Topology Reconstruction for Target Operation Network (TON): A Link Prediction Perspective

Changjun Fan, Baoxin Xiu *, Zhong Liu, Chao Chen, Yao Yang
Science and Technology on Information Systems Engineering Laboratory
National University of Defense Technology
Changsha, Hunan, P. R. China, 410073

***Abstract*—Information is crucial in the military area, how to purify the obtained information and improve its quality is of great significance. This paper aims to shed some lights from the perspective of target operation network topology reconstruction. First, formalized it as a TON structure reconstruction problem; then proposed a link reliability index, named LR, based on which to identify the spurious edges and predict the missing edges; a greedy algorithm based on LR was designed to reconstruct the observed network; finally, experiments on the public data validated the effectiveness of the reconstruction algorithm, the reconstructed network is closer to the true network in nine topological indexes, including clustering coefficient, assortativity, congestion, synchronization, propagation threshold, network efficiency, betweenness mean standard deviation and natural connectivity, compared to the observed network, which means our reconstruction algorithm could narrow the structural gap between the observed network and the true one.**

I. Introduction

In real battle field, however advanced and excellent the reconnaissance means and information collecting works are, it's still rather difficult to guarantee the information to be 100% accurate due to some objective or subjective reasons. In sensitive area, like military battlefield, a small error might cause a huge catastrophe to the outcome of a battle, therefore, a scientific method to refine the information from the battle and to make it a further step closer to the truth before putting them into real uses will be of significant importance for commanders to hold the battle situation, make operational plans, and make decisions. Analysis of the battle field information processing from the perspective of target operation network could be described as the problem of TON structure reconstruction, more specificially, abstract a observed target operation network from the information collected, refine the network by identifying the false information and predicting the missing information and thus obtain reconstructed network closer to the truth than the observed one, so as to lay the foundation for the following central analysis based on network topology. This paper proposed a link prediction based reconstruction model for target operation network based on FINC operation network model, to meet the practical demand of real situations.

FINC is the abbreviation of Force, Intelligence, Networking, C2(command and control)[1][2][3][4][5]. In the process of modeling, triangular nodes represent intelligence units, such as radar, and rectangular nodes stand for operational units, such as missile positions, and the circular nodes indicate command and control units, such as command post. Meanwhile, the edge of the network is used to represent information transmission links, which can be both direct and indirect, edge weight represents the information transmission time delay. Thus we may agree that military system or military network model based on FINC model is able to analyze indexes such as system information delay, reconnaissance performance and synergy ability. The existing FINC model is a military network model based on social network theory, nodes in which could have different properties and structures, meanwhile it's also a well-behaved method to analyze military network's command and communication capability.

Link prediction refers to a prediction of existing probability of nonexist edges in the network[6][7][8]. In the area of computer science, many studies have been conducted on link prediction[9][10]. Exterior information, logically speaking, may obtain higher accuracies, yet the access to those information is difficult to get. Recently link prediction based on network structure has been a heated topic[6][7][8]. It's easier and more reliable to obtain the network structure information than the exterior information, meanwhile network structure based methods are more generative. Current methods can be divided into two main parts according to the prediction theory: similarity based methods and likelihood based methods. Similarity based methods also divide into local information based methods, path based methods and random walk based methods, likelihood based methods divide into hierarchical structure models and stochastic block models. All in all, the effect of link prediction based reconstruction methods is mainly determined by link prediction index chosen. Current link prediction indexes are mainly based on the network topology, for two main reasons, one is the easy access to the network topology information, the other is the satisfactory prediction accuracies of them. Of course there is little evidence shows the combination of exterior information could lead to worse prediction. Provided that the obtained information is correct enough, it's true that the more the available information is, the closer it will get to the truth.

The remaining parts of this paper is arranged as follows: Section 2 formalize the problem this paper aims to address; Section 3 model the link reliability, and proposed the LR index; Section 4 put forward a local greedy search algorithm for TON structure reconstruction; Section 5 conducts experiment

*Correspondence to: baoxinxiu@163.com

on the public dataset, and validates the algorithm proposed before. The last section conclude the paper.

II. Problem Formulation

A target operation network $G(V, E, A, L)$, where V stands for a set of nodes, E indicates edges, A is a set of node types, including C2 nodes, fire nodes and information nodes;L stands edge types, including reporting relationship, command relationship, and communication relationship; The problem of TON reconstruction task is to find out a strategy $\{E_{mising}, E_{spurious}\}$, and obtain $maxR(G + \{E_{mising}\} - \{E_{spurious}\})$, here $\{E_{mising}\}$,$\{E_{spurious}\}$ represent missing edges and spurious edges respectively, $R()$ is the calculation of TON reliability.

