Investment (%)	Quantitative	18	13	9	5	1

Table 4. List of Indicator Weights

Level 1 Indicator	I1	I2	I3	I4	I5	I6	Level 2 Indicator	Weight
K1	0.2	0.3-0.4	0.15-0.175	0.15-0.175	0.15-0.175	0.15-0.175	Quality of the Managers	0.4
							Quality of the Employees	0.18
							Quality of Technology	0.3
							Company Size and System	0.12
K2	0.2	0.15-0.175	0.3-0.4	0.15-0.175	0.15-0.175	0.15-0.175	Number of Patents and Proprietary Technology	0.3
							Investment Ratio in R&D	0.2
							New Product Development Capability	0.3
K3	0.2	0.15-0.175	0.15-0.175	0.3-0.4	0.15-0.175	0.15-0.175	Advance Level of Technology and Equipment	0.2
							Assets-Liabilities Ratio	0.2
							Liquidity Ratio	0.13

							Quick Ratio	0.13
							Return On Total Assets Ratio	0.14
							Return On Equity Ratio	0.2
							Total Assets Turnover Ratio	0.5
K4	0.2	0.15- 0.175	0.15- 0.175	0.15- 0.175	0.3- 0.4	0.15- 0.175	Accounts Receivable Turnover Ratio	0.3
							Sales Contract Compliance Rate	0.2
							Net Profit Growth	0.5
K5	0.2	0.15- 0.175	0.15- 0.175	0.15- 0.175	0.15- 0.175	0.3- 0.4	Sales Growth Rate	0.3
							Growth Rate of Technology Investment	0.2

Remarks: K1- Company's Basic Qualities, K2- Company's Ability in Technology Innovation, K3- Company Financial Situation, K4- Company's Operational Management Ability, K5- Business Growth and Development Capacity

Experiment 2: Analysis of the IT Company Dataset

Information on the 31 IT companies was randomly selected from the Nangang and Acheng districts of Ha'erbin, Heilongjiang Province, China, including nine hardware development companies and 14 software development companies from Nangang District and eight software development companies from Acheng District. Following the advice of experts, and our own analyses, we selected the indicators and the rating and scoring systems shown in Table 3, for the CCRELP analysis.

Following expert advice and our own analyses, we assigned six groups of floating weights for the level 1 indicators and fixed weights for the level 2 indicators (see Table 4). Then, we evaluated the accuracy of the rating system according to the different weights to examine the weights' impact on the accuracy of the rating system.

Before being rated, companies had to be grouped into communities based on the density of the connections between them, their industries, and their areas to simplify the credit evaluation for each corporate group. For this dataset, the CDLP was used to detect its community structure. Figure 2, shows the results of the community detection via CDLP. The total dataset was divided into four communities (*C1* represented hardware development companies from Nangang District, *C2* was the software companies from Acheng District, and *C3* and *C4* were the software companies from Nangang District). During the grouping process, the company with code 21, which should belong to community *C3*, was mistakenly placed in community *C2* because it had denser connections with that community. Although communities *C3* and *C4* both contained software companies from Nangang District, they were separated into two communities because each had highly dense internal connections.

