Decoding the complexity of large-scale pork supply chain networks in China

Mengsi Cai
College of Systems Engineering, National University of Defense Technology, Changsha, China
Ge Huang
College of Economy and Management, Changsha University, Changsha, China
Yuejin Tan and Jiang Jiang
College of Systems Engineering, National University of Defense Technology, Changsha, China
Zhongbao Zhou
School of Business Administration, Hunan University, Changsha, China, and
Xin Lu
College of Systems Engineering, National University of Defense Technology, Changsha, China and
School of Business, Central South University, Changsha, China

Received 21 December 2019
Revised 3 April 2020
19 May 2020
Accepted 29 May 2020

Abstract

Purpose – With the development of global food markets, the structural properties of supply chain networks have become key factors affecting the ability to evaluate and control infectious diseases and food contamination. The purpose of this paper is to describe and characterize the nationwide pork supply chain networks (PSCNs) in China and to demonstrate the potential of using social network analysis (SNA) methods for accessing outbreaks of diseases and contaminations.

Design/methodology/approach – A large-scale PSCN with 17,582 nodes and 49,554 edges is constructed, using the pork trade data collected by the National Important Products Traceability System (NIPTS) in China. A network analysis is applied to investigate the static and dynamic characteristics of the annual network and monthly networks. Then, the metric maximum spreading capacity (MSC) is proposed to quantify the spreading capacity of farms and estimate the potential maximum epidemic size. The structure of the network with the spatio-temporal pattern of the African swine fever (ASF) outbreak in China in 2018 was also analysed.

Findings – The results indicate that the out-degree distribution of farms approximately followed a power law. The pork supply market in China was active during April to July and December to January. The MSC is capable of estimating the potential maximum epidemic size of an outbreak, and the spreading of ASF was positively correlated with the effective distance from the origin city infected by ASF, rather than the geographical distance.

Originality/value – Empirical research on PSCNs in China is scarce due to the lack of comprehensive supply chain data. This study fills this gap by systematically examining the nationwide PSCN of China with large-scale reliable empirical data. The usage of MSC and effective distance can inform the implementation of risk-based control programmes for diseases and contaminations on PSCNs.

Keywords Pork supply chain, Network analysis, Maximum spreading capacity, Effective distance, National Important Products Traceability System, African swine fever

Paper type Research paper

Funding: The author acknowledges the National Natural Science Foundation of China under Grant 91846301, 71771213, and 82041020. The other team members of the author were partially supported by the National Natural Science Foundation of China under Grant 71790615, 71901067 and 71690233. This study was also supported by the Hunan Science and Technology Plan Project (2017RS3040, 2018JJ1034).

Author contributions: XL designed the research; MC, GH and JJ performed the experiments and wrote the paper and XL, ZZ and YT helped to revise the manuscript. The authors declare no conflict of interest.

Industrial Management & Data Systems
Vol. 120 No. 8, 2020
pp. 1483-1500

0263-5577
DOI 10.1108/IMDS-12-2019-0689

1. Introduction

The supply chain system has developed rapidly in recent years and has changed from a simple chain system to a complex network system with non-linear, large-scale and high-dimensional features, as well as time and space scale characteristics (Lu and Shang, 2017; Osadchiy *et al.*, 2016). Understanding the complexity of pork supply chain network (PSCN) is a key factor to evaluate and control the spread of related infectious diseases and food contaminations (Dubé *et al.*, 2008; Nöremark *et al.*, 2011; Büttner *et al.*, 2015; Rautureau *et al.*, 2012; Lu *et al.*, 2019; Kukielka *et al.*, 2017). While animal movement databases have been early established in European countries (Relun *et al.*, 2016) and the USA (Lee *et al.*, 2017), little information is available on pig trade transactions in China (Wu *et al.*, 2016).

Trade data (or animal movements) can be analysed as a network structure (directed or undirected and static or temporal), in which the firms (holdings) are treated as "nodes" and trade transactions (movements) are regarded as "links" or contacts. In recent decades, social network analysis methods have been increasingly used to characterize the topology of various supply chain networks from both static and dynamic aspects, in order to achieve a better understanding of spatio-temporal dynamics of infectious diseases, to assess the risk of spreading and to predict and prevent the transmission of infectious diseases and food contaminations (Webb, 2005; Lentz *et al.*, 2009, 2011; Büttner *et al.*, 2013, 2015; Kiss *et al.*, 2006; Martínez-López *et al.*, 2009; Vernon and Keeling, 2012; Bajardi *et al.*, 2011; Dubé *et al.*, 2008, 2009, 2011; Volkova *et al.*, 2010; Rautureau *et al.*, 2012; Robinson *et al.*, 2007; Bigras-Poulin *et al.*, 2007; Lebl *et al.*, 2016; Nöremark *et al.*, 2011; LeBlanc *et al.*, 2015; Natale *et al.*, 2009; Volkova *et al.*, 2010; Lu *et al.*, 2019; Kukielka *et al.*, 2017; Salines *et al.*, 2017).

Most of the research studies were conducted in the countries of Europe and North America, with very few empirical investigations in China. For instance, Wang *et al.* developed a modelling framework to examine the effect of market power on the food safety of the pork industry in China by using data collected from a pork firm survey (Wang *et al.*, 2019). Han *et al.* investigated interfirm exchange relationships and food quality management in China's pork supply chain with the data from 229 pork processors in eastern China (Han *et al.*, 2011). Zhou *et al.* explored the role of live bird markets biosecurity indicators and poultry movement in the avian influenza A (H7N9) affected areas in China with the data from a cross-sectional survey (Zhou *et al.*, 2015). It is clear that studies in China have been focussed on theoretical research from economic or managerial aspects due to the absence of data on comprehensive food traceability systems.