How to identify the spurious edges and predict the missing edges, and how to determine each number to obtain the optimal TON reliability, is the core problems we are to address in this paper. Inspired by link prediction research, this paper proposed link prediction based method for TON structure reconstruction. Three definitions are given first:

Definition 1. ***(Link Prediction):*** *Given network $G(V, E)$, based on the obtained network structure, to predict the occurrence possibility $f(v_i, v_j)$ for those nonexist edges.*

Definition 2. ***(Link Reliability):*** *It refers to the possibility of whether the link really exists. The higher the reliability, the more possible that there exists certain kind of links between two operation units.*

Definition 3. ***(TON Reliability):*** *It refers to the degree of structure similarity between the observed TON and the real one.*

III. Link Reliability Model for TON

As mentioned before, the observed network is a noisy incomplete network, not all the reliability of the observed edges are 1, nor the reliability of non-observed edges equal to 0. In fact, a credibility evaluation for all the edges of the observed network is needed, so as to provide basis for the subsequent spurious edges recognition and missing edges prediction.

While under the circumstance of unknown real network and other information, it remains a great challenge to calculate each edge's link reliability based on the observed network only. However, for target operation network, the analysis of its reconstruction should be well combined with the specific features.

Three hypothesis are put forward before modeling the link reliability of target operation network:

Hypothesis I: The link connection probability between between nodes is correlated with nodes type.

There are three types of nodes in TON: fire nodes, command and control nodes and information nodes. The nodes type do generate impacts on linking behaviors, for instance, there are no possible links between fire nodes, which is a reason for choosing the stochastic block model, for its basic hypothesis is whether nodes are connected in a network is determined by the cluster they belong to.

Hypothesis II: The link reliability of the exist edges may not be 1.

For the observed network contains some noise, i.e., spurious edges, link reliability of these edges are not 1.

Hypothesis III: The link connection probability between nodes is correlated with their indirect connection stregth.

Generally speaking, if there are more reachable paths between two nodes, which means their indirect connection strength is strong, it is more likely to generate direct connections between these two nodes.

As a resul, this paper takes two factors into modeling, which are nodes types, represented as T, and nodes indirect connection strength, represented as S; let R stands for the link reliability, then

$$R = f(T, S) \tag{1}$$

Now the quantitative effects of these two factors on link reliability evaluation are analyzed as follows:

A. Influence of the node type

This section quantify the nodes types' influence on link reliability based on stochastic block model.

Theorem 1: For an target operation network $G(V, E, A, L)$, let σ_x denotes the type number of node v_x , and r_{σ_x,σ_y} stands for all of the possible connections between these two types of nodes. $l^o_{\sigma_x,\sigma_y}$ represents the observed edge numbers between σ_x type nodes and σ_y type nodes. Then link reliability for $\{v_x, v_y\}$ is :

$$p(A_{xy} = 1|A^o) = \frac{l^o_{\sigma_x,\sigma_y} + 1}{r_{\sigma_x,\sigma_y} + 2} \tag{2}$$

Proof of Theorem 1 sees in Appendix.

B. Influence of the network topology

The definition of indirect connection strength between two nodes is defined as the number of reachable paths between them. Weight of the path is determined by its length, the shorter the path is, the bigger the weight is.

Definition 4. ***(Indirect Connection Strength):****Indirect Connection Strength S_{xy} between node v_x and node v_y refers to the weighted sums of the reachable paths with different lengths between these two nodes.*

According to Definition 4, the mathematical expression of S_{xy} is given as follows:

$$s_{xy} = \sum_{l=1}^{\infty} \alpha^l |paths^{<l>}_{x,y}| = \alpha A_{xy} + \alpha^2 (A^2)_{xy} + \alpha^3 (A^3)_{xy} + \cdots \tag{3}$$

where $\alpha > 0$ is an adjustable parameter controlling the path weight, and $|paths^{<l>}_{x,y}|$ is the l-length path numbers. When

α is less than the reciprocal of the maximum eigenvalue of adjacency matrix A, S_{xy} can also be expressed as:

$$S = (I - \alpha A)^{-1} - I \tag{4}$$

C. Link Reliability Index

Based on the following three points, we get the following conclusions by using the simplest form of product:

$$R = T * S \tag{5}$$

Firstly, the influences of two factors are all positive, which means, the higher value T or S is, the higher link reliability value is;

Secondly, if any value of these two factors is 0, the link reliability value is 0.

Thirdly, the value of link reliability could be adjusted through normalization.

