Research on the Partnership between Supply Chain Enterprises Based on Complex Network

Jun-jin WANG\textsuperscript{1,a}

\textsuperscript{1}School of Maritime Economics and Management, Dalian Maritime University, Dalian, Liaoning, China
\textsuperscript{a}1204703207@qq.com

Keywords: Complex Network, Energy Supply Chain, Cooperation, Link Prediction.

Abstract. Under the perspective of complex networks, supply chain is abstracted as complex network which takes enterprise as node, and takes the relation between enterprises as edge. Taking energy supply chain (ESC) for example, this paper analyses the accuracy of prediction by combining with other seven real networks. The result shows that the predictive effect of half part information index is obviously better than the index of node part information and global information, the predictive effect among some individual global information index is higher while the accuracy are volatile; under the network with smaller cluster coefficient, the node part information has poor predictive effect but more stableness accuracy. Distinguish the accuracy of prediction clearly by comparing with seven real networks, which are to predict the “unknown links” and “future links” more accurately.

1. Introduction

In the complicated and changeable market environment, supply chain is a dynamic network, which consisting of supplier, manufacturer, distributor, retailer and consumer. It’s a typical complex system, because its complexity cannot only be found in its constitution entities and structure but also in the complex interaction between entities and the market environment. With the rapid development of the economy, the scale and structure of the supply chain network becomes larger, and the environment that all nodes of enterprises in the network should face becomes much more complex. Meanwhile, management problems are more prominent. Therefore we need to survey the problems from a new perspective and use new theory and knowledge to do research on the current supply chain. There are many researches that domestic and foreign scholars have studied on the supply chain based on the complex network by now. Choi et al. \cite{1} viewed the supply chain system as an adaptive system and the emergence of the supply chain network is not a destination site. Venkata et al. \cite{2} analyzed the complex network structure characteristics of the supply chain and the relationship between supply chain members and the system goal by using the self organization theory of evolution. Pathak et al. \cite{3} used the simulation method of Multi-agent to simulate the order fulfillment in the supply chain network based on the CAS theory. Kühnert et al. \cite{4} proved that the supply chain network of urban material obeys the scale distribution, and there is a small number of core supply nodes in the urban material supply network, which plays a decisive role in the allocation of goods and materials. Zhang Tao thought the changing of the overall operational environment is not enough to solve complex problems in today’s supply chain. He analyzed the complex adaption system and the supply chain system by using the theory of complexity science. What’s more, simulation and performance evaluation of the supply chain system is carried out on the Swarm simulation platform\cite{5}.

Link prediction uses the known nodes in the network and the properties of the network structure in order to predict the possibility of a link between two nodes that have not yet been generated. Indeed, it is to predict the “unknown link” which has not been found, but also to predict some that are not existing or possibly existing. We can obtain a good prediction effect by using the external environment information of the network nodes. But in many cases this information is hard to obtain, for its confidentiality or its authenticity. By now, using similarity or proximity prediction is the
main method for network link prediction. Of course, there are many measures to characterize the similarity, however, using the property of the node is the best way. Nodes whose properties are similar in the network are more likely to be linked[6]. A large number of studies show that using the similarity of the nodes in the network to predict the link result will also show good performance. For example, Liu Hongkun and Lü Lin [7] have researched the evolution mechanism of urban aviation network in China by using the link prediction to couple the common neighbor’s number and geographical position, population, GDP and the third industry output value of the network nodes.

In view of the cooperation between enterprises in the energy supply chain, this paper offers a new perspective that using complex network to research problems about evolution mechanism of energy supply chain cooperation, takes the new energy supply chain as an example, describes the supply chain network deeply, and uses link prediction to research cooperation problems of node enterprise.

2. Supply Chain Network Link Prediction

We can use the link prediction to predict an energy supply chain cooperation network connection. At first, the evaluation index of predicting connection accuracy is introduced, and the node structure information is quantified in the energy supply chain network. Then we quantify the network after the edge of the partition (divided into training and test set). By using 17 kinds of similarity indexes, the forecast edge (the edge of the test set) can be scored, and the accuracy of the prediction is visualized by the evaluation index, and the comparison is made.