China is the largest pork producer and consumer in the world (Shao *et al.*, 2018); however, empirical evidence informing the PSCN in the Chinese context is scarce and there is a large research gap in this field. In recent years, spurred by food contamination scandals like Shuanghui sausage involving clenbuterol in 2011 and the outbreak of African swine fever (ASF) in 2018 (Wang *et al.*, 2018), food safety is becoming one of China's largest social concerns and has attracted global attention (The Lancet, 2012). A better understanding of the PSCN in China is required to inform the implementation of food safety improvement and disease control.

Based on the above considerations, this paper analysed the topology of a large-scale national PSCN in China and the spread of ASF outbreak in 2018. In the analysis, the static and dynamic structures of the PSCN were characterized, the metrics of maximum spread capacity (MSC) and maximum infection chain were proposed to quantify the epidemic size and infection risk of an outbreak of a disease or contamination. Finally, the spatio-temporal pattern of the ASF outbreak in China was investigated.

The structure of the paper is organized as follows. Section 2 reviews the extant literature on studies on PSCNs in different countries. Section 3 introduces the data source and network analysis methods. In Section 4, the results are presented. The conclusions and future works are discussed in Section 5.

2. Literature review

The development of animal movement databases and food traceability systems has contributed a lot to the empirical investigations on the networks of cattle (Dubé *et al.*, 2008; Vernon and Keeling, 2012; Lentz *et al.*, 2009; Bajardi *et al.*, 2011; Natale *et al.*, 2009; Robinson *et al.*, 2007; Christley *et al.*, 2005a, b), sheep (Webb, 2006; Kiss *et al.*, 2006; Volkova *et al.*, 2010) and pig movements (Lentz *et al.*, 2011; Büttner *et al.*, 2013, 2015; Lee *et al.*, 2017; Relun *et al.*, 2016; Kukielka *et al.*, 2017; Salines *et al.*, 2017; Lebl *et al.*, 2016). The social network analysis (SNA) is becoming the methodology of choice to describe PSCN structures (Kim *et al.*, 2011). The major relevant studies in different countries are presented in Table 1. In European countries like Sweden, France and Germany, livestock movements are mandatorily registered to ensure traceability, which allows scientists to characterize the structural patterns of trade networks in the livestock industry. In the USA, multisite systems are also established to provide animal movement data. However, in some developing countries such as Georgia (Kukielka *et al.*, 2017) and China (Ji *et al.*, 2012), data collection is generally limited such that small-scale data are often obtained by questionnaires or simulations in studies.

In the field of epidemiology, the SNA also has been used to identify potential superspreaders or superreceivers of diseases that may affect swine as well as other livestock species. A number of network measures are used to investigate the potential spread of infectious diseases, such as degree measures (Christley *et al.*, 2005a, b) and components (Kao *et al.*, 2006; Büttner *et al.*, 2013). The infection chain is also used to estimate the potential maximum epidemic size of an outbreak (Nöremark *et al.*, 2011; Büttner *et al.*, 2015; Rautureau *et al.*, 2012). The infection chain of one farm is defined as the number of farms that are directly and indirectly connected to the farm through animal movements, with the consideration of time order of the movements (Dubé *et al.*, 2008). In addition, Lebl *et al.* demonstrated that the

Country	Data	Node type	Nodes	Edges	Estimation metrics	Reference
Sweden	National data (2006–2008)	Nucleus or multiplier; sow pool; farrow to grow; farrow to finish and fattening	2,506–2,977	18,138–20,807	In-degree; out-degree; ingoing infection chain and outgoing infection chain	Nöremark *et al.*, (2011)
France	National data (2010.01–2010.06)	Breeding; farrow-to-grow; farrow-to-finish; growing; grow-to-finish and finishing herd	13,968	155,154	Degree, betweenness and ingoing infection chain	Rautureau *et al.*, (2012)
Germany	A producer community (2006–2009)	Multiplier; farrowing farm; farrow-to-finishing farm; finishing farm and abattoir	483	926	Weakly connected components; strongly connected components; out-degree and outgoing infection chain	Büttner *et al.*, (2013)
The USA	One multisite system (2012–2014)	Sow farm; gilt development unit; boar stud; nursery; wean-to-finish and finishing farm	500	109,868	Centrality measures	Lee *et al.*, (2017)

Table 1. Studies on pork supply chain networks in different countries

network activity is an important factor in evaluating the effects of a disease spread in the German pig trade network (Lebl *et al.*, 2016). Besides SNA methods, epidemiological models, such as the susceptible–infected–recovered (SIR) model, are also used for assessing the spread of infectious diseases in PSCNs (Büttner *et al.*, 2016; Lebl *et al.*, 2016).

China is the biggest producer and consumer of pork (Shao *et al.*, 2018), simultaneously, food safety problems linked to pork products have been repeatedly reported (The Lancet, 2012). The outbreak of ASF in 2018 has also resulted in devastating economic impacts on the pork market as well as on food safety issues. The current pork industry in China is characterized by the dominant position of smallholding pig producers and slaughterhouses (Han *et al.*, 2011), and the structure of pig farmers is simple and non-professional, which is different from that of European countries and the USA (see Table 1). The organization of such a fragmented pork chain in China induces difficulties in tracking food contamination and diseases (Wu *et al.*, 2016).

To mitigate risk and improve food safety in China, studies on risk management, food safety management systems, quality management and cost-effectiveness are conducted with theoretical models, simulations and questionnaires (Han *et al.*, 2011; Yu *et al.*, 2013; Ji *et al.*, 2012). However, very few studies have presented an empirical analysis of real-world PSCNs due to the difficulties in data collection. Our study on the national PSCN in China based on large-scale reliable empirical data fills these gaps. We present an empirical study on the use of SNA methods for large-scale PSCN analysis in China, to quantify the structural characteristics, to estimate the potential epidemic size of an outbreak and to investigate the spatio-temporal spread pattern of ASF on the network.