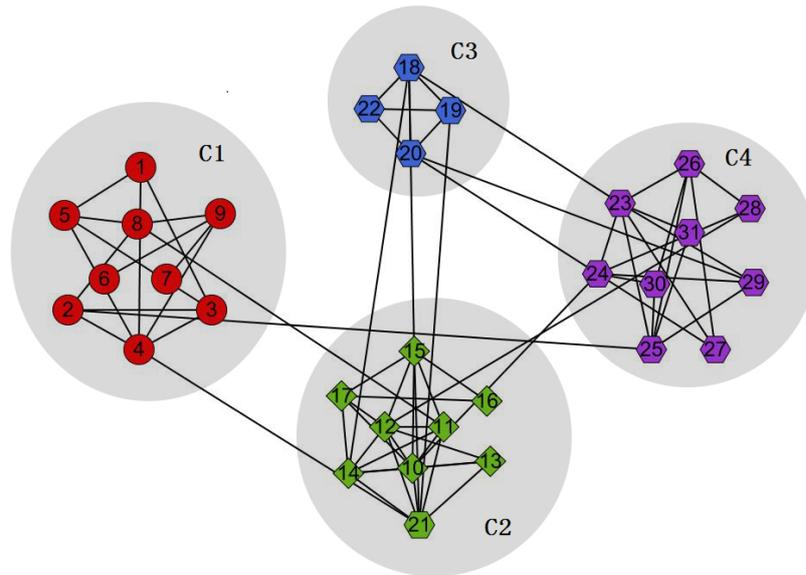


Figure 2. Community Detection Results of the IT Company Dataset

After the computing process, the dataset was divided into four communities, within each of which there was denser connections. Next, we conducted a corporate credit rating for each community group. In Figures 3(a), and (b), I1 indicates the accuracy level of the rating system when 0.2 was assigned as the weight for $K1$ to $K5$ and when the system was applied to companies in community $C1$. In Figure 3(a), I2 to I6 indicate the accuracy level of the rating system when 0.3 was assigned as the weight for $K1$ to $K5$ and when the system was applied to companies in community $C1$. In Figure 3(b), the indications of I2 to I6 are the same as in Figure 3(a), but the weight for $K1$ to $K5$ was 0.4. The graph shows that, for companies in community $C1$, raising the weight of $K1$, $K3$, and $K4$ could increase the accuracy of the rating; however, raising the weight of $K2$ and $K5$ would not only fail to improve the rating accuracy but would also sometimes reduce it, suggesting that the Company's Basic Qualities, Company's Financial Situation, and Company's Operational Management Ability played key roles in the corporate credit rating of the hardware companies from Nangang District.

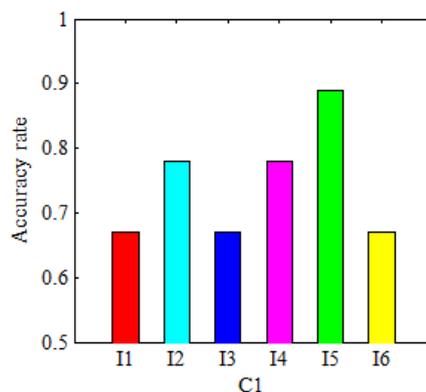


Figure 3(a). Changes in Rating Accuracy for C1 when Different Weights Were Assigned to Indicators (0.3 was Assigned as Weight for $K1$ to $K5$)

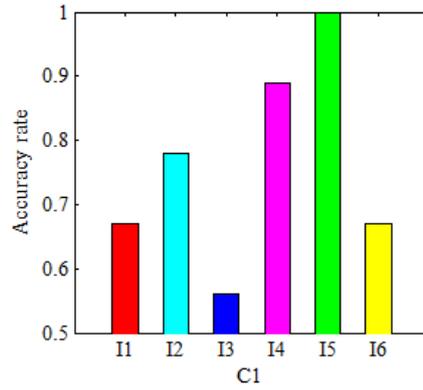


Figure 3(b). Changes in Rating Accuracy for C1 when Different Weights Were Assigned to Indicators (0.4 was Assigned as Weight for K1 to K5)

In Figure 4, I1 indicates the accuracy level of the rating system when 0.2 was assigned as the weight for *K1* to *K5* and when the system was applied to companies in community *C2*; I2 to I6 indicate the accuracy level of the rating system when 0.3 (Figure 4[a]), and 0.4 (Figure 4[b]), were assigned as the weights for *K1* to *K5*. As the figure shows, the increase in weight for *K1*, *K2*, and *K3* could improve the accuracy of the rating for community *C2*, indicating that Company’s Basic Qualities, Company’s Ability in Technology Innovation, and Company’s Financial Situation played important roles in the corporate credit rating of the software companies from Acheng District.

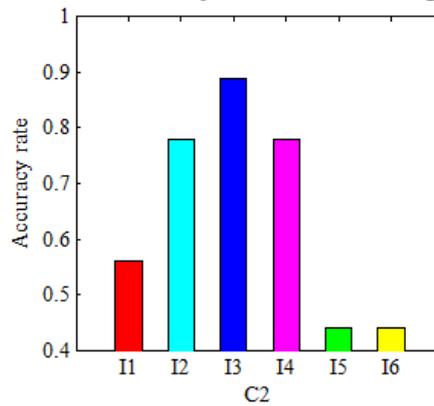


Figure 4(a). Changes in Rating Accuracy for C2 when Different Weights Were Assigned to Indicators (0.3 was Assigned as Weight for K1 to K5)

