Chin. Phys. B Vol. 29, No. 8 (2020) 088901

Manufacturing enterprise collaboration network: An empirical research and evolutionary model*

Ji-Wei Hu(胡辑伟)[1,2], Song Gao(高松)[1,2], Jun-Wei Yan(严俊伟)[1,2], Ping Lou(娄平)[1,2,†], and Yong Yin(尹勇)[1]

[1] School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China
[2] Hubei Key Laboratory of Broadband Wireless Communication and Sensor, Wuhan University of Technology, Wuhan 430070, China

(Received 20 February 2020; revised manuscript received 22 May 2020; accepted manuscript online 27 May 2020)

With the increasingly fierce market competition, manufacturing enterprises have to continuously improve their competitiveness through their collaboration and labor division with each other, *i.e.* forming manufacturing enterprise collaborative network (MECN) through their collaboration and labor division is an effective guarantee for obtaining competitive advantages. To explore the topology and evolutionary process of MECN, in this paper we investigate an empirical MECN from the viewpoint of complex network theory, and construct an evolutionary model to reproduce the topological properties found in the empirical network. Firstly, large-size empirical data related to the automotive industry are collected to construct an MECN. Topological analysis indicates that the MECN is not a scale-free network, but a small-world network with disassortativity. Small-world property indicates that the enterprises can respond quickly to the market, but disassortativity shows the risk spreading is fast and the coordinated operation is difficult. Then, an evolutionary model based on fitness preferential attachment and entropy-TOPSIS is proposed to capture the features of MECN. Besides, the evolutionary model is compared with a degree-based model in which only node degree is taken into consideration. The simulation results show the proposed evolutionary model can reproduce a number of critical topological properties of empirical MECN, while the degree-based model does not, which validates the effectiveness of the proposed evolutionary model.

Keywords: manufacturing enterprise collaboration network, complex network, topological properties, fitness preferential attachment

PACS: 89.75.–k, 89.75.Fb, 05.65.+b, 05.90.+m **DOI:** 10.1088/1674-1056/ab96a8

1. Introduction

Nowadays, manufacturing enterprises are facing great challenges. Market demands are increasingly diversified and customized, and customers have growing requirements and expectations to products.[1] A single enterprise is difficult to respond promptly to the ever-changing market. Therefore, numerous enterprises collaborate to improve competitiveness so as to meet today's challenges. The collaboration among enterprises mainly refers to the supply-and-demand relation. A great number of enterprises tend to produce some specialized product parts, which brings the deepening of specialization and the division of labor. The collaboration and division of labor among manufacturing enterprises establish a manufacturing enterprise collaboration network (MECN). In MECN, only the core business of an enterprise is maintained whereas other businesses (*e.g.*, design, manufacture, *etc.*) are outsourced, thus gaining higher profits via its core competence as well as meeting the market requirements for lower cost and quick response.[2]

Recently, there has been considerable interest in the study of collaboration networks.[3,4] Scholars have proposed topological analysis methods to explore such networks.[5–8] For example, Ramasco *et al.*[5] studied the collaboration networks in terms of evolving and self-organizing bipartite graph models. Basole[8] examined the topological characteristics of interfirm collaboration networks in the global electronics industry. It was found the topological structure of networks plays an important role in learning network. As an effective approach to exploring the topology of complex systems, the complex network theory has received ongoing attention and it has been put into application.[9] The MECN is a typical collaboration network and it is getting more complicated with the enlargement of outsourcing. Some researches on MECN based on complex network theory have been proposed.[10–12] Specifically, enterprises in MECN were abstracted as nodes and inter-enterprise collaborations were abstracted as edges. With complex network theory approach, they modeled the MECN structures from the macroscopic view, analyzed the relations between their macrostructure and functional property, such as robustness, resistance to attack and risk spreading, and focused on the relations between two entities from the microscopic view.[13]

However, due to the difficulties in acquiring large-scale empirical data sets, most of these related studies are theoretical, and there is a lack of empirical studies to validate these theoretical studies. Additionally, some empirical studies on real-life networks have shown that many networks exhibit common topological properties.[14–18] For example, Gang *et al*

*Project supported by the National Natural Science Foundation of China (Grant Nos. 51475347 and 51875429).
†Corresponding author. E-mail: louping@whut.edu.cn

al.[14] concluded the urban supply chain network of agricultural products has scale-free and high disassortative properties. Liao *et al.*[17] found that the smartphone supply chain network had small-world feature with scale-free degree distribution. Scholars also have proposed a great many of network models to capture the properties of real networks, such as Watts–Strogatz model, Barabási-Albert model, *etc*.[19,20] Based on these network models, the evolutionary process of real networks is available. Therefore, to explore the topology and evolutionary process of MECN, it is necessary to investigate an empirical MECN and construct an evolutionary model to reproduce the topological properties found in the empirical network.

In reality, the Barabási-Albert model has limitations in explaining late-comers acquiring links relatively quickly. The fundamental reason is that the model only takes the node degree as the only metric to drive network evolution, without considering the intrinsic properties of the nodes. In network science this property is called the fitness of the node.[21] The concept of node fitness can be thought of as the amalgamation of all the attributes of a given node that contribute to its propensity to attract links.[22] It has been widely accepted that the fitness-based attachment models (such as the Bianconi–Barabási model, *etc*.) are more realistic than the classical Barabási-Albert model in capturing the growth process of real world networks.[23–28] In the MECN, manufacturing enterprises have various intrinsic attributes, such as geographic location, scale, goodwill, *etc*. The amalgamation of these empirical attributes can be defined as fitness. It plays a very active role in attracting collaboration when constructing the evolutionary model. Additionally, fitness in most network models is randomly allocated from a specified probability distribution, such as exponential and log-normal distribution.[23,29] Therefore, it is more reasonable to obtain the fitness based on the collected empirical attributes and allocate the fitness to nodes.

From the above, in this paper we empirically investigate MECN to explore the topology and construct an evolutionary model to depict the evolutionary process of MECN. Firstly, a great many of empirical data are collected to construct an MECN. From the viewpoint of complex network theory, the topological properties of MECN are analyzed. Then, the evolutionary process of MECN is investigated. An evolutionary model based on fitness preferential attachment and entropy technique for order preference by similarity to an ideal solution (entropy-TOPSIS) is proposed. Finally, simulation results validate the effectiveness of the proposed network model.

The remainder of this paper is organized as follows. In Section 2 an empirical investigation of MECN is carried out. In Section 3, the MECN evolutionary model based on fitness preferential attachment and entropy-TOPSIS is proposed. In Section 4, the simulation experiment and model analysis are described. Finally, some conclusions are drawn from the present study in Section 5.