2.1. Evaluating Index

For any non-direction network $G$, the total number of the network is $M$ and the total number of edges is $U$, so the $U = M(M - 1)/2$. But there is no link between each node in the actual situation. Based on known network properties and a given link prediction algorithm, we give every pair of nodes without edges $v_i$ and $v_j$, which represents the similarity of a node’s relationship. It is related to the probability of the link between the nodes, the greater the value, the greater the possibility of two link. We sort the value of all the nodes connected from large to small. The edge before the score has the great probability of a genuine link.

In order to test the accuracy of similarity index prediction, we divide the edge of $E$ into two parts: a training set $E^T$ and test set $E^P$. Obviously, $E = E^T \cup E^P$, $E^T \cap E^P = \Phi$. Here, we call the edge that belongs to the $U$ but does not belong to the edge of the $E$ as the not existed edge $J$ and that the edge which belongs to the $U$ but does not belong to the edge of the $E^T$, as the unknown edge $H$.

There are some main indexes for measuring the accuracy of the algorithm: AUC, Precision and Ranking Score. Of course, Different indicators to measure the emphasis are different. AUC is the most common index which measure the accuracy of the algorithm on the whole[8]. Precision only considers whether the side of the front L bit is accurate[9]. But Ranking Score gives more consideration to the ranking of the predicted edge [10]. In this paper we choose AUC as the index which AUC is that the probability of a random selection of a boundary value in $E$ is higher than that of a randomly selected not existed edge [11]. We randomly select one side of the $E^P$, and then randomly select one side of the $J$. If the score of $E^P$ is bigger than that of $J$, the score plus 1. If they are equal, we plus 0.5, in otherwise, we plus 0. We do this independent comparison $n$ times. In the $m$ times is $E^P$ bigger than $J$ and $z$ times is $E^P$ equal to $J$, AUC can be defined as $AUC = (m + 0.5z)/n$.

Assume that all scores are generated randomly, so the AUC equal to about 0.5, while AUC is greater than or equal to 0.5, the surpassing measure the degree of precise that algorithm to the method of random selection. In order to illustrate the meaning of AUC, assuming the energy supply chain network was constructed of five node enterprises, the structure diagram is shown in Fig.1.
Fig.1(a) is a completely supply chain network structure, solid line shows the training set, the dotted line represents the test set. This network have five node, the network’s existence of the amount of connected-edges in theoretically is \(5 \times (5-1)/2 = 10\), 7 of them are known, the last are not exist. To test the accuracy of the algorithm of link-prediction, it's need to choose some of edges in seven known edges to formed \(E^p\), supposing choose edges like \(\{2, 4\}\) and \(\{2, 5\}\) as testing edges, which shown in Fig.1(b), then the last five known edges forms \(E^t\), under the method of link prediction, just using the information of edges in \(E^t\), the edges of \(E^p\) in the above are choose from known edges randomly, if there is an another chance to test the accuracy, the testing edges are not necessarily \(\{2, 4\}\) and \(\{2, 5\}\), but the any edges in \(E^t\) are chosen as testing edges, the edges of \(E^t\) are undetermined temporarily, according to the link prediction, the scores of last five unknown edges are \(s_{25} = 4, \ s_{24} = 4.5, \ s_{14} = 3.8, \ s_{34} = 4.6, \ s_{35} = 4\), thus, before calculate the value of AUC we need to sort the unknown edges by its score, \(s_{14} < s_{25} = s_{35} < s_{24} < s_{34}\), then according to the calculating formula of AUC, the score of AUC is 0.583.

This could be explained that when choosing the edges of with the change of amount of edges, the value of AUC will change accordingly, even if \(E^p\) are all the same, the value of AUC will difference as well. Because the \(n\) in the calculating formula of AUC is sample time, sampling method have many types such as random sampling, traverse by each item, snowball sampling etc, \(n\) is equal to six in the energy supply chain constructed by five nodes, which mean that test edges and non-existent edges are compared for six times. But while using computer program to calculate the value of AUC in the paper, the compared edges which meaning compared between prediction edges and non-existent edges are chosen randomly, thus, even the information of \(E^t\) are all the same, the difference of sampling times and sampling comparison object will lead to the change of AUC, sampling times \(n\) is the bigger the better in case of the value of AUC will close to the reality.

