

Received July 23, 2020, accepted August 6, 2020, date of publication August 11, 2020, date of current version September 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3015910

# Topology of International Supply Chain Networks: A Case Study Using Factset Revere Datasets

MAHENDRA PIRAVEENAN<sup>1</sup>, HONGZE JING<sup>1</sup>, PETR MATOUS<sup>1</sup>, AND YASUYUKI TODO<sup>2</sup>

<sup>1</sup>Faculty of Engineering, The University of Sydney, Sydney, NSW 2006, Australia

<sup>2</sup>Graduate School of Economics, Waseda University, Shinjuku City 169-8050, Japan

Corresponding author: Mahendra Piraveenan (mahendrarajah.piraveenan@sydney.edu.au)

The work of Yasuyuki Todo was supported in part by the KAKENHI under Grant 18KK0045 and Grant 18J03642.

**ABSTRACT** International supply chain networks play a prominent role in shaping the economic outlook of the world. It has been a recent trend to analyse the topology of supply chain networks in order to gain a wholistic understanding about the interdependencies of firms in this regard. In this work, we undertake an extensive structural and topological analysis of the supply chain networks constructed from the Factset Revere dataset. The dataset is provided by FactSet Research Systems Inc. that captures global supply chain relationships between companies. The dataset consists of 154, 862 companies from 216 countries, with 1,571, 949 supply relationships among them. In addition to considering the global network, we also analysed country-specific networks of ten countries, which are the most significant nations represented in the dataset. The analysis revealed that all supply chain networks studied were relatively sparse scale-free networks, with scale-free exponents ranging from 1.0 to 2.0. In terms of centrality analysis, quite predictably, large multinational corporates dominated. Comparing the centrality values of firms in terms of the global vs the country-specific networks, two classes of firms were found where the difference in centrality was significant. The first group was small firms with locally-centered business operations, such as Volunteers of America, New York State Teachers Retirement System, CarePlus Health Plan etc, where the country-based centrality scores and the rankings based on them were significantly more prominent than the global equivalent. The second group was firms with specific countries of origin which register themselves in other countries, such as China Shengda Packaging Group Inc (registered in US), Chinacast Education Corps (registered in the US), and China Biologic Products Inc (registered in the US). These firms all had significantly higher global centrality scores compared to country-based centrality scores. Overall, however, it was found that there was strong correlation between global centrality-based ranking and country-specific centrality ranking of firms. This indicated that in general, firms which are important to the global supply chain network are also important to the supply chain networks of individual countries. Studying the community structure of the supply chain networks, we identified twelve dominant communities, many of which had significant correlations with particular industries or countries. Some of these communities were made of firms primarily from a pair of countries, or had other interesting features. Therefore, the topological analysis of the supply chain networks created from this large dataset gives interesting insights about how the international supply chain networks are structured, and how they operate.

**INDEX TERMS** Supply chains, internationalisation, complex networks, domestic protectionism.

## I. INTRODUCTION

International supply chains are becoming increasingly interconnected, forming Supply Chain Networks (SCNs), which

The associate editor coordinating the review of this manuscript and approving it for publication was Bohui Wang<sup>1</sup>.

display features of complex networks [1]–[10]. Therefore, it becomes necessary to study the topological features of such networks, in order to understand their interdependency, evolution, robustness and resilience. A number of recent studies have looked at various supply chain networks, analysing their topological structure and growth patterns [2]–[5], [7]–[10].

In this paper also we attempt such a task, with a specific Supply Chain Network (and a number of country specific subnetworks) whose topological properties have not been studied extensively before.

The SCNs considered here are generated from the Factset Revere dataset, which is a global dataset containing information about mainly publicly listed firms. We analyse the global supply chain network and ten country specific networks generated from it, considering basic topological metrics, such as centrality measures, clustering, path length, measures of scale-freeness etc. We also consider the relationship between country-specific and global rankings of firms based on these metrics, identifying firms which have a global importance not reflected on the country-specific subnetwork, and vice-versa. Further, we undertake a community structure analysis of the global network and country-specific subnetworks, highlighting the interplay between specific industries, trade clusters, and international financial and political relationships. Our results show that the sparse scale-free networks created by the supply chains, at international as well as domestic levels, display strong correlations in terms of global and domestic centrality rankings; but there are groups of firms which do not adhere to this pattern, and display a higher prominence in the domestic network compared to the global network or vice versa. Our analysis further highlights some interesting features in terms of community formation between firms belonging to different economic power houses, such as China, the US, and Japan. The conducted analysis provides useful insights about the interconnectedness of the international supply chain networks, how they operate, and how they evolve.

This paper makes three important contributions to the modelling and analysis of supply chain networks: firstly, it provides insight into the topology of supply-chain networks, and establishes that they are typically scale-free networks with scale-free exponents which are smaller than those found in most other real world scale-free networks. Secondly, it undertakes comparative centrality analysis which compares the local and global importance of firms, and establishes that there is strong correlation between these in most cases, and highlights the nature of companies which violate this general rule: that is, the paper sheds light into the features needed for companies to be more locally central, or be more globally central, in supply chains. Thirdly, it undertakes community analysis to highlight the interplay between country-based community formation and industry-based community formation among firms. Overall, the contribution of the paper is to shed light on the topological features of typical supply and inter-firm networks, and to demonstrate how such structural features are necessary for these networks to perform their intended functions.

The rest of the paper is organised as follows: section II provides a general theoretical understanding of, and state-of-the-art for, the topological analysis of supply chain networks. In section III, a description of the dataset is provided (subsection III-A), followed by a description of the network

creation (subsection III-B), and a list of definitions of the topological metrics and measures used in this work (subsection III-C). Section IV describes the analysis that was conducted and the results obtained. Section V offers a broad discussion, including the advantages and shortcomings of the presented approach, and the novelty and importance of the presented results. Finally, section VI offers a summary of conclusions, and a list of potential future research directions.

## II. BACKGROUND

Two important streams of theoretical views can be found in the literature on systems of industrial production. The first stream is based on conventional market theory and posits that markets consist of arms-length transactions between economic actors and it is efficient for them to distribute their transactions across independent partners. This position is also consistent with resource-dependency theory [11], which recommends maintaining unconstrained access to a large number of competing and substitutable partners.

As a reaction to such individualistic views of interorganizational relationships, economic sociologists have proposed a social network view of systems of industrial production. Social network literature focuses on relational considerations between organizations, which were neglected in the original approaches or considered only in their dyadic form in transaction cost economics. These network conceptualizations of systems of industrial production have become increasingly popular as it has become more accepted that organizational behaviour and performance are not well explained by atomistic individualized approaches.

No business link exists in isolation. Two firms may be more likely to establish a partnership if they already have some trusted partners in common. White [12] proposed that in competitive production markets, firms gravitate toward dense cliques of producers watching each other. According to Uzzi [13], it is beneficial for firms to embed their transactions in a dense network of partnerships with other organizations. Close-knit groups can allow for the emergence of trust, free information flows, resource pooling, and collective problem-solving. Uzzi [13] also predicted that within the same context, firms will imitate successful networking strategies therefore converge toward similar arrangements. He *et al.* [14] showed that dense networks of businesses in China help diffuse shocks and decrease firm's risk of default.

The emergence of clustering in business networks has not been thoroughly explored in the international context in which different mechanisms and constraints may be present compared to the domestic partners. This can be linked to a lack of suitable data and traditional predominance of firm-centric approaches to studies of international business. However, no firm is in complete control of their international networks, let alone the networks of their partners. International business relationships emerge from interactions of all interdependent actors within their business environment. Chandra and Wilkinson [15] argued that because of complex feedback loops, international interfirm partnerships

are impossible to predict separately but aggregate structural patterns of internationalization can be explained.

It is in this backdrop that we analyse the topology of international supply chain networks using the Factset Revere dataset, which is more fully described in the next section. Building on network-based theories of international relationships, in this study, we pay particular attention to the complex interactions between firms in the world's three largest economies: (1) the United States, (2) China, and (3) Japan, as these have potentially large effects on the rest of the global economy. It should be noted here that a recent paper by Piraveenan *et al.* [10] analysed one particular aspect - namely assortativity - of the topology of supply chain networks generated from the Factset Revere dataset. They defined and employed a range of customised assortativity measures in their analysis, and showed that these networks show assortativity in terms of country and level of internationalisation: that is, firms have a slight preference to make supply chain relationships with other firms from the same country, as well as firms which have undergone a similar level of internationalisation. The current study extends the work of Piraveenan *et al.* by analysing the supply chain networks generated from Factset Revere data using multiple topological metrics and concepts. The broad analysis is meant to shed light on clearly identifiable patterns in the topologies of the global as well as country-specific supply chain networks, and act as a catalyst for generating and analysing in-depth questions related to specific topological features, such as the one about assortativity explored already in Piraveenan *et al.* [10].

It should be noted that a number of recent studies have done work related to this paper: nevertheless, this study offers a unique perspective as described below. For example, Perera *et al.* [7] undertook a detailed topological analysis of supply chain networks, but focussed on the manufacturing industry only. They focused on comparing the differences in topologies of undirected contractual relationships (UCR) and directed material flow (DMF) supply chain networks, and used datasets collected by Willems [16] and Parhi [17]. While their analysis focussed on a range of basic topological metrics, they have not focussed on global vs local centrality or community structure as this study does. Similarly, Pathak *et al.* [18] also presented a basic topological analysis, however their focus was on how these topological metrics evolve over time. Hearnshaw *et al.* [19] also focussed on topological analysis, but their work was more focussed on providing a groundwork for, and justifying the use of, complex network science to analyse supply chain networks. Nayar and Vidal [20] focussed specifically on the robustness of supply chain networks against 'disruptions', and used a multi-agent system to model the supply networks. Thadakamaila *et al.* [1] also focussed on the 'survivability' (robustness) of supply chain networks modelled through multi-agent systems. Hou *et al.* [21] also analysed the robustness and resilience of supply chain network topologies using multi-agent system modelling, but they specifically looked

at how trust between firms affects the topological evolution of such networks. The focus of Mari *et al.* [4] was similar, in that they analysed how classical complex network topologies could be adapted to design resilient (robust) supply chain networks. Similarly, Zhao *et al.* [22] also focussed on the robustness aspect of supply chain networks, analysing their resilience against random and targeted disturbances. Pero *et al.* [23] focussed on the relationship between structure and function in the supply chain network context. Compared to such existing works, which primarily focus on robustness and resilience and/or the relationship between form and function, the current work is unique in that it focusses, after a succinct topological analysis, mainly on the questions of international importance vs local importance, and understanding the features of, and driving forces behind, community and cluster formation. In short, the current work is unique because it is presented in the backdrop of increasing internationalisation, and attempts to shed light on those topological aspects of supply chain networks which can only be interpreted by comparing international, country-based, and industry-based viewpoints.

### III. METHODS

#### A. DATASET

This study uses 2016 FactSet Revere data [24] that includes 154,862 publicly listed firms and public institutions, including those in the United States (38,708), China (140,580), Japan (74,110), the United Kingdom (68,140), Canada (42,870), India (37,480), Australia (31,550), France (30,050), Singapore (17,360), and Russia (12,020). FactSet collects interfirm relationship data from primary public sources such as investor reports and SEC 10-K annual filings, investor presentations, and press releases. Both relationships disclosed by the company and reverse relationships which are reported by their partners are captured in the data. FactSet Analysts continuously monitor and review the quality of the data.

The FactSet dataset includes 129 categories classifying each firm's main industry type. To investigate the general trends in production network structures across major industrial categories, following Piraveenan *et al.* [10], we manually matched the original categories with the high level categories of the Standard Industrial Classification system, which is commonly used in the US and the UK, especially by government agencies [25]. The system has ten broad industrial "divisions", which in turn are composed of 202 "industry groups", according to which we manually matched the FactSet industry categories to the high level divisions. These ten divisions are: (1) agriculture, forestry and fishing, (2) mining, (3) construction, (4) manufacturing, (5) transportation, communications, electric, gas and sanitary services, (6) whole sale trade, (7) retail trade, (8), finance, insurance and real estate, (9) services, and (10) public administration. This broad industrial classification of the Factset Revere dataset is shown in Fig. 1.

## B. NETWORK CREATION

As mentioned, the Factset Revere dataset that we considered consists of 154, 862 companies from 216 countries, with 1,571, 949 supply relationships among them. We constructed a global network, as well as ten country specific networks from this dataset, representing the US, Russia, China, India, Australia, Japan, Singapore, France, Canada and the UK. We also constructed an EU network, which considered all European Union countries at the time of purchase of this dataset, including the UK. Each country specific network would only include firms from that particular country and the supply relationships among them. Thus, the sum of all links from the country specific networks would be a lot smaller than the total number of links in the global network, even if we consider the country specific networks of all countries represented in the dataset. In our case we only consider 10 countries, therefore the total number of firms in these ten networks is 62,456, while the total number of links is 701,277. Thus, roughly about half the nodes and links in the global network are represented in the ten country-specific networks. Some properties of the ten country specific networks we constructed are shown in Fig. 2.

