CSCanada
Canadian Social Science
Vol. 14, No. 9, 2018, pp. 58-63
DOI:10.3968/10607

ISSN 1712-8056[Print]
ISSN 1923-6697[Online]
www.cscanada.net
www.cscanada.org

Study of Exploring Hidden Relationship Among Commercial Bank Customers Based on Complex Network Theory

SU Qian[a],*; YAN Guanghui[b]

[a]Management Faculty, Chengdu University of Information Technology, Sichuan, China.
[b]School of Electronic and Information Engineering, Lanzhou JiaoTong University, Gansu, China.
*Corresponding author.

Received 16 June 2018; accepted 30 August 2018
Published online 26 September 2018

Abstract

Commercial banks are very keen to understand the relationship among their customers, but it is difficult to get directly. This paper proposes a method using the transaction behavior data between customers, combined with complex network theory to establish a network on relationship of customer transactions, so as to obtain the hidden relationship among customers. Based on real bank customers 3 months of transfer transaction records, we established Transaction complex network model of bank customers. The quantitative analysis of the complex network model shows that it satisfies the characteristics of scale-free and small world networks. This study demonstrate an approach by applying complex theory to solve customer relationship management problems, and the findings are helpful for banks' in depth analysis of their customers.

Key words: Complex network; Customer relationship; Transfer record

Su, Q., & Yan, G. H. (2018). Study of Exploring hidden relationship among Commercial Bank Customers Based on Complex Network theory. *Canadian Social Science, 14*(9), 58-63. Available from: http://www.cscanada.net/index.php/css/article/view/10607
DOI: http://dx.doi.org/10.3968/10607

INTRODUCTION

With the incessant improvement of Internet technology, prompt development of global economy and decreasing discrepancy between production and service, management concept focused on production-sale mode has been progressively taken the place of oncustomer-service pattern. Customers of commercial bank on the basis of their own economic demands accomplish the financial transaction via the plane established by financial systems themselves (Zhang, 2012, pp. 68-72).At present, customers relationship analysis technology are not well equipped to response the analytical requirement of complicated concern of amass transaction data due to unilateralism, directness and redundancy, and accordingly lack a uniform and effective management mode to legitimately serve and manage complex and diverse customer transaction information.

Vase customer base is the group served by bank, consist of numerous different individuals, and formed network structure naturally. Research on the topology of customer transaction network may be of great importance at improving agent management of bank. Research on topological structural Characteristic may and should guide financial systems to optimize business pattern, transform intricate customer management to clear, goal-directed, organized complex network management, reinforce agent relation regulation of bank, enhance service quality and customer satisfaction, and acquire more client resource to raise company's influence.

1. LITERATURE REVIEW

Study of complex networks started at the famous Swiss mathematician Eular pioneered graph theory. After a long period of time, the subject was limited to theoretical research level. The late 1950s, the network model proposed by Erdos and Renyi - random network (ER random network) created a precedent for the systematic study of complex networks, and triggered a new situation of scientific research(Liu, 2009). But until the last century, researches scholars expended were still mainly based on the complete rules and complete random network structure assumptions. Until 1998 Watts, etc. published small-world network model in Nature(Watts & Strogatz, 1998, pp. 440-442). Followed, Barabasi and other researchers published paper constructed scale-free network model in 1999 in Science(Barabasi & Albert, 1999, pp. 509-512), leading the complex network to a new historical stage.

The core idea of complex network is to turn relations in all reality system entities and entities into nodes and edges of the network, to describe the relationship between the various parts of the system in the form of network, and to facilitate in-depth analysis of the topological characteristics of the system structure, revealing essential law of reality system. Mr. Qian had given a more strict definition: If a network which has some or all of the features of self-similarity, self-organization, attractor, small world, scale can be called complex network (Qian & Dai, 2006. p. 94).

Based on this, the important research trends of the current complex system and network analysis is using mathematical analysis, calculation, simulation models and other methods to analyze the network model established on large-scale complex systems with complex structure and dynamic behavior. In recent years, based on the classical field of complex networks such as information network, biological network has received a considerable amount of research results, however, the study of complex economic network is still in its initial stage. In the financial sector, most foreign scholars govern the operation of the financial system or the economy using random graph theory model laws, and study the health of the financial markets modeling on financialsystem using statistical physics approach. Representative studies include CB model constructed by Rama Cont and Bouchaud for describing financial market volatility(Rama & Bouchaud, 1998, pp. 170-196), which is based on random graph model, showing the long tail of the return distribution characteristics; the Sznajd model built by Sznajd-Weron for describing the stock market price formation(Sznajdweron & Weron, 2000, pp. 115-123), usually used in the rules of the network Ising-type model. Moreover, Challet and Zhang put forward the famous game MG model(Challet, Marsili, & Zhang, 2005) which reflect the overall performance of the market by analyzing the microscopic behavior of individual decisions. These studies focused on small-scale networks while the traditional random network model has been difficult to make an objective description for large-scale network, which has a complex structure.

