Optimization-based topology identification of complex networks*

Tang Sheng-Xue(唐圣学)[a)†], Chen Li(陈 丽)[a)], and He Yi-Gang(何怡刚)[b)]

[a)] *Province–Ministry Joint Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability, Hebei University of Technology, Tianjin* 300130, *China*

[b)] *College of Electrical and Information Engineering, Hunan University, Changsha* 410082, *China*

(Received 28 August 2010; revised manuscript received 23 June 2011)

In many cases, the topological structures of a complex network are unknown or uncertain, and it is of significance to identify the exact topological structure. An optimization-based method of identifying the topological structure of a complex network is proposed in this paper. Identification of the exact network topological structure is converted into a minimal optimization problem by using the estimated network. Then, an improved quantum-behaved particle swarm optimization algorithm is used to solve the optimization problem. Compared with the previous adaptive synchronization-based method, the proposed method is simple and effective and is particularly valid to identify the topological structure of synchronization complex networks. In some cases where the states of a complex network are only partially observable, the exact topological structure of a network can also be identified by using the proposed method. Finally, numerical simulations are provided to show the effectiveness of the proposed method.

Keywords: complex networks, topology identification, optimization, particle swarm

PACS: 05.45.–a, 05.45.Xt **DOI:** 10.1088/1674-1056/20/11/110502

1. Introduction

There exist widely complex networks in the real world. Typical examples are the power grid network,[1] the World Wide Web,[2] metabolic networks,[3] and so on.[4−6] Since the scale-free networks and "small-world" networks were found,[7,8] complex networks have made significant progress in many fields, such as network modeling, synchronization, structure optimization and control.[9−12] However, sometimes the exact topology structure of a complex network is unknown or uncertain in advance, and the identification of the topological structure is of benefit to study the properties of complex networks. For instance, if faults or malfunctions occur in links of networks, it is significant to locate or detect quickly the positions of faults. Thus identifying the topological structure of unknown network becomes a new and hot research field.[13−19]

Recently, researchers[13−19] have mainly investigated the adaptive feedback control method based on synchronization to identify topology of networks, which has been extended to identify uncertain complex networks and time-delay networks. However, there are mainly two disadvantages of the adaptive synchronization-based method. Researchers in Ref. [16] pointed out that the topological structure of synchronization complex networks cannot be estimated correctly by this method. This is the first disadvantage. Using the developed or revised adaptive synchronization-based methods in Refs. [16] and [18] to identify the topological structure on synchronization networks is a difficult choice because the implementation of the presented schemes must change the network coupling modes. The second disadvantage is that all of the nodes of the network must be observable. This condition is very rigorous in some complex networks, such as neural networks and metabolic networks.[3,6] In this paper, we propose a new scheme for identifying the topological structure of networks, in which the topology identification is converted into a minimal optimization problem. On exploring the topological structure parameters of the estimated network, the objective error values between the states of the driving network and the estimated (response) network are minimized, and then the topology iden-

*Project supported by the National Natural Science Foundation for Distinguished Young Scholars of China (Grant No. 50925727) and the National Natural Science Foundation of China (Grant No. 60876022).

†Corresponding author. E-mail: tangsx@yahoo.cn

© 2011 Chinese Physical Society and IOP Publishing Ltd

tification is implemented effectively. The scheme is valid not only for the non-synchronization complex networks, but also for the synchronization networks and the networks with unobservable nodes.

Inspired by the particle swarm optimization (PSO) algorithm and quantum mechanics theories, the quantum-behaved PSO (QPSO) has been developed in recent years.[20,21] Similarly, QPSO is an extremely simple algorithm. However, QPSO is a global optimization algorithm due to the uncertainty principle of quantum behaviors.[22] In the present work, we present an improved QPSO for efficiently exploring the parameter values of complex network.

The remainder of this article is organized as follows. The topological structure identification of complex networks is converted into a minimal optimization problem in Section 2. In Section 3 the simple and efficient QPSO-based algorithm to solve the minimal optimization problem is presented. In Section 4 numerical simulations and exciting results are described. Finally, in Section 5 some conclusions are drawn from the present study.