3. Materials and methods

3.1 Data description

To improve food safety and promote the Chinese government's "Internet and Agriculture Act", the National Important Products Traceability System (NIPTS), a national food traceability system in China, was initiated under the instructions of the State Council and was constructed by the Chinese Ministry of Commerce in 2014. To examine the effectiveness of the NIPTS, 58 pilot cities divided into five batches were involved in the system in different time periods. A pilot city is a city selected by the Chinese Ministry of Commerce to be one of the first to be involved in the NIPTS. The NIPTS covers a wide diversity of products, such as meat, vegetables, Chinese medicinal crops, wine, etc., and the System involves more than 13,500 corporations and 200,000 shops and contains billions of transaction records (Ministry of Commerce of the People's Republic of China, 2017).

We used the pork trade data from the first batch of ten pilot cities (Dalian, Qingdao, Nanjing, Wuxi, Suzhou, Shanghai, Ningbo, Hangzhou, Chongqing, Chengdu and Kunming) in the NIPTS covering the period from 1 January 2015 to 31 December 2015. Most of the pilot cities were located in eastern China, which is one of the most important pig production and pork consumption regions in China. And Chengdu city is located in the largest pig production province, Sichuan. The data contain detailed records of pig trade transactions from farms to slaughterhouses ($F \rightarrow S$) and pork trade transactions from slaughterhouses to retailers ($S \rightarrow R$). These transactions are daily recorded, and the data items include the date of transaction, volume, price and other specific information. There are 6,686,175 pork transactions between 17,582 supply chain participants in 254 cities, which allow us to construct a large-scale nationwide PSCN in China.

3.2 Network construction

In China's pork supply chain, there is a lack of a professional classification of pig farms, such as core groups, breeding groups, etc. In the NIPTS, three major sections of China's pork supply chain are considered, i.e. upstream pig raising, midstream slaughtering and processing. Major pork supply chain participants include farms, slaughterhouses and retailers. Pigs produced by different farms are processed by slaughterhouses then the pork in slaughterhouses is delivered to retailers. Based on the trade data in the NIPTS, we thus used a three-layer model to describe how products move from productions to consumptions in the PSCN (Figure 1).

Setting participants as nodes and transactions as edges, we transformed trading data in the NIPTS into a network. Representing the network as a directed graph, $G = (V, E)$, in which V denotes the nodes and E denotes the set of edges. $N = |V|$ is the number of nodes in G and $M = |E|$ represents the number of edges. Thereby, an annual network including $N = 17,582$ nodes and $M = 49,554$ edges is constructed using the trade data in 2015. The weight of an edge is counted by the sum of pork trade volume on this edge, i.e. between the pair of two nodes. In addition, we used the daily transactions to construct temporal networks, in which nodes and edges are added or removed by day. Hence, an edge in the temporal network can be represented by triple variables $(i, j, t_{i,j})$, where i and j denote the nodes and $t_{i,j}$ represents the forming time of the transaction between i and j. We set the time window Δt = one month to slice the network, then the PSCN is separated into 12 monthly subnetworks.

3.3 The network analysis

A variety of global and local social network metrics are used to characterize the topology of the annual network and 12 monthly networks. In addition, inspired by the definition of infection chain (Dubé *et al.*, 2008), we propose the MSC as a new metric to quantify the maximum spreading range of a farm and to estimate the potential maximum epidemic size of an outbreak, we also use maximum infection chain to evaluate the risk of a retailer of being infected. All metrics are introduced as follows.

At the network level, average degree indicates the average number of links for each node connected within the network, calculated by $< k > = (\sum k_i/N)$. A higher average degree implies increased interconnectivity amongst nodes (Freeman, 1978). Density is the proportion of the links present M in comparison to the number of all possible links $N(N-1)/2$, which ranges between 0 and 1 (Wasserman and Faust, 1994). Considering the network in this paper is directed and has three layers, then the number of all possible links is calculated by $N_1 \times N_2 + N_2 \times N_3$, in which $N_i\,(i = 1, 2, 3)$ represents the number of nodes in the layer i of the network. Degree assortativity is a preference for a network's nodes to attach to others that are similar in some way and often operationalized as a correlation ranging between -1 (completely disassortative) and 1 (perfectly assortative) (Newman, 2002). Degree assortativity is helpful to explore the cooperation pattern of supply chain firms, for example, an assortative pattern to a certain extent indicates a situation of win–win cooperation. Network centralization provides a value between 0 (if all nodes have the same connectivity in the network) and 1 (if the network has a star topology), which explains whether the network has a

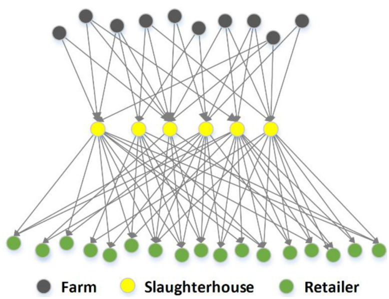


Figure 1.
The three-layer pork supply chain network framework in the empirical National Important Products Traceability System data

structural concentration trend (Kim *et al.*, 2011). Network heterogeneity is the coefficient of the variation in connectivity, a supply chain network with high-network heterogeneity exhibits hub nodes that have a large number of contractual connections (Perera *et al.*, 2017). A weakly connected component (WCC) is a part of a network where all nodes are connected by at least one path through the network, not taking the direction of the contact into account (Kao *et al.*, 2006). A strongly connected component (SCC) is a part of the network where all nodes can reach each other through directed links, either directly or via other nodes (Kao *et al.*, 2006).

At the node level, *in-degree* (k_i^{in}) and *out-degree* (k_i^{out}) denote the numbers of ingoing (predecessor nodes) and outgoing edges (successor nodes), respectively, and measure how well connected a node i is (Freeman, 1978). Betweenness centrality $C_B(i)$ measures the extent to which a node is located on paths between other nodes and is a measure of centrality in a graph based on the shortest paths (Wasserman and Faust, 1994). In supply chain networks, a firm with high-betweenness centrality indicates the higher importance of its influence on product flows.