Of course, the country specific networks had a lot of singleton nodes, but most of these are not singletons in the global supply chain network.

## C. NETWORK TOPOLOGY METRICS

### 1) SCALE-FREENESS MEASURES

Scale-free networks are ubiquitous in real world [26]. In a scale-free network, the degree distribution follows a power law, and the probability of a node to have a degree of  $k$  is given by [27]–[30]:

$$p_k = Ak^{-\gamma} \quad (1)$$

where  $A$  is a constant and  $\gamma$  is the power law exponent (also referred as scale-free exponent). A higher value of  $\gamma$  results in a degree distribution with a steeper slope, while a lower value of  $\gamma$  results in a flatter degree distribution.

To quantitatively measure the ‘scale-freeness’ of a particular network, the  $R^2$ -correlation of the degree distribution to a power law can be used. To compute this, the degree distribution of the given network (in log-log scale) should be plotted and a straight-line should be fitted to this distribution (in the form of  $\log(p_k) = -\gamma \log(k) + \log(A)$ ) and the  $R^2$ -correlation (also called the *correlation of determination*) of the fit should be computed. The  $R^2$ -correlation is computed as:

$$R^2 = 1 - \frac{\sum_i (y_i - \bar{y})^2}{\sum_i (y_i - f_i)^2} \quad (2)$$

where  $y_i$  are y-values of the data points,  $\bar{y}$  is the mean y-value of the data points, and  $f_i$  are the values returned by the fitted function for data points  $i$  [31]. This quantity is briefly called ‘scale-free correlation’ elsewhere in the text, to mean that it is the  $R^2$ -correlation measuring the scale-freeness of a given network.

### 2) PATH LENGTH AND CLUSTERING MEASURES

The average path length, or the characteristic path length, of a network is simply the average of the length of all paths in that network, computed in terms of number of nodes on each path.

The clustering coefficient of a node represents the ratio between the number of links between the neighbours of that node, and the number of all possible links between those neighbours. It is defined as [32]:

$$c_i = \frac{2y_i}{k_i(k_i - 1)} \quad (3)$$

where  $k_i$  is the degree of node  $i$ , and  $y_i$  is the number of edges between the neighbours of node  $i$ . The network clustering coefficient  $\bar{C}$  is defined as the average of the clustering coefficients of all nodes in that network.

### 3) CENTRALITY MEASURES

A host of centrality measures have been proposed to analyze complex networks, especially in the domain of social network analysis. The simplest of these perhaps is the degree centrality, sometimes just called degree, of a node. A node’s degree is simply the number of links it has with other nodes in the network, and therefore gives some indication about how important that node is to the network.

A family of betweenness measures have been proposed [33]–[40] to measure a node’s importance as a conduit of information flow in a network. The first and perhaps the most well known measure of these is the classical betweenness centrality measure proposed by [33]. Betweenness Centrality measures the fraction of shortest paths that pass through a given node, averaged over all pairs of node in a network. It is formally defined, for a directed graph, as

$$BC(v) = \frac{1}{(N-1)(N-2)} \sum_{s \neq v \neq t} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}} \quad (4)$$

where  $\sigma_{s,t}$  is the number of shortest paths between source node  $s$  and target node  $t$ , while  $\sigma_{s,t}(v)$  is the number of shortest paths between source node  $s$  and target node  $t$  that pass through node  $v$ .

Closeness Centrality [34], [41] is a measure of how close a network is, on average, to the rest of the nodes in terms of shortest paths. It essentially measures the average geodesic distance between a given node and all other nodes in the network. It is defined as

$$CC(v) = \frac{1}{\sum_{i \neq v} d_g(v, i)} \quad (5)$$

where  $d_g(v, i)$  is the shortest path (geodesic) distance between nodes  $v$  and  $i$ . Note that the average is ‘inverted’ so that the node which is ‘closest’ to all other nodes will have the highest measure of closeness centrality.

### 4) COMMUNITY STRUCTURE

The community structure of a network is not a topological metric as such, but a network when partitioned into communities can yield useful information about its structure and

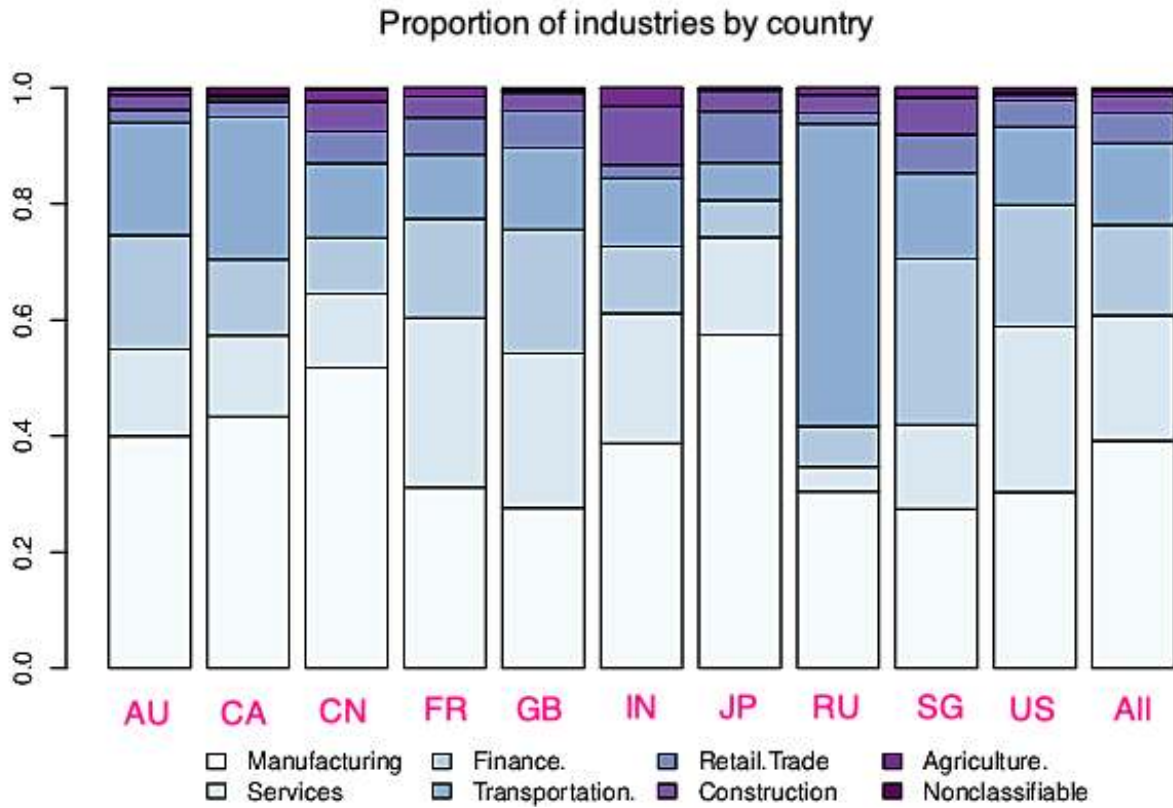


FIGURE 1. The industry classification of the firms represented in the Factset Revere dataset based on countries.

	AU	CA	CN	FR	UK	IN	JP	RU	SG	US
Nodes (global)	3155	4287	14058	3005	6814	3748	7411	1202	1736	38708
Nodes (country network)	2133	2406	9764	1830	3400	2926	6388	722	879	32008
Links (country network)	7953	10111	30214	8964	16677	8939	47745	3815	2689	564170
Average links per node(country network)	7.487	8.408	6.189	9.797	9.81	6.11	14.948	10.868	6.118	35.282
Average degree (country network)	4.067	3.834	3.032	4.115	4.491	3.735	3.735	4.335	3.197	11.57
Network Clustering coefficient	0.066	0.074	0.02	0.074	0.064	0.051	0.051	0.117	0.064	0.121
Average (characteristic) path length	4.344	5.118	5.18	4.243	4.257	4.383	4.383	3.92	4.09	4.06
Scale-free exponent	1.349	1.349	1.812	1.079	1.272	1.389	1.308	0.953	1.18	1.25
Scale-free fitness (R <sup>2</sup> value)	0.873	0.882	0.871	0.894	0.84	0.87	0.909	0.791	0.825	0.86

FIGURE 2. Basic topological metrics of the supply chain network for the ten countries of interest. Note well: Nodes (global) represents the number of firms belonging to the given country in the dataset, while Nodes (country network) represent the number of non-singleton nodes in the country-specific subnetwork. Therefore the difference between these quantities represent firms which do not have supply relationships with another firm from the same country. Links (country network) represents the number of links in the country network. The “average links per node (country network)” represent the average number of supply relationships a firm has with other firms from the same country, while the “average degree (country network)” represents the average number of firms that a firm has supply relationships with in the relevant country. The former is always higher than the later because sometimes a pair of firms may have more than one supply relationship between them in the dataset (multiple links between a pair of nodes are possible). Note well also that the “average degree (global)” and “average links per node (global)” are not shown in this table, because they are not country-specific.

function. A network is said to have community structure if the nodes of the network can be easily grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally [42]–[44]. Communities can be non-overlapping or overlapping, but in both cases, the general idea is that pairs of nodes are more likely to be connected if they are both members of the same community or communities, and less likely to be connected if they do not share communities.

There are several algorithms which can be used to partition a network into communities, among which the Louvain method [45] is prominent. The Louvian method optimises the modularity of a network. Modularity is formally defined, for a weighted graph, as:

$$Q = \frac{1}{(2m)} \sum_{i,j} [A_{i,j} - \frac{k_i k_j}{2m}] \delta(c_i, c_j) \tag{6}$$

Company name	Betweenness Centrality	Company name	Closeness Centrality	Company name	Degree
General Electric Co	0.73406532	General Electric Co	0.35677742	United States of America Government	16375
United States of America Government	0.53010315	International Business Machines Corp IBM	0.34742228	International Business Machines Corp IBM	9403
International Business Machines Corp IBM	0.3296225	Hewlett-Packard Co	0.34250908	Microsoft Corp	8451
Microsoft Corp	0.24981269	United States of America Government	0.34146434	General Electric Co	8402
Hewlett-Packard Co	0.21653895	Microsoft Corp	0.33605926	Hewlett-Packard Co	7938
Citigroup Inc	0.19768674	Oracle Corp	0.33591942	Oracle Corp	5615
JPMorgan Chase & Co	0.19213377	Honeywell International Inc	0.3318313	Johnson & Johnson	5392
Wal-Mart Stores Inc	0.18227083	3M Co	0.32852123	Time Warner Inc	5001
Oracle Corp	0.14566676	Intel Corp	0.32693531	Pfizer Inc	4724
Honeywell International Inc	0.12145201	Boeing Co	0.32319649	Cisco Systems Inc	4642

**FIGURE 3.** Top ten firms in the United States with highest degree, betweenness and closeness centrality values. The centrality values are based on the global interfirm network, not the local country network.

where  $A_{i,j}$  represents the edge weights between nodes  $i, j$ ,  $k_i, k_j$  are the sum of edge weights attached to nodes  $i, j$  respectively,  $m$  is the sum of all edge weights in the graph, and  $c_i, c_j$  are node communities. The Louvian method is a greedy optimisation method that optimises the quantity  $Q$ , and runs with a time complexity of  $O(N \log(N))$  where  $N$  is the number of nodes in the network considered.

## IV. EXPERIMENTS AND RESULTS

### A. BASIC TOPOLOGY

We first considered the basic topology of the global network, as well as the country networks of the ten countries we have selected as mentioned above. We analysed the network in terms of basic topological metrics, namely average degree, characteristic path length, network clustering coefficient, scale-free fitness as well as scale-free exponent (naturally, the scale-free exponent is meaningful only in networks where the scale-free fitness is relatively high). These metrics have already been defined in section II. The results of this basic topological analysis are shown in Fig. 2.

