WORKING PAPER

ITLS-WP-16-22

A network science approach to analysing manufacturing sector supply chain networks: Insights on topology

By
Supun S. Perera[1], Michael G.H. Bell[1], Mahendrarajah Piraveenan[2], Michiel C.J. Bliemer[1]

[1] Institute of Transport and Logistics Studies (ITLS), The University of Sydney Business School, Sydney, Australia
[2] Complex Systems Research Group (CSRG), School of Civil Engineering, University of Sydney, Australia

November 2016

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-16-22

TITLE: **A network science approach to analysing manufacturing sector supply chain networks: Insights on topology**

ABSTRACT: Due to the increasingly complex nature of the modern supply chain networks (SCNs), a recent research trend has focussed on modelling SCNs as complex adaptive systems. Despite the substantial number of studies devoted to such hypothetical modelling efforts, studies analysing the topological properties of real world SCNs have been relatively rare, mainly due to the scarcity of data. This paper aims to analyse the topological properties of twenty-six SCNs from the manufacturing sector. Moreover, this study aims to establish a general set of topological characteristics that can be observed in real world SCNs from the manufacturing sector, so that future theoretical work modelling the growth of SCNs in this sector can mimic these observations. It is found that the manufacturing sector SCNs tend to be scale free with degree exponents below two, tending towards hub and spoke configuration, as opposed to most other scale-free networks which have degree exponents above two. This observation becomes significant, since the importance of the degree exponent threshold of two in shaping the growth process of networks is well understood in network science. Other observed topological characteristics of the SCNs include disassortative mixing (in terms of node degree as well as node characteristics) and high modularity. In some networks, we find that node centrality is strongly correlated with the value added by each node to the supply chain. Since the growth mechanism that is most widely used to model the evolution of SCNs, the Barabasi - Albert model, does not generate scale-free topologies with degree exponent below two, it is concluded that a novel mechanism to model the growth of SCNs is required to be developed.

KEY WORDS: ***Network science, modelling supply network growth, empirical supply network analysis, complex network theory***

AUTHORS: **Perera, Bell, Piraveenan and Bliemer**

Acknowledgements: The authors acknowledge the Australian Research Council (ARC) for funding this work under grant DP140103643.

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (H

The Australian Key Centre in Transport and Logistics Manageme

The University of Sydney NSW 2006 Australia

Telephone: +612 9114 1813

E-mail: business.itlsinfo@sydney.edu.au

Internet: http://sydney.edu.au/business/itls

DATE: November 2016

1. Introduction

Traditionally, supply chains have been modelled as multi agent systems, in order to represent explicit communications between various autonomous entities involved (Thadakamalia et al., 2004). However, due to the increasingly complex and interconnected nature of the global supply chain networks (SCNs), recent research has focused on modelling supply chains as complex adaptive networks (Choi et al., 2001).

In order to understand, quantify and ultimately control various dynamical processes operating on SCN topologies, the first step is to construct realistic network models with tractable topological properties. Such models can be used to generate an ensemble of networks which can be studied analytically or by using numerical simulations (Bianconi, 2016). This can indeed provide remarkable insights into how the topological structure of a SCN can influence various dynamical processes, such as how efficiently information can be exchanged between individual entities, or how quickly can the overall system return to normal operations after a perturbation.

Following on from the influential work published by Thadakamalla et al. in 2004, which utilised network science to investigate the topological robustness of SCNs, a large number of theoretical research papers have appeared in this area (Xuan et al., 2011; Zhao et al., 2011(a); Zhao et al., 2011(b); Wen and Guo, 2012; Li et al., 2013; Yi et al., 2013; Li, 2014; Xu et al., 2014; Mari et al., 2015; Kim et al., 2015; Perera et al., 2016). Most of these studies have theoretically formulated plausible and generalizable growth mechanisms underlying the firm partnering process in SCN formation. Subsequently, the network topologies generated based on various growth models have been studied in depth for their topological characteristics, such as robustness and efficiency.

Despite the large number of theoretical papers published within the past few years, on network modelling of SCNs, the effort on empirical validation of the theoretical findings has been limited. This is mainly due to difficulty in obtaining information about supplier/customer relationships, which is often proprietary and confidential. Hitherto, majority of the research effort in this area has been focused on developing a generalizable network growth model which can generate topologies which can mimic the real world SCNs. Papers which topologically analyse real world supply chain data, the conclusions of which can then be used to inform modelling efforts, have been relatively scarce.

In light of the above, this study presents a comprehensive network and node level analysis of twenty-six SCNs across various manufacturing industry sectors, based on the data presented in Willems (2008). It is noted that all the SCNs considered include the full depth in terms of tiers (from suppliers to retailers).

In summary, this study aims to answer the following key research questions;

1. What common network level characteristics (if any) can be expected from real world SCNs in the manufacturing industry sector?
2. Do cost and time, at each stage of each SCN, correlate with any node level centrality metrics?
3. What are the similarities and/or differences between the real world SCNs analysed in this study and other real world complex (particularly scale-free) networks, such as the world-wide web (WWW), power grids, social and biological networks?
4. What key features should a generalised growth mechanism for the emergence of SCNs, through firm partnering, reflect?

The remainder of this manuscript is structured as follows. Section two will elaborate in the background to this study, and introduce key concepts in terms of topological analysis. Section three presents data analysis and results. In section four, we present a discussion of the results. Section five concludes the paper.

2. Background

2.1 Complex adaptive system nature of SCNs

Complex systems theory is a field of science that is used to investigate how the individual components and their relationships give rise to the collective behaviour of a given system (Ladyman et al., 2013). Complex adaptive systems (CAS) are dynamic systems which are capable of adapting and evolving with the environment within which they are embedded. It is important to note that in CAS, there exists little or no distinction between the system and the environment in which the system finds itself (Chan, 2001). Due to this, the adaptation of the system with respect to its environment is referred to as 'co-evolution' since the term 'evolution' is typically used to refer to a process that occurs independent of the environment.

Recent papers such as Choi et al. (2001) have argued that large scale supply networks are complex adaptive systems, where an interconnected network of multiple entities exhibit adaptability in response to changes in both the environment and the system itself. The collective system behaviour emerges as a nonlinear and dynamic function of the large number of activities made in parallel by interacting entities (Pathak et al., 2007). Therefore, from the point of view of a single firm, the overall SCN is a self-organising system which comprises individual entities engaging in localised decision-making. Given this distributed nature of decision making, the configuration of the final SCN structure is beyond the realm of control of one central organisation. Due to its self-organising nature, the actual structure of the SCN is probabilistic rather than deterministic, where the local choices are combined to create a stochastic structure. Indeed, individual firms may obey deterministic selection processes (Choi and Hartley, 1996) to account for self-interests and to promote their own fitness criteria. However, the final organisation of the overall SCN eventually emerges over time through the natural process of order and spontaneity (Choi et al., 2001).

As is evident from the above discussion, the characteristics of modern large scale SCNs unambiguously demonstrate CAS features, where the overall SCN, comprising dynamic entities, is embedded in a market environment. Co-evolution can therefore be observed when both the SCN and the market of which it is a part, evolve together over time.

2.2 Network modelling of SCNs

Traditionally, supply chains have been modelled as multi agent systems, in order to represent explicit communications between various entities involved (Swaminathan et al., 1998; Gjerdrum et al., 2001; Julka et al., 2002; Macal and North, 2005). Such agent-based models provide autonomy to each constituent entity and define behaviours in terms of observables accessible to the individual agent, which leads away from reliance on system-level information (Parunak et al., 1998). These models are generally effective in specifying how macro level patterns may emerge from the bottom-up (Jennings, 2000; Tesfatsion, 2002; Erez and Gati, 2004). However, due to the complex adaptive nature of modern supply networks, predicting the behaviour of the overall system based on its constituent components is extremely difficult (sometimes impossible) because of the strong possibility of emergent behaviour (Kremers, 2013). As such, adopting a macro model that would allow investigation of the overall system from a top-down view, through network analysis, is justified.

Although complex systems represent a vast range of real world systems, there exists no commonly established approach to investigate such systems. However, any complex system invariably consists of individual components interacting with each other – therefore such a system can be approximated as a network of components (Mitchell, 2006).

Broadly classified, the modelling of SCNs has mainly focused on the following network models for benchmarking purposes;

1) Random graphs (Erdȍs and Rényi, 1959) - where nodes are randomly connected to each other.
2) Small-world networks (Watts and Strogatz, 1998) - where most nodes are not neighbours of one another, but most nodes can be reached from every other node by a small number of steps.
3) Scale-free networks (Barabási and Albert, 1999) – where node degrees are in a power distribution, at least asymptotically.

The key characteristics of the above mentioned network structures are presented in Figure 1.