The MSC defines the number of retailers that are indirectly connected to one farm through pig and pork trade transactions, taking the sequential order of transactions into account. The maximum infection chain (MIC) defines the number of firms (including farms and slaughterhouses) that are directly or indirectly connected to one retailer through pig and pork trade transactions, and it also takes the sequential order of transactions into account.

4. Results and discussion

4.1 Static topologic characteristics of the network

The annual PSCN is illustrated in Figure 2, noting that the coordinates of nodes are not related to its geographical position. Table 2 shows the topological characteristics of the annual network and 12 monthly networks. At the network level, the low network density (0.0209) indicates that only 2.09% of possible edges were present in the annual network. The PSCN was found to be disassortative with negative degree assortativity, indicating that firms

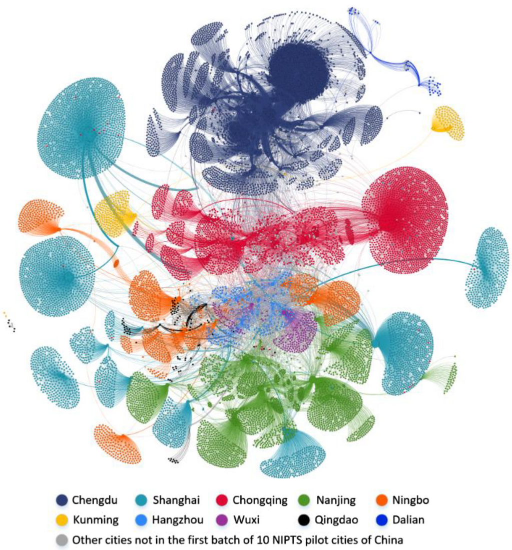


Figure 2.
Visualization of the pork supply chain network with 17,582 nodes and 49,554 links, using trading data for the first batch of ten pilot cities in the National Important Products Traceability System in china from 1 january 2015 to 31 december 2015

	Annual network	Monthly networks mean (range)
Nodes		
Farms	1,198	576 (497–632)
Slaughterhouses	136	124 (120–131)
Retailers	16,248	9,117 (7,877–11,546)
Edges		
$F \rightarrow S$ links	4,322	1,356 (1,130–1,510)
$S \rightarrow R$ links	45,232	13,425 (8,702–29,060)
Average degree	5.64	2.92 (2.27–4.97)
Average in-degree		
Farms	0	0
Slaughterhouses	31.78	10.87 (8.83–11.98)
Retailers	2.78	1.42 (1.08–2.52)
Average out-degree		
Farms	3.61	2.35 (2.27–2.43)
Slaughterhouses	332.59	106.69 (71.33–225.27)
Retailers	0	0
Density	0.0209	0.0118 (0.0093–0.0195)
Degree assortativity		
in-in	−0.207	−0.265 (−0.413to −0.176)
in-out	−0.091	−0.135 (−0.170 to −0.060)
out-in	−0.290	−0.235 (−0.411–0.216)
out-out	−0.243	−0.184 (−0.211 to −0.154)
Betweenness centrality	0.000029	0.0021 (0.0013–0.0025)
Weakly connected component		
Number of WCC	3	7 (4–9)
Size of largest WCC	17,573	9,684 (8,414–12,167)
Size of second largest WCC	7	53 (25–70)
Strongly connected component		
Number of SCC	17,582	9,818 (8,567–12,264)
Size of largest SCC	1	1
Size of second largest SCC	1	1
MSC of farms	789 (0–5,797)	312 (243–388)
MIC of retailers	61 (1–369)	21 (17–24)

Table 2.
Topological characteristics of the annual network and 12 monthly networks

tend to connect to others with a degree different from themselves; similar results can be found in the study by Perera *et al.* (2017). It also suggests that with hub nodes not connecting to each other, the network could achieve resilience to the cascading impacts of food contamination or disease transmission. The low-network centralization (0.12) indicates the widely distributed and decentralized nature of modern PSCNs. Moreover, the high-network heterogeneity (10.14) implies that the pork supply chain contains hub nodes and can realize centralized control through very few firms. In addition, the PSCN had three WCCs, with the largest one including almost all firms (17,573 and 99.9%).

At the node level, since slaughterhouse is the only bridge between farm and retailer in the PSCN, it is obvious that the slaughterhouses had non-zero average betweenness centrality (0.000029). In addition, the slaughterhouses had on average 332.59 downstream partners (out-

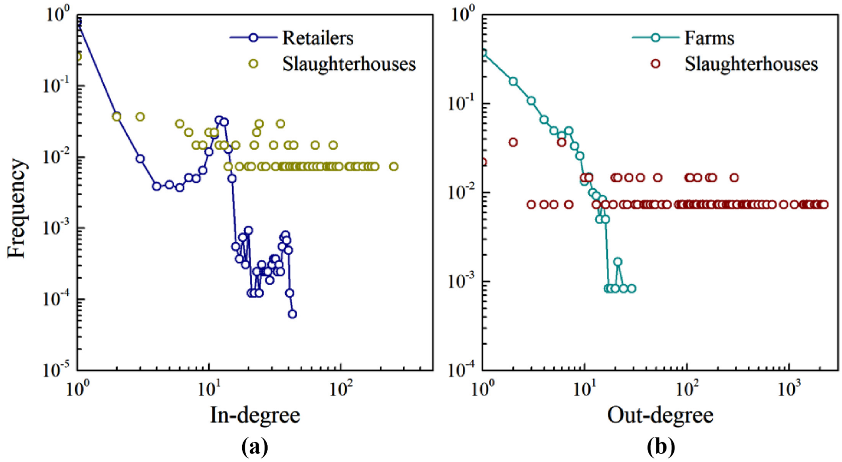


Figure 3.
The in-degree and out-degree distributions of farms, slaughterhouses and retailers in the pork supply chain network

degree), which were nearly 100 times greater than the farms (3.61) and had on average 31.78 upstream partners (in-degree), which were 11 times greater than the retailers (2.78). It is straightforward to conclude that the slaughterhouses should be continuously supervised as they had higher risks for both spreading and getting infected by diseases or contaminations. In comparison with the PSCNs in other countries, where the degree distributions of all firm types was correctly skewed (Nöremark *et al.*, 2011; Büttner *et al.*, 2013, 2015), in the PSCN of China only the out-degree distribution of farms approximately followed a power law (Figure 3), indicating that the spreading capacity is highly resilient towards random quarantine of farms (Albert *et al.*, 2000). However, with strategic removal of the most central farms, e.g. by trade restrictions or selective vaccination or culling, a rapid fragmentation of the trade network could be expected. From an epidemiological perspective, such a fragmentation of the PSCN will be important to control a spread.