2.2. Experimental Problem Description

Using computer analysis, we use the neighbor matrix to describe the network. The neighbor matrix is \(A\), which is a 132x132 matrix. If there is a cooperation between the node \(v_i\) and \(v_j\), the element on the \(A\) line \(j\) column of the \(i\) is 1, otherwise 0. There are 431 edges in the energy supply chain network and there is no self loop (no separate points). But the number of the edges in the network should be \(132 \times (132-1)/2 = 8646\). At present, the number of links is known to be 431, and the corresponding link number is 8646-431=8215. The proportion of selected test set is 10% and the training set is 90%. The test set and the training set are divided into one time, and the connectivity of the training set should be guaranteed at the same time. The number of test side of the test set is \(431 \times 10\% \approx 44\). The most accurate method is to compare all the test sites and not existed side one by one, and the number of times is \(44 \times 8215 = 361460\) (Times). After the network neighbor matrix is constructed, the data set should be partitioned. There are many kinds of data partitioning (random sampling, random sampling, each traversal, K- fold cross etc.) Because of the complexity of the computation, the computer program code is designed to be a random sample. In order to make the AUC as close as possible to the true value, this paper takes the sampling number of 100000 (Times), and carries on the 200 (Group) independent experiment, and the final AUC value is equal to 10000 (the second) sampling 200 (Group) the average value of the independent experiment.
2.3. Link Prediction Based on Similarity Index

Link prediction based on similarity index can be divided into the local information, the path and the random walk. But because there is no comparability between them, this paper re-divided the similarity index: Based on the local information of nodes, the local semi information, the global information. Through the introduction of several similarity indexes, this paper predicts the link of the supply chain network in reality, and compares the accuracy of different index prediction. Node local information index is: CN index, Salton index, Jaccard index, Sorensen index, HPI index, HDI index, LHN-I index, AA index, RA index. The local semi information index has: LP index, LRW index, SRW index. The global information index is: Katz index, LHN-II index, ACT index, RWR index, LHN-I index, TS index.

Based on the above 17 kinds of network structure similarity index, we use the following MATLAB program segment to calculate the evaluation index AUC. In order to form a vertical and horizontal comparison, we will calculate the above 7 kinds of network and ESC network prediction accuracy. The AUC average calculated results of the 8 independent experiments of 200 different real networks are shown in Table 1 where values in brackets represent the size of the adjustable parameters.

Table 1. Accuracy of similarity index in 8 real networks.