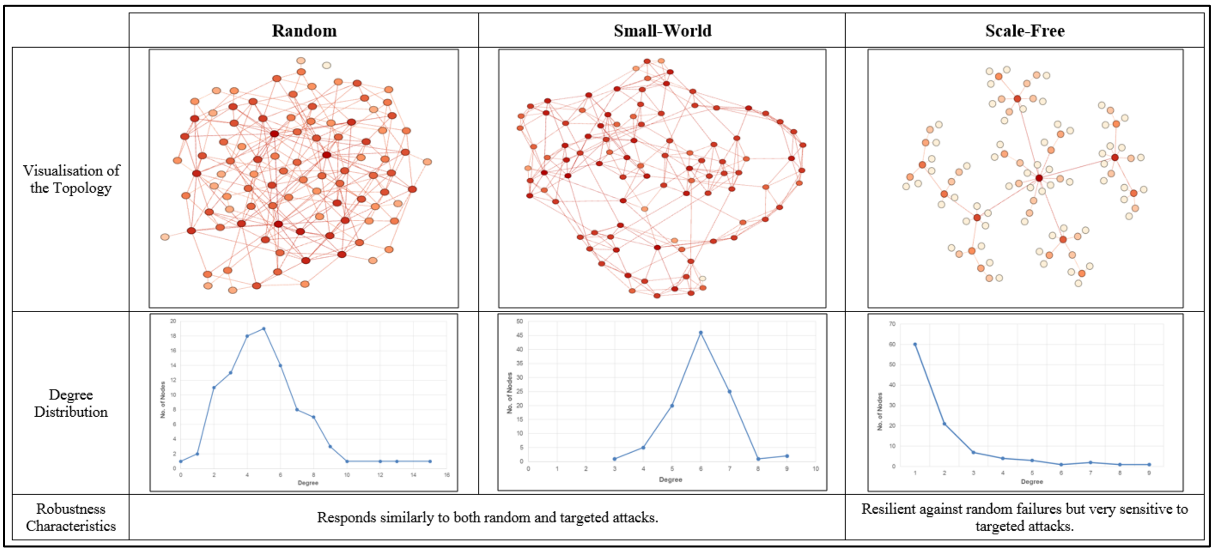

Figure 1. Comparison of random, small-world and scale free network - typical topologies

2.3 Modelling SCN evolution

In the context of supply networks, the concept of growth represents how firms join together to form SCNs. As new entrants join the supply network, they select partners from within the network. This partner selection is indeed a multi-objective problem and involves numerous factors, such as price, performance, quality, goodwill, etc (Jain et al., 2009; Li et al., 2013). Nevertheless, most research work in modelling SCN growth has hitherto given primary consideration to variants of degree based preferential attachment model, where the probability of attracting a connection from a new node is proportional to the number of existing connections possessed by each node within the network and/or other topological factors. Examples of such customised preferential attachment rules include the Hierarchy +/Ad-Hoc attachment model proposed by Thadakamalla et al. (2004), the Degree and Locality based Attachment (DLA) model proposed by Zhao et al. (2011a), and the Randomised Local Rewiring (RLR) model introduced by Zhao et al. (2011b). Each of the aforementioned growth models, over time, generate networks with distinct topologies. The topological characteristics arising from networks behaving according to the aforementioned growth model, can then be compared with the known features of other network benchmark models (as illustrated in Figure 1).

2.4 Data driven studies

Even though most modelling efforts of SCN focused on variants of preferential attachment as mentioned above, there have indeed been some studies which too have adopted a data-driven approach. For instance, Kim et al. (2011) have undertaken a node and network level topological analysis, using three case studies of automotive supply networks (namely, Honda Accord, Acura CL/TL, and Daimler Chrysler Grand Cherokee) presented by Choi and Hong (2002). Although the SCNs used in this study are complete, the SCNs are rather small in size (with a maximum of 34 firms in a given network), which limits the observations of emergent network topological properties. Kito et al. (2014) have constructed a SCN for Toyota using the data available within an online database operated by Marklines Automotive Information Platform. By analysing the SCN topology, the authors have identified the tier structure of Toyota to be barrel-shaped, in contrast to the previously hypothesized pyramidal structure. Another fundamental observation reported in this study is that Toyota SCN topology was found to be not scale free (even with finite-size effects taken into account). Although the dataset used in this study is sufficiently large (with 3,109 firms), it is limited to only the top three tiers of the overall SCN.

More recently, using Bloomberg data, Brintrup et al. (2015) and Orenstein (2016) have undertaken topological analysis of various SCNs. Brintrup et al. (2015) have studied the SCN of Airbus and have reported that this SCN displays assortative mixing and communities based on geographic locations of the firms. Orenstein (2016) has undertaken topological analysis of retail and food industry SCNs by considering the suppliers within the top three tiers. The SCNs considered in this study were found to have scale free topologies with degree exponent below 2. Although the dataset used in this study is sufficiently large and allows observation of temporal variations to the SCN topology, consideration of only a part of the SCN depth in terms of tiers has limited the generalisability of the results.

It should be noted that the key limitation in using the Bloomberg database, for constructing SCNs, is that the data are not exhaustive since the database only includes publicly listed firms. Therefore, the SCNs constructed using Bloomberg data may only provide a part of the full picture.

Although the above studies have provided a number of insights about the topological structure of various SCNs, no study to date has systematically investigated a large collection of SCNs in any particular industry using a comprehensive set of node and network level metrics. By considering a collection of twenty-six SCNs from the manufacturing industry, this study is able to investigate and establish the general topological properties of such networks in this sector. This effort will complement the large body of theoretical literature on modelling SCN topologies through various growth models, by revealing what specific topological characteristics are needed to be captured in an appropriate growth model. In addition, the correlation analysis presented in this study, between various node level centrality measures and two exogenous factors (stage cost and stage time), can be powerful in demonstrating how the position of firms can influence the overall functionality of the SCN.

Finally, this study has used a reliable dataset provided in published work by Willems (2008), which includes a large sample of manufacturing industry SCNs. Our study offers distinct insights from previous studies because (i) it is based on a large collection of real word SCNs belonging to a particular sector (ii) most of the SCNs are large (have a relatively high number of nodes) so that various emergent properties can be sufficiently demonstrated (iii) we extract topological properties to specifically compare them with the topological properties of networks generated by widely used SCN growth models, thus being able to comment on the suitability or otherwise of existing growth models for supply chain networks (iv) our correlation analysis of node centrality measures with exogenous factors provides insights into the impact of the position of firms in the dynamics of supply chains. Thus, this study is unique in several counts from previous studies described above.

2.5 Characterising complex networks

The metrics used to characterise the topology of complex networks can be broadly classified into node and network level metrics. See Costa et al. (2007) and Rubinov and Sporns (2010) for a comprehensive range of measurements used for characterization of complex networks. In this section, we describe the metrics we used in this study to analyse the topological properties of SCNs.

2.5.1 Mathematical Representation of Networks

A supply chain network can be represented as an unweighted, undirected graph $G = (N, L)$, with N nodes and L links. In the SCN models constructed in this paper, the nodes are assumed to represent individual firms, which are connected through links which represent contractual relationships between the firms. This network can also be represented by an adjacency matrix (*A*). Any element of the adjacency matrix $A = [a_{ij}]$ is given as:

$$a_{ij} = \begin{cases} 1, & \text{if } i \neq j \text{ and } i \text{ and } j \text{ nodes are connected by an edge} \\ 0, & \text{if } i \neq j \text{ and } i \text{ and } j \text{ nodes are not connected} \\ 0, & \text{if } i = j. \end{cases}$$

2.5.2 Network level metrics

Table 1 presents the list of network level metrics used for analysis in this study, and their implications within a SCN context.

Table 1: Network level metrics used and their SCN implications

Mathematical representation	**SCN Implication**
Average degree (<*k*>)	
$$< k > = \frac{\sum_i k_i}{N}$$ where *N* is the total number of nodes in the network	Indicates, on average, how many connections a given firm has. Higher average degree implies good inter-connectivity among the firms in the SCN.
Network diameter	
$$\text{diameter} = \max_{i,j} \; l(i,j)$$ where *l* is the number of hops traversed along the shortest path from node *i* to *j*.	The diameter of a SCN is the largest distance between any two firms in the network. More complex manufacturing processes can include large network diameters (i.e. many stages of production) indicating difficulty in governing the overall SCN under a centralised authority.
Network density (D)	
$$\mathrm{D} = \frac{< k >}{N - 1}$$ where <*k*> is the mean degree of all the nodes and *N* is the total number of nodes, in the network	Density of a SCN indicated the level of interconnectivity between the firms involved. SCNs with high density indicate good levels of connectivity between firms which can be favourable in terms of efficient information exchange and improved robustness due to redundancy and flexibility (Sheffi and Rice, 2005).
Network centralisation (C)	
$$\mathrm{C} = \frac{N}{N-2}\left(\frac{\max(k)}{N-1} - \text{Density}\right)$$ where *N* is the total number of nodes in the network and max(*k*) is the maximum degree of a node within the network. Density is determined as per the equation below.	Network centralisation provides a value for a given SCN between 0 (if all firms in the SCN have the same connectivity) and 1 (if the SCN has a star topology). This indicates how the operational authority is concentrated in a few central firms within the SCN. Highly centralised SCNs can have convenience in terms of centralised decision implementation and high level of controllability in production planning. However, highly centralised

	SCNs lack local responsiveness since relationships between firms in various tiers are decoupled (Kim et al., 2011).
Network heterogeneity (H)	
$\mathrm{H} = \frac{\sqrt{variance(k)}}{<k>}$ where <*k*> is the mean degree and *variance (k)* is the variance of the degree, of all the nodes in the network.	Heterogeneity is the coefficient of variation of the connectivity. Highly heterogeneous SCNs exhibit hub firms (i.e. firms with high number of contractual connections). In extreme cases, there may be many super large hubs (winner take all scenario, indicating centralised control of the overall SCN through a very few firms).
Average clustering coefficient (<C>)	
$<C> = \frac{\sum_i C_i}{N}$ where N is the total number of nodes in the network and C_i is the number of triangles connected to node i divided by the number of triples centered around node i.	Clustering coefficient indicates the degree to which firms in a SCN tend to cluster together around a given firm. For example, it can indicate how various suppliers behave with respect to the final assembler at the global level (Kim et al., 2011). Therefore, the higher the clustering coefficient, the more dependent suppliers are on each other for production (Brintrup et al., 2016).
Degree exponent (γ) (Barabasi and Albert, 1999)	
The degree distribution P_k of a scale free network is approximated with power law as follows; $P_k \sim k^{-\gamma}$ where k is the degree of the node and γ is the degree exponent (also known as the power law or scale free exponent).	SCNs with γ<2 include very large hubs which acquire control through contractual relationships with other firms at a rate faster than the growth of the SCN in terms of new firm additions. As γ continues to increase beyond 2, the SCNs include smaller and less numerous hubs, which ultimately leads to a topology similar to that of a random network where all firms have almost the same number of connections.
Assortativity (ρ) (Newman, 2002)	
Assortativity is formally defined as a correlation function of excess degree distributions and link distribution of a network.	Positive assortativity means that the firms with similar connectivity would have a higher tendency to connect with each other (for example, highly connected firms could be managing sub-communities in certain areas of production and then connect to other high-degree firms