In the annual network, the MSC of all 1,198 farms ranged between 0 and 5,797, with an average of 789. A farm with the largest MSC was indirectly connected to 5,797 retailers, which took up 35.7% of the retailers throughout the whole PSCN. If this farm becomes the contamination source, there is a high risk of probably infecting over a quarter of the retailers in the network. The MIC of all 16,248 retailers ranged between 1 and 369, with an average of 61. Surveillance efforts should be focussed on the farms with large MSC and retailers with large MIC, when an outbreak of diseases or contaminations is observed.

4.2 Dynamic characteristics of the network

4.2.1 Temporal dynamics of transactions. In comparison with the annual network, there were 9,818 (55.8%) firms active in the monthly networks with 14,781 (29.8%) edges. The variation in the number of active nodes is shown in Figure 4, which indicates a seasonal pattern of the pork supply market in China. The pork trading was significantly active between April and July and inactive during Chinese New Year. In addition, the pork supply market was clearly affected by summer vacation and festivals. The pork trade volume held a higher level on the International Labour Day (May), summer vacation (July and August) and the National Day (October). The numbers of farms, slaughterhouses and retailers all sharply decreased in February, when people were celebrating the national holiday of the Spring Festival (Chinese New Year), before which the pork preparation was done. The numbers of active farms and slaughterhouses were also low in August and September as it was probably because people

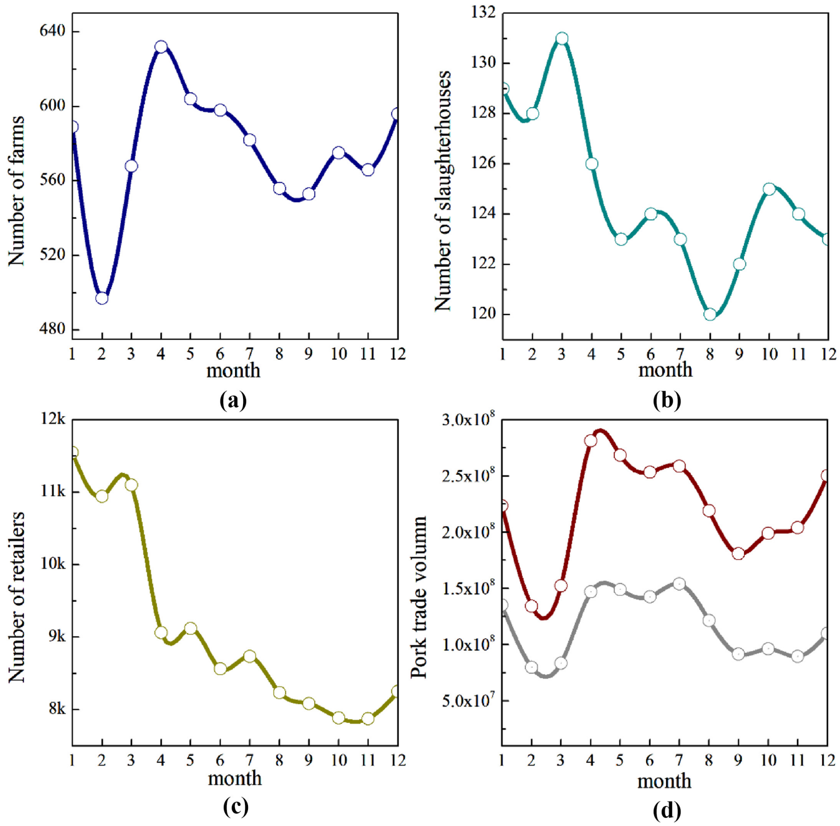


Figure 4.
Variation in the numbers of active (a) farms, (b) slaughterhouses and (c) retailers, as well as (d) the weight of links (or pork trade volume). In (d), the above and below lines represent $F \rightarrow S$ and $S \rightarrow R$ links, respectively

usually purchased enough pork before August for summer vacation and for the celebration of graduation or matriculation. This result is similar to Ontario, in which shipments are found to be less active in the summer months (Dubé *et al.*, 2008).

In comparison to the seasonal variation in livestock transactions in other countries, there are some specific characteristics of the variation in China. First, a high proportion of farms and slaughterhouses remained active in April. Second, there were two active trading periods: (1) From April to July: Since the outbreak of an epidemic after last winter, pigs were in short supply. The event of "black hoof pig" in March, where black-hoofed pigs with foot-and-mouth disease entered circulation links, which made many qualified pigs unable to enter the market for the strict control of government and companies. After this decline in pork consumption, there appeared a great recovery in the beginning of April, which is similar to the findings on the increase of transactions during spring in Italy (Natale *et al.*, 2009), Sweden (Nöremark *et al.*, 2011) and Great Britain (Robinson *et al.*, 2007). (2) December and January: Another period of active trades occurred at the beginning and in the middle of winter, which is similar to Great Britain (Christley *et al.*, 2005b). However, as a comparison, there are also some countries with no variation, such as Denmark (Bigras-Poulin *et al.*, 2007) and Germany (Büttner *et al.*, 2015), which were explained by planned pig production throughout the year or incomplete pork trade data.