| Index | USAir | PB | Football | FFWF | Power | Jazz | Karate | ESC |
|-------|-------|----|----------|------|-------|------|--------|-----|
| CN    | 0.9542 | 0.9234 | 0.8497 | 0.6053 | 0.6250 | 0.9570 | 0.6901 | 0.6581 |
| Salton| 0.9258 | 0.8783 | 0.8601 | 0.5263 | 0.6249 | 0.9664 | 0.6239 | 0.6097 |
| Jaccard| 0.9151 | 0.8768 | 0.8630 | 0.5263 | 0.6250 | 0.9619 | 0.6158 | 0.5910 |
| Sorensen| 0.9150 | 0.8768 | 0.8609 | 0.5263 | 0.6246 | 0.9622 | 0.6150 | 0.5900 |
| HPI   | 0.8820 | 0.8548 | 0.8570 | 0.5236 | 0.6250 | 0.9485 | 0.7050 | 0.6485 |
| HDI   | 0.9084 | 0.8729 | 0.8611 | 0.5251 | 0.6250 | 0.9528 | 0.5929 | 0.5870 |
| LHN-I | 0.7769 | 0.7628 | 0.8635 | 0.4000 | 0.6248 | 0.9028 | 0.5912 | 0.5744 |
| AA    | 0.9659 | 0.9268 | 0.8479 | 0.6065 | 0.6249 | 0.9631 | 0.7337 | 0.6610 |
| RA    | 0.9722 | 0.9279 | 0.8506 | 0.6083 | 0.6248 | 0.9723 | 0.7404 | 0.6576 |
| LRW (0.001) | 0.9522 | 0.9362 | 0.8624 | 0.6219 | 0.6985 | 0.9527 | 0.7500 | 0.6883 |
| LRW (3) | 0.9730 | 0.9475 | 0.8797 | 0.8991 | 0.6983 | 0.9737 | 0.7648 | 0.7145 |
| LRW (4) | 0.9722 | 0.9406 | 0.8780 | 0.6801 | 0.7568 | 0.9699 | 0.7601 | 0.7122 |
| LRW (5) | 0.9707 | 0.9446 | 0.8773 | 0.7914 | 0.8007 | 0.9645 | 0.7589 | 0.7378 |
| SRW (3) | 0.9737 | 0.9355 | 0.8744 | 0.7060 | 0.6983 | 0.9743 | 0.7599 | 0.6985 |
| SRW (4) | 0.9742 | 0.9384 | 0.8788 | 0.7006 | 0.7569 | 0.9788 | 0.7568 | 0.7095 |
| SRW (5) | 0.9742 | 0.9407 | 0.8744 | 0.7223 | 0.8007 | 0.9801 | 0.7528 | 0.7206 |
| Katz (0.01) | 0.9502 | 0.9329 | 0.8652 | 0.6766 | 0.9636 | 0.9438 | 0.7481 | 0.6862 |
| Katz (0.001) | 0.9519 | 0.9359 | 0.8641 | 0.6215 | 0.9636 | 0.9532 | 0.7559 | 0.6948 |
| LHN-II(0.9) | 0.6104 | 0.6380 | 0.8758 | 0.5359 | 0.9631 | 0.6716 | 0.4237 | 0.4626 |
| LHN-II(0.95) | 0.5889 | 0.5824 | 0.8525 | 0.5407 | 0.9628 | 0.6272 | 0.3776 | 0.4699 |
| LHN-II(0.99) | 0.5678 | 0.5292 | 0.6266 | 0.5423 | 0.9625 | 0.5549 | 0.3144 | 0.4667 |
| ACT | 0.9012 | 0.8925 | 0.5856 | 0.7235 | 0.8924 | 0.7957 | 0.6579 | 0.6449 |
| RWR (0.85) | 0.9682 | 0.9421 | 0.8743 | 0.7440 | 0.9755 | 0.9734 | 0.7145 | 0.7449 |
| RWR (0.95) | 0.9529 | 0.9291 | 0.8414 | 0.7386 | 0.9777 | 0.9612 | 0.7012 | 0.7540 |
| TSCN (0.01) | 0.6042 | 0.4923 | 0.5737 | 0.4994 | 0.9522 | 0.5077 | 0.6988 | 0.3074 |
| TSRWR (0.01) | 0.9611 | 0.9101 | 0.8717 | 0.5463 | 0.9680 | 0.9672 | 0.7836 | 0.6696 |

The following conclusions can be drawn from the data in the table, according to the supply chain network (ESC), and the prediction accuracy of the other 7 real networks.

(a) For ESC network, based on the local semi information index effect is generally better than the local information index and the global information index, the overall index performance is better than the local information index. In addition to the Power network, the other 6 network local semi information indicators perform as well as the ESC network, and the Power overall performance was
the best. Because the average path in the Power network is larger, and its average degree is smaller, so from the whole it will have better prediction effect.

(b) In the local node information index AA and RA as well as CN index performance is more prominent, and the ESC network AA is the most eye-catching, which shows that the AA indicators to weaken the degree node has played a certain role. The precision difference between the 8 network prediction in the local information index is not very large, and it is relatively stable when the network is predicted by the local node information index. The local semi information index is the most stable and the least volatile. And the prediction accuracy of global information is obviously different. For example, the smallest of global indicators in PB network is TSCN=0.4923, and the largest is RWR=0.9421, so there is a large difference between them, and the performance of the ESC network is also more obvious.

(c) In the global information index, the prediction accuracy of some network is obviously less than 0.5. For example, the LHN-II index in the Karate network and the LHN-II index in the ESC network, which means that some of the global information index in the network’s predictive effect are lower than the accuracy of random connection. Global information index in the Katz index, RWR index forecast effect is still more eye-catching. And TSCN indicators for most of the network prediction results are not good, but the Power prediction results are the best, because Power network density is low.

(d) The prediction effect of the local node information index in ESC, Power, FWFW is just so(between 0.6-0.7). But in the Jazz, USAir it has a good performance (in the 0.9-1), because the statistical characteristics of ESC, Power, FWFW clustering coefficient are smaller, but USAir, Jazz, the clustering coefficient is bigger. Therefore, the local information of nodes in the network such as ESC, which has a smaller clustering coefficient, is less effective, and the results of [12] are verified.