2. Empirical investigation of MECN

Automobile is a complex product composed of exceeding thousands of parts. Correspondingly, numerous manufacturing enterprises are involved to form a huge automotive MECN. In this section, an empirical MECN is investigated. First, the data collection is described. Then, the empirical MECN is constructed. Finally, the topological properties of MECN are analyzed from the viewpoint of complex network theory.

2.1. Data collection

We use the online database operated by the Gasgoo Automotive Community as our main source (http://i.gasgoo.com/). This database is an authoritative automotive industry chain supply and demand platform in China. It provides information about numerous firms in Chinese automotive industry, and holds information about various attributes of each firm, such as name, geographical location, product list, supporting client firms, *etc*. The database allows users to search for firms by name, products, and clients.

The data acquisition was conducted during March–May 2018. First, original equipment manufacturer (OEM) information for each car brand was collected from Sohu News. Then, using the OEM named input item in the Gasgoo Automotive Community, related part supplier information was collected. Specific enterprise information includes name, found year, location, registered capital and supporting manufacturers. Finally, due to the incompleteness of some data, enterprise attribute data were supplemented from https://xin.baidu.com/. Additionally, the number of intellectual properties and the number of risk warnings were collected for enterprises. By following this procedure, information about 245 OEMs and 6790 part suppliers were acquired. The number of connections between them was 20938. The initial part supplier information (a portion) is shown in Table 1. The detailed information is illustrated below:

(i) No: Each manufacturing enterprise is represented by a unique number started from 1.

(ii) Found year: Each manufacturing enterprise has its own found year. That is an integer value.

(iii) Location: The province in which each manufacturing enterprise is located. The average distance from the enterprise's province to each province is calculated and used.

(iv) Registered capital: It represents the registered capital of each manufacturing enterprise. The data of different units are uniformly converted into CNY unit.

(v) Number of intellectual properties: It is the total number of patents, copyrights, trademarks, *etc*. That is an integer value.

(vi) Number of risk warnings: It is the total number of administrative penalties, judgment documents, stock freeze, *etc*. That is an integer value.

Table 1. Information about part suppliers (a portion).

No	Part suppliers	Found year	Location	Registered capital	Number of intellectual properties	Number of risk warnings
1	Guangzhou Qicheng Auto Parts Co., Ltd.	2013	Guangdong	2.0×10^6 CNY	1	6
2	Shenyang Pinghe Valeo Automotive Transmission System Co., Ltd.	2013	Liaoning	3.5×10^7 USD	16	4
3	Shanxi Zhongjin Industrial Machinery Co., Ltd.	1996	Shanxi	8.0×10^6 CNY	29	1
4	Conde Ryan electromagnetic technology (China) co., Ltd.	2005	Jiangsu	9.8×10^6 EUR	23	0
5	Shanghai saks powertrain components system co., Ltd.	2001	Shanghai	1.4×10^7 USD	89	4

2.2. Construction of MECN

The empirical MECN is composed of two sets of nodes (OEMs and part suppliers), and links between these two kinds of nodes (but no links between the same kind of nodes). Such a network is called a bipartite network. A bipartite network structure is shown in Fig. 1. Generally, MECN can be denoted as $G=\{M,S,E\}$, where $M=\{m_1,m_2,\ldots,m_i,\ldots,m_{n_M}\}$ and $S=\{s_1,s_2,\ldots,s_j,\ldots,s_{n_S}\}$ are the disjoint sets of the OEM nodes and part supplier nodes respectively, n_M and n_S are the number of OEM nodes and part supplier nodes respectively, and E is the set of edges connecting OEM nodes and part supplier nodes. The OEM nodes only connect to part supplier nodes and *vice versa*. Mathematical analysis of a network requires it to be represented through an adjacency matrix. The MECN can be described using an adjacent matrix $R=\{a_{ij}\}_{n_M\times n_S}$, where a_{ij} is the number of connections between OEM nodes and part supplier nodes,

$$a_{ij}=\begin{cases}1, & \text{if there is a connection between } m_i \text{ and } s_j,\ i\in n_M \text{ and } j\in n_S,\\ 0, & \text{otherwise.}\end{cases} \tag{1}$$

From the collected dataset, MECN is empirically constructed (Fig. 2). The number of OEM nodes and the number of part supplier nodes are 245 and 6790, respectively. The number of edges is 20938.

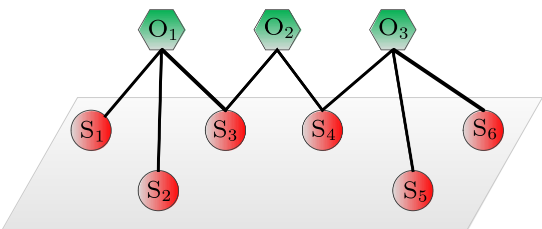

Fig. 1. Bipartite network structure.

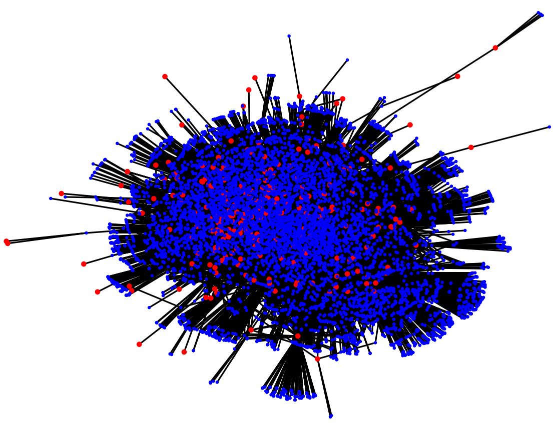

Fig. 2. MECN visualization. Red nodes are OEMs and blue nodes are part suppliers.

2.3. Topological properties of MECN

Real world complex networks exhibit various interesting topological properties. These sets of topological measurements provide a meaningful explanation for a network's dynamical properties. In this subsection, a few important topological properties are considered, viz. average path length, clustering coefficient, degree distribution, and assortativity. The topological properties of empirical MECN are shown in Table 2. Then, each topological property is introduced in detail.

Table 2. Topological properties of empirical MECN.

Topological properties	Result
OEMs	245
Part suppliers	6790
Edges	20938
Part suppliers per OEM	85.46
OEMs per part suppliers	3.08
Average path length L	3.755
Clustering coefficient c	0.253
Assortativity coefficient r	−0.480
(one-mode projection on OEMs) Assortativity coefficient r_M	−0.231
(one-mode projection on part suppliers) Assortativity coefficient r_S	−0.021
OEM degree	Max: 963 and Min: 1
Part supplier degree	Max: 23 and Min: 1

2.3.1. Average path length

The average path length presents an approach to characterizing the spread of a network by calculating the average distance between any pair of nodes.[30] The distance d_{ij} between node i and node j in the network is defined as the number of edges on the shortest path connecting the two nodes. The average path length L is defined as the average of distance between all pairs of nodes in the network,

$$L=\frac{2}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}d_{ij}, \tag{2}$$

where N represents the number of nodes in the network.