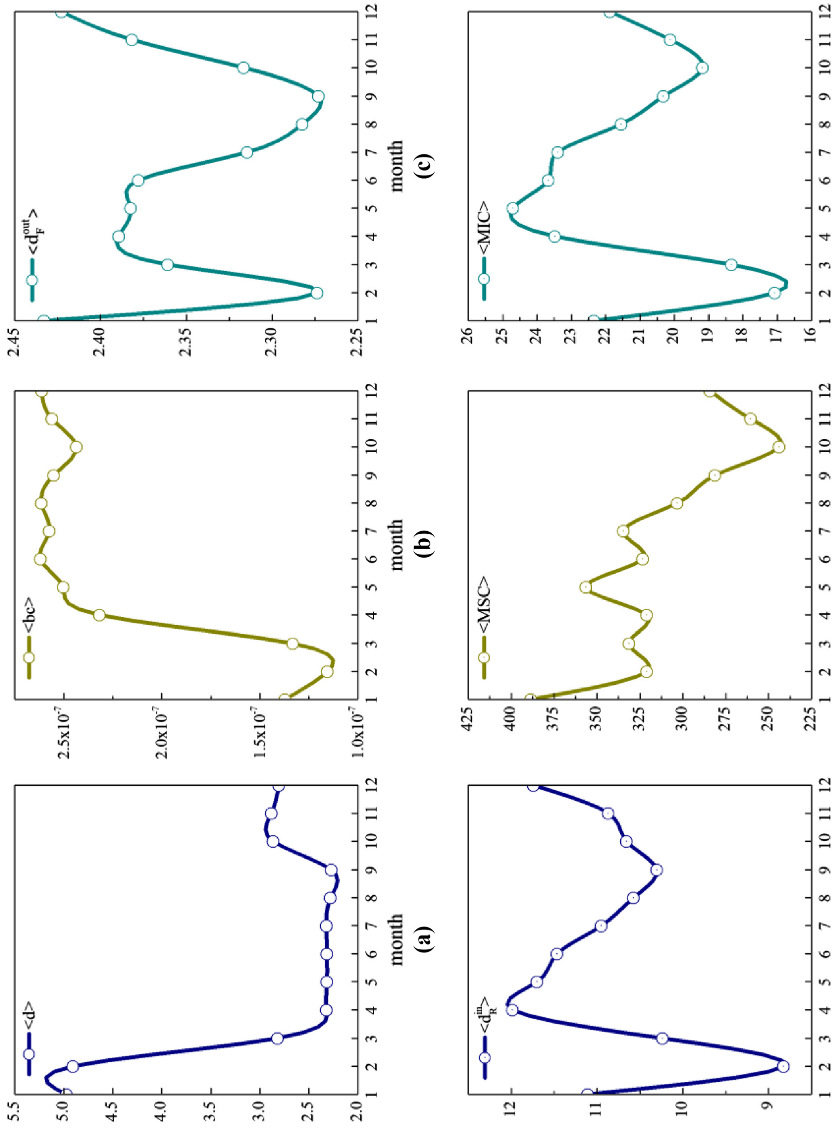


Figure 5.
Temporal properties of (a) average degree $< d >$, (b) average betweenness centrality $< bc >$, (c) average out-degree of farms $< d_F^{out} >$, (d) average in-degree of retailers $< d_R^{in} >$, (e) average maximum spreading capacity of farms $< MSC >$ and (f) average maximum infection chain of retailers $< MIC >$ for the 12 monthly networks

4.2.2 Temporal characteristics of network properties. As shown in Figure 5, when the pork trade volume reached a peak between April and July (as shown in Figure 4), the average degree $< d >$ reached an off-peak between April and September. The large pork trade volume of firms that had a low average network degree indicates that many productive firms sold large volumes of products to their trading partners. In contrast to $< d >$, the average betweenness centrality $< bc >$ reached a peak between April and August, when the pork trade volume reached its peak but the number of supply chain participants showed a decreasing trend. With higher betweenness centrality between April and December, slaughterhouses played an important role in connecting farms and retailers in the PSCNs, enabling the entire pork trading network to operate efficiently.

Spearman's rank correlation was performed to measure the relationships amongst pork trade volume, pig price, pork price, $< d >$ and $< bc >$. As shown in Table 3, a strong positive correlation was found between pig/pork prices and $< bc >$, indicating that the supply chain participants tended to develop new trading relationships when the price increases. A notable negative correlation between $< d >$ and $< bc >$ indicates that with an increase in the number of trading partners, a supply chain participant will have less influence on the trading flow within the whole supply chain network and have difficulty in building a so-called bridge between two trading components.

4.3 Estimating the potential epidemic size

From an epidemiological perspective, we used MIC to evaluate the risk for retailers of being infected by diseases or contaminations. As shown in Figure 5, the in-degree of the retailers $< d_R^{in} >$ was not equivalent to the MIC, the values of $< MIC >$ were about two times larger than the values of $< d_R^{in} >$, indicating that there were retailers with limited direct contacts but a large number of indirect contacts. The $< MIC >$ showed a similar seasonal pattern with the dynamics of the pork trade volume and the average betweenness centrality. A higher $< MIC >$ between April and July indicates that retailers have a higher risk of being infected in these months than in other periods. Taking the indirect contacts and the chronological order of trade transactions into account, the MIC could be used to support diseases or contaminations control for risk-based surveillance, the retailers with high MIC receive many contacts from other supply chain participants.

