Network Topology Connection Optimization Control Algorithm Based on Network Efficiency and Average Connectivity

Xianwang Li*, Yuchuan Song*, Ping Yan, and Xuehai Chen
State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing, China
*Corresponding author, Email: lxw023@163.com, syc@cqu.edu.cn, yp@cqu.edu.cn, cxh@cqu.edu.cn

Abstract—In order to achieve optimization control of complex network topology connection and to identify the network topology which with the greatest network connection gain, this paper proposes the network topology connection optimization control algorithm based on network efficiency and average connectivity. The algorithm uses network efficiency to characterize network connectivity gains, uses average network connectivity to characterize network connection costs, and presents its calculational optimization algorithm. When a network has the small-world characteristics, and \( m \leq n \), the time complexity of the algorithm can reach \( O(n^2) \). Experimental results show that: the proposed algorithm can increase the connection gain and reduce connection costs.

Index Terms—Evaluation Model; Network Size; Network Efficiency; Connection Costs

I. INTRODUCTION

Many systems in nature and human society can be expressed by adopting the complex network model, in which the nodes represents the entities in the system, and the sides stand for the connections between the entities. Such as the interpersonal relationship network, the paper collaboration network, the movie star cooperation network, the E-mail network, the blog reference network, the telephone network and so on. In recent years, the complex network has become one of the hot topics in the study of a variety of subjects. Many complex networks have community structure, which can be treated as a sub-figure of the network. In the internal of the community, the nodes connect closely, while between the communities, the nodes connect sparsely [1]. Research shows that the social network and the biochemical network both have obvious community structure. The structure of the network has a close connection with the function of the network, thus the research on the community structure of the network can reveal the rules hidden in the complex network, and it contributes to the prediction and control of the behavior, such as the community division of web, the analysis and function prediction of the protein network and so on.

Complex network is the abstraction of complex system, it is everywhere, many practical systems can be abstracted as network models to be studied. One of the core issues in complex network research is the relationship between complex system structure and function, so optimizing network structure, improving the system function has attracted the attention of many scholars in the field of complex networks [2].

On the research of current complex network optimization, different scholars have optimization goals aim at different needs for different networks. Some scholars have focused on improving network robustness: Jing WP etc. proposed the wireless sensor network topology optimization algorithm based on Kleinberg model, by optimizing the network significantly improved the fault tolerance and reliability; Fan W, etc. through key nodes’ discovery of existing complex network topology and improvement and analysis of elimination algorithm to improve the P2P network’s connectivity and robustness [2-4].

Some scholars have focused on improving the network topology characteristics: Wang LF and others studied the network edge reconnection algorithm, through edge reconnection to improve network synchronization; Holme P etc. [5-7], presented a heterogeneous connection strategy to optimize the dynamic wireless network’s topology structure, made it has the small world property; Hu JM etc., put forward an wireless sensor network optimization algorithm for urban traffics, by optimizing the network to make it has the small world property [8-10].

Some scholars have focused on ways to improve network efficiency: Rafiee M and others studied the network topology optimization problem on information-sharing in multi-agent systems, and used mixed integer programming algorithm to obtain optimized network model and edge weights [11]; Xue YH etc., in order to improve network transmission efficiency, proposed an algorithm which used uniform and non-uniform network topology to improve the efficiency of the network transmission [12-16]; Ouveysi I and others in order to reduce the traffic flow and the network congestion problems of fiber-optic network, optimize network robustness, they considered the network topology optimization and survival strategies and proposed a LCM-WP algorithm to optimize the optical fiber network [17-21]. Tang min and others put forward a new optimization algorithm for hierarchical
network that solves the maximum matching of the bipartite graph. On the basis of building the priority matching rules of the hierarchical network, the algorithm generates the hierarchical network system by combining the breadth-first search strategy, and then finds the maximum matching one by one, according to the reversed order of the network, namely start from the bottom up. In the aspect of the efficiency of the algorithm, the preprocessing operations of the bipartite graph and the solution strategy with the reversed order of the network both greatly accelerate the speed of matching.