It is found that although the number of nodes in many real networks is huge, the average path length of the network is quite small. That is, the average path length will increase to some extent as the network size increases. From Table 2, the average path length L is 3.78. It is quite small relative to the scale of the network.

2.3.2. Clustering coefficient

For one-mode networks, the definition of clustering coefficient c is the ratio between the number of triangles observed in one network and the total number of possible triangles which may appear. For node i with k_i neighbors the total number of possible triangles is just the number of pairs of neighbors given by $k_i(k_i-1)/2$. The clustering coefficient c for node i is

$$c=\frac{2t_i}{k_i(k_i-1)}, \tag{3}$$

where t_i is the number of triangles. While there are no triangles in bipartite networks since no connections are allowed between nodes of the same type. Some scholars transform bipartite networks into one-mode networks in order to be able to analyze them. However, it generally loses some information about the original bipartite network, and brings an inflation of the number of edges and other drawbacks caused by projection. Hence, some definitions of the clustering coefficient in the original bipartite network have been proposed. The measuring of the overlap between neighborhoods of pairs of nodes has been proposed in Ref. [31]. And the bipartite clustering coefficient of node u is defined as

$$c_u=\frac{\sum_{v\in N(N(u))}c_{uv}}{|N(N(u))|}, \tag{4}$$

where $N(u)$ represents the set of neighbors of u, $N(N(u))$ is the second order neighbors of u in G excluding u, and c_{uv} is the pairwise clustering coefficient between node u and node v and c_{uv} is defined as

$$c_{uv}=\frac{|N(u)\cap N(v)|}{|N(u)\cup N(v)|}. \tag{5}$$

The clustering coefficient for the whole graph is the average as given below.

$$c=\frac{1}{n}\sum_{u\in G}c_u, \tag{6}$$

where n is the number of nodes in G.

Clustering coefficient is used to describe the proportion that the neighbors of a certain node are also neighbors. The clustering coefficient of the network reflects the cliquishness of the mean closest neighborhood of the network. From Table 2, clustering coefficient c is 0.2526.

2.3.3. Degree and degree distribution

Degree is not only the simplest and most intuitive attribute of network nodes, but also an important statistical parameter. The degree of a node is defined as the number of nodes connected to it. Therefore, intuitively speaking, the greater the degree of a node, the more the nodes are associated with it, and the greater the role of its nodes in the network. Degree distribution is defined as the random selection of a point with a probability of k edges. Many empirical studies show that the degree distribution of a large number of actual networks conforms to a power law distribution, and the degree distribution of one or two kinds of nodes of bipartite networks also conforms to a power law distribution. The network with power law distribution of degree distribution is called scale-free network. Researchers often plot the degree distributions of complex networks on a double logarithmic scale and look for the evidence of a linear curve by using the log-transformed data.[32]

From Table 2, the average degree of OEMs is 85.46 and the average degree of part suppliers is 3.08. There is a significant gap between the average degree of OEMs and part supplier. This also reflects a larger number of part suppliers than OEMs'. Figure 3 shows the complementary cumulative distributions of OEM degree and part supplier degree. It is clear even from a superficial visual inspection that the degree distributions on a log–log scale do not look linear, and MECN is therefore not scale-free. Maximum-likelihood method (MLE) can be used to fit a range of possible heavy-tailed distributions: power law, truncated power law, exponential and stretched exponential. The MLE fit is used from the *powerlaw* package.[33] The fit curves of each distribution are also shown in Fig. 3. The fitting results are shown in Table 3. The Kolmogorov–Smirnov distance between the data and the fit is denoted as D. The smaller the value of D, the better the fitting is.

Actually, most enterprise nodes have lower degrees and only a few enterprise nodes have higher degrees. The finding is consistent with the current automotive industry market. Most of part suppliers are small-sized enterprises, and only can collaborate with several OEMs. Few firms with high degree are considered to be hub firms, and they are essential to ensure the whole network remains functional. If they encounter failures, the stability of the network will be greatly reduced. While some low degree nodes' failures may have little influence on the whole network performance. It is suggested that managers should pay more attention to protecting the enterprises with more collaborators.

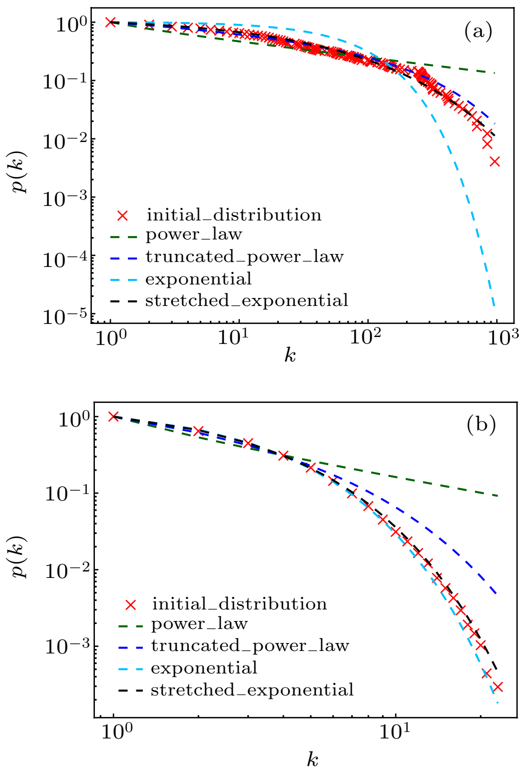

Fig. 3. MLE fit of degree distribution of (a) OEMs and (b) part suppliers.

Table 3. Degree distribution fitting results of empirical MECN.

Distribution	$p(k)$	D (OEM)	D (part supplier)
Power law	$k^{-\alpha}$	0.2107	0.1309
Truncated power law	$k^{-\alpha}\,\mathrm{e}^{-\lambda k}$	0.0545	0.0372
Exponential	$\mathrm{e}^{-\lambda k}$	0.3260	0.0301
Stretched exponential	$(\lambda k)^{\beta-1}\,\mathrm{e}^{-(\lambda k)^{\beta}}$	0.0649	0.0203

From Table 3, the degree distribution of OEMs conforms to truncated power law with $\alpha = 1.0000$ and $\lambda = 0.0015$. The degree distribution of part suppliers conforms to stretched exponential with $\lambda = 0.4260$ and $\beta = 0.9185$. The MECN is not a scale-free network.