Therefore, with formula (2) and (4), the value of link reliability for v_x, v_y could be calculated as:

$$r_{xy} = g_{xy} \times s_{xy} = \frac{l^o_{\sigma_x \sigma_y} + 1}{r_{\sigma_x \sigma_y} + 2} \times [(I - \alpha A)^{-1} - I]_{xy} \tag{6}$$

And we call the link reliability index LR.

IV. TON Reconstruction Algorithm based on Link Reliability Index

Section 3 put forward the Link Reliability index for TON, the following definitions of the missing edge and the spurious edge are given based on this index:

Definition 5. ***(Missing Edge):****The missing edges are those nonexist edges with higher LR values, referring to the actual missing edges due to incomplete information.*

Definition 6. ***(Spurious Edge):****The spurious edges are those exist edges with lower LR values, referring to the actual nonexist edges due to noisy information.*

Rank all the edges in TON (including exist edges and nonexist edges) in the descending order, and identify missing edges and spurious edges according to Definition IV and IV, the biggest problem here is to determine the number: the missing edge number and the spurious edge number. This section aims to address the problem.

Definition II explains the physical meaning of TON Reliability, here give its calculation:

$$Reliability_{A^o} = \prod_{A^o_{ij}=1, i \leq j} r_{ij} \tag{7}$$

Formula (7) is the product of reliability of all the observable links, i.e., the similarity degree of the topology of an observed target operation network to the topology of a real network is determined by the reliability of all the observable links; if all the observable links are more credible, the topology of the observed network is closer to the real one.

This is a typical combinational optimization problem: solution space is all possible combinations of missing edges number and spurious edges number, objective function is TON reliability. Traditional enumeration algorithm is of $O(|E|^2)$ time complexity, which is computational prohibitive for networks with large amount of edges. This section put forward a local greedy search algorithm (1),which is based on the hypothesis that the number of missing edges equal to the spurious one.

Algorithm 1 TON Reconstruction Algorithm Based on Equal Numbers of Missing Edges and Spurious Edges

1: **Step 1:** Divide the edges in the observed network A^o into two categaries, exist edges E_{exist} and non-exist edges $E_{nonexist}$, calculate each LR values for all edges,and rank the edges of E_{exist} in an ascending order(thus edges ranked higher indicate higher probability of the spurious edge),rank the edges of $E_{nonexist}$ in a descending order(thus the edges ranked higher indicate higher probability of the missing edge);
2: **step 2:** Get each edge from E_exist and $E_nonexist$ one by one, e_i^{exist} and $e_j^{nonexist}$, reliability values are r_i^{exist} and $r_j^{nonexist}$ respectively, calculate $\delta = r_j^{nonexist}/r_i^{exist}$;
3: **step 3:** If $\delta > 1$, means the reconstruction strategy (remove the spurious edges e_i^{exist} and add the missing edges $e_j^{nonexist}$) increases the reliability of TON, thus accept it; drop it if not, and repeat Step 2;
4: **step 4:** Repeat the previous three steps until the continuous drop number reaches 5, stop the algorithm.

As a local greedy search algorithm, Algorithm 1 just output a near-optimal solution, and it is based on the assumption that the number of missing edges equals to the spurious one, which is always not accurate in reality. As a result, a more reliable reconstruction algorithm deserves a further research.

An incomplete target operation network with noise can be refined after Algorithm 1, its topological structure is reconstructed to be closer to the real one, which is beneficial to the subsequent analysis based on TON topology.

V. Experiments

This section validates the effectiveness of algorithm 1.

A. Data Description

A public reported data is chosen as the experimental data. It contains 89 entities, including 12 C2 units, 26 fire units and 51 information units, and 150 observable links,including 16 F-I(Fire-Information) links, 26 F-C(Fire-C2) links, 51 I-I(Information-Information) links, 30 I-C(Information-C2) links and 17 C-C(C2-C2) links, the relationship of these entities are shown as Figure 1.

Among them, blue nodes represent C2 units, such as control center and command post; red nodes stands for fire units, such as various missile positions; green nodes represents information units, such as various optical fiber stations and radar stations. Edges with different colors stand for different types of links, purple edges stand for optical fiber communication links, red edges represent command and control links and blue edges stand for reporting links.