From Fig. 2 we may observe that the country networks are relatively sparse: except the US country network, every other country network has an average number of links per node within six to fifteen, depending on the size of the network, which translates to a more or less consistent network density of between 0.104% and 0.117%, which is quite sparse. The US country network has a domestic average number of links per node of 35.25 which corresponds to a network density of 0.055%, given the relatively large number of nodes, which is even more sparse. We may also observe that all the country networks are scale-free, with scale-free fitness values ranging from 79% to 91% (in section II we have described how the scale-free fitness can be calculated). The scale-free exponents are typically around unity, which is smaller than most real world networks, which have a scale-free exponent between 2.0 and 3.0 [46]. In the case of the country network of Russia, the scale-free exponent is less than one, therefore this network

does not have the typical power law distribution associated with scale-free networks [46]. The characteristic path lengths of the networks are between 3.92 and 5.18, which are on the order of the logarithm of the network size. However, the clustering coefficients are not high enough to attribute small-world properties to these networks.

### B. CENTRALITY AND RANK-BASED CORRELATIONS

We then computed the centrality distributions of all the country networks. For this, we considered (i) node degree (ii) betweenness centrality (iii) closeness centrality. Figures 3, 4, and 5, show the ten companies that had the highest centrality values by each of these measures in the US, Japan and Australia, which are among the largest country networks in the dataset (the corresponding results for the rest of the countries considered are similar, but have not been presented due to space restrictions). Note well that even though the countries were considered individually, the centrality values were calculated from the single global interfirm network. From these figures we could identify the companies which are most 'central' (and thus, important) in each country. For example, the most influential companies in the United States, according to this dataset, are General Electric, the US Government, IBM, Microsoft Corporation, and Hewlett-Packard corporation when betweenness centrality is used. It might be noted that, considering the size of the dataset (38708 US companies are present in the dataset), the top ten list changes little when different centrality measures are employed. General Electric, the US Government, IBM, Microsoft Corporation, Hewlett-Packard Corporation, and Oracle make the top-ten list regardless of the centrality measure used. A similar observation can be made from Fig. 4 representing Japan, and Fig. 5 representing Australia. For example, Hitachi, Mitsubishi, Sony, Sumimoto, Toyota, Toshiba, Panasonic, and Fujitsu make the top ten list of Japanese firms regardless of the centrality metric used, and BHP Billiton, Australian Government, Worley Parsons, Rio Tinto, Telstra, all make

Company name	Betweenness Centrality	Company name	Closeness Centrality	Company name	Degree
Hitachi	0.13271028	Hitachi	0.32739339	Sony Corporation	4688
Mitsubishi Corp	0.12668559	Sony Corporation	0.32343764	Hitachi	3707
Sony Corporation	0.11953511	Toyota Motor Corporation	0.32149607	Fujitsu Limited	3257
Sumitomo Corporation	0.10544705	Panasonic Corp	0.32049125	Panasonic Corp	3188
Mitsubishi Heavy Industries	0.10120329	Toshiba Corporation	0.32004019	Toshiba Corporation	3028
Mitsui & Co Ltd	0.09840937	Sumitomo Corporation	0.31920529	NEC Corp	2788
Toyota Motor Corporation	0.09634139	Fujitsu Limited	0.31900975	Toyota Motor Corporation	2204
Toshiba Corporation	0.09080531	Mitsui & Co Ltd	0.31767344	Mitsubishi Corp	1703
Panasonic Corp	0.08782789	NEC Corp	0.31642241	Sumitomo Corporation	1657
Fujitsu Limited	0.08488108	Mitsubishi Corp	0.31448287	Honda Motor Co.	1543

FIGURE 4. Top ten firms in Japan with highest degree, betweenness and closeness centrality values. The centrality values are based on the global interfirm network, not the local country network.

Company name	Betweenness Centrality	Company name	Closeness Centrality	Company name	Degree
BHP Billiton Ltd	0.03868966	UGL Ltd	0.30379225	BHP Billiton Ltd	1530
Government - Australia	0.03594304	WorleyParsons Ltd	0.30336643	Government - Australia	1026
WorleyParsons Ltd	0.01918358	Government - Australia	0.2988782	Telstra Corporation Ltd	870
Rio Tinto Ltd	0.01503404	BHP Billiton Ltd	0.2977944	Rio Tinto Ltd	723
UGL Ltd	0.01251468	Telstra Corporation Ltd	0.29660422	WorleyParsons Ltd	525
Telstra Corporation Ltd	0.0115444	Platinum Capital Limited PMC	0.29579922	Genetic Technologies Ltd	498
Cimic Group Ltd	0.00885768	Data3 Limited	0.2942119	Santos Ltd	392
Macquarie Group Ltd	0.00736878	Stockland	0.2930265	Australia & New Zealand Banking Group Ltd	346
Austal Limited	0.0067882	Rio Tinto Ltd	0.29283417	Woodside Petroleum Ltd	323
Stockland	0.00672413	UXC Ltd	0.28983159	Westpac Banking Corp	284

FIGURE 5. Top ten firms in Australia with highest degree, betweenness and closeness centrality values. The centrality values are based on the global interfirm network, not the local country network.

the top ten list of Australian firms regardless of the centrality metric used. It could be noted that these firms are all well-known multi-national corporates or government controlled bodies.

What happens if we consider the country specific subnetworks for each country, rather than the single global interfirm network, to calculate the centrality values of firms? Does this significantly change the top-ten list of firms for each country, and does it make the top ten list more diversified across different centrality measures? The answers to these questions can be gleaned from figures 6, 7 and 8, which show the list of top ten firms based on centrality value for United States, Japan and Australia. Obviously, when country-specific networks are considered, most of the networks are fragmented, so we only consider the largest components of each country specific network. Comparing the top-ten list for United States based on the global network (Fig 3) and based on the country-based network (Fig 6), it is clear that the lists are very similar. Out of 22 firms represented in either list, 10 firms find mention in both lists at least once. and most of these firms find mention multiple times, so that the top-ten lists look very similar. A similar scenario could be observed

with respect to Japan and Australia (and other countries that we considered, though the results corresponding to which are not shown), and therefore, it is fair to say that whether global or country-based network is considered in analysing the centrality of companies does not change the top-ten list of companies based on centrality in each country by much.

However, we were interested in not just identifying the firms with the highest centralities in the supply network, but also in identifying firms which have a relatively higher local or relatively higher global importance in each country. Therefore, we ranked the firms in each country based on their 'local' centrality value (centrality values calculated by considering the country networks), as well as the 'global' centrality value (centrality values calculated by considering the single global supply chain network), and plotted global rank vs local rank for each country. This is shown in figures 9, 10 and 11 respectively for the US, Japan and Australia (again, the plots for the other countries are not shown due to space restrictions). We also calculated the correlation coefficient between the global ranks and local ranks for each country. Note well that, even though the global network has firms belonging to all countries, the ranking was done only

Company name	Betweenness Centrality	Company name	Closeness Centrality	Company name	Degree
United States Government	0.15935002	United States Government	0.417087	United States Government	14544
General Electric Co	0.07846588	General Electric Co	0.40051699	International Business Machines Corp IBM	6503
International Business Machines Corp IBM	0.03518577	International Business Machines Corp IBM	0.39335009	Microsoft Corp	5719
Microsoft Corp	0.0328112	Microsoft Corp	0.38858993	General Electric Co	5690
Wal-Mart Stores Inc	0.02615416	Hewlett-Packard Co	0.38486489	Hewlett-Packard Co	5268
Berkshire Hathaway Inc	0.02402699	Oracle Corp	0.38074675	Johnson & Johnson	3999
Federal Government Agencies	0.02366932	Federal Government Agencies	0.3729174	Time Warner Inc	3985
Hewlett-Packard Co	0.02279432	AT&T Inc	0.37187464	Oracle Corp	3920
Johnson & Johnson	0.02149758	Berkshire Hathaway Inc	0.36876401	AT&T Inc	3434
Bank of America Corp	0.01915805	Wal-Mart Stores Inc	0.36758761	Pfizer Inc	3381

FIGURE 6. Top ten firms in the United States with highest degree, betweenness and closeness centrality values. The centrality values are based on the the local country network.

Company name	Betweenness Centrality	Company name	Closeness Centrality	Company name	Degree
Mitsubishi Heavy Industries	0.12105783	Mitsubishi UFJ Financial Group Inc	0.39216544	Mitsubishi Heavy Industries	691
Government - Japan	0.06541028	Mizuho Financial Group	0.37808698	Mitsubishi Corp	650
Mitsubishi UFJ Financial Group Inc	0.0643764	Mitsubishi Heavy Industries	0.37797314	Mitsubishi UFJ Financial Group Inc	556
Mizuho Financial Group	0.05314985	Sumitomo Corporation	0.37627383	Toyota Motor Corporation	547
Sumitomo Mitsui Trust Holdings Inc	0.05013374	Government - Japan	0.37456737	Sumitomo Corporation	511
Mitsubishi Corp	0.04545042	Sumitomo Mitsui Trust Holdings Inc	0.37168404	Government - Japan	497
Itochu Corp	0.03075187	Mitsubishi Corp	0.3712883	Nippon Steel & Sumitomo Metal Corporation	484
Sumitomo Corporation	0.0306628	Mitsui & Co Ltd	0.36802298	Hitachi	478
Sumitomo Mitsui Financial Group	0.02942455	Sumitomo Mitsui Financial Group	0.36598449	Sumitomo Mitsui Trust Holdings Inc	474
Mitsui & Co Ltd	0.02802429	Panasonic Corp	0.36581386	Mizuho Financial Group	473

FIGURE 7. Top ten firms in Japan with highest degree, betweenness and closeness centrality values. The centrality values are based on the local country network.

Company name	Betweenness Centrality	Company name	Closeness Centrality	Company name	Degree
Government - Australia	0.23461082	Government - Australia	0.37607316	Government - Australia	428
National Nominees Ltd.	0.14067399	National Nominees Ltd.	0.3474737	BHP Billiton Ltd	228
BHP Billiton Ltd	0.09999698	Stockland	0.34604156	Rio Tinto Ltd	164
Santos Ltd	0.05512314	BHP Billiton Ltd	0.34473909	Milton Corporation Ltd	154
Wesfarmers Ltd	0.0448847	Diversified United Investment Ltd	0.3419311	Diversified United Investment Ltd	136
Origin Energy Ltd	0.04126695	Milton Corporation Ltd	0.34106297	National Nominees Ltd.	130
Diversified United Investment Ltd	0.04077973	Australian United Investment Co Ltd	0.33905435	Westpac Banking Corp	129
Commonwealth Bank of Australia	0.04040256	Commonwealth Bank of Australia	0.33888328	Telstra Corporation Ltd	126
Milton Corporation Ltd	0.04015406	AGL Energy Ltd	0.33690018	Santos Ltd	123
Stockland	0.0377761	Origin Energy Ltd	0.33667502	National Australia Bank Ltd	120

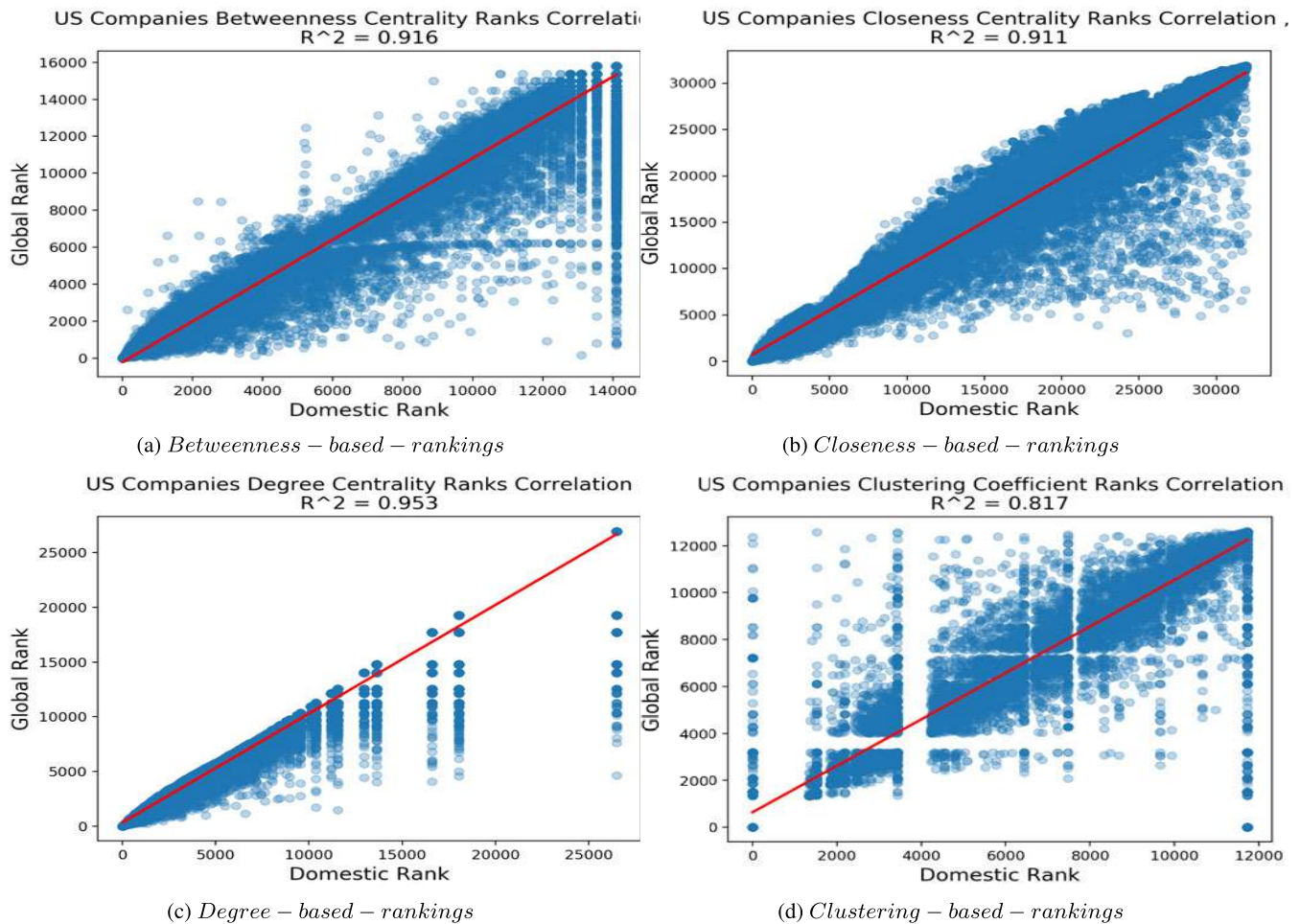
FIGURE 8. Top ten firms in Australia with highest degree, betweenness and closeness centrality values. The centrality values are based on the local country network.