For undirected networks, when degree distribution is denoted as p_k and excess degree (remaining degree) distribution is denoted as q_k, one can introduce the quantity $e_{j,k}$ as the joint probability distribution of the remaining degree distribution of the remaining degrees of the two nodes at either end of a randomly chosen link. Given these distributions, the assortativity of an undirected network is defined as; $$\rho = \frac{1}{\sigma_q^2}\left[\sum_{jk} jk(e_{j,k} - q_j q_k)\right]$$ where σ_q is the standard deviation of q_k.	undertaking the same function). This structure can lead to cascading disruptions – where a disruption at one leaf node can spread quickly within the network through the connected hubs (Brintrup et al., 2016). In contrast, a negative assortativity indicates that it is the firms with dissimilar connectivity that tend to pair up in the given network.
Modularity (Q) (Newman and Girvan, 2004)	
$$Q = \sum_{s=1}^{k}\left[\frac{l_s}{L} - \left(\frac{d_s}{2L}\right)^2\right]$$ where k is the number of modules, L is the number of links in the network, l_s is the number of links between nodes in module s, and d_s is the sum of degrees of nodes in the module s. To avoid getting a single module in all cases, this measure imposes Q=0 if all nodes are in the same module or nodes are placed randomly into modules.	SCNs with high modularity contain pronounced communities – i.e. partially segregated subsystems or modules embedded within the overall SCN system (Ravasz et al., 2002; Newman, 2003).
Percolation threshold for random node removal (f_c) (Cohen et al., 2000)	
The percolation threshold for random node removal is given as; $$f_c = 1 - \frac{1}{\frac{<k^2>}{<k>} - 1}$$	The percolation threshold of a SCN indicates the percentage of firms needed to be randomly removed prior to the overall SCN breaks into many disconnected components. In summary, this indicates the number of random firm failures that would drive the SCN from a connected state to a fragmented state (loss of overall interconnectivity).

where <k> is the mean degree and <k^2> is the second moment of the degree, of all the nodes in the network. It is noted that in the above formulation represents the proportion of nodes required to be removed from the network before the giant component ceases to include all the nodes.	

2.5.3 Node level metrics

Node level metrics characterize, in various ways, the importance of a particular node for the functionality of the overall network, based on its embedded position in the broader relationship network. Most node level metrics therefore relate to node centrality. Depending on the context, various centrality measures can be adopted to identify the key players of a given network. Table 2 presents the list of node level metrics used in this study, and their implications within a SCN context.

Table 2: Node level metrics used and their SCN implications

Mathematical representation	**SCN Implication**
Degree (k)	
The degree k_i of any node i is represented by; $k_i = \sum_j a_{ij}$ where a_{ij} is any element of the adjacency matrix A.	Represents the number of direct neighbours (connections) a given firm has. For instance, in a given SCN, the firm with the highest degree (such as the integrators that assemble components) is deemed to have the largest impact on operational decisions and strategic behaviours of other firms in that particular SCN. Such a firm has the power to reconcile the differences between various other firms in the SCN and align their efforts with greater SCN goals (Kim et al., 2011).
Betweenness centrality (normalised) (Freeman, 1977)	
The betweenness centrality of a node n is defined as; $C_b(n) = \frac{2}{(N-1)(N-2)} \sum_{s \neq n \neq t} \frac{\sigma_{s,t}(n)}{\sigma_{s,t}}$ where s and t are nodes in the network, which are different from n, $\sigma_{s,t}$ denotes the number of	Betweenness centrality of a firm is the number of shortest path relationships going through it, considering the shortest path relationships that connect any two given firms in the SCN. Therefore, it indicates the extent to which a firm can intervene over interactions among other firms in the SCN by being a gatekeeper for relationships.

shortest paths from *s* to *t*, and $\sigma_{s,t}(n)$ is the number of shortest paths from *s* to *t* that *n* lies on.	
Closeness centrality (Sabidussi, 1966)	
The closeness centrality of a node n is defined as; $C_c(n) = \frac{1}{< L(n,m) >}$ where *<L(n,m)>* is the length of the shortest path between two nodes *n* and *m* (note that for unweighted graphs with no geodesic distance information, each link is assumed to be one unit of distance). The closeness centrality of each node is a number between 0 and 1.	Closeness centrality is a measure of the time that it takes to spread the information from a particular firm to the other firms in the network. While it is closely related to betweenness centrality, closeness more relevant in situations where a firm acts as a generator of information rather than a mere mediator/gatekeeper. For example, due to various hindrances, the market demand information can easily be distorted when it flows from the downstream firms towards upstream firms. Such distortions can lead to undue deviation between production plans of manufacturers and supply plans of suppliers, leading to a phenomenon known as the bullwhip effect. Firms with high closeness centrality levels therefore play a major role in sharing the actual market demand information with upstream firms in the SCN, thus diminishing the adverse impacts arising from bullwhip effect (Xu et al., 2016).
Eigenvector centrality (Ruhnau, 2000)	
If the centrality scores of nodes are given by the matrix X and the adjacency matrix of the network is A, then x can be defined iteratively as; $x \propto Ax$ i.e. $\lambda x = Ax$ The eigenvector centrality scores are obtained by solving this matrix equation. It can be shown that, while there can be many values for λ, only the largest value will result in positive scores for all nodes.	Eigenvector centrality measures a firm's influence in the SCN by taking into account the influence of its neighbours. It assumes that the centrality score of a firm is proportional to the sum of the centrality scores of the neighbours. A firm with a high eigenvector centrality is assumed to derive its influential power through its highly connected neighbours.

3. Data analysis and Results

3.1 Data source and structure

Willems (2008) provides a dataset of real world multi echelon supply chains, used for inventory optimization purposes. The overall dataset includes a total of 38 multi echelon supply chains, from various industries. The chains described in this paper comprise actual supply chain maps created by either company analysts or consultants. Since these maps have been implemented in practice, they demonstrate how users have modelled actual supply chains.
The above-mentioned dataset includes the following key information;

- The industry sector of each supply chain network;
- For each supply chain;
    - The stages representing each firm involved; and
    - The arcs representing precedence relationship between stages.
- For each stage;
    - Its classification and tier based on its function within the overall supply chain;
    - The direct cost added at the stage (stage cost); and
    - The average processing time at the stage (stage time).

Please note that, even though in the context of an individual supply chain, the terminology 'stage' makes sense (manufacturing stage, retail stage etc), when we represent them as supply chain networks, it does not. Each 'stage' in a supply chain therefore simply represents a node, which is also a firm, within the SCN.

From the original dataset, the networks with more than 100 firms (i.e. nodes) were selected for our analysis, and there were twenty-eight such large networks. Smaller networks were omitted in this analysis since they do not offer any interesting insights into emergence of various complex topological features. Then, using the industry sector information, these SCNs were categorised into six main groups as illustrated in Table 3. As can be seen, the set of SCNs considered by us vary in size (with a minimum of 108 to a maximum of 2025 nodes).

Table 3: Classification of SCNs considered in the study

Group	Industry Sector	SCN Ref #	No. of Nodes	No. of Links
Aircraft Parts (2)	Aircraft Engines and Engine Parts	1	468	605
		2	2025	16225
Chemicals (7)	Perfumes, Cosmetics, and Other Toilet Preparations	3	186	359
		4	844	1685
		5	976	1009
	Paints, Varnishes, Lacquers, Enamels, and Allied Products	6	271	524
	Industrial Organic Chemicals, Not Elsewhere Classified	7	1479	2069
	Soap and Other Detergents, Except Specialty Cleaners	8	133	164
	Pharmaceutical Preparations	9	253	253
Electrical (6)	Electro medical and Electrotherapeutic Apparatus	10	145	224
		11	482	941
		12	1386	1857
	Telephone and Telegraph Apparatus	13	1206	4063
	Primary Batteries, Dry and Wet	14	617	753
	Power-Driven Hand tools	15	334	1245
Computer (6)	Computer Peripheral Equipment	16	152	211
		17	154	224
		18	156	263
		19	156	169
	Computer Storage Devices	20	577	2262
	Semiconductors and Related Devices	21	108	452
Arrangement of Transportation of Freight and Cargo (2)		22	116	119
		23	626	632
Farm Machinery and Equipment (3)		24	409	853
		25	706	908
		26	1451	4812

3.2 Limitations of the dataset

A key limitation of the available dataset is the lack of information in relation to the geographical locations of individual firms. This information was not provided in the original dataset in Willems (2008) due to confidentiality reasons. Unlike the virtual networks (such as WWW or social networks), the SCN structure is largely influenced by geographical aspects (since the congregation or dispersion of the suppliers depend on the raw material distribution over various geographic regions). Therefore, if geographic location information was available, in depth conclusions could have been made about various observed structural features of the SCNs.

In addition, this study is unable to investigate the dynamic nature of the SCNs since the dataset does not provide any information pertaining to temporal changes in the SCN topology. Lastly, the relationship strength between firms are not captured in the dataset in terms of the amounts material flow. Although specific production capabilities of firms within each tier are known, no information is available in relation to how much each upstream firm supplies to the downstream firms.