2. Optimization formulation of identification

Consider the following weighted complex network consisting of N nodes, with each node being a different dynamical system, called the driving network and described by[16]

$$\dot{X}_i = F_i(X_i) + \sum_{j=1}^{N} H(X_j), \quad i = 1, 2, \ldots, N, \qquad (1)$$

where $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})^{\mathrm{T}}$ is the state vector of the i-th node dynamical system F_i, and $H(\cdot)$ is the inner coupling function. In addition, c_{ij} is the unknown weighted topology connection from node i to node j. The coupling matrix $C = (c_{ij})_{N\times N}$, where $c_{i,i} = \sum_{j=1, i\neq j}^{N}$, does not need to be symmetric and irreducible.

In order to identify the unknown coupling configuration matrix C, the node dynamical systems F_i and inner coupling function $H(\cdot)$ are assumed to be known in advance, and thus the estimated (response) network can be described as

$$\dot{Y}_i = F_i(Y_i) + \sum_{j=1}^{N} H(Y_j), \quad i = 1, 2, \ldots, N, \qquad (2)$$

where $\hat{c}_{ij}$ is the estimated parameter of c_{ij}.

Therefore, the problem of identifying topology matrix C can be converted into a minimal optimization problem

$$\min J(\hat{c}_{ij}) = \sum_{k=1}^{M} |X(k) - Y(k)|^2, \qquad (3)$$
$$\hat{c}_{ij}\in\Gamma$$

where $J(\hat{c}_{ij})$ is the objective evaluation function; Γ is the predefined searching spaces; $X(k) = (X_1^{\mathrm{T}}(k), X_2^{\mathrm{T}}(k), \ldots, X_N^{\mathrm{T}}(k))^{\mathrm{T}}$, $Y(k) = (Y_1^{\mathrm{T}}(k), Y_2^{\mathrm{T}}(k), \ldots, Y_N^{\mathrm{T}}(k))^{\mathrm{T}}$ are state vectors of the driving and the estimated networks at discrete time k, respectively; M denotes the length of data used for topology parameters estimation; $Y(k)$ is obtained by simulating the response network with initial value $Y(0) = X(0)$ and the estimated configuration matrix $\hat{C}_{N\times N}$.

To identify the synchronization networks, the state vector $X(k)$ used for evaluating the objective function $J(\hat{c}_{ij})$ is selected from the state vectors of driving network in non-synchronization. The non-synchronization state vectors of synchronization network in this paper can be obtained by setting the non-synchronization initial value $X(0)$. To identify the networks with unobservable states of nodes or unreachable nodes, state vector $X(k)$ is composed only of observable states variables, which will be explained in detail by two examples in Subsection 4.3.

3. QPSO-based identifying algorithm

PSO is a swarm intelligent optimization algorithm, in which the trajectory of a particle is determined by its position vector x_i and velocity vector v_i.[20] The particle moves along a determined trajectory in Newtonian mechanics. According to the uncertainty principle, the term trajectory in the quantum setting is meaningless because x_i and v_i cannot be determined simultaneously. Therefore, if individual particles in a PSO system have quantum behaviour, the particle can search for the optimal goal in the whole space,[20] and as a result the population with a probability of one converges at the globally optimal position.[22]

In order to improve the convergent performance of QPSO algorithm in this work, the position x_i of the i-th particle moves according to the following iterative equation:

$$\begin{cases} x_i(t+1) = p + \alpha \cdot |\text{Mbest}_i - x_i(t)| \cdot \ln(1/u), & \text{if } u \geq 0.5, \\ x_i(t+1) = p - \alpha \cdot |\text{Mbest}_i - x_i(t)| \cdot \ln(1/u), & \text{if } u < 0.5, \end{cases} \tag{4}$$

where $i = 1, 2, \ldots, N$ indicates the number of the particles of population (swarm); $t = 1, 2, \ldots, t_{\max}$ represents the iteration number; $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,d})^{\mathrm{T}}$ denotes the position of the i-th particle of population; u is the value generated according to a uniform probability distribution in the range [0 1]; p is the local attractor of the i-th particle; Mbest_i is the global point called *Mainstream Thought* or *Mean Best* of the population; α is the contraction–expansion coefficient. Mbest_i and α are explained and calculated in the following.

The local attractor p of the i-th particle of population is defined as a stochastic position that lies in a hyper-rectangle with $p_i(t)$ and $p_g(t)$ being two ends of its diagonal and calculated as

$$p = (c_1 p_i(t) + c_2 p_g(t))/(c_1 + c_2), \tag{5}$$

where $p_i(t)$ (*pest*) indicates the personal best position of the i particle and $p_g(t)$ (*gbest*) the global best position of population, and c_1, c_2 are two random numbers in [0 1].