to those firms which belonged to the country under consideration. Thus, the ‘lowest’ rank would equal to the size of the country network, for both local and global ranks.

We may note from figures 9, 10 and 11 that the correlation between global and local centrality ranks is high. For the

three countries under consideration, across different centrality measures, this correlation ranges from 95.3% to 58.1%, though it is typically above 80%. This correlation between global and local centrality ranks is also high for other countries that we studied.





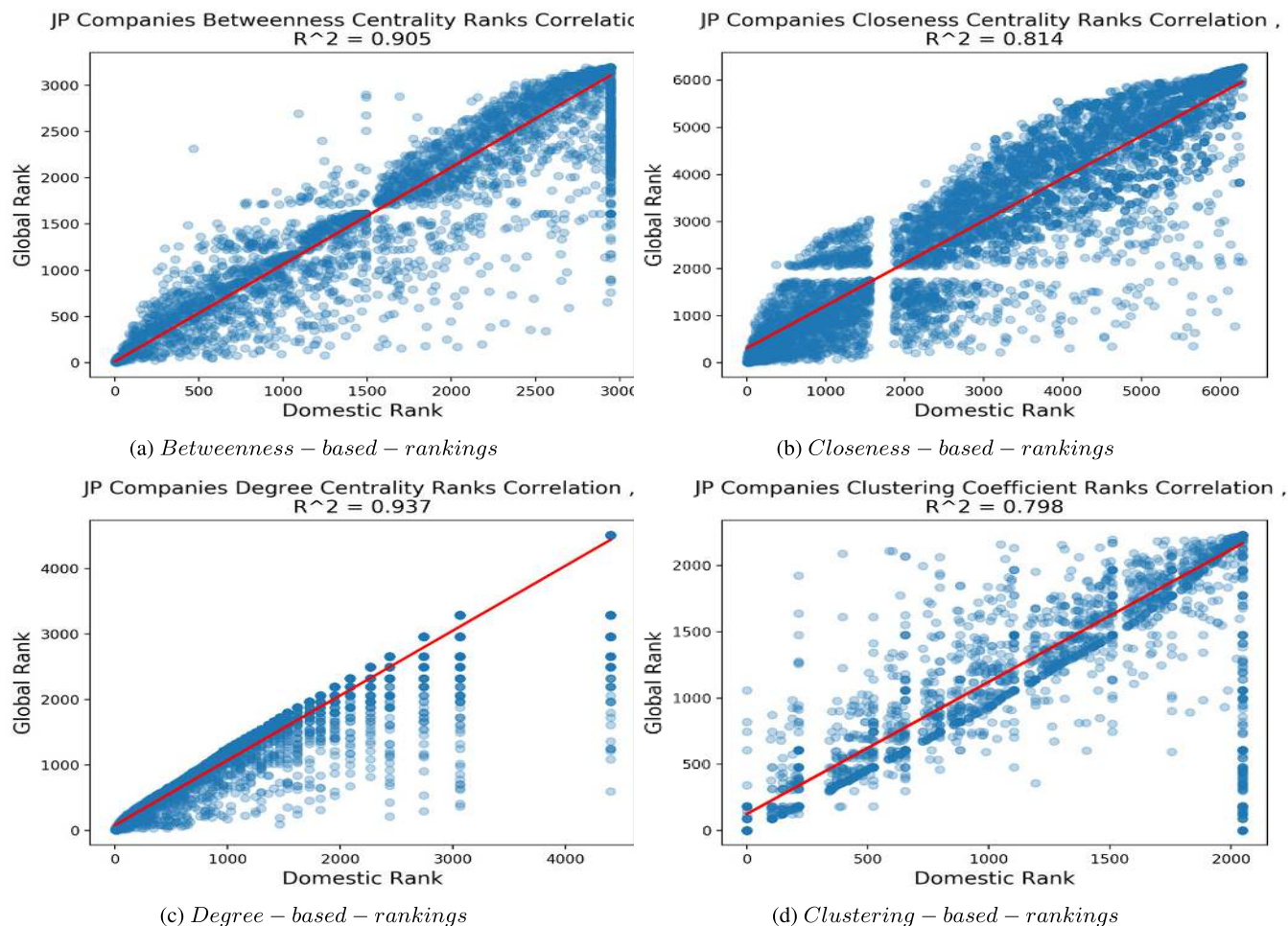
**FIGURE 9.** Correlation between global and local centrality-based ranks for the US firms: (a) Betweenness-centrality based rankings (b) Closeness-centrality based rankings (c) Degree-centrality based rankings (d) Clustering-coefficient based rankings.

**C. OUTLIER FIRMS BASED ON CENTRALITY MEASURES**

Which are the firms which are the ‘outliers’ in figures 9, 10, and 11 - that is, among the firms belonging to each country (US, Japan and Australia) which have the most difference in terms of their local and global importance? In figures 12, 13, 14, we show firms which have the highest difference in rank based on global-network based centrality measures and country-network based centrality measures. For comparison, clustering coefficient was also considered, which it can be argued is a sort of centrality measure. The ‘error’ in centrality rankings was calculated by fitting a straight line to the global rank - local rank plot for the relevant centrality measure, and calculating the difference between the global rank and the fitted value for the given firm. Thus, this ‘error’ can be a decimal number, though it is calculated from ranks which are integer numbers. Firms which have a negative ‘error’ - that is, much lower centrality ranking based on the global network compared to the norm for firms in that country, are much more important in the global interfirm network and less important in their domestic network. Similarly, firms which have a positive ‘error’ - that is, much higher centrality

ranking based on the global network compared to the norm for that country, are very prominent firms domestically but not so in the international scene. We were interested in finding which are the ‘outlier’ firms in this respect in all countries that we considered, and whether there are any similarities between these outliers, and some of the results for the prominent countries that we study are shown in figures 12, 13, 14. Note well here that, since we consider ranks based on centrality in figures 9, 10, and 11, the lower the rank, the more important the firm is with respect to the metric in question. However, in figures 12, 13, 14, the errors were constructed as the difference between actual global rank and ‘fitted’ global rank (indicating the ‘expected’ global rank for each firm based on its local rank), so if the error is positive, the firm is more locally important, while if the error is negative, the firm is more globally important.

In the case of United States which has the highest representation in the dataset, we could see from Fig. 12 that the ‘outlier’ firms which are comparatively much more important in global scene are, mainly of Chinese origin, or from other overseas origins, but considered US companies because they



**FIGURE 10.** Correlation between global and local centrality-based ranks for the Japanese firms: (a) Betweenness-centrality based rankings (b) Closeness-centrality based rankings (c) Degree-centrality based rankings (d) Clustering-coefficient based rankings.

are registered in the US. Naturally, they do most of their business with Chinese or other overseas firms, and thus their global importance is significantly higher than their domestic importance in the US interfirm network. On the other hand, firms which have relatively higher domestic importance compared to global importance are typically small-sized firms, which serve a niche market in the domestic population, such as Volunteers of America, New York State Teachers Retirement System, CarePlus Health Plan, AQR Capital management holdings PLC etc. Big Corporates or Internationalised firms are not represented among these outliers, which makes sense.

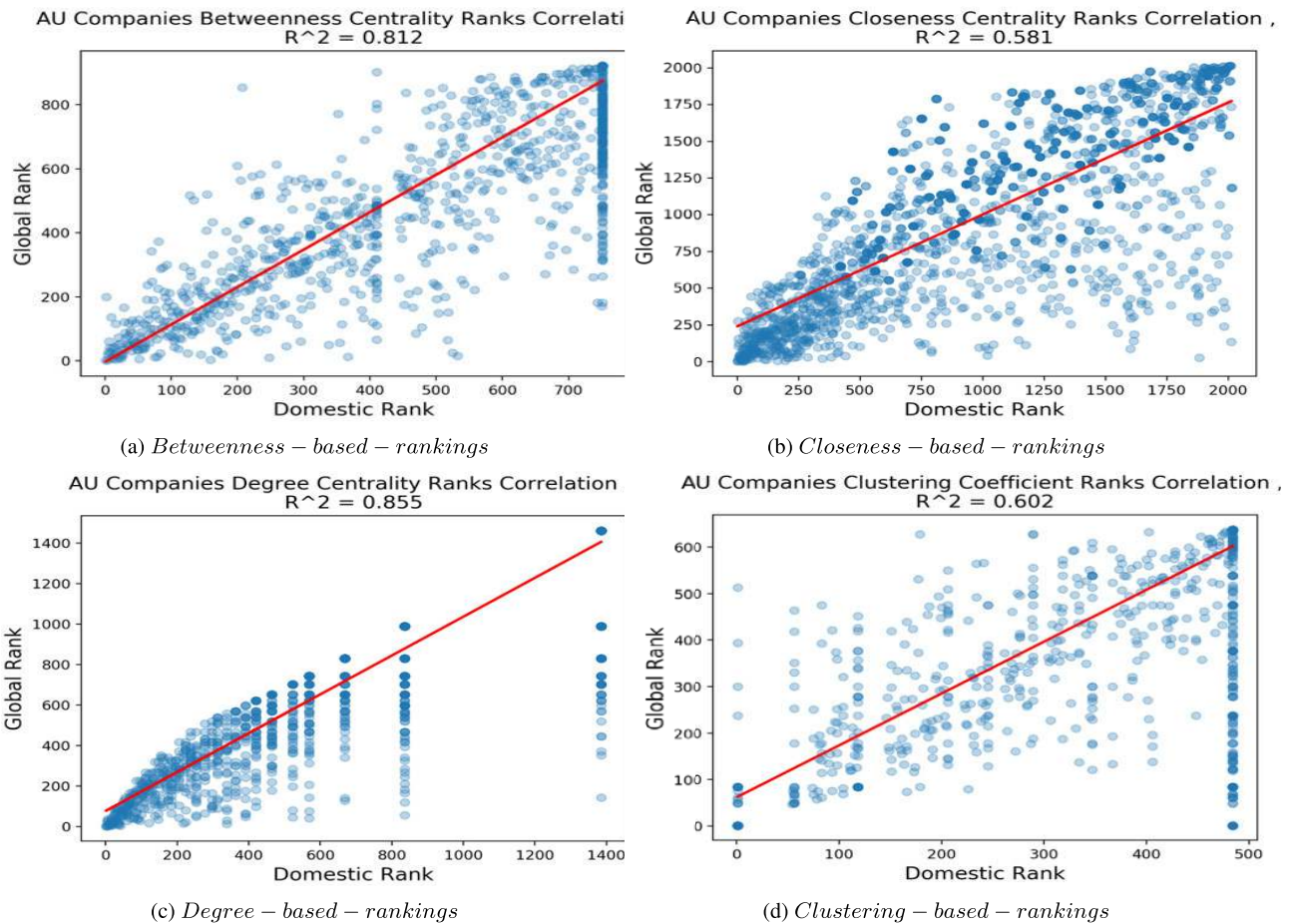
A similar situation exists in Japan, though less pronounced, as shown in Fig. 13. The Japanese firms which are locally-central outliers are all small firms which appear to serve some local niche market, such as Sakai moving service, Nishi Nippon City Bank Employee Stock Ownership plan, Asahimatsu foods etc. No large and/or well known Japanese corporate is represented among them. On the other hand, the firms which are the globally central outliers include several companies with apparently non-Japanese roots, even though they are

registered in Japan. Examples of these are EPS Corp, Roland Corp, Innotech Corp, Cresco Ltd etc. Therefore, it seems that considering the outliers based on the local-global centrality profiles is an effective way to determine the nature of relative local/global importance of a firm in a supply chain network. This pattern is repeated in most other countries that we have analysed.