Complexity science research proposes a new vision, which studies characteristic exhibited in complex systems interactions between the entities. For complex customer trading network, the traditional random network model is difficult to make a precise description of its topological properties, and further, whether bank customers trading network meets random network model is also questionable. Therefore, assist the bank more scientific, reasonable and effective business relationship with customers, improve customer loyalty and satisfaction, we try to introduce complex network theory on the basis of previous studies into the field of data analysis of customer transactions, build a commercial banking customers financial transactions network based on customer transaction information, further quantitative analysis of the topology of the network characteristic parameters, such as degree of customer transactions, distribution the average path length, average clustering coefficient, etc., verify the customer's transaction network scale-free and small-world characteristics.

2. METHODOLOGY OF BUILDING COMMERCIAL CUSTOMER TRANSFER COMPLEX NETWORK

The Commercial Bank Information Technology Division has a mass of financial transactions data, which are generated by reception bank transaction terminal or online banking transaction terminal and recorded in the business table of backstage transaction database of trading platform. Each data record represents a transaction between customers. From the perspective of complex network data analysis, a transaction can be abbreviated by a graph G_k consisted of a vertex collection V_k={v_{k1},vk_2}, wherein v_{k1} represents transaction account and v_{k2} represents related transaction account, and an edge collection E_k connected those two vertices. According to the research needs, the transaction amount, time could represent the weight of edge, by which constructed complex networks further enrich the relationship between customer transaction information. In order to facilitate research, this article focuses on the existence of transactions between the customer contact within a certain period of time, not to consider the transaction amount and time and other relevant information.

2.1 Definition of Customer Transaction Network

Definition 1: To construct an undirected graph as shown in (1), in which the vertices represent bank customer transaction account and edges represent relationship between customer financial trading:

$$G=(V, E) \tag{1}$$

Wherein, $V=\{v_1,v_2,...,v_n\}$ is the set of vertices in customer transaction network, consisting of transaction account in customer transaction data. $E=\{e_1,e_2,...,e_n\}$ is the set of edges, which represents transaction relationship between v_i and v_j.

Definition 2: Using adjacency matrix A shown as (2) to express topology of customer transaction network.

$$A_{n\times n}=\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}, \quad (2)$$

Wherein, $a_{i,j}=\begin{cases}1, \text{if } v_i \text{ connects } v_j \\ 0, \text{others}\end{cases} (1\le i,j\le n)$

2.2 Analysis of Degree distribution characteristics

Degree k_i of node i is the number of connections or edges the node has to other nodes(Joshi, Bertino, Latif et al., 2001).In the network data analysis, the degree is a simple but important concept.

Degree distribution is an indispensable part of complex networks, Used to describe the degree distribution of the nodes in the network. When randomly selecting a node from the network, we describe the fraction of nodes in the network with degree k as p(k). We observe degree distribution of the nodes by drawing p(k) ~ k, or calculating moment of distribution. p(k) of the n-order moments defined as (3)

$$<k^n>=\sum_k k^n p(k) \quad (3)$$

Simply, the greater the degree of a node is, the more important it would be to some extent(Cajueiro & Tabak, 2007, pp. 6825-6836). The corresponding transactions on the client network, the greater the degree is indicate that the customer has more trading relationship with others, have a larger impact on other customers in the selection of services, and is relatively more important for bank.

2.3 Analysis of Average Path Length and Clustering Coefficient

The distance d_{ij} between node i and node j in network is defined as the shortest path between two nodes. The maximum distance between any two nodes in network is called a diameter of network, referred to as D, that is

$$D=\max_{i,j} d_{ij} \quad (4)$$

The average path length, the average shortest distance between all nodes, is

$$L=\frac{1}{\frac{1}{2}N(N+1)}\sum_{i\ge j} d_{ij} \quad (5)$$

Wherein, N is the number of vertices.

The average path length is one of important feature measures in complex network analysis, which show the connectivity and transmission performance between the customers in customer transactions network(Barabasi & Alber, 1999, pp. 509-512).

For any node ion the network, its clustering coefficient C_i is given by the proportion of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them, that is

$$C_i=\frac{2M_i}{k_i(k_i-1)} \quad (6)$$

Wherein, k_i is the number of neighbors of a vertex, $\frac{1}{2}k_i(k_i-1)$ is the maximum number of edges that could exist among the vertices within the neighborhood, and M_i represents the number of edges that actually exist.