2.3.4. Assortativity

Assortativity is defined as the tendency for nodes in the network to form connections preferentially to others similar to them. Generally, assortativity is important, for something that affects a single high-degree node can quickly cascade to other high-degree nodes.[34] A network is said to be assortative when high degree nodes are, on average, connected to other nodes with high degree and low degree nodes are, on average, connected to other nodes with low degree. On the contrary, it is said to be disassortative. This mechanism has been proposed as the key ingredient for the formation of communities in networks. To characterize assortativity, the behavior of the average nearest neighbor's degree of the firms of degree k is studied,

$$k_{nn}(k) \equiv \sum_{k'} k' P\left(k' | k\right), \tag{7}$$

where $P\left(k' | k\right)$ is the conditional probability with which a firm of degree k is connected to a firm of degree k'. Besides, assortativity can be simply measured by the assortative coefficient, which is defined as the Pearson's correlation coefficient of degrees at either side of an edge.[35] The assortative coefficient r can be express as

$$r = \frac{M^{-1}\sum_i j_i k_i - \left[M^{-1}\frac{1}{2}\sum_i (j_i + k_i)^2\right]}{M^{-1}\sum_i \frac{1}{2}(j_i^2 + k_i^2) - \left[M^{-1}\frac{1}{2}\sum_i (j_i + k_i)\right]^2}, \tag{8}$$

where j_i and k_i are the degrees of the vertices at the ends of the i-th edge, with $i = 1, \ldots, M$, and M is the total number of edges in the network. If k_{nn} increases with k or $r > 0$, the network is assortative. If k_{nn} decreases with k or $r > 0$, the network is disassortative. If $r = 0$, the network is non-assortative.

Indeed, the assortativity coefficient $r = -0.4803$ of the original MECN is calculated, which indicates the empirical MECN is disassortative. The reason for this result is that the MECN is a bipartite network and the number of OEMs is much smaller than the number of suppliers limited by the actual situation. In order to better understand its assortativity, the MECN bipartite network is transformed into one-mode network. In the previous section, MECN is denoted as $G = \{M, S, E\}$, where $M = \{m_1, m_2, \ldots, m_i, \ldots, m_{n_M}\}$ and $S = \left\{s_1, s_2, \ldots, s_j, \ldots, s_{n_S}\right\}$ are disjoint sets of the OEM nodes and part supplier nodes respectively, n_M and n_S are the number of OEM nodes and part supplier nodes respectively, and E is the set of edges connecting OEM nodes and part supplier nodes. Specifically, a projection onto the OEM nodes M results in a one-mode network where node m is connected to m', m, $m' \in M$, only if there exists a pair of edges (m, s) and (m', s) in E such that m and m' share a common neighbor $s \in S$, in the bipartite MECN. Similarly, in a projection onto the part supplier nodes S, a node s is connected to a node s' in the projection if they share a neighbor $m \in M$. Even though a pair of OEM nodes m and m' can share many common neighbors from the part supplier set S in G, there will be only a single link connecting such nodes in the projected version. Finally, an unweighted one-mode network with no multiple edges between node pairs and no self-loops is created.[36,37] Figure 4 shows the change of k_{nn} as k grows in the one-mode networks.

In Fig. 4(a), it can clearly be seen that k_{nn} decreases as k grows in the OEM one-mode network. But the change of k_{nn} as k grows cannot be judged in Fig. 4(b). Actually, the assortativity coefficient of the one-mode projection network on OEM is $r_M = -0.231$, the assortativity coefficient of the one-mode projection network on part suppliers is $r_S = -0.021$. It is can be seen that both of them are disassortative. In short, empirical MECN is a disassortative network. Enterprises in MECN with high link number can lead their communities in production. They tend to connect to low-degree enterprises in order to expand the scale of production. And the disassortative property indicates that the risk spreading is fast and the coordinated operation is difficult in MECN.

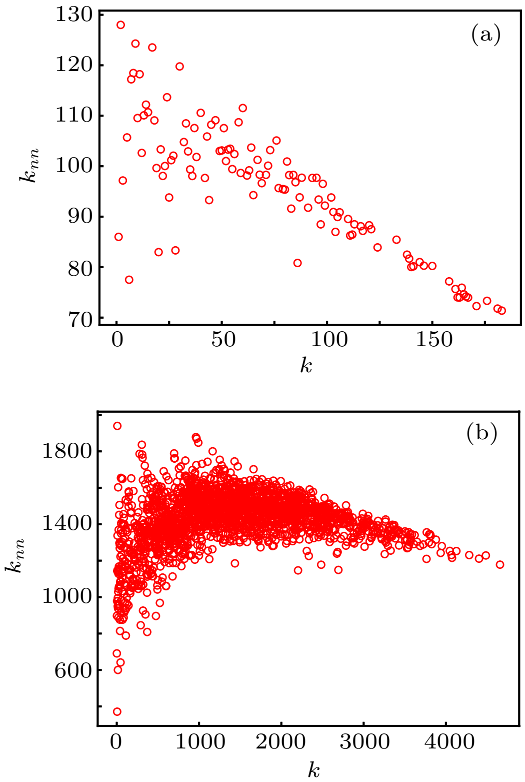

Fig. 4. Degree correlation in one-mode projection network on (a) OEMs and (b) part suppliers.

Based on these findings, the MECN is a complex network. It has small average path length and large clustering coefficient, which indicates small world property. The MECN is neither a scale-free network nor a random network. What is more, the network is disassortative.

3. MECN evolutionary model

A great number of empirical studies of networks have shown that many networks exhibit common topological properties. Actually, the MECN is a complex network with small-world and disassortative properties. In order to explore the evolutionary process of MECN and learn how these topological properties arise, the MECN evolutionary model is proposed.

Generally, the preferential attachment is an essential mechanism to construct the evolutionary model. The most well-known element is the degree preferential attachment in the Barabási-Albert model, which stipulates that the probability of a new node making a link with an existing node is proportional to the number of links (degree) of the existing node. That is, the probability p_i with which a new node makes a connection to an existing node i with degree k_i is given by

$$p_i = \frac{k_i}{\sum_{j\in N} k_j}, \tag{9}$$

where N is the set of nodes to which the new node can connect. This simple degree-based attachment is a rich–get–richer mechanism, where nodes with already high degree are more likely to acquire more links. However, the Barabási-Albert model also has some limitations introduced in Section 1. Correspondingly, the fitness preferential attachment mechanism taking the intrinsic property into account (also including the node degree) has proved more reasonable. A prominent example of fitness-based models is the Bianconi–Barabási model. The fitness preferential attachment mechanism is

$$p_i = \frac{k_i f_i}{\sum_{j\in N} k_j f_i}, \tag{10}$$

where k_i represents the degree of node i and f_i represents its fitness. In contrast to the Barabási–Albert model, it is possible in the Bianconi–Barabási model for a relative newcomer to overtake an older node in terms of the number of links.