When we consider Australia though, as shown in Fig. 14 the picture is less clear. For example, universities are represented as both globally central outliers and locally central outliers. However, this could be due to some universities having overseas campuses, and thus being forced to do business with lots of overseas firms, while others do not. Both globally central outlier firms and locally central outlier firms appear to be mostly small and relatively unheard-of firms, within which it seems hard to form a distinction.

**D. COMMUNITY STRUCTURE OF INTERFIRM NETWORKS**

Next we analysed the community structure of the global interfirm network, in order to understand firms from which countries or which industries form closest supply connections



**FIGURE 11. Correlation between global and local centrality-based ranks for the Australian firms: (a) Betweenness-centrality based rankings (b) Closeness-centrality based rankings (c) Degree-centrality based rankings (d) Clustering-coefficient based rankings.**

within themselves. To this end, we applied the Louvain method [45] on the global interfirm network, which has 154,862 nodes and 1,571,949 links, to create a community-based partition. We found a total of 1014 communities, of which 12 communities had more than 2000 nodes and 18 communities had more than 50 nodes. In the following analysis, we choose to focus on the largest 12 communities (which each had more than 2000 nodes), and any company which did not belong to any of these communities is denoted as belonging to the ‘other’ community.

Fig. 15 shows the percentage of firms from each country we considered in the largest 12 communities generated by the Louvain method (with the 13th community representing the “other”). Conversely, Fig. 16 shows the percentage of firms from each community present in each of the countries we considered. In other words, Fig. 15 shows the country distribution among communities, whereas Fig. 16 shows the community distribution among countries. From Fig. 15, we may see that US firms dominate communities 0, 1, 5 and 6, while Japanese firms dominate community 7, Chinese firms dominate community 9, and all other communities have firms from the ‘rest of the’ countries (countries outside the list of

specifically considered countries in this paper), cumulatively, as the majority. On the other hand, from Fig. 16, we may see that the highest proportion of US firms are in community 0, the highest proportion of Chinese firms are in community 9, the highest proportion of Canadian, Australian, Russian and Indian firms are in community 2, the highest proportion of Singaporean firms are in community 11, the highest proportion of French, and European Union (within which France is also counted) firms are in community 4, and the highest proportion of Japanese firms are in community 7. So it is important to note that if a community has firms from a certain country in the majority, this does not imply that the majority of firms from that country are in that community. Conversely, if the majority of firms from a particular country are in a community, this also does not imply that firms from that country are a majority in that community. That statement is true only for community 0 - US, community 7 - Japan, and community - 9, China. Even in these communities, ‘majority’ simply implies the highest proportion, and not necessarily more than half.

Fig. 17 shows the industries present in each of these communities, according to the industry classification presented



**FIGURE 12.** Outlier firms in US based on centrality scores. The “centrality errors” shown represent the difference between the *global* rank of a firm based on the relevant centrality, and the fitted value of the global centrality rank - local centrality rank profile, as shown in Fig. 9. A relatively higher rank (numerically) would imply relatively lower importance. Therefore a positive “error” corresponds to higher local (country-based) importance, and a negative “error” corresponds to higher global importance: (a) Betweenness-centrality based errors (b) Closeness-centrality based errors (c) Degree-centrality based errors (d) Clustering-coefficient based errors.

earlier in Fig. 1. It could be seen from Fig. 17 that certain industries are predominantly present in certain communities. For example, transportation is predominant in community 2, manufacturing is predominant in community 3, finance in community 5, and service sector in communities 1 and 6. The other communities were not associated with predominant industries to the same extent, even though manufacturing, finance and service industries together made up most of the firms in most other communities as well. Therefore, it seems that a significant proportion of Canadian, Australian, Russian and Indian firms are in the transportation community (it could be noted that these are all large countries which require considerable transport infrastructure), whereas all the other communities which correspond to the largest share of firms from a particular country are non-descriptive in terms of industry. This could be corroborated by Fig. 1 which shows that Russia, Canada, and Australia are the three

countries where the transport sector firms make up the highest percentages.

If we consider community 2 (which is dominated by transportation), firms from the European Union (16.93%), the US (15.66%) and China ( 10.27%) are most prominent in it, as Fig. 18 a shows. Similarly, If we consider community 3 (which is dominated by manufacturing - Fig. 18 b), firms from The European Union (23.18%), the US (22.85%) and China (10.15%) are again most prominent in it. Community 5 (which is dominated by finance - Fig. 18 c) is dominated by firms from US (32.54%) and the European Union (25.33%). Community 6 ( which is dominated by service - Fig. 18 d) has more than half US firms (50.62%) and a considerable percentage of European Union firms (18.42%).

The other communities are not dominated by any particular industry, even though community 7 (Fig. 19 a) has an interesting feature. It is dominated by Japanese firms (58.52%).

Company name	betweenness error
H.I.S. Co Ltd	1810.1226
Osakafu Kokuhoren	1536.4234
Kawamoto Corp	1321.1034
Roland D.G. KK	1291.1034
Kitanotatsujin Corporation	1094.1034
...	...
YKK Corporation	-2250.4774
Anritsu Corp	-2252.0598
Lotte Group	-2344.709
Japan Oil	-2361.2018
Nexon Co Ltd	-2430.7438

(a) Betweenness – based – outliers

Company name	closeness error
Nishi Nippon City Bank Employee Stock Ownership Plan	1688.004
NTT Data NCB KK	1688.004
Kuze Co Ltd	1672.986
Asahimatsu Foods Co Ltd	1630.316
Itami Sangyo Co. Ltd.	1608.29
...	...
EPS Corp.	-4637.878
Comtec KK	-4726.862
KU Holdings Co. Ltd.	-4785.26
Cresco Ltd	-5050.73
Roland Corp.	-5112.986

(b) Closeness – based – outliers

Company name	degree error
Totenko Co Ltd	201.6847
KK Hokuetsu Ginko	201.6847
Hasegawa Co Ltd	201.6847
Sakai Moving Service Co. Ltd.	201.6847
Chubu Suisan KK	201.6847
...	...
IPS Academia Japan Inc	-3189.3811
CBC Company	-3189.3811
Roland Corp.	-3221.3811
Innotech Corp	-3352.3811
D&M Holdings	-3843.3811

(c) Degree – based – outliers

Company name	clustering error
YKK Corporation	1586.087
KK Kozo Keikaku Kenkyujo	1571.4128
Topre Corp	1397.7602
Toyo Corporation	1385.7836
Valuecommerce Co Ltd	1301.087
...	...
Nippon Pet Food Co Ltd	-2166.6154
Mitsubishi Densen Kogyo KK	-2166.6154
Honshu Chemical Industry Co Ltd	-2166.6154
Sazaby League Ltd	-2166.6154
AvanStrate Inc.	-2166.6154

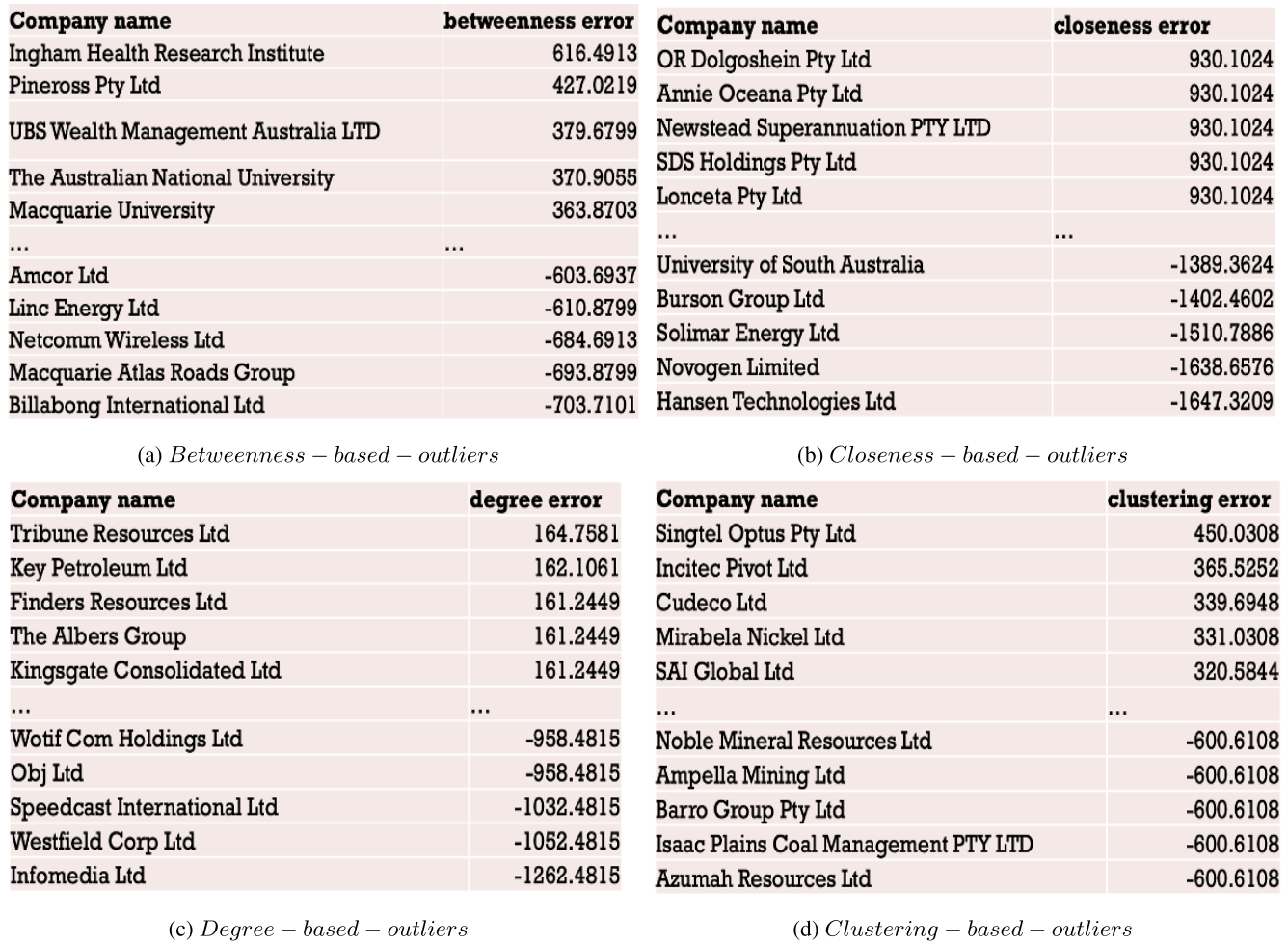
(d) Clustering – based – outliers

**FIGURE 13.** Outlier firms in Japan based on centrality scores. The “centrality errors” shown represent the difference between the *global* rank of a firm based on the relevant centrality, and the fitted value of the global centrality rank - local centrality rank profile, as shown in Fig. 10. A relatively higher rank (numerically) would imply relatively lower importance. Therefore a positive “error” corresponds to higher local (country-based) importance, and a negative “error” corresponds to higher global importance: (a) Betweenness-centrality based errors (b) Closeness-centrality based errors (c) Degree-centrality based errors (d) Clustering-coefficient based errors.

Moreover, the firms with the next highest presence are Chinese firms (13.20%), and it appears that firms from these two countries in this community are very tightly connected to other firms from their own nation. Similarly, community 9 (Fig. 19 b) is dominated by Chinese firms (65.4%), and firms from no other single country have significant presence. Conversely, we may note that 45.7% of all Japanese firms are present in community 7 (thus, even though Japanese firms are a majority in community 7, community 7 firms are not a majority among Japanese firms), while 29.37% of all Chinese firms are present in community 9 (again a majority of Chinese firms are not present in community 9, even though a majority of community 9 firms are Chinese).