The generation of the network topology is a very important function in the management system of the network. The topological graph of the network provides an intuitive means to understand the connection situation of the global network. Through the topological graph of the network, the network administrator can grasp the overall situation of the whole network. And it is also a very good entrance for the network administrator to implement the network management functions. To realize the generation of the network topological graph, all necessary information to construct the network topological graph must be collected firstly.

From the current research situation, many scholars focus only on the purposes of complex network optimization, while ignoring the costs of network optimization. To this end, this paper presents the optimization control algorithm for complex network topology connection based on considering network gains and connectivity costs.

Minimum spanning tree and fully connected network are two extremes of connected network, although in the actual network construction they are rarely used, there is huge network topological space between them.

In order to achieve the optimization control of complex network topology connection and identify the network topology which with maximum network connection gain, through defining network information connection gain, connection costs and network gain, using network efficiency and average connectivity degree to characterize network connection gain and connection costs, this paper presents a evaluation algorithm for complex network connection’s optimization and design, provides the basis for the actual complex networks system optimization and control.

This paper mainly make expanding and innovative work in the following areas:

(1) In order to increase the actual network system’s connection gain, reducing the cost of network connection, this paper proposes the network topology connection optimization control algorithm based on network efficiency and average connectivity. The algorithm uses network efficiency to characterize network connectivity gains, uses average network connectivity to characterize network connection costs, and presents its calculational optimization algorithm. When a network has a small-world characteristic, and $m \ll n$, the time complexity of the algorithm can reach $O(n_\lambda)$. Experimental results show that: the proposed algorithm can increase the connection gain and reduce connection costs.

(2) In order to further validate the efficiency and correctness of the proposed algorithm, the simulation experiments on the effectiveness of the model, the efficiency of the algorithm and the optimal control of the network topology connection are carried out. When the average degree of the network is definite, due to the cost of the network connection remains the same, the rewiring probability $p$ of the sides in the network increases, which will shorten the length of the shortest path between the node pairs in the network, improve the whole efficiency of the network, and increase the connection gains of the network. The simulation results show that: we can take some ways to optimize practical complex network’s topology connection control, fully connections between network nodes will not necessarily improve the network connection gain, by appropriately control the information connectivity degree and the connection way in complex network, we can get the optimal performance of the network connection; Optimal network average value which can enable the network to obtain the optimal network connections gain exists in both small-world and scale-free networks.

II. OPTIMIZED CONTROL ALGORITHM FOR NETWORK TOPOLOGY CONNECTION

A. Theory

Suppose figure $G=(s,1)$ is a non-self-loop undirected connected network, where $s=\{s_1, s_2, ..., s_s\}$ is the set of all nodes in the network; $m=(m_1, m_2, ..., m_m)$ and $m \subseteq s \times s$ is the set of edges between nodes. $\alpha_{ij}$ indicates the weights of $m_{ij}$ which connects node $s_i$ and $s_j$.

Definition 1 For figure $G=(s,1)$, $f_{ij}$ represents the weights of edge between nodes $s_i$ and $s_j$ in $G$, when $v_i$ and $v_j$ directly are adjacent to each other, $f_{ij}=\alpha_{ij}$, otherwise $f_{ij}=\infty$, then n square matrix $A=(f_{ij})_{n \times n}$ is called G’s edge weight matrix. Edge weight matrix $A$ is used to represent the distance that a network node can reach with one step. when $f_{ij}=\infty$, indicating the nodes $s_i$ and $s_j$ can not step up, they are not adjacent. Network edge weight is used to indicate the degree of difficulty of information flow between nodes, the greater the value, the more resources consumed.

Definition 2 For Figure $G=(s,1)$, use $f_{ij}$ to denote the adjacent situation between nodes $s_i$ and $s_j$ in $G$, when $s_i$ and $s_j$ are directly adjacent to, $f_{ij}=1$, otherwise $f_{ij}=0$, then n square matrix $A'=(f_{ij})_{n \times n}$ is called the adjacency matrix of $G$.