Nevertheless, the size of the dataset, both in terms of the number of networks available and in terms of number of nodes in each network, as well as the cost and time data associated with nodes (called stage cost and stage time by Willems (2008)), make this a very attractive dataset to study.

3.3 Data analysis and results

Using the stage and arc data from the dataset, we constructed SCNs, where the nodes represent the individual firms and the links represent the contractual relationships between firms (undirected). Cytoscape software and JAVA programming language were used to visualise and analyse the SCNs. The results are presented below.

3.3.1 Network level metrics

Table 4 represents the results of network level metrics for each SCN analysed. The three key observations that can be made from the table are (i) most SCNs are scale-free, and most of them have a power-law exponent less than 2.0 (ii) most SCNs are disassortative, in terms of both degree assortativity, as well as assortativity calculated based on stage cost and stage time, and (iii) all SCNs include high levels of modularity indicating presence of closely knit communities. These observations have important implications, which are discussed in more detail in the following sub-sections.

Table 4: Summary of results obtained for each SCN analysed

Industry Group	Industry Sector	SCN Ref#	No. of Nodes	No. of Links	Average Degree	Network Diameter	Characteristic Path Length	Network Centralization	Network Density	Network Heterogeneity	Degree Exponent	Power Law Correlation	Assortativity (Degree based)	Assortativity (Cost based)	Assortativity (Time based)	Modularity	Communities	Percolation Threshold for Random Failures
Aircraft Parts	Aircraft Engines and Engine Parts	1	468	605	2.59	9	4.07	0.20	0.06%	3.03	1.09	0.96	-0.46	0.22	-0.34	0.78	17	96.1%
		2	2025	16225	16.03	6	4.11	0.11	0.80%	2.45	0.85	0.87	-0.72	-0.55	-0.98	0.28	5	99.1%
Chemicals	Perfumes, Cosmetics, and Other Toilet Preparations	3	186	359	3.86	8	4.30	0.08	2.10%	1.06	1.20	0.97	0.25	0.13	0.06	0.61	11	86.1%
		4	844	1685	3.99	13	5.43	0.07	0.50%	1.56	1.38	0.98	-0.14	-0.03	-0.03	0.63	8	92.1%
		5	976	1009	2.07	18	9.43	0.02	0.02%	0.97	1.81	0.85	0.00	-0.02	0.02	0.90	35	66.9%
	Paints, Varnishes, Lacquers, Enamels, and Allied Products	6	271	524	3.87	8	4.06	0.08	1.40%	1.21	1.49	0.99	-0.31	-0.16	-0.07	0.59	10	88.2%
	Industrial Organic Chemicals	7	1479	2069	2.80	14	6.49	0.03	0.02%	1.57	1.50	0.89	-0.06	-0.19	-0.46	0.76	15	88.5%
	Soap and Other Detergents, Except Specialty Cleaners	8	133	164	2.47	6	3.88	0.20	1.90%	1.09	1.23	0.91	0.03	-0.40	0.20	0.70	11	77.2%
	Pharmaceutical Preparations	9	253	253	2.00	10	5.78	0.13	0.08%	1.68	1.29	0.88	-0.25	0.47	0.05	0.79	7	84.9%
Electrical	Electromedical and Electrotherapeutic Apparatus	10	145	224	3.09	8	4.96	0.08	2.10%	0.92	1.33	0.95	-0.08	-0.09	-0.29	0.66	14	78.8%
		11	482	941	3.91	8	4.36	0.18	0.80%	2.68	1.07	0.94	-0.61	-0.62	-0.58	0.56	8	96.8%
		12	1386	1857	2.68	20	9.82	0.04	0.20%	2.06	1.49	0.96	-0.29	-0.04	-0.20	0.88	16	92.3%
	Telephone and Telegraph Apparatus	13	1206	4063	6.74	5	3.65	0.12	0.60%	2.75	1.09	0.99	-0.73	0.12	-0.55	0.50	12	98.2%
	Primary Batteries, Dry and Wet	14	617	753	2.44	10	6.48	0.04	0.40%	1.51	1.56	0.97	-0.64	-0.62	-0.12	0.81	19	85.7%
	Power-Driven Handtools	15	334	1245	7.46	7	3.50	0.18	2.20%	1.83	0.89	0.95	-0.60	-0.03	-0.24	0.31	27	96.8%
Computer	Computer Peripheral Equipment	16	152	211	2.78	6	3.46	0.20	1.80%	1.72	1.13	0.89	-0.18	0.19	-0.22	0.52	29	90.0%
		17	154	224	2.91	12	4.95	0.07	1.90%	1.02	1.31	0.96	0.16	-0.04	0.06	0.60	9	79.8%
		18	156	263	3.37	12	4.28	0.16	2.20%	1.28	1.23	0.65	-0.25	-0.22	-0.02	0.59	10	87.3%
		19	156	169	2.17	13	5.38	0.20	1.40%	1.65	1.21	0.80	-0.28	-0.17	-0.11	0.74	19	85.8%
	Computer Storage Devices	20	577	2262	7.84	10	3.52	0.34	1.40%	2.47	0.91	0.59	-0.47	-0.03	-0.75	0.60	17	98.2%
	Semiconductors and Related Devices	21	108	452	8.37	4	2.06	0.74	7.80%	1.83	0.69	0.44	-0.82	-0.76	-0.82	0.12	3	97.2%
Arrangement of Transportation of Freight and Cargo		22	116	119	2.05	8	4.88	0.07	1.80%	0.99	1.59	0.97	-0.06	-0.11	-0.17	0.79	13	67.4%
		23	626	632	2.02	8	5.45	0.07	0.30%	1.17	1.54	0.69	0.06	-0.05	0.03	0.91	29	73.5%
Farm Machinery and Equipment		24	409	853	4.17	4	3.24	0.34	1.00%	1.91	1.26	0.99	-0.12	0.02	-0.12	0.62	8	94.6%
		25	706	908	2.57	6	5.38	0.04	0.40%	1.78	0.93	0.85	-0.81	-0.26	-0.71	0.87	32	89.7%
		26	1451	4812	6.63	5	3.18	0.40	0.50%	2.68	1.22	0.98	-0.11	-0.05	-0.01	0.71	8	98.1%

3.3.1.1 Scale-freeness and degree exponent (γ)

Interestingly, most SCNs analysed indicate scale-free topology – i.e. the degree distribution follows the power law. More specifically, 22 out of the 26 networks analysed display 80% or higher correlation with a power-law fit, and these can undoubtedly labelled as scale-free networks. Moreover, all of these scale-free networks display a degree exponent which is less than 2.0.

Indeed, most real world networks have been observed to be scale-free, including technological, social, and biological networks (Barabasi et al., 2000). However, in most cases, it has been found that scale-free networks have degree exponents between 2 and 3 (Barabasi, 2016). The growth mechanisms underlying such networks have been related to some form of preferential attachment, most notably the Barabasi-Albert (BA) model, which is known to generate scale-free networks with approximately γ =3. Therefore, despite its elegance and simplicity, networks with a γ below 3 cannot be generated by the BA growth model (Nguyen and Tran, 2012).

Many properties of a scale-free network depend on the value of the degree exponent, γ (Barabasi, 2016). Therefore, it is interesting to establish how the network properties vary with γ. For a scale-free network, the expected maximum degree $k_{\max}$ (also known as the natural cut-off) which represents the expected size of the largest hub is estimated as follows (Barabasi, 2016);

$$k_{\max} \approx k_{\min} N^{\frac{1}{\gamma-1}} \qquad \textbf{(Eq. 1)}$$

where $k_{\max}$ and $k_{\min}$ are the expected maximum and minimum degree of a node, respectively. N is the system size, in terms of the number of nodes.

Based on Eq. 1 above, the link acquisition rate of the biggest hub (i.e. node with $k_{\max}$) behaves as follows;

- When γ<2, the exponent $\frac{1}{g-1}$ is larger than 1. Therefore, the link acquisition rate of the largest hub is faster than the growth of the network in terms of the number of nodes present.
- When γ=2, the exponent $\frac{1}{\gamma-1}$ is 1. Therefore, the biggest hub acquires links linearly with the network size, thus forcing the network towards a hub and spoke configuration.
- When γ>2, the exponent $\frac{1}{\gamma-1}$ is less than 1. Therefore, as γ increases beyond 2, the degree distribution decays faster, thus making the hubs smaller and less numerous.

Based on equation 1, when $\gamma<2$ the link acquisition rate of the largest hub is faster than the growth of the network in terms of the number of nodes it contains (indeed no large networks can exist in this regime, since the largest hub will eventually run out of nodes to connect to). In this scenario, the high-degree nodes are disproportionately attractive. This winner-takes-all dynamic leads to a hub-and-spoke network topology in which all nodes are within a short distance from each other.

As can be seen from the above discussion, the threshold for degree exponent $\gamma = 2.0$ is critical, and decidedly influences link acquisition rates and the resultant evolution of the network. Therefore, our observation that most SCNs that we analysed have degree exponents less than 2.0 implies that they cannot be successfully modelled by a mechanism like the Barabasi-Albert (BA) model.

3.3.1.2 Network centralisation

All the SCNs included have relatively lower network centralisation values – indicating the largely distributed and decentralised nature of modern SCNs. This lower centralisation could also be due to the recent supply chain practice known as modular assembly, where manufacturers obtain pre-assembled modules from a reduced base of suppliers (such as through intermediate sub-assemblers), as opposed to the traditional approach in which individual components are procured and assembled by the manufacturer (Hu et al., 2008).

However, it was noted that in general, more complex and specialised manufacturing and assembly processes, such as aircraft parts, computer equipment, electrical and farm machinery, SCNs were more centralised than other SCNs such as cargo and chemicals.