Network clustering coefficient is the average of all the nodes in the clustering coefficient, that is

$$C=\frac{1}{N}\sum_1^N C_i \quad (7)$$

Clustering coefficient reflects the extent of the network group. Corresponding to bank customers transaction network, it reflects the degree of polymerization tendency among customers.

3. RESULTS AND DISCUSSION

3.1 Results of Customer Transaction Network

We selected customers' personal transactions records of the western provincial capital city bank's 2015 first quarter as sample data, extracting trading hours, the account, the account related transactions, transaction amount and other raw data, and decrypting for confidential data.

The bank customer transaction data used to study is a total of 77,502. Tissue paper in accordance with the time period of data facilitate data analysis from the perspective of evolutionary time, the basic information as shown in Table 1.

Table 1
Network Statistics of 1th Quarter 2015

Object	Nodes	Edges
January	29395	27367
February	13997	24037
March	14332	26098

Using Gephi to build customer trading network (January) as shown in Figure 1

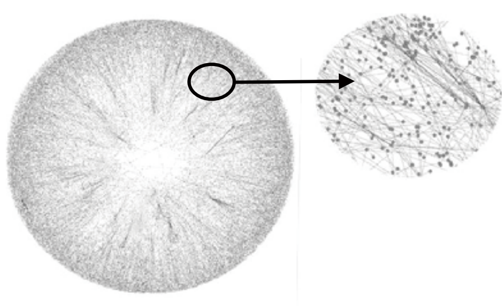

Figure 1
Network Constructed on Transaction Data of January

Any node in Figure 1 represents an efficient trading account, edge between nodes represents the financial relationship between the associated accounts

Taking into account the large number of customers trading network nodes and edges, in time-varying characteristics of complex architecture, it is difficult to obtain intuitive topology features, and accordingly can not determine which network model met, and therefore can not directly use the existing network model to analysis its topological nature objectively and systematically. Thus, the follow-up andin-depthstudy of the smooth progress of the foundation is to quantitative analysis the characteristic parameters of the network topology, such as degree, degree distribution, average path length, clustering coefficient average and to determine the network model of customer transactions network.

3.2 Degree Distribution Characteristics of Transaction Network

To analyze the trend of relevant time period of the plurality of transaction network topology, the paper respectively carries out the probability distribution statistics of the sample data in January, February and March, which draw a scatter plot of the double logarithmic coordinates, and do the linear regression, experimental results shown in Figure 2, Figure 3 and Figure 4. For comparison in each figure, the left chart is the customer trading network of distribution, with node degree K as abscissa and the number of nodes with corresponding degree as ordinate, and the right chart is degree distribution - double logarithmic scatter plot fitting drawn, with node degree K as abscissa and the logarithm of probability of the number of nodes with K degree on the total number.

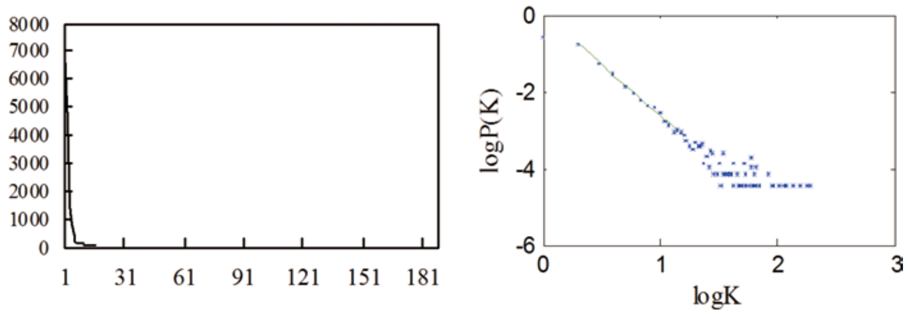

Figure 2
January Customer Transaction Network Degree Distribution Graph (Above) and Degree Distribution - Double Logarithmic Scatter Plot

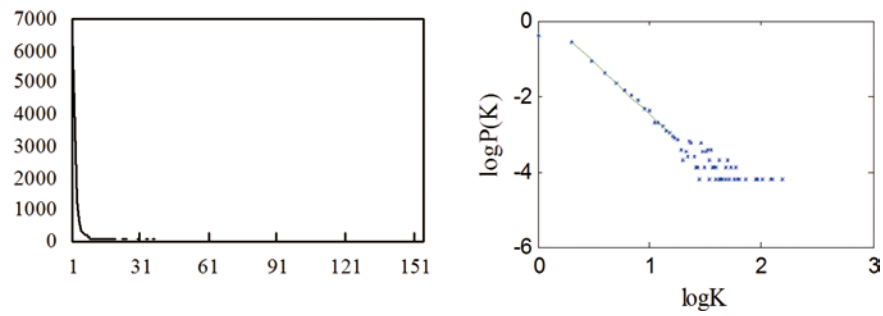

Figure 3
February Customer Transaction Network Degree Distribution Graph (Above) and Degree Distribution - Double Logarithmic Scatterplot