Definition 3 For figure $G=(s,1)$, $r_{ij}$ represents node $j$ which is connected to the node $s_i$ in $G$, if...
\[ f_y = 1(j=1..n) \text{, then } r_k = s_j, \text{ and } k = k+1 \text{; The } \]
\[ n \times p_{max} \] \text{is called } G \text{'S connected matrix.}

Definition 4 Node distance is the minimum value of the edge weights of all paths between two nodes, represented by \( d \). Between \( s_i \) and \( s_j \), if the path does not exist, then \( d_{\text{max}} \). The maximum distance between nodes in the network is diameter of the network, indicated by \( E \). When the network is unentitled network, \( d \) represents the number of edges on the shortest path between two nodes.

\[
d_y = \min \sum_{m=1}^{\infty} \sigma_m, s_i, s_j \in r_k \in E_i \tag{1}
\]

\( d_y \) represents the distance between nodes \( s_i \) and \( s_j \), \( r_k \) represents a certain path connecting nodes \( s_i \) and \( s_j \), \( E_i \) indicates the set of path of connection node \( s_i \) and \( s_j \).

Definition 5 network efficiency \( Q \) is the mean of the reciprocal of the shortest path between any two nodes, which is used to indicate the average difficulty of information flow in the network. The higher the network efficiency, the easier the network information flow.

\[
Q = \frac{1}{k(k-1)} \sum_{i<j} \frac{1}{d_y}
\tag{2}
\]

where in, \( k \) is the number of nodes in the network, \( d_y \) is the distance between node \( i \) and \( j \). From the definition of network efficiency \( Q \), network efficiency \( Q \) expresses the average proximity between all node pairs in the network. In the network, the closer between nodes and the shorter the distance, the greater the value of the network efficiency.

B. Network Connectivity Evaluation Model

In the set up process of actual network system, the communication needs between different nodes, location and requirements for system functions are different, so the construction cost of node communication lines is different.

In this paper, the path length between nodes is used to measure the cost of building links and the comfort level of information dissemination. The shorter path length between nodes, the lesser the link building cost and the easier the dissemination of information.

For the actual network, the shorter the average path length can make the network information quickly spread, saving network resources consumed by the construction, improve the overall efficiency of the network, so by increasing network connectivity can reduce the average path length of the network, save resources, increase network connection revenue; however, due to the increase of average network connectivity, the node will lead to information overload and redundancy for communication links, and result in higher costs of information congestion and redundancy.

Simple communication network as shown by Figure 1, nodes with the same number may exist various kinds of connections (network A, B, C represent weighted networks, Network D, E, F, respectively represents the untitled network’s connection model corresponding to network A, B, C, the network G, H, I respectively denotes distinct connection models of ring network, chain network and the star network .Network A and network D’s communication connection is very full, the information can spread quickly in all nodes to obtain high gain which is related to information high-speed dissemination, but also makes the network connection cost increase; compared to network A and network D, network B, C, and network E, F has lesser network connection costs, each node can achieve effective dissemination of information. Therefore, for networks of any size, we can take a certain way to optimize and control its connection mode, to obtain a higher gain of the network connection.

As can be seen from Figure 1, the high gain of network connection is related to the information connection with the appropriate number: with the increase of the degree of average network connection, it will make the length of the shortest path between the network nodes decrease and increase the gain of the network connection; Meanwhile, excessive network connectivity will generate higher network connection costs which is caused by excessive information connections.

Therefore, this article considering the network connection level and the shortest path length of each node, defines the network connections gain \( w \):

\[
w = h(Q) \times (1-l) \tag{3}
\]

where in, \( h(Q) \) is the connection gain of complex network, \( l \) is the complex network connection cost, \( Q \) is the network efficiency.

For the information transmitted in the network, if the distance between two nodes is large, and meanwhile
maintain the integrity of the transmission of information, then the information transmission between nodes consumes a lot of resources; on the contrary, it is easy to transmit information between nodes. Because the network efficiency characterizes the average network proximity of the nodes in the network, so the network efficiency to some extent reflects the connection gains of the entire network.