We used three categories of measures (components, out-degree and MSC) to estimate the maximum potential epidemic size of an outbreak. The components measure uses the size of the largest SCC as the lower bound of the maximum epidemic size and the size of the largest WCC as the upper bound of the maximum epidemic size (Kao *et al.*, 2006). The out-degree measure uses the 99th percentile and the maximum of the out-degree to estimate the lower and upper bounds of the potential maximum epidemic size, respectively. The MSC measure estimates the lower bound of the potential maximum epidemic size with the 90th percentile of the MSC and estimates the upper bound with the maximum value. These measures offer straightforward quantification on the potential extent of contamination or disease at the

Table 3. The correlation between pig and pork price, pork trade volume, average degree $< d >$ and average betweenness centrality $< bc >$

Coefficient	Pork trade volume	Pig price	Pork price	$< d >$	$< bc >$
Pork trade volume	1				
Pig price	−0.076	1			
Pork price	0.12	0.967**	1		
$< d >$	−0.427	−0.384	−0.537	1	
$< bc >$	0.544	0.667*	0.817**	−0.812**	1

Note(s): ** and * denote statistical significance at the 1% and 5% confidence levels, respectively

Table 4.
Comparison of three network measures to estimate the potential maximum epidemic size for 12 monthly networks

	Components		Maximum spreading capacity		Out-degree	
Month	Lower	Upper	Lower	Upper	Lower	Upper
Jan	1	12,167	1,045	2,499	7	2,032
Feb	1	11,421	917	2,290	7	2,036
Mar	1	11,636	918	2,211	8	1,439
Apr	1	9,790	902	2,737	8	998
May	1	9,739	977	2,448	8	1,005
June	1	9,122	899	2,277	9	1,001
July	1	9,242	897	2,381	10	957
Aug	1	8,735	808	2,034	10	939
Sep	1	8,646	767	2,147	9	995
Oct	1	8,474	703	2,141	16	958
Nov	1	8,414	647	1,784	17	937
Dec	1	8,832	710	2,035	15	880
Summary	1 (1–1)	9,684 (8,414–12,167)	849 (647–1,045)	2,248 (1,784–2,737)	10 (7–17)	1,181 (880–2,036)

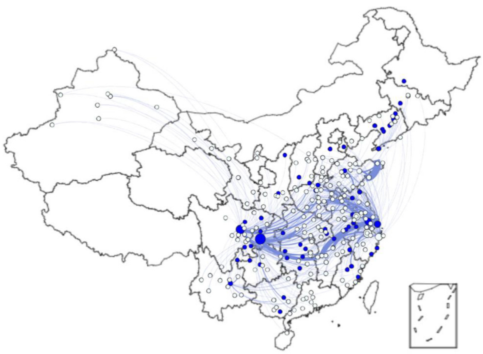
beginning of an investigation and can be used to guide decision-makers in the choice and application of control measures.

The estimation results are shown in Table 4. The components measure provided identical lower bound values in all months as no SCC was observed in 12 monthly networks. And the components overestimated the upper (4–10 times higher) bound of maximum epidemic size than that of MSC and out-degree. Similar results can be found in the studies by Büttner *et al.* (2015) and Dubé *et al.* (2008). The MSC measure provided higher, lower and upper bounds of potential maximum epidemic size than the out-degree measure. The estimates provided by the MSC indicate that an outbreak of diseases or contaminations could infect 849–2,248 firms in the PSCN. From a policy perspective, the estimates should be taken into consideration when trying to determine the strength of control measures and control measures can be adopted first in the farms with the high MSC since they have many contacts with other supply chain participants. In addition, all these three estimation measures decrease from the beginning to the end of the year, which indicates a decline in the spreading capacity of the PSCN through the year.

4.4 Quantifying the outbreak of African swine fever

ASF has resulted in unprecedented disasters and challenges to the Chinese swine industry (Wang *et al.*, 2018). On August 3, 2018, China reported the first outbreak of ASF in Shenyang, a northeastern city in Liaoning Province, China. By the end of 2018, the outbreaks have been reported in 23 provinces and municipalities across China. Over 100 farms and slaughterhouses in 97 cities were infected between 3 August 2018 and 20 December 2019 and more than 200,000 pigs were culled.

To quantify the transmission of ASF amongst Chinese cities, we constructed a directed city network as shown in Figure 6, using the geographical location information of each firm in the PSCN. There were 254 nodes (cities) and 712 directed edges (transactions between cities) in the city network and 58 cities in the network were infected with ASF. We used the effective distance proposed in the study by Brockmann and Helbing (2013) to describe these national dynamics of ASF. Given the flux fraction $0 \leq P_{ij} \leq 1$, i.e. the fraction of pork trade volume leaving city j and arriving at city i through the directed city network, we defined the effective distance d_{ij} from a city j to a connected city i as $d_{ij} = (1 - \log P_{ij}) \geq 1$. A small fraction of pork trading volume $j \rightarrow i$ is effectively equivalent to a large distance and vice versa. Based on the



Lines represent product flows along direct connections between 254 cities in China. Cities infected with African swine fever are marked blue. The larger the node is, the larger the degree of the node

Figure 6. Contagion phenomena of African swine fever in China based on the city network converted from the pork supply chain network with geographical locations

effective distance, d_{ij}, we define the effective distance D_{ij} from an arbitrary reference node j to another node in the directed city network by the length of the shortest path from j to i, as follows: $D_{ij} = \min\lambda(\Gamma)$, where $\lambda(\Gamma)$ is the sum of directed effective lengths along the legs of the ordered path $\Gamma = n_1, n_2, \ldots, n_m$.

Figure 7 presents the correlation of epidemic arrival times T_{ASF} with effective distances D_{eff} and geographical distance D_g from the origin city on the basis of 2018 data on ASF in China. Arrival time is defined as the date of the first ASF virus-confirmed case in a given city after the initial outbreak on 3 August 2018. The origin city is the city where the first confirmed case occurred. On a national scale, T_{ASF} weakly correlates with geographical distance $D_g(R^2 = 0.19)$. The ASF data exhibit a linear relationship ($R^2 = 0.73$) between arrival time and effective distance from the origin city, though there are only six available samples of effective distance. Therefore, we consider that the spreading speed and arrival times of ASF can be calculated through the effective distance if the network data are available and complete. The transmission analysis results of effective distance are helpful for governments to effectively control the transmission of ASF by blocking transports of live pigs and pig products according to the PSCN structures. Also, for an outbreak of other diseases or contaminations, control measures can be adopted in the firms connected to the origin of the outbreak with short effective distances.