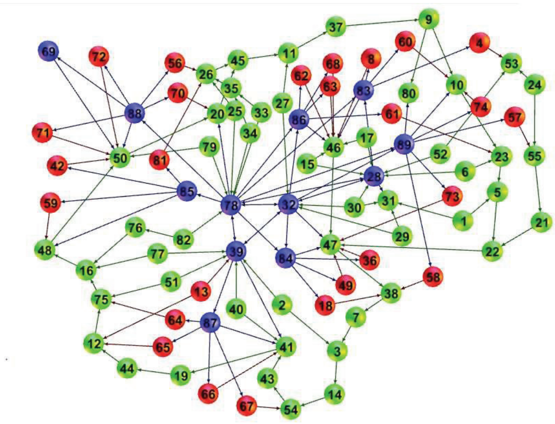

Fig. 1. TON Topology of the Experimental Data

B. Evaluation Indexes

Before evaluating our method's ability in reconstructing TON structure, there is a basic hypothesis here that the experimental data is the real battle field situation, this paper made it as the true network, and assumed the observed network by randomly adding or removing some edges. Evaluation indexes are put forward to quantify the algorithm's ability in narrowing the structure gap between the reconstructed network (from observed network) an the true one.

1) Evaluation Index for Link Reliability: For an observed network A^O, calculate each edge's LR value (including both exist edges and nonexist edges). When evaluating algorithm's ability in predicting missing edges, divide all the nonexist edges ($A^O_{ij} = 0$) into two types, one is the missing edge N_{01} ($A^O_{ij} = 0, A^T_{ij} = 1$); the other is the true nonexist edge $N^0 0$ ($A^O_{ij} = 0, A^T_{ij} = 0$). Now randomly select one edge in N_{01} and in N_{00}, if the edge from N_{01} has a higher LR value than the one from N_{00}, add one score; if two values are equal, add 0.5 score. After comparing $||N^{01}|| \times ||N^{00}||$ times, if one score exists n' times, and 0.5 score exists n'' times, we got $n' + 0.5n''$ scores, the algorithm's ability in predicting missing edges is calculated as:

$$AUC\_Missing = \frac{n' + 0.5n''}{||N^{01}|| \times ||N^{00}||} \quad (8)$$

Similarly, the algorithm's ability in identifying spurious edges could be calculated as:

$$AUC\_Spurious = \frac{n' + 0.5n''}{||N^{11}|| \times ||N^{10}||} \quad (9)$$

where N^{11} ($A^O_{ij} = 1, A^T_{ij} = 1$) represent true exist edges; N^{10} ($A^O_{ij} = 1, A^T_{ij} = 0$) represent spurious edges.

Obviously, if all scores are generated randomly, $AUC \approx 0.5$. Therefore, AUC value measures the extent our LR index is superior to other link prediction index.As a matter of fact, AUC is equivalent to Mann-Whitney U statistical test and Wilcoxon rank-sum statistical test[11] in its form.

2) Evaluation Index for TON Reconstruction: How to measure the structural gap between the reconstructed network A^R and the true network A_T, reference [12] utilized five network structural indexes, clustering coefficient, assortativity, cogestion, synchronization and propagation threshold respectively. Here we adopted four more topological indexes: network efficiency, average betweenness, betweenness mean square deviation and natural connectivity. Following are the further explanations of them:

Network efficiency[13]: $L = \frac{1}{N(N-1)\sum_{i \geq j} \frac{1}{d_{ij}}}$, defined as average reciprocal of the distance between any two nodes, reflecting the network connectivity;

Mean betweeness[14]: $\overline{B} = \frac{1}{N}\sum_{i=1}^{N} b_i$, defined as the average betweenness for all nodes, reflecting the network congestion;

Betweenness mean square deviation[15]: $\sigma = \sqrt{\frac{1}{N}\sum_i (b_i - \overline{B})^2}$, defined as the mean square deviation of the betweeness for all nodes;

Clustering coefficient[14]: $C = \frac{1}{N}\sum_{i=1}^{N} \frac{E_i}{1/2k_i(k_i-1)}$, defined as the average connecting ratio of neighboring nodes in the network, reflecting the clustering degree of nodes in the network;

Assortativity[16]][17]: $r = \frac{W^{-1}\sum_k u_k v_k - [W^{-1}\sum_k a/2(u_k v_k)]^2}{W^{-1}\sum_k \frac{1}{2}(u_k^2 + v_k^2) - [W^{-1}\sum_k 1/2(u_k + v_k)]^2}$, where u_k, v_k are the two nodes' degrees of the edges k, W the total edge number. Value of r ranges between [-1,1], when $r > 0$, the network is assortative; when $r < 0$, the network is disassortative; and when $r = 0$, the network is not related. $|r|$ reflects the assortativity degree of networks.

Natural Connectivity[18]: $\overline{\lambda} = ln(\frac{1}{N}\sum_{i=1}^{N} e^{\lambda_i})$, and λ_i is the eigenvalue of adjacent matrix $A(G)$ in G , reflecting the network vulnerability.