In MECN, enterprises have various intrinsic attributes, such as geographic location, scale, goodwill, *etc*. These intrinsic attributes play a critical role in attracting collaboration. That is, enterprises associated with greater economic benefit, larger scale and more numerous collaborations, are thus naturally more attractive to others.[38] The amalgamation of these attributes can be defined as fitness. Therefore, constructing an evolutionary model in which enterprise node fitness is taken into consideration is more persuasive. Additionally, considering the fact that the fitness in most of network models is randomly allocated from a specified probability distribution, we introduce the entropy-TOPSIS method to calculate fitness based on the collected empirical attributes and allocate them to nodes. In this section, firstly, the MECN evolutionary model based on fitness preferential attachment and entropy-TOPSIS is proposed. Then, the entropy-TOPSIS method is introduced into calculate fitness.

3.1. Model description

The MECN evolutionary model is denoted as $G = (M, S, E)$, where G is the MECN evolutionary model, $M = \{m_1, m_2, \ldots, m_i, \ldots, m_{n_M}\}$ and $S = \{s_1, s_2, \ldots, s_j, \ldots, s_{n_S}\}$ are the OEM node set and the part supplier node set respectively, and is the set of edges connecting OEM nodes and part supplier nodes. Generally, in the early stages, the network contains relatively few enterprises. Then, owing to the entry of enterprises and the establishment and ending of cooperative relationship, the network evolves with time. Correspondingly, there are the growth of nodes and the link construction and link deletion in the evolutionary model. Besides, considering enterprises have various intrinsic attributes which conduce to attracting collaboration, the node fitness is introduced into the evolutionary model.

Therefore, an evolutionary model that integrates together the node growth, fitness preferential attachment, link establishment, and link deletion is proposed. Before going further, we give notations used in the model in Table 4. The process of MECN evolutionary model construction is shown in Fig. 5, and its specific steps are explained below.

Table 4. Notations used in evolutionary model.

Notation	Meaning
t	time step of evolution
u_0, v_0, e_0	initial number of OEM nodes, part supplier nodes, and edges respectively
e	number of edges
n_M, n_S, n_E	maximum number of OEM nodes, part supplier nodes, and edges respectively
k_i, k_j	degree of nodes i and j respectively
f_i, f_j	fitness of nodes i and j calculated by the entropy-TOPSIS method respectively
$d_{\max\_s}$, $d_{\max\_m}$	maximum number of part supplier nodes that OEM nodes can connect with maximum number of OEM nodes that part supplier nodes can connect with
S_{ij}	probability of successful connection between nodes i and j
q	parameter that affects where the link is deleted (from local-world or global network)
Q_i	probability of being selected as local-world network for node i

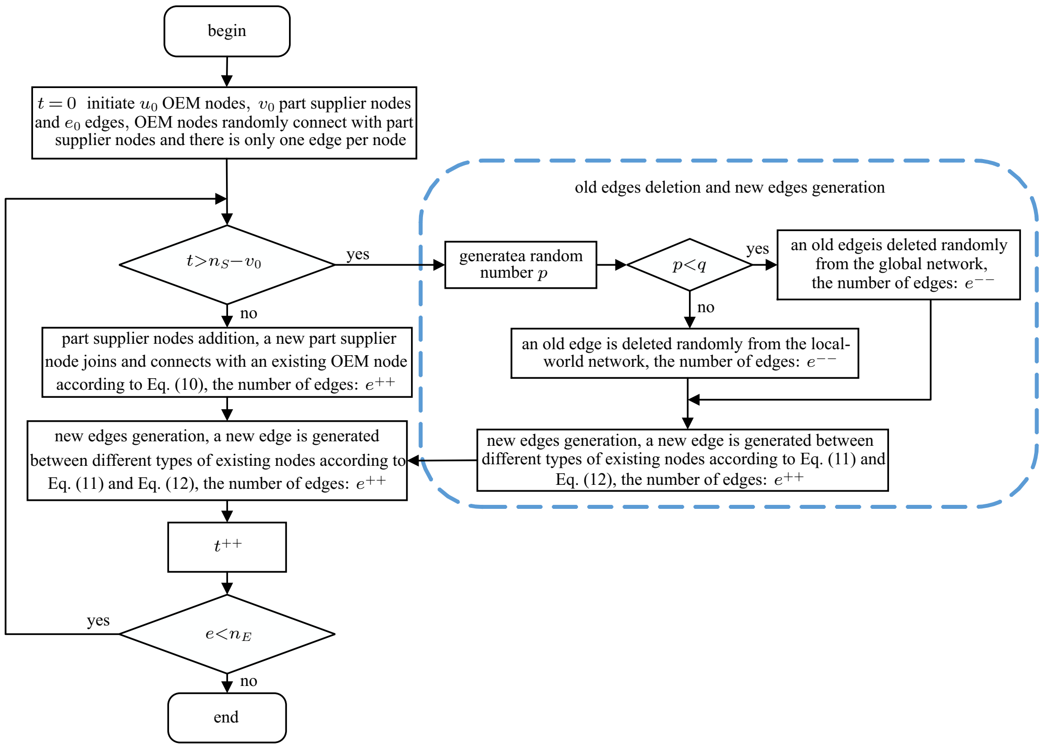

Fig. 5. MECN evolutionary model construction process.

(I) Initialization: At $t = 0$, there are u_0 OEM nodes,v_0 part supplier nodes and e_0 edges between them $(u_0 = v_0 = e_0)$. The OEM nodes randomly connect with part supplier nodes and there is only one edge per node.

(II) Part supplier nodes addition: At the t-th time step $(t \leqslant n_s - v_0)$, a new part supplier node joins and connects with an OEM node. The OEM node is selected according to the fitness preferential attachment defined by Eq. (10). The probability S_{ij} of successful connection between the new part supplier node j and the existing OEM node i is

$$S_{ij} = 1 - \frac{k_j}{d_{\max\_n}}. \tag{11}$$

If the connection fails, an OEM node is reselected and connected.

(III) Generation of new edges: At each time step, a new edge is constructed between different types of nodes. The selection of the two nodes is based on fitness preferential attachment defined by Eq. (10) and the probability of successful connection given by Eq. (11). If the connection already exists or fails, the two types of nodes are reselected and connected.

(IV) Detection of old edges and eneration of new edges: At the t-th time step $(t > n_s - v_0)$, an existing edge is deleted randomly from the global network with probability q or from the local-world network with probability $1-q$. The local-world network includes an OEM node i and its partners, and the selected probability Q_i according to Eq. (12). Then, a new edge is constructed as indicated in step (III).

$$Q_i = \frac{d_{\max\_m} - k_i}{\sum_1^{u_0} (d_{\max\_m} - k_m)}. \tag{12}$$

After t $(t \leqslant n_s - v_0)$ time steps, there are n_M $(n_M = u_0)$ nodes in M, $v_0 + t$ nodes in S, and $e_0 + 2t$ edges in E. After t $(t > n_s - v_0)$ time steps, there are n_M nodes in M, n_s nodes in S, and $n_s + t$ edges in E. When the number of edges in the network reaches the target number n_E, the evolutionary model is constructed to end.