3.3.1.3 Network heterogeneity

Network heterogeneity is the coefficient of variation of a given network's degree distribution (Dong and Horvath, 2007). Past research has shown that heterogeneity in the degree distribution can inhibit global synchronisation of networks (Nishikawa et al., 2003). However, heterogeneity can be a desirable property in SCNs, since global synchronisation of a SCN could facilitate cascading failures due to disruptions (Pereira, 2010).

Jun et al (2007) has shown that scale free networks with $\gamma \approx 1.7$ is most heterogeneous and the scale-free networks become more homogeneous as γ increases beyond 1.7. The SCNs analysed in this study indicate generally high levels of heterogeneity (between 0.9 and 3). This is indeed attributable to the γ of these networks (generally in the vicinity of 1.7). As noted by Ou et al., 2009, empirical studies indicate that most real world complex networks are generally highly heterogeneous with the γ ranging from 1 to 3.

3.3.1.4 Assortativity

In terms of degree assortativity, Majority of the SCNs (23 out of 26) were found to be slightly or strongly disassortative, where the hubs tend to avoid each other, instead linking to lower-degree nodes. As a result, the network structure of these SCNs tends to display hub and spoke character (as opposed to core periphery structure observed in assortative networks). Some networks such as 'Perfumes, Cosmetics, and Other Toilet Preparations' and 'Semiconductors and Related Devices' were strongly disassortative. We found no SCN which was strongly assortative.

The SCNs mostly displayed slight or strong disassortative tendencies in terms of stage-cost and and stage-time as well. That is, firms which contribute high stage-cost are on average more likely to be connected with firms that contribute low stage-cost, and vice-versa, and the same is true for stage-time. No SCN that we studied displayed strong assortative tendency in terms of these attributes.

Disassortative mixing has been observed commonly in economic systems (Barabasi, 2016). For example, in economic settings, trade typically takes place between individuals or organisations of different skills and specialties. The above is certainly true in the SCN – where a supplier is most likely to link with a manufacturer, rather than to another supplier. This inherent functional property, in the SCN context, is likely to be responsible for forcing the SCNs towards disassortative mixing.

An unfavourable implication of the disassortativity observed in the SCNs (particularly in terms of degree) is that since high degree nodes are less connected to one another, many paths between nodes in the network are dependent on high degree nodes. Therefore, failure of a high degree node in a disassortative network would have a relatively large impact on the overall connectedness of the network (Noldus and Van, 2015). On the other hand, disassortative networks are generally resilient against cascading impacts arising from targeted attacks – since hub nodes are not connected with each other, the likelihood of disruption impacts cascading from one hub node to another is minimised (Song et al, 2006).

3.3.1.5 Modularity and Communities

Majority of the SCNs analysed were found to have moderate to high levels of modularity, which indicates the presence of partially segregated subsystems or modules embedded within a SCN system (Ravasz et al., 2002; Newman, 2003). This reveals the organisational intricacies underlying the SCN structure. Indeed, the individual firms which form SCNs, tend to be well partitioned based on the heterogeneity of their respective functions.

At first glance, it may appear that since SCNs have tiered structures, each tier could represent a 'community' or module, resulting in high modularity. However, it was noted that in the SCN dataset considered, no horizontal connections were observed (i.e. no connections between firms within the same functional tier). Therefore, the communities or modules observed in these SCNs are due to *vertical connections*. This could be due to more complex manufacturing processes, which bring together many parts, demanding connections with multiple suppliers. Another reason could be that firms intentionally link with multiple firms at upper and lower tiers (known as multi-sourcing), as to improve the reliability of maintaining production levels over the demand threshold through redundancy and flexibility (Sheffi, 2001).

Indeed, more efficient SCNs will possess communities that allow for improved information flow and innovation diffusion (Hearnshaw and Wilson, 2013). Past research measuring network modularity (Newman, 2003) reveal that networks with communities which are fuzzy in their segregation are better at diffusing and transmitting information across the entire network than networks with a more distinct community structure (Danon et al., 2008).

Hearnshaw and Wilson (2013) note that for supply chain systems to function efficiently from initial suppliers to final consumers it is necessary that vertical connections between communities are formed and maintained. However, such inter-tier connections are likely to be costly compared to intra-tier connections, particularly due to differing interests and functions between each tier. Accordingly, the leader firms (i.e. hubs), within each tier, are more likely to initiate and maintain inter-tier connections, due to their enhanced capacity to link across different functions, and their ability to tackle riskier exchange relationships given their greater resources (Goyal, 2012).

3.3.1.6 Structural robustness to random failures

A crude indicator of the structural robustness of a given network is its integrity in terms of the presence of a giant component. For a network to have a giant component, most nodes that belong to it must be connected to at least two other nodes (Barabasi, 2015). The presence of a giant component within a given network can be established using the Molloy-Reed criterion (Reed, 1995), as shown below;

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2 \qquad \textbf{(Eq. 2)}$$

Based on the above equation, networks with $\kappa>2$, are deemed to include a giant component. On the other hand, when $\kappa<2$, the overall network is composed of many disconnected clusters. Almost all of the SCNs analysed in this paper include κ values well above 2, indicating the presence of a giant connected component, in which all firms belong to a single component which is the SCN itself.

Based on the insights provided by the Molloy-Reed criterion, using Eq. 3 below, one can predict the percentage of nodes required to be randomly removed from the overall network in order to rid the network of its overall integrity by destroying the giant component (Cohen et al., 2000).

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \qquad \textbf{(Eq. 3)}$$

It is noted that f_c in the above formulation represents the proportion of nodes required to be removed from the network before the giant component ceases to include all the nodes.

In fact, the above equation can also be written as;

$$f_c = 1 - \frac{1}{\kappa - 1} \qquad \textbf{(Eq. 4)}$$

Where κ is defined as Eq. 2 above

The f_c values calculated for the majority of SCNs indicate very high level of robustness against random node removals (on average, about 88% of nodes should be randomly removed before the giant connected component disintegrates).

Note that for networks with γ <3, the second moment of the degree distribution $\langle k^2 \rangle$ diverges as N tends to infinity limit. This in turn forces f_c to converge to 1, implying that in order to fragment a scale free network of infinite size, one must randomly remove all of its nodes (Barabasi, 2015).

Indeed the enhanced robustness of SCNs with respect to random node removals derives from the hub structure of these networks. Random node removals, by definition, affect nodes irrespective of their degrees. Since scale-free networks with $g > 0$ comprise mainly less connected nodes and a few hubs, the chance of randomly removing a hub is almost negligible. Therefore, random node removals are likely to affect mainly the less connected nodes, which although numerous, play a limited role in maintaining a network's integrity (Barabasi, 2015).

3.3.2 Node level metrics

Since node level metrics themselves provide information about individual nodes rather than networks as a whole, here we choose to primarily study some correlations between them. In particular, we studied the correlation coefficients between node level centrality metrics (namely, degree centrality (DC), betweenness centrality (BWC), closeness centrality (CC) and eigenvector centrality (EVC)) and (1) stage cost and (2) stage time for each node. The results of this assessment are presented in Figure 2 and 3 below.

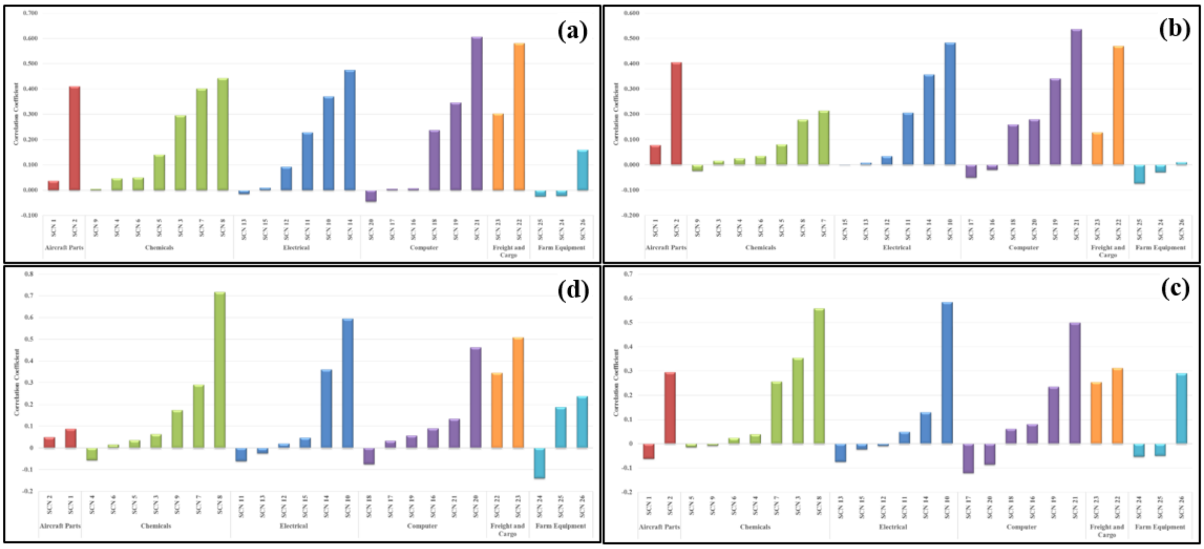

Figure 2: Correlation between stage cost and; (a) degree centrality, (b) betweenness centrality, (c) closeness centrality, and (4) eigenvector centrality, for each SCN