Congestion[19][20]: $Congestion = max(b_i)$, defined as the biggest betweenness of of nodes in the network, reflecting the biggest congestion in the network;

Synchronization[21][22]: $Synchro = max(\lambda_i)/min(\lambda_i)$, defined as the ratio of the biggest eigenvalue and the smallest negative eigenvalue of the network Laplacian matrix, reflecting the network's synchronization ability.

Propagation threshold[23]: $Spreading = < k > / < k^2 >$, k is the network degree distribution defined as the ratio of first-order moment and the second-order moment, reflecting the propagation threshold of the network.

After calculating the above indexes value for the real network, the observed network and the reconstructed network, the following formulas are used to measure these networks' structural gaps:

Relative error of the observed network and the real network is defined as:

$$RE^O = (X(A^O) - X(A^T))/X(A^T) \quad (10)$$

Relative error of the reconstructed network and teh real network is defined as:

$$RE^R = (X(A^R) - X(A^T))/X(A^T) \tag{11}$$

$X(A)$ stands for one index value from the above nine indexes.

As a result, comparison of two values RE^O and RE^R could be used to see whether the reconstructed network is closer to the true network than the observed network.

C. Results Analysis

1) Evaluation Results for LR: We choose Katz, the index performing best among structural similarity based link prediction indexes, and SBM, the index performing best among likelihood based link prediction indexes, as comparisons with the index LR, calculate the values of $AUC\_Missing$ and $AUC\_Spurious$ for each index, to measure their ability in predicting missing edges and identifying spurious edges.

First we evaluate their ability in predicting network's missing links, adjust the ratio of edges randomly removed (missing edges), and calculate $AUC\_Missing$ value each time, results shown as Figure 2 (each point in the figure is the average value for 100 times computation).

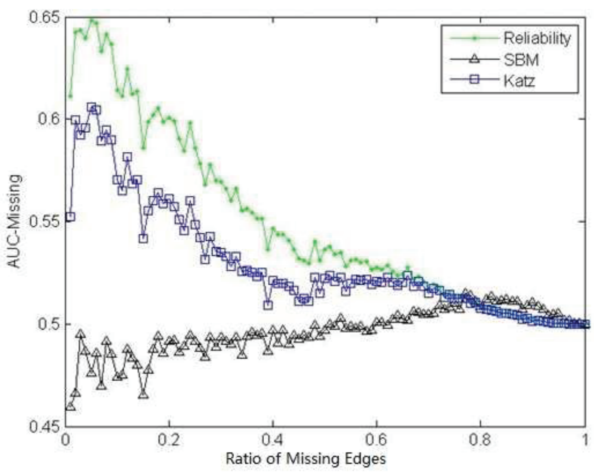

Fig. 2. Comparisons between Reliability and other indexes in predicting missing edges.

It can be seen from Figure 2 that LR index performs better than other two indexes as the ratio of missing edges is less than 0.77; as the ratio is over 0.77, LR performs a little worse than SBM, and the same as Katz. Overall, LR performs no worse than other two indexes for 97.03% times.

Similarly, when comparing ability in identifying spurious edges, adjust the ratio of edges randomly added (spurious edges), and calculate $AUC\_Spurious$ value each time, results shown as Figure 3 (each point in the figure is the average value for 100 times computation).

It can be seen from Figure 3 that LR index always performs better than other two indexes,

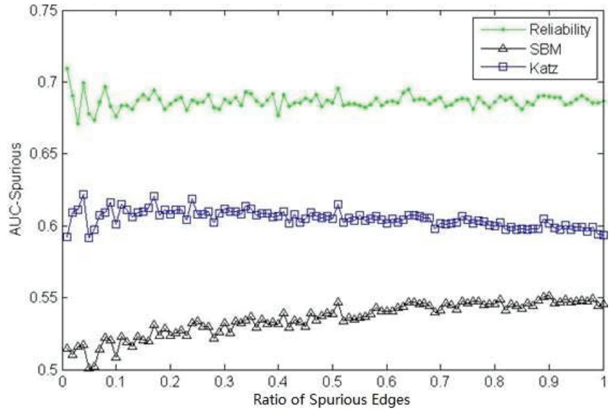

Fig. 3. Comparisons between Reliability and other indexes in detecting false edges.

Combined with Figure 2 and Figure 3, LR index is a better link prediction index in purifying the network noise, i.e., identify spurious edges and predict missing edges.

2) Evaluation Results for TON Reconstruction: Here we validate the effectiveness of the reconstruction algorithm. First, adding some noise on the true network, and obtain the observed network A^O, assume the ratio of missing edges is α, the ratio of spurious edges is β, let $\alpha = \beta = p$, and call p as the observed error. After the reconstruction algorithm, the reconstructed network A^R is got, measure the closeness of A^O, A^R and A^T by calculating RE^O and RE^R. If $RE^R < RE^O$, means the A^R is closer than A^O to A^T, the reconstruction algorithm is validated then.