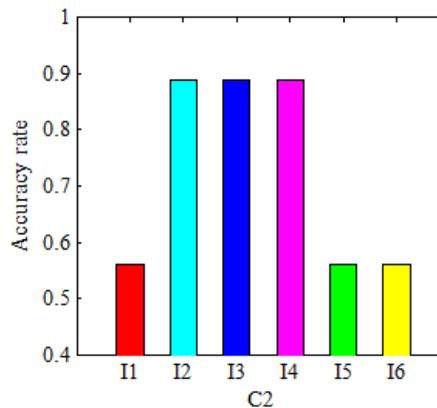


Figure 4(b). Changes in Rating Accuracy for C2 when Different Weights Were Assigned to Indicators (0.4 was Assigned as Weight for K1 to K5)

Figure 5, shows the changes in rating accuracy for community *C3* with different indicator weights. Increasing the weights for *K3* and *K5* increased the accuracy of the rating for community *C3*, suggesting that the Company's Financial Situation and Business Growth and Development Capacity played important roles in the corporate credit rating of companies in community *C3*.

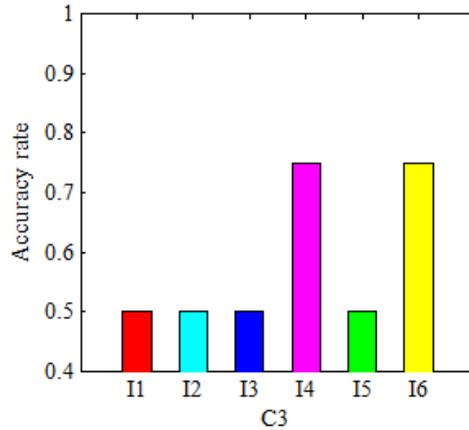


Figure 5(a). Changes in Rating Accuracy for C3 when Different Weights Were Assigned to Indicators (0.3 was Assigned as Weight for K1 to K5)

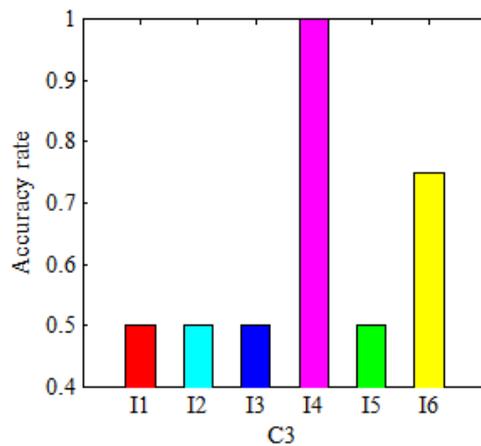


Figure 5(b). Changes in Rating Accuracy for C3 when Different Weights Were Assigned to Indicators (0.4 was Assigned as Weight for K1 to K5)

Figure 6, shows how the rating accuracy changed when different indicator weights were applied during the evaluation of companies from community *C4*. Increasing the weights for *K2*, *K4*, and *K5* could increase the accuracy of the rating for community *C4*; thus, the Company's Ability in Technology Innovation, Company's Operational Management Ability, and Business Growth and Development Capacity played important roles in the corporate credit rating of companies in community *C4*.

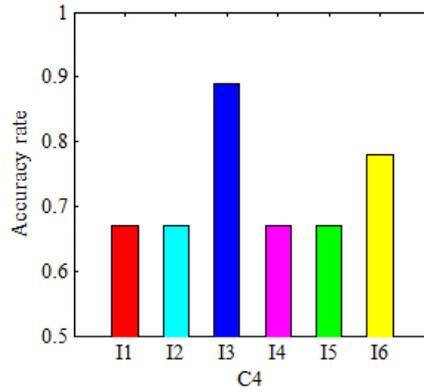


Figure 6(a). Changes in Rating Accuracy for C4 when Different Weights Were Assigned to Indicators (0.3 was Assigned as Weight for K1 to K5)

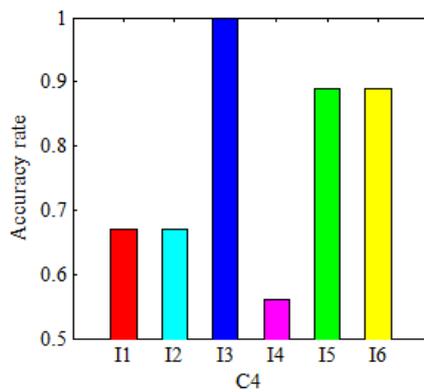


Figure 6(b). Changes in Rating Accuracy for C4 when Different Weights Were Assigned to Indicators (0.4 was Assigned as Weight for K1 to K5)

This experiment shows that it is feasible to conduct corporate group credit rating evaluation by dividing companies into different communities based on the density of their connections and their industries and locations. The influence of each indicator on the communities was different when there was a difference in the industry classification (such as *C1* and other communities), in area allocation (such as *C2* and other communities), or even within the same industry and area when there was a difference in the density of connections between companies (such as *C3* and *C4*). The difference when each rating indicator was applied to different communities could be calculated, making it possible to develop a more reasonable evaluation system for companies with similar characteristics and making it easier to collect data for horizontal comparisons across corporate groups.