The global point Mbest_i of the population is defined as the randomized weighted mean of the *pest*- and the *gbest*-positions of all particles and is given by

$$\text{Mbest}_i = \frac{1}{N} \sum_{i=1}^{N} \beta p_i + (1-\beta) p_g, \tag{6}$$

where N is the size of the population, and β is a random number in [0 1].

The contraction–expansion coefficient α is adaptively adjusted according to the gathering h and convergent speed s of the particles and given by

$$\alpha = \alpha_0 + \alpha_s s - \alpha_h h, \tag{7}$$

where α_0, α_s, α_h are constants and the gathering h and convergent speed s in minimal optimization problem are defined as

$$\begin{cases} h = J(gbest(t-1))/J(gbest(t)), \\ s = J(gbest(t))/\frac{1}{N} \sum\limits_{i=1}^{N} J(x_i(t)). \end{cases} \tag{8}$$

To assure converging at a globally optimal position with a probability of one, the maximum $\alpha_{\max}$ of coefficient α must be less than 1.781.[23] The *pbest* position of each particle and the *gbest* position of population are updated in the same way as that in the PSO system. The procedure for identifying topological structure of network, based on the QPSO in this work, is given by the following steps.

Step 1 Identify the dimension d of independent variable according to the topology properties of complex network and unknown parameters c_{ij}, and then initialize a population of particles with random positions x_i in the d-dimensional space using a uniform probability distribution function. Set $s(0) = 0$, and $h(0) = 0$.

Step 2 Simulate the estimated complex network (2) for each particle with initial value $Y(0) = X(0)$ and obtain the state vectors $Y(k)$ at the same time points of network (1).

Step 3 Evaluate the objective function value of each particle according to Eq. (3), and set the *pbest* to be equal to the objective function value and set the *pbest* location to be equal to the personal location. Find the globally best objective function value of the population and set the *gbest* to be equal to the globally best objective function value and set the *gbest* location to be equal to the globally best position.

Step 4 Calculate α, Mbest_i, and p using Eqs. (7), (6), and (5), respectively, and update the position of each particle according to Eq. (4).

Step 5 Simulate the estimated complex network (2) for each particle at the current generation with initial value $Y(0) = X(0)$ and obtain the state vectors $Y(k)$ at the same time points of network (1), and then evaluate the objective function value of each particle.

Step 6 Compare each particle objective function value with the present particle *pbest*. If the current value is better than the previous *pbest*, then set the value *pbest* to be equal to the current value and set the *pbest* location to be equal to the current location.

Step 7 Compare the objective function value with the population overall previous *pbest*. If the current value is better than the *gbest*, then reset *gbest* to be the current particle objective function value and position.

Step 8 Calculate gathering h and convergent speed s according to Eq. (8).

Step 9 Go to Step 4 until a user-predefined stop criterion is satisfied, which is usually a sufficiently good fitness or a maximum number of iterations (generations).

4. Numerical simulation and result

To show the effectiveness of the proposed method, five illustrative examples are shown and compared with the cases of the adaptive synchronization-based method. Lorenz chaotic attractor is taken as the node of networks, and the dynamical system of the i-th node is described as $\dot{X}_i = F(X_i) = [(a(x_{i,2} - x_{i,1}); bx_{i,1} - x_{i,1}x_{i,3} - x_{i,2};\ x_{i,1}x_{i,2} - cx_{i,3}]$ with attractor parameters $a = 10$, $b = 28$, and $c = 8/3$.

In the numerical simulations successive $M(= 100)$ vector states of both the original network and the estimated network are used to calculate the objective function value J. In QPSO, the parameters are set as $\alpha_0 = 1$, $\alpha_{\max} \leq 1$, $\alpha_s = 0.25$, $\alpha_h = 0.25$. The searching space Γ is [–10 10].