Make p changes from 0 to 1, and the step is set 0.05, calculate the nine topological indexes above, results are shown as Figure 4(each point in the figure is the average value for 100 times computation).

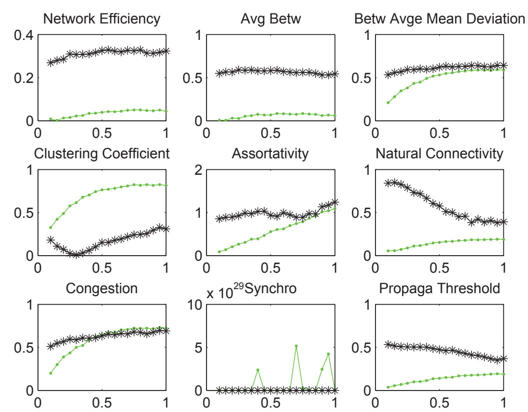

Fig. 4. Comparisons of RE^O and RE^R in nine topological indexes

In Figure 4, black '.' stands for RE^O value, and green '*' stands for RE^R value, it can be seen that except the congestion index and Clustering Coefficient index, RE^R is less than RE^O all the time. For congestion index, as p less than 0.5, RE^R is less than RE^O; for clustering coefficient index, A^R seems farther to A^T than A^O;for synchronization index, since there existing some abnormal points, which stretches the vertical coordinate axis in a large scale, and make it difficult to judge. All in all, comparisons on nine topological indexes show the reconstructed network is closer to the true one than the observed network.

VI. Conclusion

In this paper, we put forward a target operation network structure reconstruction algorithm and validate it in the public data, experimental results indicate the algorithm is able to effectively narrow the structure gap between the observed network and the real one to some certain extent.Firstly, a link reliability index, called LR, was proposed based on the hypothesis that the existing possibility between two nodes is mainly determined by their types and topological locations, and verified that LR could better identify spurious edges and predict missing edges, compared with SBM and Katz index; then network reliability index, based on LR, was put forward, we designed a locally greedy search algorithm based on the assumption that the spurious edge number equals to the missing edge number to reconstruct the TON structure. The experimental data consists 89 units and 155 links, let it be the true network, and assumed an observed network through adding some noise (randomly remove some missing edges and add some spurious edges). Nine indexes were utilized to measure the algorithm performance, experiments on the observed network showed that the reconstruction algorithm could effectively narrow its gap with the true one for each index, since there are both static structural index and dynamic structural index in these nine indexes, they comprehensively represent all the structural characteristics for a network, as a result, the reconstructed network is closer to the real one in topology.

Acknowledgment

This work is partially supported by National Natural Science Foundation of China (No.71471176, No.61303266) . All authors acknowledge the National University of Defense Technology.

References

[1] H. C. White, S. A. Boorman, and R. L. Breiger, "Social structure from multiple networks. i. blockmodels of roles and positions," *American journal of sociology*, pp. 730–780, 1976.

[2] C. J. Anderson, S. Wasserman, and K. Faust, "Building stochastic blockmodels," *Social networks*, vol. 14, no. 1, pp. 137–161, 1992.

[3] P. Doreian, V. Batagelj, and A. Ferligoj, *Generalized blockmodeling*, vol. 25. Cambridge university press, 2005.

[4] Y. Guo-li, H. Jin-cai, and Z. Wei-ming, "Combat system quantitative assessment method based on finc model," *Computer Engineering*, vol. 37, no. 10, 2011.

[5] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing, "Mixed membership stochastic blockmodels," in *Advances in Neural Information Processing Systems*, pp. 33–40, 2009.

[6] L. Lin-yuan, "Link prediction on complex networks," *Journal of University of Electronic Science and Technology of China*, vol. 39, no. 5, pp. 651–661, 2010.

[7] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.

[8] L. Lin-yuan and Z. Tao, *Link Prediction*. Higher Education Press, 2013.

[9] R. R. Sarukkai, "Link prediction and path analysis using markov chains," *Computer Networks*, vol. 33, no. 1, pp. 377–386, 2000.

[10] A. Popescul and L. H. Ungar, "Statistical relational learning for link prediction," in *IJCAI workshop on learning statistical models from relational data*, vol. 2003, Citeseer, 2003.

[11] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[12] R. Guimerà and M. Sales-Pardo, "Missing and spurious interactions and the reconstruction of complex networks," *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22073–22078, 2009.