Now let us consider the reverse scenario: that is, identifying communities in individual country networks. For this purpose, again we applied the Louvian algorithm, but this time on country supply chain networks rather than the global supply chain network. The results for the United States,

Japan, Australia, and China are shown in figures 20 a, 20 b, 20 c, and 20 d respectively. We could see from Fig. 20 a that the US interfirm network is dominated by communities 0, 1 and 4 (which are themselves dominated by manufacturing, again manufacturing, and service industries, respectively), while communities 2, 3 and 5 also have more than 10% representation. This tallies well with Fig. 1 which shows that the US interfirm network is dominated by manufacturing and service industries. Similarly, we may observe from Fig. 20 b that the Japanese interfirm network is dominated by communities 7 and 3 (where community 7 represents multiple industries, community 3 is dominated by manufacturing). This tallies well with Fig. 1 which shows that the Japanese interfirm network is dominated by manufacturing industry. Further, Fig. 20 c shows that the Australian interfirm network is dominated by communities 2, 4, and 5 (which are themselves dominated by transportation, manufacturing and finance respectively). This again corresponds to Fig. 1, which shows



**FIGURE 14.** Outlier firms in Australia based on centrality scores. The “centrality errors” shown represent the difference between the *global* rank of a firm based on the relevant centrality, and the fitted value of the global centrality rank - local centrality rank profile, as shown in Fig. 11. A relatively higher rank (numerically) would imply relatively lower importance. Therefore a positive “error” corresponds to higher local (country-based) importance, and a negative “error” corresponds to higher global importance: (a) Betweenness-centrality based errors (b) Closeness-centrality based errors (c) Degree-centrality based errors (d) Clustering-coefficient based errors.

	Nodes	Links	US	CN	CA	AU	IN	RU	SG	EU	JP	Other
0	20128	331209	7320 36.37%	984 4.89%	532 2.64%	163 0.81%	271 1.35%	91 0.45%	162 0.80%	2970 14.76%	712 3.54%	6923 34.39%
1	14143	192671	6205 43.87%	494 3.49%	422 2.98%	233 1.65%	191 1.35%	52 0.37%	31 0.22%	2386 16.87%	403 2.85%	3776 26.35%
2	25468	185532	3989 15.66%	2615 10.27%	1497 5.88%	1093 4.29%	1057 4.15%	533 2.09%	363 1.43%	4313 16.53%	248 0.97%	9760 38.32%
3	19414	132047	4436 22.85%	1970 10.15%	362 1.86%	207 1.07%	574 2.96%	187 0.96%	148 0.76%	4501 23.18%	1697 8.74%	5332 27.46%
4	23187	147736	7087 30.56%	726 3.13%	531 2.29%	598 2.58%	458 1.98%	128 0.55%	310 1.34%	5315 22.92%	554 2.39%	7480 32.26%
5	16047	86654	5222 32.54%	285 1.78%	290 1.81%	458 2.85%	420 2.62%	86 0.54%	63 0.39%	4064 25.33%	67 0.42%	5092 31.73%
6	2904	11734	1470 50.62%	46 1.58%	107 3.68%	141 4.86%	30 1.03%	5 0.17%	15 0.52%	535 18.42%	10 0.34%	545 18.77%
7	5788	22717	212 3.66%	764 13.20%	19 0.33%	25 0.43%	39 0.67%	2 0.03%	50 0.86%	190 3.28%	3387 58.52%	1100 19.00%
8	2001	7023	642 32.08%	68 3.40%	77 3.85%	18 0.90%	110 5.50%	4 0.20%	15 0.75%	343 17.14%	9 0.45%	715 35.73%
9	6313	16182	206 3.26%	4129 65.40%	36 0.57%	22 0.35%	81 1.28%	18 0.29%	27 0.43%	152 2.41%	46 0.73%	1596 25.28%
10	5460	16658	728 13.33%	209 3.83%	137 2.51%	38 0.70%	90 1.65%	49 0.90%	15 0.27%	2038 37.33%	21 0.38%	2135 39.10%
11	4309	10303	200 4.64%	133 3.09%	50 1.16%	46 1.07%	206 4.78%	11 0.26%	470 10.91%	363 8.42%	43 1.00%	2787 64.68%
12[other]	9700	19437	991 10.22%	1635 16.86%	727 2.34%	113 1.16%	221 2.28%	36 0.37%	67 0.69%	1550 15.98%	214 2.21%	4646 47.90%

**FIGURE 15.** The percentage of firms from each country we considered in the largest 13 communities generated by the Louvain method.

that the Australian interfirm network is dominated by firms from finance, manufacturing and transport industries, besides the services industry which is the third dominant industry overall, but does not seem to be represented by any specific large community in Australia. Finally, Fig. 20 d shows that the Chinese interfirm network is dominated by communities 9, 2 and 3 (which are themselves dominated by manufacturing, transportation, and again manufacturing respectively).

This again corresponds to Fig.1, which shows that the Chinese interfirm network is dominated by firms from manufacturing, though transport is not that dominant.

**V. DISCUSSION**

The extended network analysis of interfirm networks undertaken in this work provided several interesting results. In terms of the basic topology, we observed that the global

	Nodes	0	1	2	3	4	5	6	7	8	9	10	11	12(other)
US	38708	18.91%	16.03%	10.31%	11.46%	18.31%	13.49%	3.80%	0.55%	1.66%	0.53%	1.88%	0.52%	2.56%
CN	14058	7.00%	3.51%	18.60%	14.01%	5.16%	2.03%	0.33%	5.43%	0.48%	29.37%	1.49%	0.95%	11.63%
CA	4287	12.41%	9.84%	34.92%	8.44%	12.39%	6.76%	2.50%	0.44%	1.80%	0.84%	3.20%	1.17%	5.30%
AU	3155	5.17%	7.39%	34.64%	6.56%	18.95%	14.52%	4.47%	0.79%	0.57%	0.70%	1.20%	1.46%	3.58%
IN	3748	7.23%	5.10%	28.20%	15.31%	12.22%	11.21%	0.80%	1.04%	2.93%	2.16%	2.40%	5.50%	5.90%
RU	1202	7.57%	4.33%	44.34%	15.56%	10.65%	7.15%	0.42%	0.17%	0.33%	1.50%	4.08%	0.92%	3.00%
SG	1736	9.33%	1.79%	20.91%	8.53%	17.86%	3.63%	0.86%	2.88%	0.86%	1.56%	0.86%	27.07%	3.86%
EU	28720	10.34%	8.31%	15.02%	15.67%	18.51%	14.15%	1.86%	0.66%	1.19%	0.53%	7.10%	1.26%	5.40%
JP	7411	9.61%	5.44%	3.35%	22.90%	7.48%	0.90%	0.13%	45.70%	0.12%	0.62%	0.28%	0.58%	2.89%
Other	51837	16.82%	10.11%	24.73%	16.78%	20.07%	14.59%	1.53%	2.32%	1.80%	3.29%	7.53%	5.84%	11.05%

FIGURE 16. The percentage of firms from each community present in each of the countries we considered by the Louvain method.

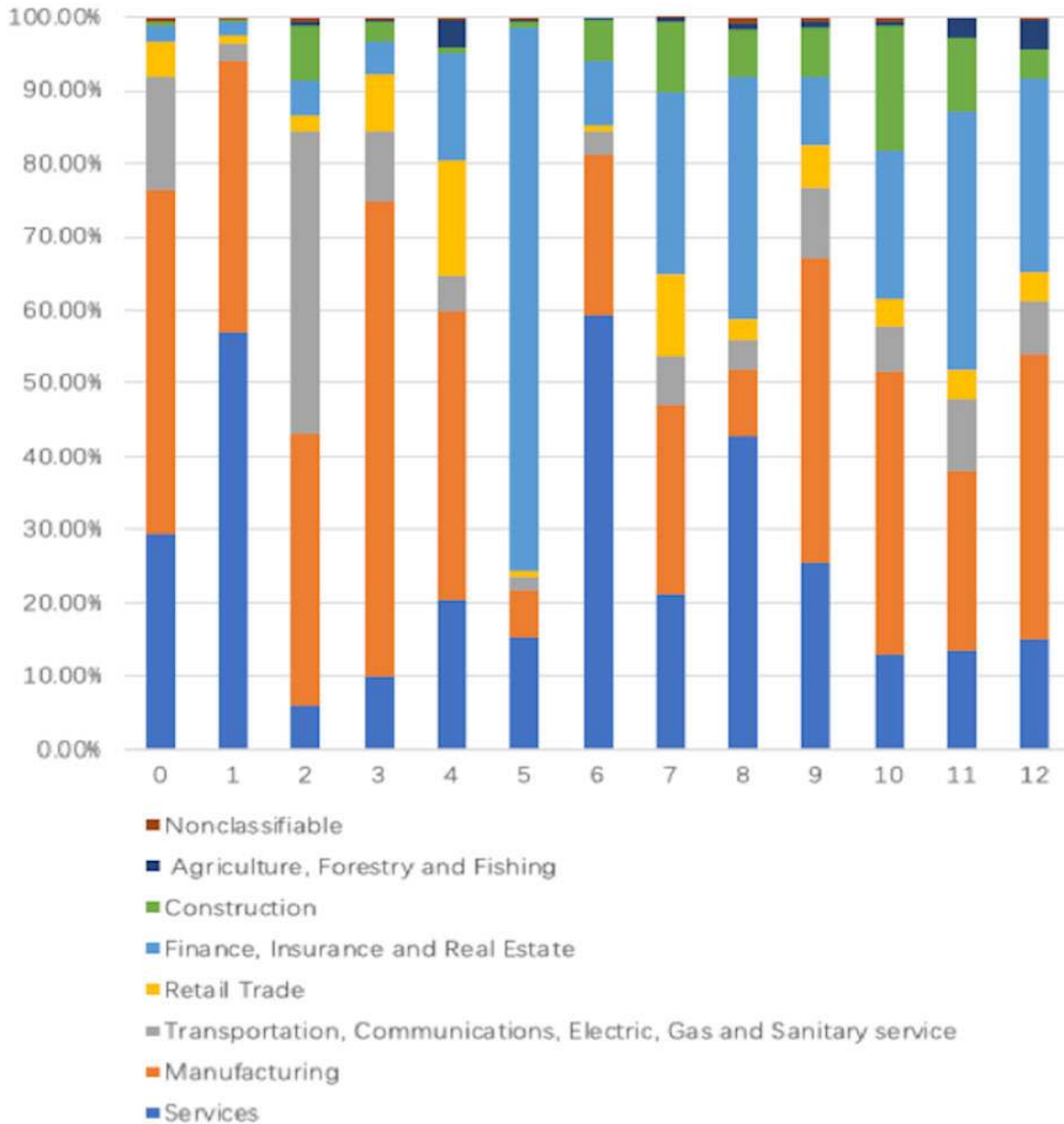


FIGURE 17. The industry classification of the communities derived.

as well as the ten country-specific networks considered are all scale-free, even though their scale-free exponents ranged from 1.0 – 2.0, rather than the range of 2.0 – 3.0 observed

in most other real world networks. The networks were also relatively sparse, and displayed some small-world properties. Then we identified the firms which were the most central