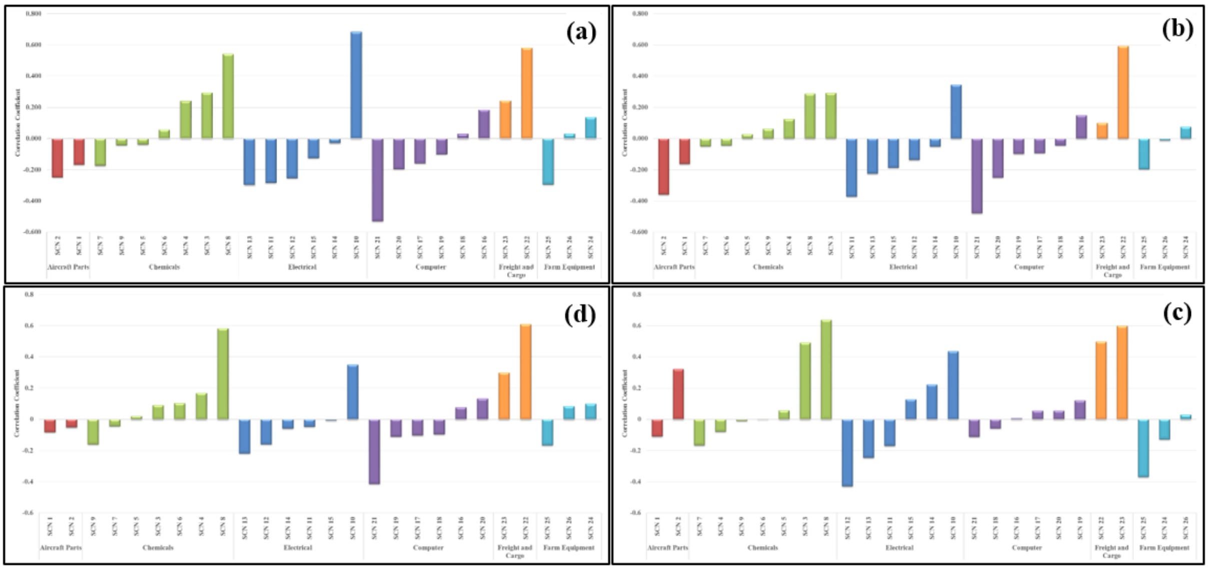

Figure 3: Correlation between stage time and; (a) degree centrality, (b) betweenness centrality, (c) closeness centrality, and (4) eigenvector centrality, for each SCN

These figures provide important insights and demonstrate how, from a SCN point of view, the position of an individual firm with respect to the others, can influence both strategy and behavior (Borgatti and Li, 2009). In terms of stage cost, we found that the correlation between stage cost and centrality metrics is mostly positive or neutral for most firms. In particular, some computer, and freight and cargo networks, the correlation was strongly positive, implying that the more centrally placed the firm is, the higher cost in contributes to the SCN. In most other SCNs the correlation was in the indifferent band, and no SCN displayed a strongly negative correlation. No qualitative difference was observed between different centrality metrics used overall, though in terms of individual SCNs, some centrality metrics returned stronger correlations than others.

In terms of stage time however, the observations were slightly different. While for some firms stage-time was positively correlated with centrality metrics, for others it appeared that it was negatively correlated with centrality metrics. For example, one of the computer SCNs (SCN 21) displays strong negative correlation between stage time and centrality metrics, except closeness centrality. Interestingly, SCNs also seem in general more sensitive to the centrality metric used, when stage time is considered. Therefore, we may conclude that there is no overall tendency in terms of correlation between stage time and node centrality, and the context and characteristics of individual SCNs determine whether more centrally located firms are likely to have relatively longer or shorter stage times.

Figure 4 and 5 represent the same set of results in a different representation, where sector based differences are highlighted. From these figures, we could see that cargo sector displays the most positive correlation between sage time and centrality, while aircraft, electrical, computer and farm equipment sectors display slightly negative correlations between stage time and centrality. Of course, each entry in Fig 4 and 5 was obtained by averaging correlation values for all networks in a particular sector, which has limited meaning. One network with high positive or negative correlation might cause an entire sector to be misrepresented, where in reality most other SCNs might be insensitive to node centrality. Nevertheless, taken together with figures 2 and 3, figures 4 and 5 represent a useful insight. Another interesting observation is that where there is negative correlation observed, closeness centrality often bucks the trend. That is, even when some SCNs have negative correlation between stage time and all other centrality measures, they have slightly positive correlation between closeness centrality and stage time. Therefore, it appears that in some sectors, like aircraft parts and computers, firms with relatively high stage time a peripherally placed in terms of all other centrality measures but centrally placed in terms of closeness centrality. This is a catalyst for further research, but beyond the scope of this broad preliminary study.

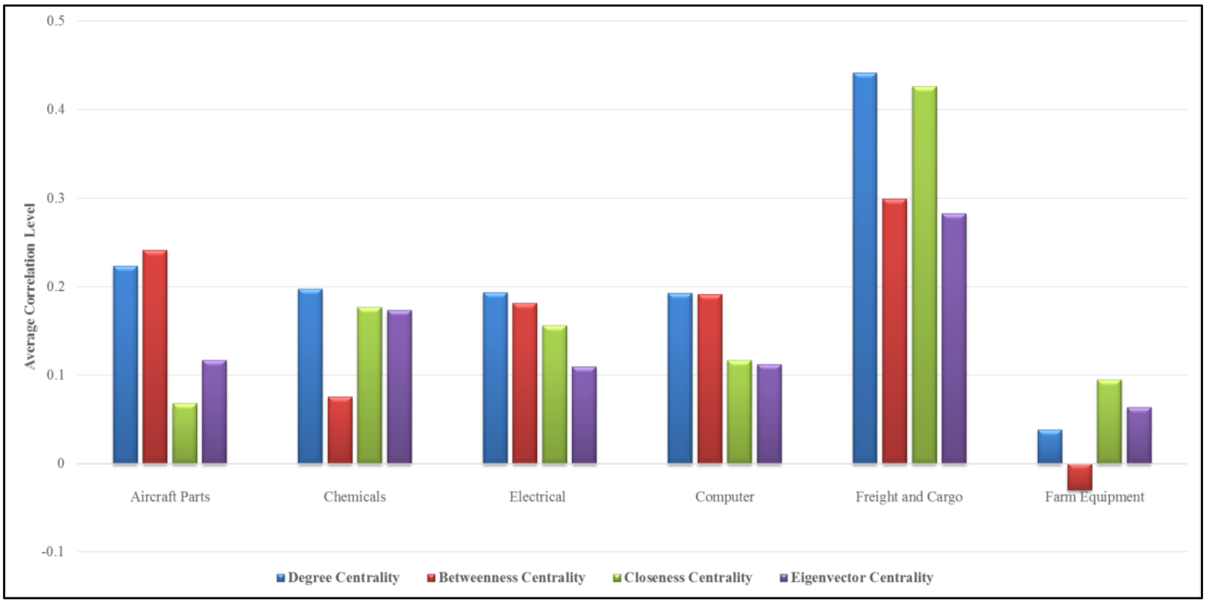

Figure 4: Average correlation levels of stage cost with node level centrality metrics for SCNs in each sector

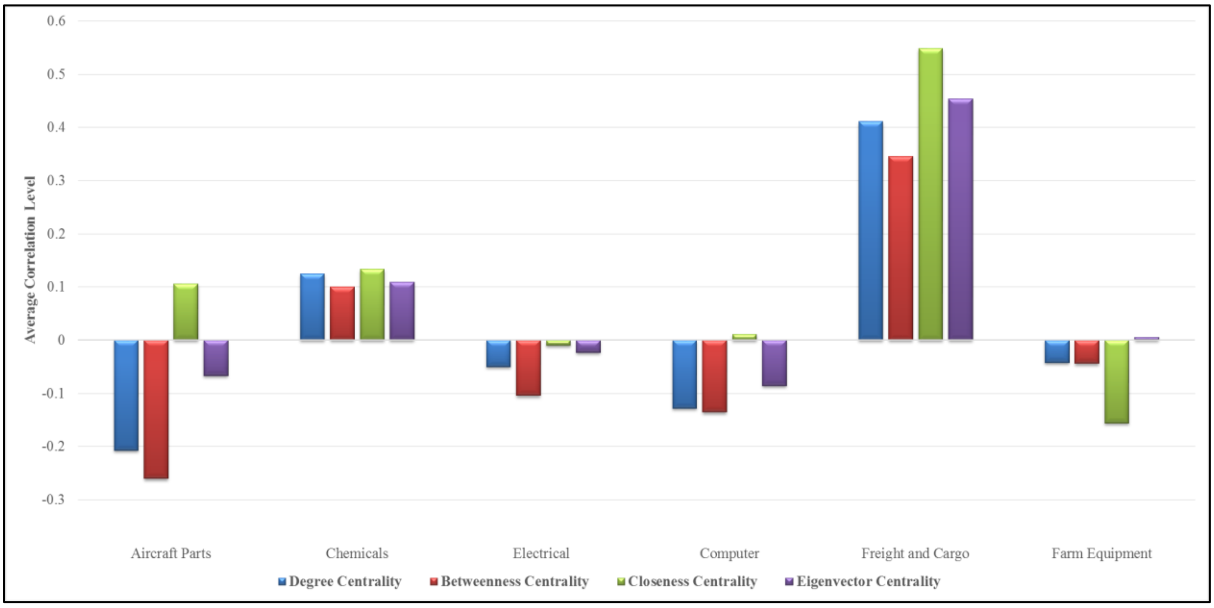

Figure 5: Average correlation levels of stage time with node level centrality metrics for SCNs in each sector

4. Discussion of results

The results that we have obtained have some important implications, which we discuss below.

4.1 What are the similarities and/or differences between the real world SCNs analysed in this study and other real world instances of complex networks?