[13] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-worldnetworks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[14] X. WANG, X. LI, and C. Guanrong, "Network science: An introduction," 2012.

[15] T. Nishikawa, A. E. Motter, Y.-C. Lai, and F. C. Hoppensteadt, "Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize?," *Physical review letters*, vol. 91, no. 1, p. 014101, 2003.

[16] M. E. Newman, "Assortative mixing in networks," *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.

[17] M. E. Newman, "Mixing patterns in networks," *Physical Review E*, vol. 67, no. 2, p. 026126, 2003.

[18] J. Wu, M. Barahona, Y.-J. Tan, and H.-Z. Deng, "Robustness of regular ring lattices based on natural connectivity," *International Journal of Systems Science*, vol. 42, no. 7, pp. 1085–1092, 2011.

[19] G. Yan, T. Zhou, B. Hu, Z.-Q. Fu, and B.-H. Wang, "Efficient routing on complex networks," *Physical Review E*, vol. 73, no. 4, p. 046108, 2006.

[20] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, "Optimal network topologies for local search with congestion," *Physical Review Letters*, vol. 89, no. 24, p. 248701, 2002.

[21] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, "Synchronization in complex networks," *Physics Reports*, vol. 469, no. 3, pp. 93–153, 2008.

[22] M. Barahona and L. M. Pecora, "Synchronization in small-world systems," *Physical review letters*, vol. 89, no. 5, p. 054101, 2002.

[23] R. Pastor-Satorras and A. Vespignani, "Epidemics and immunization in scale-free networks," *Handbook of graphs and networks: from the genome to the internet*, pp. 111–130, 2005.

APPENDIX
PROOF OF THEOREM 1

The stochastic model is introduced first before the proof.

Stochastic block is one of the most pervasive networks [1][2][3][5]. This model divides the nodes in the network into several groups, and whether the two nodes are linked is determined by the group they belong to. In other words, the roles of all nodes in the same group are identical. The stochastic block model is particularly suitable for situations when nodes' roles exert significant influences on their linking behaviors. A stochastic block model is composed by two parts: one is the plans that the network is divided into a number of groups; the other is the linking probability matrix between two nodes that come from two different groups. Let Ω represents all the grouping plans, given a specific plan $P \in \Omega$ and a specific linking probability matrix Q, a stochastic block model $M = (P, Q)$ is then determined. The current structure from an observable network can be regarded as from an unknown stochastic block model. Suppose the current network structure is A^o , psi represents a specific property, then the probability that this network with the property is:

$$p(\psi|A^o) = \int_{\Theta} p(\psi|M)p(M|A^o)dM \tag{12}$$

where Θ stands for the set of total stochastic block models, $M \in \Theta$ refers to a specific stochastic block model, and $p(\psi|M)$ represents the probability that the observed network generated by model M has property psi, and $p(M|A^o)dM$ stands for the probability that the observed network is actually generated by model M. According to Bayes Theory,we have:

$$P(M|A^o)p(A^o) = p(A^o|M)p(M)$$

since:

$$p(\psi|A^o) = \frac{\int_{\Theta} p(\psi|M)p(A^o|M)p(M)dM}{\int_{\Theta} p(A^o|M')p(M')dM'} \tag{13}$$

Next, formula(13) is utilized to prove the theorem.

Proof. Let the property ψ be $A_{xy} = 1$, probability $p(A_{xy} = 1|A^o)$ represents the link reliability of $\{v_x, v_y\}$, according to formula(13):

$$p(A_{xy} = 1|A^o)$$
$$= \frac{1}{Z}\sum_{P\in\Omega}\int_0^1 |Q|p(A_{xy} = 1|P,Q)p(A^o|A,Q)p(P,Q)dQ$$

where $|Q|$ stands for the numbers of matrix elements, which are equal to the square of the number of groups in the network, and

$$Z = \sum_{P\in\Omega} |Q|p(A^o|P,Q)p(P,Q)dQ \tag{14}$$

since $p(A_{xy} = 1|P,Q) = Q_{\sigma_x,\sigma_y}$, then

$$p(A^o|P,Q) = \prod_{\alpha\leq\beta} Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}} \tag{15}$$

Take formula (15) into formula (14),and get

$$Z = \sum_{P\in\Omega}\prod_{\alpha\leq\beta}\int_0^1 Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}}dQ_{\alpha\beta} \tag{16}$$

Let

$$H = \prod_{\alpha\leq\beta}\int_0^1 Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}}dQ_{\alpha\beta} \tag{17}$$

Next we prove:

$$H = exp\{-\sum_{\alpha\leq\beta}[ln(r_{\alpha\beta}+1) + ln\tbinom{r_{\alpha\beta}}{l^o_{\alpha\beta}}]\} \tag{18}$$