4.1. Example 1

Firstly, consider a non-symmetric and non-synchronization dynamical diffusive network consisting of three identical nodes with topology matrix C, whose elements are $c_{1,2} = 3$, $c_{2,1} = 4$, $c_{2,3} = 2$, $c_{3,2} = -1$ the other elements $c_{i,j} = 0$ $(i \neq j)$, and $c_{i,i} = \sum_{j=1,i\neq j}^{N} c_{i,j}$. The inner coupling $H(\cdot)$ is an identical matrix I_3 and the initial states $X(0) = [-5, 2, 8; -13, 2, -3; 5, -2, 3]$. Obviously, the topology matrix C is not symmetric, and there are only 6 independent variables, therefore the searching space dimension of particles is 6. The simulation interval is set to be 0 s–0.4 s. The population size is set to be 120, and the maximum generation number used as the stop criterion is set to be 1000.

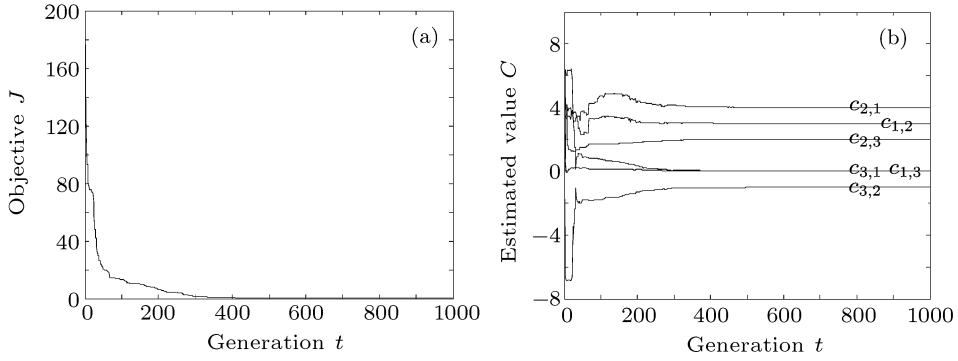


Fig. 1. Evolutions of identification processes of non-synchronization network, showing (a) the objective function value J and (b) the topology matrix C versus generation t.

Figure 1 shows that the value of the objective function value J decreases with the increase of generation t, which indicates that PSO converges to a globally optimal value, and the topology matrix C converges to the true values rapidly. From Fig. 1, the matrix C has been identified correctly.

4.2. Example 2

In this illustrative example, consider a symmetric synchronization complex network[16] with topology parameters $c_{1,2} = c_{2,1} = 4$, $c_{1,3} = c_{3,1} = 7$, $c_{1,4} = c_{4,1} = 3$, $c_{2,3} = c_{3,2} = 5$, $c_{3,4} = c_{4,3} = 2$, the others $c_{i,j} = 0$ $(i \neq j)$ and $c_{i,i} = \sum_{j=1,i\neq j}^{N} c_{i,j}$. Simulation parameters are given as the maxim generation is 5000, the population size is 240, and the initial states are the same as those in Example 1. The network arrives at synchronization at 1.8 second, therefore the states in interval [0 0.2] are selected for computing the objective function value J. Figure 2 shows the evolving process of the objective function value J and the searching progress of topology matrix C. From Fig. 2 the topological structure of synchronization network is identified correctly.

In order to correctly identify the topological structure of network by using the adaptive synchronization-based approach, the necessary conditions for linear independence must be satisfied in Refs. [16] and [19]. In fact, the independence conditions are very difficult to verify. In Refs. [16] and [18] to avoid synchronization of network being identified, the coupling strength η or the inner-coupling mode $H(\cdot)$ is adjusted so that the network is desynchronized. The limitations of desynchronization method is that the behaviours of networks are changed, and it is applicable only to the adjustable or controllable networks. In the proposed

optimization-based method, the topological structure of synchronization network can be identified by selecting the non-synchronization information. If the network has been in a synchronization state, any way of destroying synchronization states, including changing the coupling mode or directly changing the partial or all states of nodes, is beneficial to the topology identification.