It is interesting to note that although most real world scale free networks tend to have γ between 2 and 5, all the SCNs analysed in this study possess $\gamma < 2$. Similar findings were also reported by Orenstein (2016), where SCNs in food industry (General Mills, Kellogg's and Mondelez) and in retail markets (Nike, Lowes and Home Depot), included γ between 1.2-1.8. Although there are many other instances of real world networks which have $\gamma<2$, such cases have received much less attention in the network science literature (Seyed-Allaei et al., 2006).

Table 5 illustrates some instances of real world networks and their respective degree exponents, as reported by various studies.

Table 5: Real world instances of complex networks and their respective degree exponents

Network	Degree Exponent γ	Reference
Power-grid	4	Chung et al (2003)
Citations	3.0	Redner (1998)
WWW	2.7/2.1	Broder et al (2000)
Internet	2.5	Medina et al (2000)
Actors	2.3	Watts and Strogatz (1998)
Phone calls	2.1-2.3	Chung et al (2003)
E-mails	1.8	Ebel et al (2002)
Yeast Protein-Protein Net	1.5, 1.6, 1.7, 2.5	Chung et al (2003)
E.coli Metabolic Net	1.7, 2.2	
Yeast Gene Expression Net	1.4-1.7	
Gene functional interactions	1.6	
Dependency of software packages	1.6/1.4	Newman (2003)
Word web	1.5	i Cancho and Solé (2001)
Gnutella	1.4	Annexstein et al (2001)

* In the case of directed networks, the exponents is shown in the form of in/out

It is evident from the information presented in the above table, that most biological networks include γ <2. Indeed, this observation was first made by Aiello et al. (2000). In addition, some other virtual networks, such as emails, software packages and Gnutella also indicate γ<2.

Seyed-Allaei et al. (2006) reproduce the properties of networks with γ<2 by a simple prototype model. They note that the key feature of such networks is that their average degree grow linearly with the system size (see **Eq. 1**), which implies that link creation between nodes is inexpensive. The above is certainly true for almost all the networks with γ<2, outlined in the table above. For example, the nodes in Gnutella networks represent computers and the links are simply the logical connections between two computers which is essentially costless to create.

Seyed-Allaei et al. (2006) also notes that the model which generates networks with γ<2, also involves global interaction of nodes which require some sort of a global information exchange mechanism that is not part of the network itself. Using software package networks as an example, the authors stipulate that information exchanges in such a scenario occurs among programmers, where these programmers are responsible for the evolution of the system although they do not exchange information only through that system. Finally, the authors postulate that global interaction and information diffusion plays an essential role in establishing dense collaboration networks.

Indeed the above argument also holds in the case of SCNs – where contract establishment cost between firms, which are already within the SCN is almost negligible compared to involving a new firm into the system. Also, SCNs include global interaction of the firms involved – in terms of delivering the right amount of product at the right time to the right location. In fact, this global information exchange mechanism is not facilitated by the SCN itself – rather it relies upon other networks such as email, telephone and the internet.

Recent momentous advancements in network science has encouraged researchers to move beyond understanding and quantifying towards controlling complex networks. For instance, Nacher and Akutsu (2012), and Molnar et al. (2013) have examined the dynamical control of a network by considering a model of reduced complexity, where a minimum set of possible nodes dominates the whole system, called the minimum dominating set (MDS). An important finding on this front suggests that only a few nodes are needed to control the entire network if the power law degree exponent of the network is less than 2, whereas many nodes are required if it is larger than 2. When $\gamma < 2$, the number of connections in the network increase faster than the number of nodes, resulting in a highly heterogeneous network connectivity. Such networks tend to be dense and centralised with small average shortest path lengths, and therefore are inherently easy to dominate.

Given the vital role of coordination and control in SCNs, particularly due to largely unpredictable market demand conditions, it could be that SCNs self-organise themselves towards hub and spoke topologies, where $\gamma < 2$, so as to minimize the size of the MDS. Indeed, being able to control the overall network through control of a handful of firms can have remarkable advantages in an economic context.

So far, the concept of MDS has been applied to the design and control of various network systems such as mobile ad hoc networks, transportation routing and computer communication networks (Hanson et al., 2016). Given these advancements, it is tempting to envision the future SCNs being easily synchronized and controlled through a very few firms, owing to various applications of the MDS concept.

4.2 The viability of the Barabasi - Albert (BA) and similar as growth models for supply chain networks

Past studies of supply chain networks have relied upon BA model for benchmarking purposes (Thadakamalia et al., 2004, Zhao et al., 2011(a)). However, based on the results presented in this paper, it can understood that the BA model cannot sufficiently represent the growth mechanisms underlying SCNs, due to the following reasons:

1) The BA model or any generalised preferential attachment mechanisms generate networks with γ between 2 and 3. The SCNs presented in this study consistently display γ <2. As explained before, the threshold of $\gamma = 2$ has important characteristics which influence the evolution dynamics of networks.
2) The BA model cannot generate networks with communities which are prevalent across all SCNs presented in this paper.
3) The predominantly dis-assortative mixing as observed in the SCNs considered in this paper, is not a feature of networks generated by the BA model.

While SCNs in real world may not evolve through a single mechanism, it is possible to infer general growth and design principles from the global properties of existing SCNs. While most real world networks have been convincingly modelled with preferential attachment mechanism (Barabasi et al., 1999, Albert et al., 1999), it cannot explain exponents of power-law graphs less than 2.0. For biological networks, Chung et al. (2003) has demonstrated that partial duplication can produce power-law graphs with exponents less than 2, consistent with data on biological networks. However, such a growth mechanism in a SCN context, is yet to be comprehensively formulated, which needs to incorporate features such as dis-assortativenes and high modularity. Our study has highlighted the necessity of such a mechanism.

5. Conclusions and future directions

This study has presented an investigation of 26 manufacturing sector SCNs using network theoretic measurements. We have found that manufacturing SCNs are generally scale free with scale-free exponent <2. Also, these SCNs tend to demonstrate high modularity and disassortative mixing. We have also shown that most SCNs show weak correlation between stage cost and centrality, indicating that the more central the firm is in the SCN, the higher its stage cost contribution. Whereas, in terms of stage time, both positive and negative correlations with node centrality was observed, depending on the network under scrutiny.

Our work for the first time attempted to generalize the topological features of a large number of SCNs from a particular sector. It is notable that since we only considered relatively large networks, finite size effects are minimal. While some topological features were indeed network specific, the topological similarities between the networks was striking. Our work could be used as a bench mark for developing generalized growth mechanisms for supply chain networks in future.

6. References

1. A. Barabási, *Network Science*. (Cambridge University Press, 2016).
2. Aiello, W., Chung, F., & Lu, L. (2000, May). A random graph model for massive graphs. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing* (pp. 171-180). Acm.
3. Albert, R., Jeong, H., & Barabási, A. L. (1999). Internet: Diameter of the world-wide web. *Nature*, *401*(6749), 130-131.
4. Annexstein, F. S., Berman, K. A., & Jovanović, M. A. (2001, July). Latency effects on reachability in large-scale peer-to-peer networks. In *Proceedings of the thirteenth annual ACM symposium on Parallel algorithms and architectures* (pp. 84-92). ACM.
5. Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *science*, *286*(5439), 509-512.
6. Barabási, A. L., Albert, R., & Jeong, H. (2000). Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: Statistical Mechanics and its Applications*, *281*(1), 69-77.
7. Barabási, Albert-László, & Albert, Réka. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509-512.
8. Bianconi, Ginestra. "Processes on Networks". 2016. Lecture.
9. Borgatti, S. P., & Li, X. (2009). On social network analysis in a supply chain context*. *Journal of Supply Chain Management*, *45*(2), 5-22.
10. Brintrup, A., Ledwoch, A., & Barros, J. (2016). Topological robustness of the global automotive industry. *Logistics Research*, *9*(1), 1-17.
11. Brintrup, A., Wang, Y., & Tiwari, A. (2015). Supply networks as complex systems: a network-science-based characterization.
12. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A. & Wiener, J. (2000). Graph structure in the web. *Computer networks*,*33* (1), 309-320.
13. Chan, S. (2001, November). Complex adaptive systems. In *Research seminar in engineering systems* (Vol. 31).
14. Choi, T. Y., & Hartley, J. L. (1996). An exploration of supplier selection practices across the supply chain. *Journal of operations management*, *14*(4), 333-343.
15. Choi, T. Y., & Hong, Y. (2002). Unveiling the structure of supply networks: case studies in Honda, Acura, and DaimlerChrysler. *Journal of Operations Management*, *20*(5), 469-493.
16. Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of operations management*, *19*(3), 351-366.
17. Chung, F., Lu, L., Dewey, T. G., & Galas, D. J. (2003). Duplication models for biological networks. *Journal of computational biology*, *10*(5), 677-687.
18. Cohen, R., Erez, K., Ben-Avraham, D., & Havlin, S. (2000). Resilience of the Internet to random breakdowns. *Physical review letters*, *85*(21), 4626.
19. Costa, L. D. F., Rodrigues, F. A., Travieso, G., & Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics*, *56*(1), 167-242.
20. Danon, L., Arenas, A., & Díaz-Guilera, A. (2008). Impact of community structure on information transfer. *Physical Review E*, *77*(3), 036103.
21. Dong, J., & Horvath, S. (2007). Understanding network concepts in modules.*BMC systems biology*, *1*(1), 24.
22. Ebel, H., Mielsch, L. I., & Bornholdt, S. (2002). Scale-free topology of e-mail networks. *Physical review E*, *66*(3), 035103.
23. Erdős, P.; Rényi, A. (1959). "On Random Graphs. I". *Publicationes Mathematicae 6*: 290–297.