3.2. Entropy weight and TOPSIS

TOPSIS is a technique for solving multiple criteria decision making problems. Applying this method to enterprise evaluation can obtain the comprehensive evaluation value of each enterprise. However, the weight of each evaluation index is predetermined in the calculation, which is strong subjective and seriously affects the evaluation results. Entropy is a measure that uses probability theory to measure the uncertainty of information.[39] The entropy weight method can be adopted to determine the weight to avoid the effect of subjective factors. The entropy-TOPSIS method have been applied to the comprehensive assessment of empirical study.[30,41]

In the MECN evolutionary model, each enterprise has an intrinsic fitness to attract links. Considering fitness in most models is randomly allocated from a specified probability distribution, and obtaining fitness based on empirical enterprise attributes is more reasonable. Therefore, the entropy-TOPSIS method is introduced to calculate the fitness. First, the enterprise attributes are regarded as evaluation indexes. Then, the corresponding weight of each evaluation index is obtained by the entropy weight method. Finally, the TOPSIS method is used to obtain the comprehensive evaluation value of enterprises. The evaluation value can be defined as enterprise fitness. And the final calculated fitness can be randomly allocated to nodes in the evolutionary model. The calculation steps of this method are described as follows.

Step 1 Identify a decision matrix $X = [x_{ij}]_{m\times n}$, $i \in (1,m)$, $j \in (1,n)$. Assuming that there are m enterprises and n enterprise attributes, the evaluation value of attribute j in enterprise i is x_{ij}. Initial matrix X is as follows:

$$\boldsymbol{X} = [x_{ij}]_{m\times n} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}. \tag{13}$$

Step 2 Construct the normalized matrix $X' = \left[x'_{ij}\right]_{m\times n}$. There are two types of standardized methods. When the attributes are of positive-type, the calculation for normalization can be expressed as Eq. (14). When the attributes are of negative-type, the calculation for normalization can be expressed as Eq. (15),

$$x'_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}}, \tag{14}$$

$$x'_{ij} = \frac{x_{\max} - x_{ij}}{x_{\max} - x_{\min}}. \tag{15}$$

Step 3 Translate matrix and obtain matrix $H = [h_{ij}]_{m\times n}$. The calculation of the entropy method involves the calculation of the logarithm and the value cannot be 0, therefore, the value of all attributes is shifted rightwards by 1 unit,

$$h_{ij} = x'_{ij} + 1. \tag{16}$$

Step 4 Construct the standardized matrix $H' = \left[h'_{ij}\right]_{m\times n}$. The formula used is as follows:

$$h'_{ij} = \frac{h_{ij}}{\sum_{i=1}^{m} h_{ij}}. \tag{17}$$

Step 5 Calculate the information entropy of each attribute. Let the entropy value of the j-th attribute be e_j, the matrix of each attribute entropy is $\boldsymbol{E} = [e_j]_{1\times n}$, where

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^{m} h'_{ij} \ln h'_{ij}. \tag{18}$$

Step 6 Calculate the entropy weight. Each attribute entropy weight can be obtained and composed into a weight vector $\boldsymbol{W} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}$ and the weight of the j-th attribute is expressed as w_j,

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{n} (1 - e_j)}. \tag{19}$$

Step 7 Calculate the weighted normalized matrix $\boldsymbol{V} = [v_{ij}]_{m\times n}$,

$$v_{ij} = h_{ij} \times w_{ij}. \tag{20}$$

Step 8 Determine the positive ideal solution A^+ and the negative ideal solution A^-,

$$A^+ = \left\{A_1^+, A_2^+, \dots, A_j^+, \dots, A_n^+\right\}, \tag{21}$$

$$A^- = \left\{A_1^-, A_2^-, \ldots, A_j^-, \ldots, A_n^-\right\}, \tag{22}$$

where $A_j^+ = \left\{\max\left(v_{1j}, v_{2j}, \ldots, v_{nj}\right) | j \in (1, n)\right\}$ and $A_j^- = \left\{\min\left(v_{1j}, v_{2j}, \ldots, v_{nj}\right) | j \in (1, n)\right\}$.

Step 9 Calculate the relative distance for enterprise nodes from the positive ideal solution and negative ideal solution,

$$d_i^+ = \left\{\sum_{j=1}^{n}\left(v_{ij} - A_j^+\right)^2\right\}^{1/2}, \tag{23}$$

$$d_i^- = \left\{\sum_{j=1}^{n}\left(v_{ij} - A_j^-\right)^2\right\}^{1/2}. \tag{24}$$

Step 10 Calculate the relative closeness D_i of each alternative from the following equation:

$$D_i = \frac{d_i^-}{d_i^+ + d_i^-}. \tag{25}$$

Actually, D_i is the comprehensive evaluation score of enterprise node i. The higher the value of D_i, the better the evaluation, the stronger competitiveness is, and the greater the probability with which the collaborations are obtained will be. Therefore, the static value D_i can be regard as enterprise fitness, $f_i = D_i$.

4. Simulation results

The evolutionary model can be evaluated by asking whether it reproduces certain topological properties of the empirical network. Besides, in order to prove the effectiveness of the proposed evolutionary model more persuasive, a comparative experiment is conducted. In the comparative experiment, the fitness preferential attachment is replaced by the degree preferential attachment defined by Eq. (9), whereas others remain the same. The model for comparison is called degree-based model. In this section, first of all, the parameters of the model are determined according to the empirical network and multiple experimental results. Then, the topological properties of the models are introduced and the effectiveness of the evolutionary model is proved.

4.1. Parameter setting

Most of the parameter settings of simulation experiments are set according to our collected empirical database. n_M, n_S, and n_E representing the total number of OEMs, part suppliers, and connections between them are 245, 6790, and 20938, respectively. Considering the small portion of OEMs, the initial number of these parameters are $u_0 = v_0 = e_0 = 245$. Besides, considering the maximum degree of two types of nodes in the empirical network, similar parameters are set to be $d_{\max\_m} = 950$ and $d_{\max\_n} = 30$. Enterprise node fitness is calculated by the entropy-TOPSIS method based on collected empirical enterprise attributes. And the calculated fitness can be randomly allocated to nodes in the model.

According to our evolutionary rules, the number of final edges is $n_S + t = 20938$, the evolutionary time step is calculated to be $t = 14148$. Finally, the parameter q affecting the source of the link deletion (the global or local-world network) is obtained from multiple simulation experiments. Actually, we find that q has a great influence on the degree distribution of OEMs, but little effect on the degree distribution of part suppliers. When q changes from 0 to 1, the degree distribution of OEMs transforms from the stretched exponential to the truncated power law to power law. While the degree distribution of part suppliers is always stretched exponential. According to our results, the value of q is set to be 0.6.