(e) In the ESC network, it is clear that the global information index is not stable and the effect is not good, but the RWR index is the highest, even the accuracy is greater than the performance of the stability of the local semi information. It is worth mentioning that, it does not mean the RWR index forecast in other supply chain network effect is the best. This paper is the application of link prediction method to research the energy supply chain network node enterprises cooperation connection, but the calculation of the various indicators of accuracy cannot be generalized for all supply chain network prediction.

3. Conclusion

Researching on the supply chain network composed of the cooperation relationship between node enterprises and enterprises, this paper analyses statistical characteristics, centrality and cooperation of supply chain network and empirical analysis based on an energy supply chain (ESC) network and previous researches. We provide a new perspective for the research of supply chain management. Results show that: 1) Based on the local semi information index, the predictive effect is better than that of the local information and global information index, and the performance of the global index is better than that of the local information index. 2) In the overall index, the prediction accuracy is high, but the overall performance is not stable. 3) The local information of nodes is less effective in the network with a smaller coefficient of clustering. 4) For the local information of nodes, the prediction accuracy is small, on the contrary, the global information index has great volatility.

Under the market competition, cooperation and complex dynamic environment, each enterprise in supply chain network has their own management strategies, it’s important to study the partner rule and dynamic formation and change process of supply chain network for revealing the whole macro nature of the supply chain network and analysis of the stability and anti-risk ability of supply chain network. This research is helpful to reduce the operation cost of supply chain network and to improve the network efficiency and competitiveness of a single node enterprise, and also provides a new method supply chain for macro management. It can create a good environment for enterprise cooperation and improve the competitive behavior of enterprises which has a certain guiding significance and practical application value. It has an important theoretical significance for
enriching the management of supply chain. It also has a strong practical significance to improve the competitiveness of enterprises in the supply chain.

Reference

[1] Choi T Y, Dooley K J, Rungtusanatham M. (2001), “Supply networks and complex adaptive systems: control versus emergence”, Journal of Operations Management, Vol. 19, No. 3, pp. 351-366.

[2] Venkata V, Katare S, Patkar P R, et al. (2004), “Spontaneous emergence of complex optimal networks through evolutionary adaptation”, Computers and Chemical Engineering, Vol. 28, No. 9, pp. 1789-1798.

[3] Pathak S D, Dilts D M, Biswas G. (2007), “On the evolutionary dynamics of supply network topologies”, Engineering Management, Vol. 54, No. 4, pp. 662-672.

[4] Kühnert C, Helbing D, West G B. (2006), “Scaling laws in urban supply networks”, Physica A: Statistical Mechanics and its Applications, Vol. 363, No. 1, pp. 96-103.

[5] Zhang T, Sun L Y, Sun H H, Li G. (2003), “Analysis and modeling of operation mode of supply chain—research based on CAS paradigm”, System Engineering Theory and Practice, Vol. 11, No. 11, pp. 8-13.

[6] Kolaczyk E E. (2007), “Some implications of path-based sampling on the Internet. Proceedings of a workshop on statistics of networks”, Washington: National Academies Press, pp. 207-226.

[7] Liu H K, Lü L Y, Zhou T. (2011), “Uncovering the network evolution mechanism by link prediction. Scientia Sinica Phys”, Mech & Astron, Vol. 41, No. 7, pp. 816-823.

[8] Hanley J A, McNeil B J. (1982), “The meaning and use of the area under a receiver operating characteristic (ROC) curve”, Radiology, Vol. 143, No. 1, pp. 29-36.

[9] Herlocker J L, Konstann J A, Terveen K et al. (2004), “Evaluating collaborative filtering recommender systems”, ACM Transactions on Information Systems, Vol. 22, No. 1, pp. 35-53.

[10] Zhou T, Ren J, Medo M, et al. (2007), “Bipartite network projection and personal recommendation”, Physical Review E, Vol. 76, No. 4, pp. 60-115.

[11] Fawcett T. (2006), “An introduction to ROC analysis”, Pattern Recognition Letters, Vol. 27, No. 8, pp. 861-874.

[12] Zhou T, Lü L, Zhang Y C. (2008), “Predicting missing links via local information”, The European Physical Journal B, Vol. 71, No. 4, pp. 623-630.