The higher efficiency of the network, indicates that the shorter the distance between nodes, the easier the network information flow, the smaller the load of network information transmission, the greater the benefits. Therefore, this article defines \( h(Q) \) as follows:

\[
h(Q) = \frac{Q}{Q_{\text{max}}} \tag{4}
\]

where in, \( Q \) and \( Q_{\text{max}} \) respectively represents the network efficiency and maximum value of network efficiency.

For maximum network efficiency \( Q_{\text{max}} \), we consider an idealized example: fully connected networks. As nodes in the fully connected network are fully connected with each other, the distance between the nodes can all achieve the minimum value, making the information among them can be transmitted with the highest efficiency;

When the network on the basis of fully connected network removes a number of edges, it will cause the minimum distance between a number of nodes increase and then reduce efficiency of the network. It can be seen, the fully connected network’s efficiency can achieve the maximum, \( Q_{\text{max}} \) used herein refers to the fully connected network’s efficiency.

For example, in Figure 1, the efficiency of A and D is the \( Q_{\text{max}} \) of the simple weights and untitled communication networks. From (4) we can see, \( h(Q) \in [0,1] \).

When using the network model to analyze the real network, an important variable is the network cost. We expect that when the number of network edges increases, its efficiency can be increased even more rapidly; \( Q \), when the number of network edges is fixed, by changing the network connections to improve network efficiency and therefore realistic network’s construction tends to focus on network construction cost and efficiency.

For network nodes consumed resources, the theory used simple graph can be presented by the node value, betweenness, etc. However, many practical communication networks connection cost of nodes does not depend on the degree value of network and betweenness because of its different routing policies.

In order to characterize network connection cost \( x \), this paper uses connection degree of network edge to represent it and define it as follows:

\[
x_{\text{max}} = \sum_{r \in E} q_{ij} r(q_{ij}) = \sum_{r \in E} h(q_{ij}) \tag{7}
\]

where in, \( q_{ij} \) and \( q_{ij}' \) respectively denotes elements of fully connected network adjacency matrix and edge weight matrix. From (5) it can be seen \( x \in [0,1] \).

Since \( h(Q) \), and \( x \) are respectively defined in \([0,1]\), so we can use formula (3) to study network connection cost-benefit.

When the network is fully connected network, its network connectivity and cost benefits reaches the maximum value 1, while its network connection gain is 0, this is one of the important reasons why fully connected network is rarely used in the real world, at the same time it also illustrates the reasonability of the proposed model.

C. Network Topology Connection Optimization Control Algorithm

Here are simple algorithm steps for assessing the network connection gain:

**Input:** edge weight matrix \( A = (f_{ij})_{n \times n} \), the adjacency matrix \( A' = (f_{ij}')_{n \times n} \), connection matrix \( r = (r_{ij})_{n \times n} \), \( \beta \), \( p \).

**Output:** network connection gain.

1) According to \( A = (f_{ij})_{n \times n} \), compute the shortest distance matrix \([d_{ij}]_{n \times n}\);

   / Floyd between all nodes;

2) Construct a fully connected network with the same size, according to formula (2) and (7) to respectively calculate \( E_{\text{max}} \) and \( Q_{\text{max}} \);

3) According to \([d_{ij}]_{n \times n} \), use formula (2) and (4) to calculate \( h(Q) \);

4) According to \([d_{ij}]_{n \times n} \) and \( A' = (f_{ij}')_{n \times n} \), use formula (5) to calculate \( Q \);

5) According to formula (3), output \( w \).

End

As can be seen from the above algorithm steps, for the cost-benefit assessment of the network connection, the key lies in determining the network’s shortest distance matrix \([d_{ij}]_{n \times n} \), its time complexity is \( O(n^3) \). By the analysis for the Floyd algorithm, in the process of Floyd algorithm calculating the shortest distance matrix \([d_{ij}]_{n \times n} \). The determination of shortest distance between any node needs to conduct \( n \) cycles for the distance matrix, but when all the network nodes’ shortest distance have been found, we will find that \([d_{ij}]_{n \times n} \) remain unchanged in subsequent cycles. Therefore, in solving the \([d_{ij}]_{n \times n} \), if blindly conduct \( n \) cycles, will inevitably lead to a waste of computing resources, improve the algorithm's time complexity.