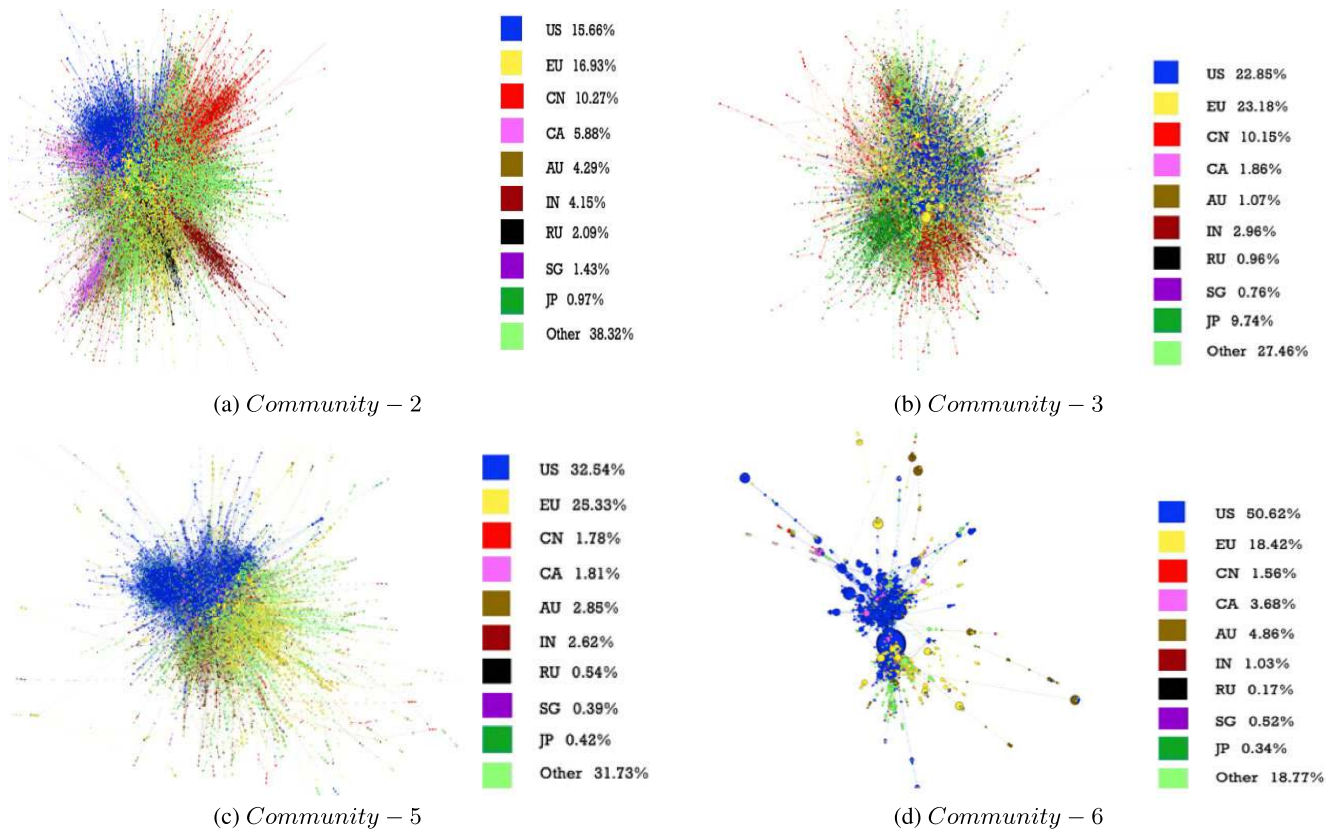


FIGURE 18. The community networks of (a) community 2 (b) community 3 (c) community 5 (d) community 6, with the country affiliation of firms highlighted in different colours.

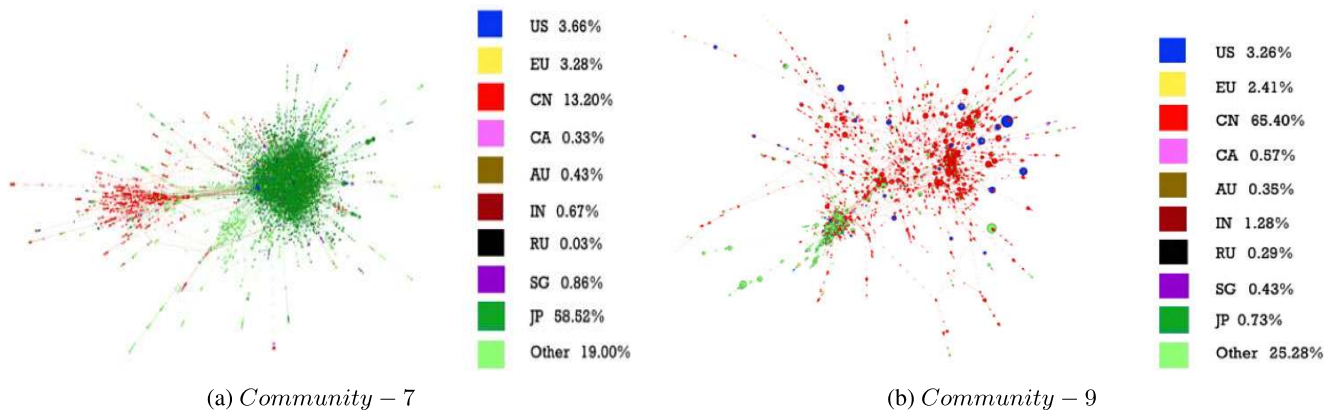
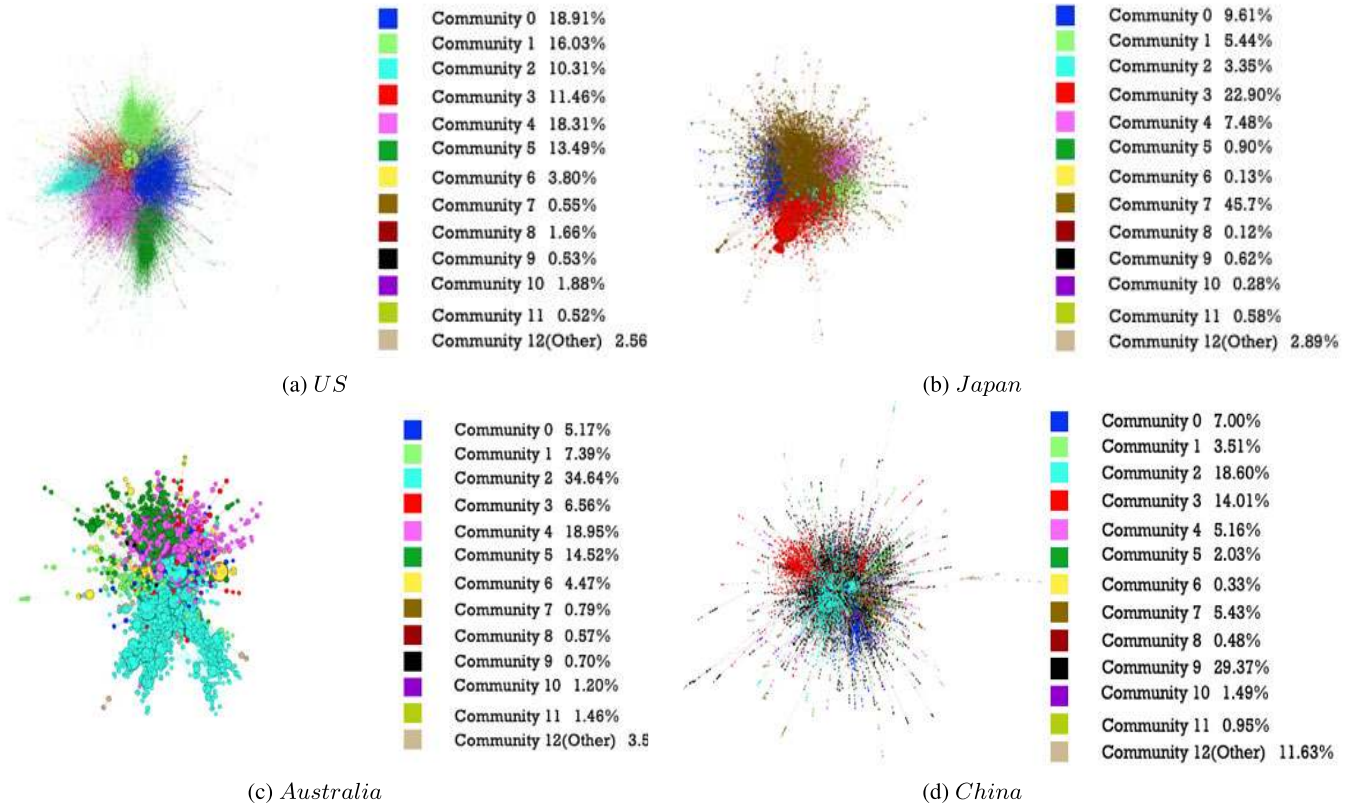


FIGURE 19. The community networks of (a) community 7 (b) community 9, with the country affiliation of firms highlighted in different colours.

in each country-networks, based on a number of centrality measures. It was found that these were, predictably, quite often large multi-national corporates and government bodies. However, certain firms which are small firms with overseas origins registered in a country, such as China Shengda Packaging Group Inc (registered in US), Chinacast Education Corps (registered in the US), and China Biologic Products Inc (registered in the US), had higher centrality when the global supply chain network was considered as opposed to their

own country-specific networks, and some other firms, such as Volunteers of America, New York State Teachers Retirement System, CarePlus Health Plan etc, had higher centrality when the local (country-specific) network was considered as opposed to the global supply chain network. To further analyse this trend, we considered the correlation between global centrality and country-specific centrality for all firms in a specific country. In general, this correlation was quite strong for all countries considered by us, ranging from 95.3%





**FIGURE 20.** Country networks of (a) US (b) Japanese (c) Australian (d) Chinese firms represented in Factset Revere, with communities colour-coded. Only the giant component of each country network is shown here.

to 58.1%, and typically above 80%. However, certain firms were clear outliers in the global centrality - local centrality profiles of each country. Quite often, these firms seemed to be either firms which had some locally focussed business model, or, on the other side of the spectrum, locally registered firms with overseas roots. Some good examples are New York State Teachers Retirement System and China Shengda Packaging Group Inc registered in the United States.

Then we analysed the community structure of the global supply chain network, and applied the communities identified in the global network to classify nodes in the country-specific subnetworks as well. We used the Louvian method for community clustering, and found that there were 12 primary communities in the global network which had number of nodes more than 2000. Some of these communities were primarily aligned with certain industry sectors, such as community 2 which is aligned with the transportation sector, and community 5 which is aligned with finance sector. Some other communities had majority representation from certain countries, such as community 6 which had 50.62% US firms, or community 7 which had 58.52% Japanese firms. Conversely, some country networks also had majority representation from certain communities. For example, Japanese country network had 45.7% of community 7, while Russian country network had 44.34% of community 2. It is important to note however that the majority representation relationships between countries and communities are not mutual: that is,

if a community has a majority representation from a certain country, this does not mean that that country will have majority representation from the corresponding community, and vice versa. For example, the sixth community network has a majority of nodes (50.62%) from the US, however the corresponding US country network has only 3.8% nodes from community 6, which is not a numerical majority, and not even the largest community represented in the US country network.

There were also other interesting features in the community structure. For example, The analysis uncovered a dense cluster among Japanese and Chinese firms (community 7). This contrasts with the trading relationships between US and China, or US and Japan. Although US and China, the two largest economies have the largest numbers of firms in the dataset, no cluster emerged with a predominant presence of firms of these two countries. Conversely to the structure of Japan-China business relationships, US and Chinese firms do not seem to form dense Sino-American business cliques. The supply relationships between firms from these two countries are possibly more of isolated dyadic nature. The same applies to interactions between US and Japanese firms.

In presenting these results, we were mindful that the analysis had some limitations. For example, we were unable to consider tie-strength, since this data was not always available, so all analysis was conducted assuming that all ties were of equal strength, which is obviously not the case. We also did

not consider the directionality of the ties, or the durability of ties (how long the ties have existed). These considerations should be taken into account while interpreting the results. The performance of community structure analysis was limited by the inherent limitations of the Louvain method [45], [47]. For example, since the method is meant to optimise modularity in the analysed network, it is not necessarily the best in identifying hierarchical communities or communities of communities. Similarly, the results of the centrality analysis have to be taken in context of the inherent weaknesses of each centrality measure used, which are well documented [40]. For example, betweenness based measures only consider ‘shortest paths’ in calculating centrality, whereas in the context of interfirm networks, non-shortest paths are also important. Even though some centrality measures could take tie-strength into account if available, we did not have this data and thus the centrality measures which we employed did not use it. Perhaps the most significant limitation is that the nature of ties (not merely tie strengths) are quite variable in supply chain networks. For example, a manufacturer-supplier relationship is inherently different from a wholesaler-retailer relationship, and this work, as well as most other works relating to topological analysis of supply chain networks [7], [19], [23], are forced to ignore it and assume that links are homogenous, even when tie strength may vary. Some of these difficulties in supply chain network analysis cannot be easily overcome [7], yet it is important to be cognisant of them in considering the results.

Yet, some of these concerns were partially mitigated by our study design. For example, in identifying locally and globally important firms and analysing the correlation between local and global importance, we used centrality-based ranking rather than centrality values themselves, which eliminated some issues arising from the limitations of centrality analysis. We also consistently used four inherently different centrality measures, so that issues related to particular centrality metrics were mitigated. We also typically had a large number of firms from each country to compute correlations, eliminating finite-scale effects. Therefore overall, the analysis resulted in reliable and significant observations. The analysis is among the first efforts which specifically focusses on the ‘international vs local importance’ aspect of firms in supply chains, thus shedding light also into the consequences of increasing internationalisation and globalisation. The topological analysis, it should be noted, is the most efficient way of observing the overall patterns which are prevalent in the structure of supply chain networks, as indirect methods such as inference, machine learning or data mining are likely to result in relatively unclear generalisations which are not easily quantifiable in terms of their significance.