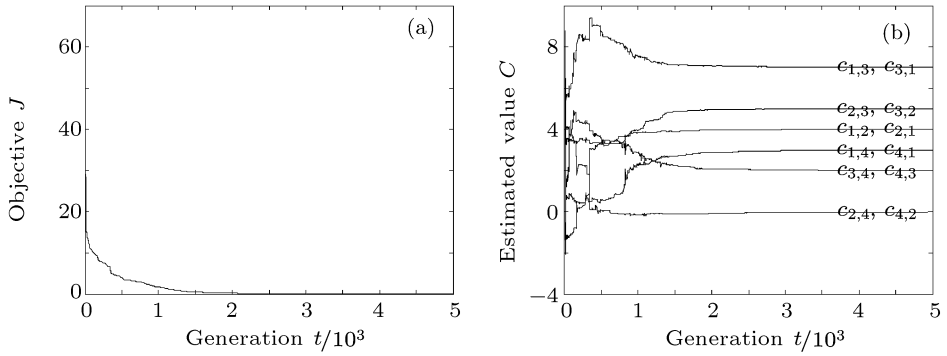


Fig. 2. Evolutions of identification processes of synchronization network, showing (a) objective function value J and (b) topology matrix C versus generation t.

4.3. Example 3

In the third example, consider the non-synchronization networks, in which only partial states of networks are observable. There are two typical cases, (i) only partial nodes are reachable or observable, while the other nodes are not reachable; and (ii) only partial states of dynamical systems of nodes are observable, although all nodes are accessible. The network structures and parameters are identical to those in Subsection 4.1. In case (i), the vector states used in the objective function are $X(k) = (X_1^{\mathrm{T}}(k), X_2^{\mathrm{T}}(k))$, $Y(k) = (Y_1^{\mathrm{T}}(k),\ Y_2^{\mathrm{T}}(k))$, and in case (ii) $X(k) = (X_{1,1}^{\mathrm{T}}(k),\ X_{2,1}^{\mathrm{T}}(k),\ X_{3,1}^{\mathrm{T}}(k))$, $Y(k) = (Y_{1,1}^{\mathrm{T}}(k),\ Y_{2,1}^{\mathrm{T}}(k),\ Y_{3,1}^{\mathrm{T}}(k))$, both not the same as $X(k) = (X_1^{\mathrm{T}}(k),\ X_2^{\mathrm{T}}(k),\ X_3^{\mathrm{T}}(k))$, and $Y(k) = (Y_1^{\mathrm{T}}(k),\ Y_2^{\mathrm{T}}(k),\ Y_3^{\mathrm{T}}(k))$ in Subsection 4.1. Figure 3 shows objective value J and topology parameters evolution progress in case (i). Figure 4 shows the evolution progresses in case (ii). From Figs. 3 and 4, the topological structures are identified with only partial states of networks.

Compared with the adaptive synchronization-based method in Refs. [13], [15]–[17], and [19], the optimization-based method proposed in this paper can identify not only the topological structure of network, where only partial states of dynamical system of node are observable and the network topological structure can be identified with the revised adaptive synchronization method in Ref. [18], but also the topological structure network where the nodes are partially reachable. Therefore, the proposed method is significant to identify the topological structure of network with unreachable nodes and to detect the isolated nodes.

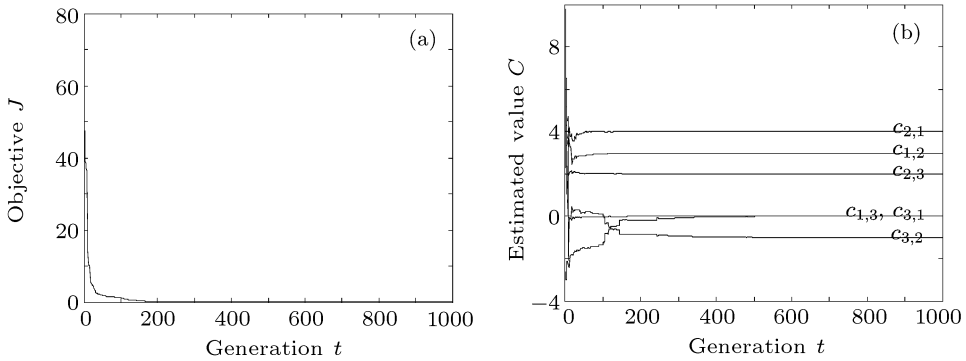


Fig. 3. Evolutions of identification processes of case (i), showing (a) objective function value J and (b) topology matrix C versus generation t.