Based on this, in the use of Floyd algorithm to calculate \([d_{ij}]_{n \times n} \), this article moves the intermediate node
loop into the innermost loop, uses network connection matrix \( r = (f_ij)_{n \times p} \) to optimize it.

The optimized complex network's shortest path calculation algorithm flowchart is shown in Figure 2.

The optimized complex network's shortest path calculation algorithm flowchart is shown in Figure 2.

It can be found through analysis that the column value \( p \) of the network connection matrix \( r = (f_ij)_{n \times p} \) is the maximum value of nodes in the network; When all the shortest distance between network nodes have the found, the maximum distance between the network nodes (network diameter \( R \)) has also been identified.

Thus, the time complexity of the algorithm shown in Figure 2 is \( O(M_{max}) \). \( M \) represents the number of cycles required to determine the diameter \( R \) of the network, \( p \) represents the maximum value of network nodes. Since most of the actual real-world networks are sparse and small-world networks, the network connection between the nodes is not very good, network nodes’ maximum value is relatively smaller, the network diameter \( R \ll n \), even the maximum value of degree of scale-free network node is relatively large, only a few nodes have a larger value. Therefore, in the calculation of the gain of the actual network connections, the proposed algorithm’s time complexity can reach \( O(n) \), and it can get a good computing power for large-scale complex network topology connections’ optimization control.

### III. SIMULATION AND ANALYSIS

#### A. Effectiveness Analysis of Models

In order to analyze the effectiveness of the proposed model, this paper takes the simple communication network shown in Figure 1 as an example, to give a comparative analysis for networks connection gain. By a simple analysis of Figure 1, it can be found that in weighted networks A, B, C: In A network, there are multiple redundant paths, such if \( e \), \( l_w \), \( l_e \), \( l_d \); In network B, there also exists redundant paths; only network C’s connection is relatively reasonable. For the untitled network in Figure 1, although we can not intuitively judge the redundant paths in the network, the network connection gain between them is different. Use the proposed model to calculate the network connection gain of Figure 1, the results are shown in Table 1 (\( \beta = 1, p = 4 \)).

As can be seen through Table 1, the network A and network D’s sufficient connection makes it possible to exchange information between the nodes with the minimum path length, and it can achieve the maximum efficiency, increasing the network connection gain.

But increase of the network connection level results in excessive redundancy of network connection and connection costs increase, leading to network A and network D with the minimum network connection gain; Though network B and network C have the same network efficiency value, network C has lower network connection degree so that it has the highest network connection gain; network F also has the same characteristics.

Calculated results are the same with the theoretical analysis, which illustrates the effectiveness of the proposed algorithm, it can effectively weights the network connection benefits and costs of the network connection during building weighted network and untitled network.

### Table I. Comparison of Different Network Connection Gain

<table>
<thead>
<tr>
<th>type</th>
<th>network</th>
<th>( Q_{max} )</th>
<th>( E_{max} )</th>
<th>( h(Q) )</th>
<th>( I )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted network</td>
<td>A</td>
<td>0.84</td>
<td>1</td>
<td>53</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.68</td>
<td>0.81</td>
<td>0.70</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.68</td>
<td>0.81</td>
<td>0.22</td>
<td>0.63</td>
</tr>
<tr>
<td>Unauthorized network</td>
<td>D</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>0.81</td>
<td>0.81</td>
<td>0.59</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.68</td>
<td>0.68</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.68</td>
<td>0.68</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>0.60</td>
<td>0.60</td>
<td>0.34</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>0.68</td>
<td>0.68</td>
<td>0.34</td>
<td>0.43</td>
</tr>
</tbody>
</table>

The shortest distance matrix \([d_{ij}]_{n \times m}\)

![Figure 2. Calculation algorithm flowchart of shortest distance matrix](image-url)
connection gain and minimum network connection cost. Its connection mode is optimal, the reason is that the star network can make path length between nodes the minimum with the least number of connections. Untitled network’s computing results show that sufficient connection between network nodes will not necessarily increase the network connection gain, properly control complex network’s information connectivity degree and connection ways, we can get the optimal performance of the network connection.