## VI. CONCLUSION

Topological analysis of global and country-specific supply chain networks is very useful to understand interdependencies in terms of supply relationships, and the resultant trade clusters and communities formed which determine the

stability and robustness of the global economy. Particularly, during times of crises like the current COVID-19 crisis, the global supply chains would be affected tremendously, so it is vitally important to understand the relative strengths and weaknesses of global and country-specific supply chain networks, and the part individual firms play in shaping these, and this paper helps address these questions. The primary contributions of this paper are that it establishes that the typical supply chain networks are scale-free with a low scale-free exponent, highlights the nature of firms which are more globally central than locally in supply networks and vice versa, and identifies which industries and countries drive the global community structure of supply and interfirm networks.

However, more data about individual firms themselves, such as their sales, income, size, human resources etc, and other relationships between them, such as shareholding relationships and patent relationships, are important to build a holistic and data-rich set of complex networks which will help us clearly understand the dynamics and evolution of the global financial system. In considering future research directions, we are cognisant of this, and our future research will focus on incorporating and analysing datasets which can give us input about the above-mentioned attributes. Specifically, our future research direction focusses on modelling the impact of crises on supply chain networks, and how such crises impact the economical and financial resilience of firms. In this sense, our future research will focus on modelling ‘cascading failures’ in interfirm networks, using large data sets with detailed firm attributes. The future research will look at ‘fire-sale’ events, and other such crisis responses of the interfirm networks as a result of crises, and how these impact the economic stability of world markets. This research has clear applications in several current contexts, such as the COVID-19 crisis threatening the world, the bushfire crisis that threatened parts of Australia last summer, and the crises triggered by trade disputes between countries.

## ACKNOWLEDGMENT

The authors would like to thank U. Senanayake for motivating discussions and help with figures.

## REFERENCES

- [1] H. P. Thadakamalla, U. N. Raghavan, S. Kumara, and A. Albert, “Survivability of multiagent-based supply networks: A topological perspective,” *IEEE Intell. Syst.*, vol. 19, no. 5, pp. 24–31, Sep. 2004.
- [2] Y. Kim, Y.-S. Chen, and K. Linderman, “Supply network disruption and resilience: A network structural perspective,” *J. Oper. Manage.*, vols. 33–34, no. 1, pp. 43–59, Jan. 2015.
- [3] Y. Li, “Networked analysis approach of supply chain network,” *J. Netw.*, vol. 9, no. 3, p. 777, Mar. 2014.
- [4] S. I. Mari, Y. H. Lee, M. S. Memon, Y. S. Park, and M. Kim, “Adaptivity of complex network topologies for designing resilient supply chain networks,” *Int. J. Ind. Eng.*, vol. 22, no. 1, pp. 1–15, 2015.
- [5] M. Xu, X. Wang, and L. Zhao, “Predicted supply chain resilience based on structural evolution against random supply disruptions,” *Int. J. Syst. Sci., Oper. Logistics*, vol. 1, no. 2, pp. 105–117, Apr. 2014.
- [6] M. Bell, S. Perera, M. Piraveenan, M. Bliemer, T. Latty, and C. Reid, “Network growth models: A behavioural basis for attachment proportional to fitness,” *Sci. Rep.*, vol. 7, no. 1, p. 42431, Mar. 2017.

- [7] S. S. Perera, M. G. H. Bell, M. Piraveenan, D. Kasthurirathna, and M. Parhi, "Topological structure of manufacturing industry supply chain networks," *Complexity*, vol. 2018, pp. 1–23, Oct. 2018.
- [8] T. Y. Choi, K. J. Dooley, and M. Rungtusanatham, "Supply networks and complex adaptive systems: Control versus emergence," *J. Oper. Manage.*, vol. 19, no. 3, pp. 351–366, May 2001.
- [9] S. D. Pathak, J. M. Day, A. Nair, W. J. Sawaya, and M. M. Kristal, "Complexity and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective," *Decision Sci.*, vol. 38, no. 4, pp. 547–580, Nov. 2007.
- [10] M. Piraveenan, U. Senanayake, P. Matous, and Y. Todo, "Assortativity and mixing patterns in international supply chain networks," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 29, no. 2, Feb. 2019, Art. no. 023124.
- [11] R. Mudambi and T. Pedersen, "Agency theory and resource dependency theory: Complementary explanations for subsidiary power in multinational corporations," in *Proc. Bridging IB Theories, Constructs, Methods Across Cultures Social Sci. Basingstoke, Palgrave Macmillan*, 2007, pp. 1–16.
- [12] H. C. White, "Where do markets come from?" *Amer. J. Sociol.*, vol. 87, no. 3, pp. 517–547, Nov. 1981.
- [13] B. Uzzi, "The sources and consequences of embeddedness for the economic performance of organizations: The network effect," *Amer. Sociol. Rev.*, vol. 61, no. 4, pp. 674–698, 1996.
- [14] J. He, X. Mao, O. M. Rui, and X. Zha, "Business groups in China," *J. Corporate Finance*, vol. 22, no. 1, pp. 166–192, Feb. 2013.
- [15] Y. Chandra and I. F. Wilkinson, "Firm internationalization from a network-centric complex-systems perspective," *J. World Bus.*, vol. 52, no. 5, pp. 691–701, Sep. 2017.
- [16] S. P. Willems, "Data set—Real-world multiechelon supply chains used for inventory optimization," *Manuf. Service Oper. Manage.*, vol. 10, no. 1, pp. 19–23, Jan. 2008.
- [17] M. Parhi, "Impact of the changing facets of inter-firm interactions on manufacturing excellence: A social network perspective of the Indian automotive industry," *Asian J. Technol. Innov.*, vol. 16, no. 1, pp. 117–141, Jan. 2008.
- [18] S. D. Pathak, D. M. Dilts, and G. Biswas, "On the evolutionary dynamics of supply network topologies," *IEEE Trans. Eng. Manage.*, vol. 54, no. 4, pp. 662–672, Nov. 2007.
- [19] E. J. S. Hearnshaw and M. M. J. Wilson, "A complex network approach to supply chain network theory," *Int. J. Oper. Prod. Manage.*, vol. 33, no. 4, pp. 442–469, Mar. 2013.
- [20] A. Nair and J. M. Vidal, "Supply network topology and robustness against disruptions—an investigation using multi-agent model," *Int. J. Prod. Res.*, vol. 49, no. 5, pp. 1391–1404, Mar. 2011.
- [21] Y. Hou, X. Wang, Y. J. Wu, and P. He, "How does the trust affect the topology of supply chain network and its resilience? An agent-based approach," *Transp. Res. Part E, Logistics Transp. Rev.*, vol. 116, pp. 229–241, Aug. 2018.
- [22] K. Zhao, A. Kumar, T. P. Harrison, and J. Yen, "Analyzing the resilience of complex supply network topologies against random and targeted disruptions," *IEEE Syst. J.*, vol. 5, no. 1, pp. 28–39, Mar. 2011.
- [23] M. Pero, T. Rossi, C. Noé, and A. Sianesi, "An exploratory study of the relation between supply chain topological features and supply chain performance," *Int. J. Prod. Econ.*, vol. 123, no. 2, pp. 266–278, Feb. 2010.
- [24] *FactSet Supply Chain Relationships: Data and Methodology Guide*. Accessed: Sep. 28, 2017. [Online]. Available: <https://open.factset.com/catalog/products/factset-supply-chain-relationships/en-us>
- [25] Government of the United Kingdom. (2008). *Standard Industrial Classification of Economic Activities (SIC)*. Accessed: Oct. 19, 2018. [Online]. Available: <https://www.gov.uk/government/publications/standard-industrial-classification-of-economic-activities-sic>
- [26] F. C. Santos and J. M. Pacheco, "Scale-free networks provide a unifying framework for the emergence of cooperation," *Phys. Rev. Lett.*, vol. 95, no. 9, Aug. 2005, Art. no. 098104.
- [27] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.
- [28] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya, "Local assortativeness in scale-free networks," *EPL (Europhys. Lett.)*, vol. 84, no. 2, p. 28002, Oct. 2008.
- [29] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya, "Assortativeness and information in scale-free networks," *Eur. Phys. J. B*, vol. 67, no. 3, pp. 291–300, Feb. 2009.
- [30] M. Piraveenan, M. Prokopenko, and A. Y. Zomaya, "Information-cloning of scale-free networks," in *Proc. Eur. Conf. Artif. Life*. Berlin, Germany: Springer, 2007, pp. 925–935.
- [31] E. Kreyszig, *Advanced Engineering Mathematics*, 9th ed. Hoboken, NJ, USA: Wiley, Dec. 2005.
- [32] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford, U.K.: Oxford Univ. Press, Jan. 2003.
- [33] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Mar. 1977.
- [34] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Netw.*, vol. 1, no. 3, pp. 215–239, Jan. 1978.
- [35] L. C. Freeman, S. P. Borgatti, and D. R. White, "Centrality in valued graphs: A measure of betweenness based on network flow," *Social Netw.*, vol. 13, no. 2, pp. 141–154, Jun. 1991.
- [36] U. Brandes, "A faster algorithm for betweenness centrality," Univ. Konstanz, Konstanz, Germany, Tech. Rep., 2001.
- [37] M. E. J. Newman, "A measure of betweenness centrality based on random walks," *Social Netw.*, vol. 27, no. 1, pp. 39–54, Jan. 2005.
- [38] K. Stephenson and M. Zelen, "Rethinking centrality: Methods and examples," *Social Netw.*, vol. 11, no. 1, pp. 1–37, Mar. 1989.
- [39] H. Wang, J. M. Hernandez, and P. Van Mieghem, "Betweenness centrality in a weighted network," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 77, no. 4, Apr. 2008, Art. no. 046105.
- [40] M. Piraveenan, M. Prokopenko, and L. Hossain, "Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks," *PLoS ONE*, vol. 8, no. 1, Jan. 2013, Art. no. e53095.
- [41] G. Sabidussi, "The centrality index of a graph," *Psychometrika*, vol. 31, no. 4, pp. 581–603, Dec. 1966.
- [42] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, Jun. 2006.
- [43] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature*, vol. 435, p. 814, Jun. 2005.
- [44] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 12, pp. 7821–7826, Jun. 2002.
- [45] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti, "Generalized Louvain method for community detection in large networks," in *Proc. 11th Int. Conf. Intell. Syst. Design Appl.*, Nov. 2011, pp. 88–93.
- [46] A.-L. Barabási, "Linked: The new science of networks," *Amer. J. Phys.*, vol. 71, p. 409, 2003, doi: 10.1119/1.1538577.
- [47] M. Bandara, S. Weragoda, M. Piraveenan, and D. Kasthurirathna, "Overlay community detection using community networks," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2018, pp. 680–687.



**MAHENDRA PIRAVEENAN** received the B.Eng. degree (Hons.) in computer systems from the Department of Electrical and Electronic Engineering, The University of Adelaide, in 2004. After a stint at the software industry, he received the Ph.D. degree in complex systems from The University of Sydney, in 2010. He is currently serving as a Senior Lecturer with the Faculty of Engineering, The University of Sydney. He has published more than 30 journal articles, and more than 25 peer-

reviewed international conference proceedings. He has served as a Reviewer for several international journals, including IEEE ACCESS. He is also an Active Assessor for the Australian Research Council. His research interests include complex networks, game theory, epidemics modeling, and multi-agent systems. He is a Review Editor of *Frontiers in Artificial Intelligence* journal.



**HONGZE JING** received the master's degree in complex systems from The University of Sydney, in 2019. His research project was in the area of supply chain management. His research interests include complex systems, complex networks, social networks, and supply chain networks.



**PETR MATOUS** is currently an Associate Dean with the Faculty of Engineering, The University of Sydney. In 2015, after 13 years at The University of Tokyo, he left his position as an Associate Professor in the Department of Civil Engineering to join The University of Sydney. He researches, lectures, and consults with Australian, Japanese, and international organizations in the fields of network science and international development projects. He also participates in research and social projects in diverse communities across Australia, Asia, and Africa to leverage community networks in contexts constrained by limited resources or environmental disasters. He was a recipient of The University of Tokyo President's Award.



**YASUYUKI TODO** received the Ph.D. degree in economics from Stanford University. He has been a Professor with the Graduate School of Economics, Waseda University, since 2014, after serving as the Department Head at the Department of International Studies, The University of Tokyo. He is currently a Faculty Fellow with the Research Institute of Economy, Trade & Industry. He has published more than 50 academic articles in refereed journals, including *Nature Sustainability*, the *Journal of Industrial Economics*, the *Journal of Regional Science*, *Ecological Economics*, *Research Policy*, and *World Development*. His research interests include international economics, development economics, and applied micro-econometrics, also focusing on the role of social and economic network in economic growth and resilience based on firm- and household-level data from various countries.

...