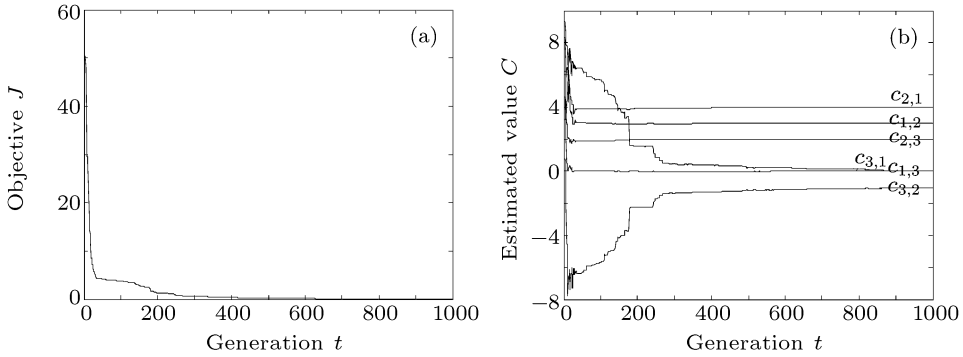


Fig. 4. Evolutions of identification processes of case (ii), showing (a) objective function value J and (b) topology matrix C versus generation t.

4.4. Example 4

Sometimes there are time-delays in the links of network, in which the identification method based on adaptive synchronization is different from that of the networks without time-delay.[17,19] Here, consider an invariant time-delay coupling network, where the inner coupling function is $H(X_j(t-\tau))$ and τ is the time-delay. Let $\tau = 0.01$, and topological structure and other parameters are the same as those in Subsection 4.1. The identification progresses of the objective function value and estimated structure values shown in Fig. 5 indicate again that the topological structure of the network is correctly identified.

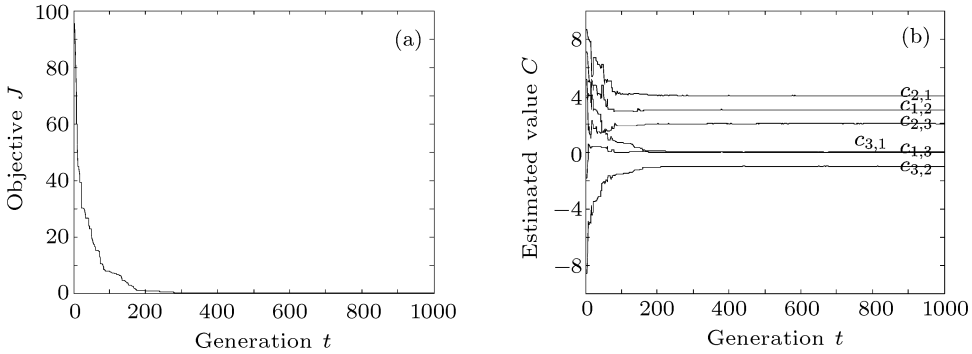


Fig. 5. Evolutions of identification processes of network with time-delay, showing (a) objective function value J and (b) topology matrix C versus generation t.

4.5. Example 5

Consider a "small-world" symmetric complex network, consisting of $N = 20$ identical Loren oscillators, whose topological structure and weighted link values are shown in Fig. 6. Assume that the topological parameters $c_{1,4}$, $c_{1,6}$, $c_{1,8}$, $c_{1,10}$, $c_{1,11}$, and $c_{1,12}$ are unknown. The identification progresses of the objective function value and the estimated values in Fig. 7 show that the topological structure of typical "small-world" complex network is correctly identified by using the proposed optimization method.

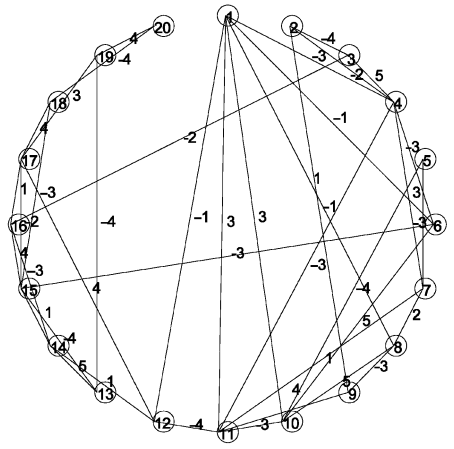


Fig. 6. Topological structure and link values of "small-world" network with 20 identical nodes.