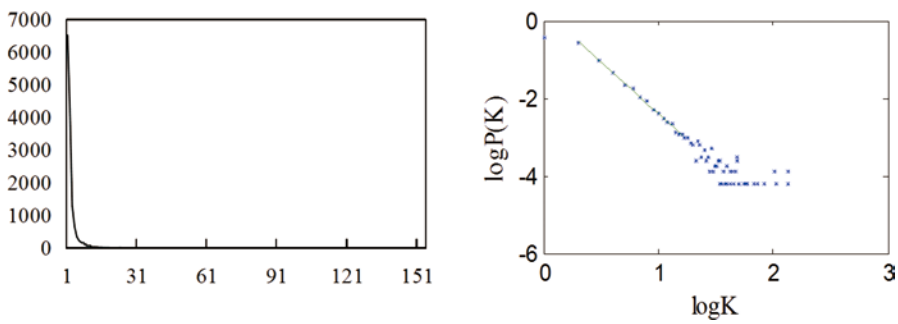

Figure 4
March Customer Transaction Network Degree Distribution Graph (Above) and Degree Distribution - Double Logarithmic Scatterplot

Degree distribution function of complex networks should comply with the power law, with the characteristics of the node degree having no obvious scale, called scale-free feature. Accordingly, we can verify whether the network node degree distribution follows power-law distribution to validate the scale-free network(WAN, CHEN, & Chen, 2007), namely:

$$P_k \sim k^{-r}$$

Therein, r is Network Scale.

By regression analysis of 3.1.2 Experiment , we calculated regression coefficients shown in Table 2.

Table 2
Experimental Result of Regression Coefficients

Month	r
January	3.3686(2.4<K<15.8)
February	3.3859(2.3<K<15.7)
March	3.3151(2.4<K<15.7)

In summary, we illustrate that January, February, March, customers transaction network follows the power law distribution, and its corresponding interval have scale-free characteristics.

Scale-free networks are characterized by highly dispersivity and self-similarity of degree distribution, that is, most node degree is very small, but few central core nodes have a high degree. This feature can indicate that a few customer in the commercial banking customer transaction network trade frequently and have a broadly economic relationship. Compared with a small degree of customers, they generate more trading relationship with others. Such customers generallyhave highly degree of loyalty and great value (ie, core customers), their related businesses larger, making a high contribution to the bank's business, giving more benefits to the financial system.Meanwhile, core customer would impact on small adjacent customer so that lead more customers to understand the bank's products or services, and ultimately bring the formation of purchase. If such customers cancel the account because of subjective or objective factors, it will likely have a negative impact on network connectivity customer transactions, and may even cause the entire network interruption or paralysis, which lead to the bank's business bound to be affected accordingly.Thus, banks should pay more attention to these customers, and take the appropriate service strategy to maintain and develop the scale of such customers.

3.3 Average Path Length and Clustering Coefficient of Customer Transaction Network

Since the distribution experiments reflecting the change of the January, February, March data results in customers transaction network, we still use these three months data as representation to do experiments. According to the formula (5)~(8), experimental results for January, February, March customer transactions network as shown in Table 3.

Wherein C, D, E are the network clustering coefficient, the network diameter, and the average path length.

Table 3
Diameter, Average Path Length and Clustering Coefficient of Customer Transaction Network

Month	C	D	L
January	0.0136	10	2.9982
February	0.0412	8	2.1901
March	0.0231	8	2.3115

Analyzing Table 3, we obtain that, for example, in the customer trading network, the network diameter is 10 in January, and the average path length is 2.9982, suggesting that a customer takes an average of three customers through other customers will be able to find transactions with which they have relations. Short average path reflects the close economic ties between the customers, and high efficiency in information exchange. If a bank or any financial enterprise in the field want to get a competitive advantage in the industry, it is important to recognize the importance of business time, take appropriate measures to improve the efficiency of the transaction between the customer, and achieve maximum customer satisfaction as far as possible. To achieve this goal, enterprises must reduce the average path length of the entire transaction network.