Proof. With Beta integral formula:

$\int_0^1 t^{a-1}(1-t)^{b-1}dt = \frac{(a-1)!(b-1)!}{(a+b-1)!}$

we have:

$\int_0^1 Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}}dQ_{\alpha\beta} = \frac{l^o_{\alpha\beta}!(r_{\alpha\beta}-l^o_{\alpha\beta})!}{(r_{\alpha\beta}+1)!}$

$= \frac{1}{r_{\alpha\beta}+1}\frac{l^o_{\alpha\beta}!(r_{\alpha\beta}-l^o_{\alpha\beta})!}{r_{\alpha\beta}!}$

and

$lnH = \sum_{\alpha\leq\beta} ln(\frac{1}{r_{\alpha\beta}+1}\frac{l^o_{\alpha\beta}!(r_{\alpha\beta}-l^o_{\alpha\beta})!}{r_{\alpha\beta}!})$

$= -\sum_{\alpha\leq\beta}[ln(r_{\alpha\beta}+1) + ln\tbinom{r_{\alpha\beta}}{l^o_{\alpha\beta}}]$

Thus

$H = exp\{-\sum_{\alpha\leq\beta}[ln(r_{\alpha\beta}+1) + ln\tbinom{r_{\alpha\beta}}{l^o_{\alpha\beta}}]$

□

According to formula (18),

$$Z = \sum_{P\in\Omega} H = \sum_{P\in\Omega} exp\{-\sum_{\alpha\leq\beta}[ln(r_{\alpha\beta}+1) + ln\tbinom{r_{\alpha\beta}}{l^o_{\alpha\beta}}] \tag{19}$$

and

$p(A_{xy} = 1|A^o)$
$= \frac{1}{Z}\sum_{P\in\Omega}\prod_{\alpha\leq\beta}\int_0^1 p(A_{xy} = 1|P,Q)Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1 - Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}}dQ_{\alpha\beta}$

$= \frac{1}{Z}\sum_{P\in\Omega}\prod_{\alpha\leq\beta}\int_0^1 Q_{\sigma_x,\sigma_y}Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}}dQ_{\alpha\beta}$

1^o. when $(\sigma_x, \sigma_y) \neq (\alpha, \beta)$,

$\prod_{\alpha\leq\beta}\int_0^1 Q_{\alpha\beta}^{l^o_{\alpha\beta}}(1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}}dQ_{\alpha\beta}$

$= exp\{-\sum_{\alpha\leq\beta,(\sigma_x,\sigma_y)\neq(\alpha,\beta)} ln(r_{\alpha\beta}+1) + ln\binom{r_{\alpha\beta}}{l^o_{\alpha\beta}}\}$

2^o. when $(\sigma_x, \sigma_y) = (\alpha, \beta)$,

$\int_0^1 Q_{\sigma_x,\sigma_y} Q_{\alpha\beta}^{l^o_{\alpha\beta}} (1-Q_{\alpha\beta})^{r_{\alpha\beta}-l^o_{\alpha\beta}} dQ_{\alpha\beta} = \int_0^1 Q_{\sigma_x\sigma_y}^{l^o_{\sigma_x\sigma_y}+1} (1 - Q_{\sigma_x\sigma_y})^{r_{\sigma_x\sigma_y}-l^o_{\sigma_x\sigma_y}} dQ_{\sigma_x\sigma_y} = \frac{(l^o_{\sigma_x\sigma_y}+1)!(r_{\sigma_x\sigma_y}-l)!}{(r_{\sigma_x\sigma_y}+2)!}$

$= \frac{l^o_{\sigma_x\sigma_y}+1}{r_{\sigma_x\sigma_y}+2} exp\{[ln(r_{\alpha\beta}+1) + ln\binom{r_{\alpha\beta}}{l^o_{\alpha\beta}}]\}$

Combine the above two situations, and reliability for link $\{v_x, v_y\}$ could be calculated as formula(20):

$$p(A_{xy} = 1|A^o) = G_{xy} = \frac{1}{Z}\sum_{P\in\Omega} \frac{l^o_{\sigma_x\sigma_y}+1}{r_{\sigma_x\sigma_y}+2} H \qquad (20)$$

And the expression of Z, H see as formula (19) and formula (**??**).

For a specific TON, node type have been fixed, which means the division plans P are also fixed, then :

$$p(A_{xy} = 1|A^o) = g_{xy} = \frac{l^o_{\sigma_x\sigma_y}+1}{r_{\sigma_x\sigma_y}+2} \qquad (21)$$

Theorem 1 has been proved. □