Besides, for the degree-based model in which only node degree is considered, the evolutionary process of the degree-based model is the same as that of the proposed evolutionary model and the parameters also remain the same.

4.2. Model analysis

To validate the effectiveness of the proposed model, a number of topological properties are considered, *e.g.*, average path length, clustering coefficient, degree distribution, and assortativity coefficient (both in bipartite network and one-mode projection network). In addition, a comparative experiment is conducted in which the fitness preferential attachment is replaced by the degree preferential attachment and others remain the same.

According to the model construction process in the previous section, the MECN evolutionary model is constructed (Fig. 6). Besides, the topological properties are explored. The results of average basic topological properties are shown in Table 5. From the table, both the evolutionary model and the degree-based model exhibit small average path length, high average clustering coefficient, and negative assortativity coefficient. These properties match well to those of the empirical network. However, comparing the results of the two models in detail, most of the topological properties are close except the OEM degree, which indicates that the degree distribution of OEMs is obviously different.

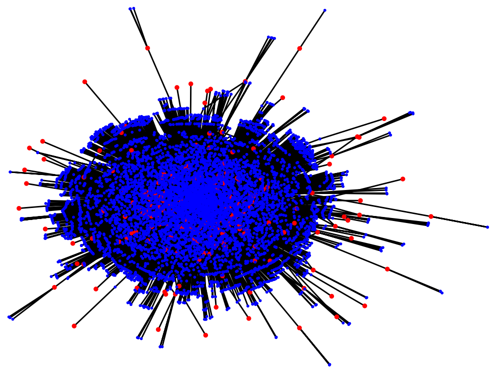

Fig. 6. MECN evolutionary model visualization. Red nodes are OEMs and blue nodes are part suppliers.

Table 5. Results of topological properties.

Topological properties	Empirical network	Evolutionary model	Degree-based model
OEMs	245	245	245
Part suppliers	6790	6790	6790
Edges	20938	20938	20938
Part suppliers per OEM	85.46	85.46	85.46
OEMs per part supplier	3.08	3.08	3.08
Average path length L	3.755	3.760 ± 0.015	3.854 ± 0.015
Clustering coefficient c	0.253	0.224 ± 0.002	0.219 ± 0.02
Assortativity coefficient r	-0.480	-0.417 ± 0.030	-0.534 ± 0.040
Assortativity coefficient r_M (one-mode projection on OEMs)	-0.231	-0.265 ± 0.020	-0.203 ± 0.020
Assortativity coefficient r_S (one-mode projection on part suppliers)	-0.021	-0.044 ± 0.015	-0.034 ± 0.015
OEM degree	Max: 963 Min: 1	Max: 949 ± 1 Min: 1	Max: 600 ± 150 Min: 1
Part supplier degree	Max: 23 Min: 1	Max: 22 ± 2 Min: 1	Max: 21 ± 2 Min: 1

Then, root-mean-square deviation (RMSD) is used to evaluate the performance of the model.[42] The RMSD is frequently used to measure the differences between values generated by a model and the values actually observed. Smaller values indicate better model predictive ability. The RMSD is described as follows:

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}, \tag{26}$$

where $\hat{y}$ is the generated value, y is the empirical value, $n (= 20)$ is the number of simulation experiments. The RMSD results are shown in Table 6. It can be concluded that the fittings of these average topological properties are very well.

Table 6. RMSD results of topological properties.

Topological properties	Evolutionary model	Degree-based model
Average path length L	0.021	0.099
Clustering coefficient c	0.029	0.034
Average degree $\bar{k}$	0	0
Assortativity coefficient r	0.066	0.073
Assortativity coefficient r_M (one-mode projection on OEMs)	0.036	0.032
Assortativity coefficient r_S (one-mode projection on part suppliers)	0.036	0.032

The average statistics of the models can typically be tuned by varying parameter values, however, the shapes of the distributions are likely to be invariable.[43] Here, the degree distribution of OEMs and the degree distribution of part suppliers of the evolutionary model, and degree-based model are investigated. Figure 7 shows the degree distribution of the empirical network, evolutionary model, and degree-based model. Actually, the degree distribution of OEMs in the evolutionary model conforms to truncated power law with $\alpha = 1.0000$ and $\lambda = 0.0018$, whereas in the degree-based model it conforms to stretched exponential. The degree distribution of OEMs in the evolutionary model is consistent with the empirical distribution, but not in the degree-based model.

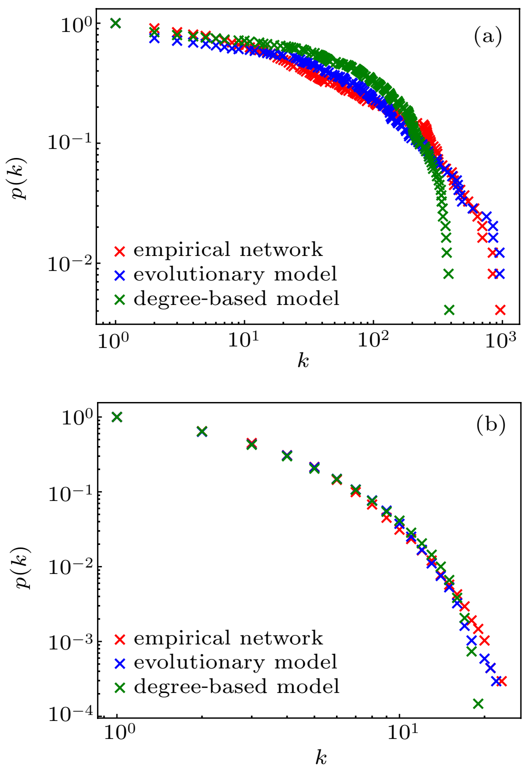

Fig. 7. Distributions of the evolutionary model (blue), compared with empirical network (red) and the degree-based model (green), showing (a) the degree distribution of OEMs, and (b) degree distribution of part suppliers.

Besides, the degree distribution of part suppliers in the evolutionary model and degree-based model have similar shapes to empirical distribution, all of them conform to the stretched exponential. In the evolutionary model, $\lambda = 0.4528 \pm 0.02$ and $\beta = 0.8706 \pm 0.03$, and in the degree-based model, $\lambda = 0.4534 \pm 0.015$ and $\beta = 0.8697 \pm 0.02$. The parameters of degree distribution obtained from the simulation experiments are slightly different. It could be an issue caused by either the rules of the model or the noises in the empirical datasets.

Overall, the proposed evolutionary model shows a good fitting to empirical network. It reproduces realistic patterns ranging from basic statistics to critical distributions. The proposed evolutionary model is effective.