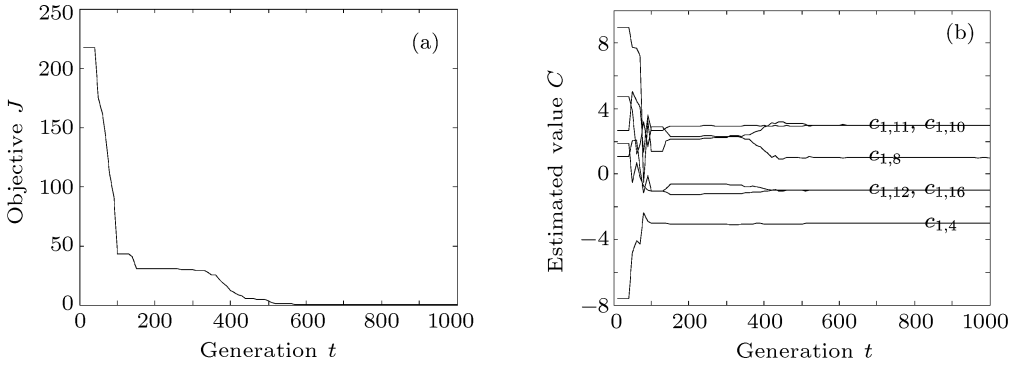


Fig. 7. Evolutions of identification processes of "small-world" network with 20 identical nodes, showing (a) objective function value J and (b) topology matrix C versus generation t.

5. Conclusions

In this paper, an optimization-based method of identifying the topological structure of a general complex network is proposed. The topology identification is converted into a minimal optimization problem in a new perspective. An improved QPSO algorithm is used to explore the globally optimal topological parameters. Compared with the synchronization-based method, the proposed method can identify not only the non-synchronization complex networks, but also the synchronization complex networks, and needs less node information, that is, only partial nodes or node states observable. Numerical simulation results verify the effectiveness of the proposed method, which sheds light on the unknown complex network properties, behaviours, and so on.

References

[1] Buldyrev S V, Parshani R, Paul G, Stanley H E and Havlin S 2010 *Nature* **464** 1025

[2] Albert R, Jeong H and Barabasi A L 1999 *Nature* **410** 130

[3] Jeong H, Tombor B, Albert R, Oltvai Z N and Barabasi A L 2000 *Nature* **407** 651

[4] Zhu Z T, Zhou J, Li P and Chen X G 2008 *Chin. Phys.* B **17** 2874

[5] Zhang Z, Fu Z Q and Yan G 2009 *Chin. Phys.* B **18** 2209

[6] Chen G, Zhou J and Liu Z 2004 *Int. J. Bifur. Chaos* **14** 2229

[7] Barabasi A L and Albert R 1999 *Science* **286** 509

[8] Watts D J and Strogatz S H 1998 *Nature* **393** 440

[9] Boccaletti S, Latora V, Moreno Y, Chavezf M and Hwang D U 2006 *Phys. Rep.* **424** 175

[10] Butts C T 2009 *Science* **325** 414

[11] Wu J S, Jiao L C and Chen G R 2011 *Chin. Phys.* B **20** 060503

[12] Pei W D, Chen Z Q and Yuan Z Z 2008 *Chin. Phys.* B **17** 373

[13] Zhou J and Lu J A 2007 *Physica* A **386** 481

[14] Liu H, Song Y R, Fan C X and Jiang G P 2010 *Chin. Phys.* B **19** 070508

[15] Xu Y H, Zhou W N, Fang J A and Lu H Q 2009 *Phys. Lett.* A **374** 272

[16] Chen L, Lu J A and Tse C K 2009 *IEEE. Trans. Circuit Syst.* II **56** 310

[17] Wu X Q 2008 *Physica* A 387 997

[18] Hong W Y, Suo W and Ping G 2010 *Acta Electron. Sin.* **38** 1065 (in Chinese)

[19] Liu H, Lu J A, Lu J H and Hill D J 2009 *Automatica* **45** 1799

[20] Kennedy J and Eberhart R 1995 *Proceedings of IEEE International Conference on Neural Networks* (Piscataway: IEEE) p. 1942

[21] Xi M L, Sun J and Xu W B 2008 *Appl. Math. Comput.* **205** 751

[22] Fang W, Sun J, Xie Z P and Xu W B 2010 *Acta Phys. Sin.* **59** 3686 (in Chinese)

[23] Sun J, Xu W B and Feng B 2005 *IEEE International Conference on Systems, Man and Cybernetics* (Hawaii: IEEE) p. 3049