24. Erez, M., & Gati, E. (2004). A dynamic, multi-level model of culture: from the micro level of the individual to the macro level of a global culture. *Applied Psychology*, *53*(4), 583-598.
25. Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 35-41.
26. Gjerdrum, J., Shah, N., & Papageorgiou, L. G. (2001). A combined optimization and agent-based approach to supply chain modelling and performance assessment. *Production Planning & Control*, *12*(1), 81-88.
27. Goyal, S. (2012). *Connections: an introduction to the economics of networks*. Princeton University Press.
28. Hanson, S. J., Mastrovito, D., Hanson, C., Ramsey, J., & Glymour, C. (2016). Scale-Free Exponents of Resting State provide a Biomarker for Typical and Atypical Brain Activity. *arXiv preprint arXiv:1605.09282*.
29. Hearnshaw, E. J., & Wilson, M. M. (2013). A complex network approach to supply chain network theory. *International Journal of Operations & Production Management*, *33*(4), 442-469.
30. Hu, S. J., Zhu, X., Wang, H., & Koren, Y. (2008). Product variety and manufacturing complexity in assembly systems and supply chains. *CIRP Annals-Manufacturing Technology*, *57*(1), 45-48.
31. i Cancho, R. F., & Solé, R. V. (2001). The small world of human language.*Proceedings of the Royal Society of London B: Biological Sciences*,*268*(1482), 2261-2265.
32. Jain, V., Wadhwa, S., & Deshmukh, S. G. (2009). Select supplier-related issues in modelling a dynamic supply chain: potential, challenges and direction for future research. *International Journal of Production Research*, *47*(11), 3013-3039.
33. Jennings, N. R. (2000). On agent-based software engineering. *Artificial intelligence*, *117*(2), 277-296.
34. Julka, N., Srinivasan, R., & Karimi, I. (2002). Agent-based supply chain management—1: framework. *Computers & Chemical Engineering*, *26*(12), 1755-1769.
35. Jun, W. U., Yue-Jin, T. A. N., Hong-Zhong, D. E. N. G., & Da-Zhi, Z. H. U. (2007). Heterogeneity of Scale-free Networks. *Systems Engineering-Theory & Practice*, *27*(5), 101-105.
36. Kim, Y., Chen, Y. S., & Linderman, K. (2015). Supply network disruption and resilience: A network structural perspective. *Journal of Operations Management*, *33*, 43-59.
37. Kim, Y., Choi, T. Y., Yan, T., & Dooley, K. (2011). Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management*, *29*(3), 194-211.
38. Kim, Y., Choi, T. Y., Yan, T., & Dooley, K. (2011). Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management*, *29*(3), 194-211.
39. Kito, T., Brintrup, A., New, S., & Reed-Tsochas, F. (2014). The structure of the Toyota supply network: an empirical analysis. *Saïd Business School WP*, *3*.
40. Kremers, E. A. (2013). *Modelling and simulation of electrical energy systems through a complex systems approach using agent-based models*. KIT Scientific Publishing.
41. Ladyman, J., Lambert, J., & Wiesner, K. (2013). What is a complex system?.*European Journal for Philosophy of Science*, *3*(1), 33-67.
42. Li, G., Gu, Y. G., & Song, Z. H. (2013). Evolution of cooperation on heterogeneous supply networks. *International Journal of Production Research*,*51*(13), 3894-3902.
43. Li, Y. (2014). Networked Analysis Approach of Supply Chain Network. *Journal of Networks*, *9*(3), 777-784.
44. Macal, C. M., & North, M. J. (2005, December). Tutorial on agent-based modeling and simulation. In *Proceedings of the 37th conference on Winter simulation* (pp. 2-15). Winter Simulation Conference.

45. Mari, S. I., Lee, Y. H., Memon, M. S., Park, Y. S., & Kim, M. (2015). Adaptivity of Complex Network Topologies for Designing Resilient Supply Chain Networks. *International Journal of Industrial Engineering*, *22*(1), 102-116.
46. Medina, A., Matta, I., & Byers, J. (2000). On the origin of power laws in Internet topologies. *ACM SIGCOMM computer communication review*, *30*(2), 18-28.
47. Mitchell, M. (2006). Complex systems: Network thinking. *Artificial Intelligence*,*170*(18), 1194-1212.
48. Molnár, F., Sreenivasan, S., Szymanski, B. K., & Korniss, G. (2013). Minimum dominating sets in scale-free network ensembles. *Scientific reports*, *3*.
49. Nacher, J. C., & Akutsu, T. (2012). Dominating scale-free networks with variable scaling exponent: heterogeneous networks are not difficult to control. *New Journal of Physics*, *14*(7), 073005.
50. Newman, M. E. (2002). Assortative mixing in networks. *Physical review letters*, *89*(20), 208701.
51. Newman, M. E. (2003). The structure and function of complex networks.*SIAM review*, *45*(2), 167-256.
52. Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113.
53. Nguyen, K., & Tran, D. A. (2012). Fitness-Based Generative Models for Power-Law Networks. In *Handbook of Optimization in Complex Networks* (pp. 39-53). Springer US.
54. Nishikawa, T., Motter, A. E., Lai, Y. C., & Hoppensteadt, F. C. (2003). Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize?. *Physical review letters*, *91*(1), 014101.
55. Noldus, R., & Van Mieghem, P. (2015). Assortativity in complex networks.*Journal of Complex Networks*, *3*(4), 507-542.
56. Orenstein, P. (2016, February). How does Supply Network Evolution and its Topological Structure Impact Supply Chain Performance?. In *2016 Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO)* (pp. 562-569). IEEE.
57. Ou, R., Yang, J., Chang, J., & Xie, W. (2009). On heterogeneity of complex networks in the real world. In *Cutting-Edge Research Topics on Multiple Criteria Decision Making* (pp. 213-219). Springer Berlin Heidelberg.
58. Parunak, H. V. D., Savit, R., & Riolo, R. L. (1998, January). Agent-based modeling vs. equation-based modeling: A case study and users' guide. In *Multi-agent systems and agent-based simulation* (pp. 10-25). Springer Berlin Heidelberg.
59. Pathak, S. D., Day, J. M., Nair, A., Sawaya, W. J., & Kristal, M. M. (2007). Complexity and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective*. *Decision Sciences*, *38*(4), 547-580.
60. Pereira, T. (2010). Hub synchronization in scale-free networks. *Physical Review E*, *82*(3), 036201.
61. Perera S, Bliemer M. C. J., and Bell M. G. H. (2016). 'Resilience Characteristics of Supply Network Topologies Generated by Fitness Based Growth Models', *95th Annual Meeting of the Transportation Research Board TRB*, Washington, D.C., United States, 14th January 2016
62. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabási, A. L. (2002). Hierarchical organization of modularity in metabolic networks.*science*, *297*(5586), 1551-1555.
63. Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, *4*(2), 131-134.
64. Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, *52*(3), 1059-1069.
65. Ruhnau, B. (2000). Eigenvector-centrality—a node-centrality?. *Social networks*, *22*(4), 357-365.

66. Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, *31*(4), 581-603.
67. Seyed-Allaei, H., Bianconi, G., & Marsili, M. (2006). Scale-free networks with an exponent less than two. *Physical Review E*, *73*(4), 046113.
68. Sheffi, Y. (2001). Supply chain management under the threat of international terrorism. *The International Journal of logistics management*, *12*(2), 1-11.
69. Sheffi, Y., & Rice Jr, J. B. (2005). A supply chain view of the resilient enterprise. *MIT Sloan management review*, *47*(1), 41.
70. Song, C., Havlin, S., & Makse, H. A. (2006). Origins of fractality in the growth of complex networks. *Nature Physics*, *2*(4), 275-281.
71. Swaminathan, J. M., Smith, S. F., & Sadeh, N. M. (1998). Modeling supply chain dynamics: A multiagent approach*. *Decision sciences*, *29*(3), 607-632.
72. Tesfatsion, L. (2002). Agent-based computational economics: Growing economies from the bottom up. *Artificial life*, *8*(1), 55-82.
73. Thadakamalla, H. P., Raghavan, U. N., Kumara, S., & Albert, R. (2004). Survivability of multiagent-based supply networks: a topological perspective. *Intelligent Systems, IEEE, 19*(5), 24-31.
74. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440-442.
75. Wen, L., & Guo, M. (2012). Statistic Characteristics Analysis of Directed Supply Chain Complex Network. *International Journal of Advancements in Computing Technology*, *4*(21).
76. Willems, S. P. (2008). Data set-real-world multiechelon supply chains used for inventory optimization. *Manufacturing & Service Operations Management*, *10*(1), 19-23.
77. Xu, M., Wang, X., & Zhao, L. (2014). Predicted supply chain resilience based on structural evolution against random supply disruptions. *International Journal of Systems Science: Operations & Logistics*, *1*(2), 105-117.
78. Xu, N. R., Liu, J. B., Li, D. X., & Wang, J. (2016). Research on Evolutionary Mechanism of Agile Supply Chain Network via Complex Network Theory.*Mathematical Problems in Engineering*, *2016*.
79. Xuan, Q., Du, F., Li, Y., & Wu, T. J. (2011). A framework to model the topological structure of supply networks. *Automation Science and Engineering, IEEE Transactions on*, *8*(2), 442-446.
80. Yi, C. Q., Meng, S. D., Zhang, D. M., & Li, J. D. (2013). Managing Disruption Risks in Supply Chain Based on Complex Networks. *Journal of Convergence Information Technology*, *8*(5).
81. Zhao, K., Kumar, A., & Yen, J. (2011b). Achieving high robustness in supply distribution networks by rewiring. *Engineering Management, IEEE Transactions, 58*(2), 347-362.
82. Zhao, K., Kumar, A., Harrison, T. P., & Yen, J. (2011a). Analyzing the resilience of complex supply network topologies against random and targeted disruptions. *Systems Journal, IEEE, 5*(1), 28-39.