5. Conclusions

In this paper, an MECN is built by using empirical data. Its topology is explored in terms of complex network theory, and its evolutionary process is studied by constructing an evolutionary model based on fitness preferential attachment and entropy-TOPSIS. Our results show that the empirical MECN is a complex network characterized by small-world and disassortative properties. Generally, the topological properties of MECN critically affect its functional properties, such as robustness, resistance to attacks, and risk spreading. The small world property indicates that the enterprises in the MECN have close relationship and high communicative efficiency among enterprises. Enterprises can respond quickly to the market. The disassortative property indicates that the enterprises with high numbers of links tend to connect the low-degree enterprises in order to expand the scale of production. However, both small-world and disassortative properties indicate that the risk spreading is fast in the MECN. Additionally, the network is generally resilient against cascading influence arising from targeted attacks, because the hub nodes are not connected with each other. Moreover, an evolutionary model based on fitness preferential attachment and entropy-TOPSIS, is proposed to depict the evolutionary process of MECN. In particular, the node fitness in the evolutionary model is randomly allocated from empirical enterprise fitness, which is calculated by the entropy-TOPSIS method based on collected enterprise attributes. It is more persuasive than being randomly allocated from a specified probability distribution. Besides, a degree-based model in which only node degree is taken into consideration, is constructed for a comparison. The simulation results reveal that the proposed evolutionary model reproduces certain topological properties of the real MECN. The effectiveness of the evolutionary model is also verified.

Our results are conducible to researchers better understanding the MECN and provide researchers with a foundation for controlling. For future studies, the attributes of enterprises we considered above are not comprehensive, and more attributes should be collected. Moreover, with the enlargement of outsourcing, the collaboration among enterprises is getting closer. Once any risk happens to the enterprises, they might rapidly spread the risk to the associated enterprises, causing the majority of enterprises to fail to run and even paralyze the whole MECN, and thus affecting the normal operation of social economy. According to the results in this paper, we will analyze and control the risk spreading in MECN, which are our future work.

Acknowledgment

The authors would like to thank all collaborators for their contribution to this project.

References

[1] Peng C and Meng Y 2016 *Rob. Comput.-Integr. Manuf.* **42** 49
[2] Parker H 2000 *Ind. Manage. & Data Syst.* **100** 255
[3] Zhang P P, Chen K, He Y, Zhou T, Su B B, Jin Y, Chang H, Zhou Y P, Sun L C, Wang B H and He D R 2006 *Physica A* **360** 599
[4] Appio F P, Martini A, Massa S and Testa S 2017 *Int. J. Production Res.* **55** 2121
[5] Ramasco J J, Dorogovtsev S N and Pastor-Satorras R 2004 *Phys. Rev. E* **70** 036106
[6] Chang H, Su B B, Zhou Y P and He D R 2007 *Physica A* **383** 687
[7] Peltomäki M, Alava M 2006 *J. Stat. Mech.: Theory & Experiment* **2006** 1010
[8] Basole R C 2016 *Decis. Support Syst.* **83** 22
[9] Choi T Y and Hong Y 2002 *J. Oper. Manage.* **20** 469
[10] Xuan Q, Du F, Li Y and Wu T J 2010 *IEEE Trans. Autom. Sci. Eng.* **8** 442
[11] Sun J Y, Tang J M, Fu W P and Wu B Y 2017 *Physica A* **473** 377
[12] Bellamy M A and Basole R C 2013 *Syst. Eng.* **16** 235
[13] Hearnshaw E J S and Wilson M M J 2013 *Int. J. Oper. & Production Manage.* **33** 442
[14] Gang Z, Ying-Bao Y, Xu B and Qi-Yuan P 2015 *Transp. Lett.* **7** 188
[15] Brintrup A, Wang Y and Tiwari A 2015 *IEEE Syst. J.* **11** 2170
[16] Feng S M, Hu B Y, Nie C, Shen X H and Ci Y S 2016 *Chin. Phys. B* **25** 030504
[17] Liao H, Shen J, Wu X T, Chen B K and Zhou M 2017 *Chin. Phys. B* **26** 110505
[18] Wu X Y, Gu R T, Pan Z Y, Jin W Q and Ji Y F 2015 *Chin. Phys. Lett.* **32** 068901
[19] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[20] Barabá A L and Albert R 1999 *Science* **286** 509
[21] Perera S, Bell M G H and Bliemer M C J 2017 *Appl. Network Science* **2** 33
[22] Bell M, Perera S, Piraveenan M, Piraveenan M, Bliemer M, Latty T and Reid C 2017 *Sci. Rep.* **7** 42431
[23] Bianconi G and Barabási A L 2001 *Europhys. Lett.* **54** 436
[24] Li G, Xuan Q and Song Z 2010 *IEEE International Conference on Intelligent Systems and Knowledge Engineering*, November 15–16, 2010, Hangzhou, China, p. 511
[25] Perera S, Perera H N and Kasthurirathna D 2017 *Moratuwa Engineering Research Conference*, May 29–31, 2017, Moratuwa, Sri Lanka, p. 159
[26] Aspembitova A, Feng L, Melnikov V and Chew L Y 2019 *PloS One* **14** 1
[27] Chattopadhyay S and Murthy C A 2017 *Inf. Sci.* **397** 219
[28] Cimini G, Squartini T, Musmeci N, Puliga M, Gabrielli A, Garlaschelli D, Battiston S and Caldarelli G 2014 *International Conference on Social Informatics*, November 10–13, 2014, Barcelona, Spain p. 323
[29] Ghadge S, Killingback T, Sundaram B and Tran D A 2010 *Int. J. Parallel Emergent Distributed Syst.* **25** 223
[30] Nair A and Vidal J M 2011 *Int. J. Production Res.* **49** 1391
[31] Latapy M, Magnien C and Del Vecchio N 2008 *Soc. Networks* **30** 31
[32] Basole R C and Bellamy M A 2014 *Decis. Sci.* **45** 753
[33] Clauset A, Shalizi C R and Newman M E J 2009 *SIAM Review* **51** 661
[34] Brintrup A, Ledwoch A and Barros J 2016 *Logist. Res.* **9** 1
[35] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701
[36] Vasques Filho D and O'Neale D R J 2018 *Phys. Rev. E* **98** 022307
[37] Vasques Filho D and O'N E A L E D R J 2019 *Networks* **11** 14
[38] Suo Q, Guo J L, Sun S and Han L 2018 *Physica A* **489** 141
[39] Liu P and Zhang X 2011 *Int. J. Production Res.* **49** 637
[40] Gu T, Ren P, Jin M and Wang H 2018 *Disc. & Continuous Dyn. Systems-S* **12** 771
[41] Li Y, Zhao L and Suo J 2014 *Sustainability* **6** 4685
[42] Sun X, Lin H and Xu K 2015 *Expert Syst. Appl.* **42** 4229
[43] Sun X, Lin H, Xu K and Ding K 2015 *Scientometrics* **104** 43