5. Conclusions

Current rating systems focus on rating single companies and lack group analyses on companies with similarities or horizontal comparisons of corporate credit rating results, hindering the development of an effective evaluation system. Aiming to solve these problems, this paper proposed the CCRELP. Considering the closeness between companies as well as their industries and geographical areas, this method first grouped the companies into communities and then calculated the group credit rating of different communities with different indicator weights. Since companies in the same community are generally closely connected and belong to the same industry and area, they tend to have more similar characteristics, making the group credit rating evaluation more valuable. The result of the experiments confirmed that the CCRELP can be used to

categorize companies with denser connections and more similarities into the same community and to calculate the influence each indicator in the evaluation system would have on rating accuracy when applied to each community. This discovery could be used to develop a more rational credit evaluation system for highly similar corporate groups. The CCRELP is thus an effective method for corporate group credit rating.

References

- [1] H. J. Sun, "Credit guarantee of small and medium-sized enterprises", Zhejiang University Publishers, Zhejiang, (2003).
- [2] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks", *Physical Review E*, vol. 69, no. 2, (2004), pp. 026-113.
- [3] M. E. J. Newman, "Networks: An Introduction", Oxford University Publishers, London, (2010).
- [4] E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, vol. 23, no. 4, (1968), pp. 589-606.
- [5] E. I. Altman, R. G. Hadelman and P. Narayanan, "Zeta analysis: a new model to identify bankruptcy risk of corporations", *Journal of Banking and Finance*, vol. 1, no. 1, (1977), pp. 29-54.
- [6] M. Miller and D. Rojas, "Improving Access to Credit for SMEs: An Empirical Analysis of the Viability of Pooled Data SME Credit Scoring Models in Brazil", Colombia&Mexico, World bank working paper, (2004).
- [7] E. Angelini, G. di Tollo and A. Roli, "A neural network approach for credit risk evaluation", *The Quarterly Review of Economics and Finance*, vol. 48, no. 4, (2008), pp. 733-755.
- [8] S. Gumparathi, S. Khatri and V. Manickavasagam, "Design and Development of Credit Rating Model for Public Sector Banks in India: Special Reference to SMEs", *Journal of Accounting and Taxation*, vol. 3, no. 5, (2011), pp. 105-124.
- [9] A. Vieira, B. Ribeiro and N. Chen, "Credit Scoring for SME Using a Manifold Supervised Learning Algorithm, Intelligent Data Engineering and Automated Learning-IDEAL", Springer Berlin Heidelberg, (2012), pp. 763-770.
- [10] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks", *Proceedings of National Academy of Science*, vol. 9, no. 12, (2002), pp. 7821-7826.
- [11] M. E. J. Newman, "Fast algorithm for detecting community structure in networks", *Physical Review E*, vol. 69, no. 6, (2004), pp. 066-133.
- [12] U. N. Raghavan, R. Albert and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks", *Physical Review E*, vol. 76, no. 3, (2007), pp. 036-106.
- [13] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure", *Proceedings of National Academy of Science*, vol. 105, no. 4, (2008), pp. 1118-1123.
- [14] G. Palla, I. Derényi and I. Farkas, "Uncovering the overlapping community structure of complex networks in nature and society", *Nature*, vol. 435, no. 7043, (2005), pp. 814-818.
- [15] Z. ZHOU, Z. Y. SUN, F. LIU and H. G. AO, "Study of Risk Grade Classification Model for Enterprises Based on Analytical Hierarchy Process", *Journal of Harbin University of Science and Technology*, vol. 1, (2010), pp. 103-107.
- [16] P. Li and Y. H. Li, "Research of Comprehensive Strength Evaluation of Cities Conference and Exhibition Economy Based on AHP", *Journal of Harbin University of Science and Technology*, vol. 2, (2010), pp. 115-119.
- [17] F. Radicchi, C. Castellano and F. Cecconi, "Defining and identifying communities in networks", *Proceedings of National Academy of Science*, vol. 101, no. 9, (2004), pp. 2658-2663.
- [18] W. W. Zachary, "An information flow model for conflict and fission in small groups", *Journal of Anthropological Research*, vol. 33, no. 4, (1977), pp. 452-473.