Small world network is the outward manifestation of the average path length of the network. The majority nodes in small-world network is not connected to each other directly, but most nodes can be reached from every other node by a small number of hops or steps. If node represent individual and edges represent connections between individuals known each other in small-world network, then the network may reflect a small world phenomenon that strangers can get to know each other by individual they commonly know. Currently two key features are available to measure whether a network has a small world: the average path length and clustering coefficient. If the network has both a large clustering coefficient and a smaller average path length, it can be called small-world networks.

Using Small world coefficient conversion inequality proposed by Sporns (Sporns & Honey, 2007), we can verify small-world characteristic of a network. Comparing the network with the same scale random network, if it satisfied (9) defined below:

$$\frac{C}{C_{ran}} \rangle \frac{L}{L_{ran}} \qquad (8)$$

then the network is a small world network, which represent clustering coefficient and the average path length of the same scale random network.

We generate ER random network with same scale as customer trading network in January, February, March by Gephi and calculate their network clustering coefficient and average path lengthseparately, comparing the results shown in Table 3. Wherein C_{ran} and L_{ran} represent the clustering coefficient and average path length of random network generated corresponding to each month customer transaction network.

Table 4
Comparison Between Customer Transaction Network and ER Random Network

Customer transaction network		ER Random Network		Comparison	
C	L	C_{ran}	L_{ran}	C/C_{ran}	L/L_{ran}
0.0136	2.9982	0.0141	3.5516	1.5531	0.8441
0.0412	2.1901	0.0123	2.9980	3.3496	0.7305
0.0231	2.3115	0.0101	3.1452	3.1089	0.7349

Table 4 shows that the three-month customer transaction network satisfies the formula (9), proving that customer transaction network has small-world characteristic. The experimental results reflect that, on the large-scale network of customer transactions, the relationship between the client body is closer, and information exchange is more efficient, protecting bank customers trading levels. Slightly higher degree of aggregation of the client node can promote the diffusion and dissemination of information and funding, but the robustness of the network is likely to diminish, thus to improve management efficiency, and strengthen customer management, the bank customer transactions in-depth study of the stability of the network should be the next step.

CONCLUSION

In this paper, we study the city bank customer transactions data using the theory and method of complex network, define customer transaction network model, and build customer trading network topology based on trading links between customers. Also, we make a statistical analysis of the topology parameters in different periods of customer transactions network, such as degree distribution, average path length, clustering coefficient, describe relevance of the complex network parameters in customer transactions network, and verify the network in line with scale-free and small world characteristic.The findings provide a scientific basis and guidance to further improve the level of banking services and customer satisfaction, but also establish the foundation for more in-depth study of customer trading network structure from the micro level.

Introduction of complex network theory is the trend of the research for bank customers network. At present, relevant research being still in its infancy, this paper merely analyzes the structure of the topological characteristics of the customer transaction network.The next step is to find the unique characteristics based on results of this study, comparing with current research on online social networks . On this basis, we can study cascade and interaction between network nodes combining with internal dynamics mechanism of complex network, to explore fund laws between nodes, which is of an social-economic importance on the development and expansion of city commercial banks in the new economy normal.

REFERENCES

Barabasi, A-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509-512.

Cajueiro, D. O., & Tabak, B. M. (2007). The role of banks in the Brazilian interbank market: Does bank type matter? *Physica A Statistical Mechanics & Its Applications, 387*(27), 6825-6836.

Challet, D., Marsili, M., & Zhang, Y. C. (2005). Minority games. *Interacting agents in financial markets.* (Reprint of the 2005 hardback ed.).

Joshi, J. B. D., Bertino, E., Latif, U. et al. (2001). TRBAC: A Temporal role -based access control model.

Liu, J. X. (2009). Complex network and review of domestic research. *Chinese Journal of Social Science, 17*(4).

Qian, X. S., & Dai, R. W. (2006). Emergence of metasynthetic wisdom in cyberspace. *Complex Systems and Complexity Science*, (02), 94.

Rama, C., & Bouchaud, J-P. (1998). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics, 4*(2), 170-196.

Sporns, O., & Honey, C. J. (2007). Identification and classification of hubs in brain networks. *Plos One, 2*(10), e1049-1-e1049-14.

Sznajdweron, K., & Weron, R. (2000). A simple model of price formation. *International Journal of Modern Physics C, 13*(1), 115-123.

Wan, Y. S., Chen, Z., & Chen, X. R.(2007). Analysis of macro structure model of complex banking. *Journal of Shanghai Jiaotong University, 41*(7), 1161-1164.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 440-442.

Zhang, C. H., & Lv, Y. (2012). A study of identification of commercial bank's customers based on social network analysis. *Finance Forum*, (08), 68-72.