B. Analysis of Algorithm Efficiency

In order to analyze the efficiency of the proposed method, this paper uses the optimization algorithm on Intel Core i5 3.10 GHz PC to run MAT-LAB program calculate the network connection gain of small-world networks whose edge weight is 1 and with different sizes (each node is connected with the surrounding 80 nodes, edges rewiring probability is 0.6). Through the simulation, the change trend of execution time can be obtained as the network size changes, shown in Figure 3 (average of several simulations).

![Figure 3. Execution efficiency graph of small world network algorithm with different sizes](image)

As can be seen from Figure 3, the proposed optimization algorithm is prior to the Floyd algorithm with higher computational efficiency, its connectivity gain is not more than 10 s when computing small-world network which with 1,000 nodes, it can be applied to optimal control algorithm of large-scale small-world network topology connection.

C. Simulation Analysis of Network Topology Connection Optimization Control

In order to optimize small-world network connection mode, the paper constructs a small-world network with size as 400, the right as 1, and the average values are 6, 12, 20, 30 to analyze it. Through simulation, we can get the change trend of small-world connection gain as the reconnection probability p changes, shown in Figure 4 ($\beta = 1, p = 2$, average of several simulations).

As can be seen in Figure 4, when the network average value reaches a certain level, because the network connection costs are not changed, therefore the increases of network side rewiring probability p will shorten the shortest path length between network nodes, improve the overall efficiency of the network, and proceeds the network connection gain. Therefore, the network connection gain increases as p increases. Meanwhile, the network connection gain will also increase as the average degree increases, but when the network average value reaches a certain value, the network connections gain’s growth slows down. In order to determine the best average degree value of small-world network, the paper takes $\beta = 1, p = 2$, edge rewiring probability $p = 0.65$, to construct a small-world network with the network size of 400 for simulation.

Then we can get the influence diagram of small-world network average value on network connection gain, as shown in Figure 5 (average of several simulations).

![Figure 4. Influence of side rewiring probability on gain network connections influence](image)

![Figure 5. Influence of average degree value on network connection gain](image)

Through Figure 5 shows that when the network average value is less than a certain value, because the network connection cost is small, increasing the average network connectivity can significantly improve the network connection gain;

But when the network average value increases to a certain extent, the increase of average network connectivity will make the network connection cost greater than the network connection gain brought by the decrease of network average path length, resulting in decrease of the network connection gain. In the small-world networks, there is an optimal network average value, which can make the connection network gain reaches optimum.

In order to optimize the scale-free network connection mode, this paper according to the algorithm provided by Barabasi AL constructs the scale-free network with size of 400, the new added nodes connect different number of nodes for comparative analysis. Through simulation, we can get the change trend of scale-free network connection gain as the number of edge nodes changes, as shown in Figure 6 ($\beta = 1, p = 2$, average of several simulations).
The net gain exists in an optimal network average value which enables network connectivity gain to be the maximum.

Integrated simulation results of Figures 4 to 6 indicate that in the complex network’s building process, we can take some ways to optimize the network connection mode. In small-world networks and scale-free networks there exist in an optimal network average value which enables the network to obtain the optimal network connections gain.

IV. CONCLUSION

In order to increase the gain of the actual network connection, reduce network connection costs, this paper proposes the network topology connection optimization control algorithm based on network efficiency and the average degree of connectivity. The algorithm uses network efficiency to characterize network connection gains and uses average network connectivity to characterize network connection costs, and presents its calculational optimization algorithm. When a network has the small-world characteristics, and \( m \ll n \), the time complexity of the algorithm can reach \( O(n^3) \).

Experimental analysis shows that: the proposed algorithm can in a certain way to conduct optimization control for the actual complex network topology connection. Optimal network average value which can enable the network to obtain the optimal network connection gain exists in both small-world and scale-free networks.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 51075414.

REFERENCES