5. Conclusions

To characterize the topology structures and their connection to the risk of disease outbreak and contamination, we constructed a directed three-layer (farms, slaughterhouses and retailers) national PSCN including 17,582 nodes and 49,554 edges, using pig and pork trade data extracted from the first ten pilot cities in the NIPTS in China in 2015. We investigated the static and dynamic characteristics of the annual network and the 12 monthly networks and proposed MSC to quantify the maximum spreading range of farms and to estimate the potential maximum epidemic size of an outbreak. Another metric MIC was used to evaluate

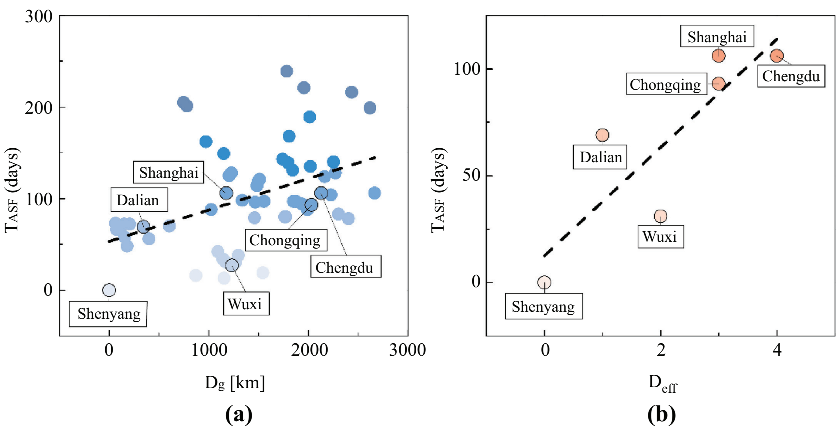


Figure 7.
Contagion phenomena of African swine fever in the pork supply chain network based on the geographical distance D_g and the effective distance D_{eff}

(a) Arrival time versus geographical distance from the source (Shenyang) for the 2018–2019 African swine fever pandemic in China. (b) The linear relationship between arrival time and effective distance for the 2018 outbreak of African swine fever in China. The effective distance is computed based on the city network converted from the pork supply chain network in the empirical National Important Products Traceability System data the risk of retailers of being infected. In addition, we analysed the spatio-temporal pattern of the ASF outbreak in China in 2018 based on the structure of the network.

China is the largest pork producer and consumer in the world; however, empirical studies on the PSCN in China are still few and there is a large research gap in this field. In comparison with the PSCNs in other countries, the PSCN in China shows distinct properties, e.g. large network size, simple classifications of pig producers, etc. The proposed measures of MSC and MIC integrate the spreading risk and supply chain structure and can be used to inform the implementation of risk-based control programmes for diseases and contaminations on PSCNs.

As MSC is capable of assessing the risk of farms of spreading disease or contamination, when the source is uncertain, control measures are suggested to be adopted in farms with high MSC since they have many contacts with other supply chain participants through trade transactions. In addition, the MSC provides the best estimates of the potential maximum epidemic size of an outbreak in the PSCN of China, in comparison with the out-degree and components measurements. The MSC considers both the chronological order and the directed nature of the contacts in the network; thus, it neither underestimates the maximum epidemic size like out-degree nor overestimates the maximum epidemic size like WCCs.

By analysing the network structure and the dynamic of the ASF outbreak, a strong positive correlation is found between the onset of ASF in each city and the effective distance from this city to the origin city with the very first confirmed ASF case. This finding can enable policymakers to understand complex contagion dynamics in the PSCN in China and inform control strategies to reduce the final outbreak size. Combining the metrics MSC and the effective distance, it is suggested that policymakers should strengthen quarantine measures on firms with high MSC in areas with short effective distances to the original outbreak place, in order to reduce the final epidemic size of an outbreak.

There are several practical implications for stakeholders in a food supply chain, e.g. farms, slaughterhouses and retailers. Usually the processing site, in this case the slaughterhouse, is connecting the producer and retailer and serving as hub nodes in the network. Therefore, they are naturally associated with higher risks of being infected and of accelerating the spread. For such sites, food safety compliance and food safety control should be seriously improved and strengthened. For retailers, they can assess the risk of infection by calculating MIC and then take appropriate precautions to minimize the effects of an outbreak. In addition, the effective distance offers a novel tool for estimating the outbreak source; when the source is known, this method can also be used to assess the risk of infection and to support the designing of effective quarantine measures, such as cutting off trade along the most risky path.

The present study made an important step in the study of nationwide PSCNs. However, we only obtained pig and pork trade data in the first ten pilot cities in 2015, though the NIPTS in China has expanded to 58 pilot cities later on. This data limitation resulted in a partial picture of the nationwide PSCN in China and a limited sample size of effective distances. In the future, we expect to obtain more recent and widely covered NIPTS data to reveal a complete and clearer picture of the food supply chain networks in China.

The following abbreviations are used in this manuscript:

Abbreviations

NIPTS	National Important Product Traceability System
PSCN	Pork supply chain network
MSC	Maximum spreading capacity

MIC Maximum infection chain
ASF African swine fever

References

Albert, R., Jeong, H. and Barabási, A.L. (2000), “Error and attack tolerance of complex networks”, *Nature*, Vol. 406 No. 6794, pp. 378-382.

Bajardi, P., Barrat, A., Natale, F., Savini, L. and Colizza, V. (2011), “Dynamical patterns of cattle trade movements”, *PloS One*, Vol. 6 No. 5, p. e19869.

Bigras-Poulin, M., Barfod, K., Mortensen, S. and Greiner, M. (2007), “Relationship of trade patterns of the Danish swine industry animal movements network to potential disease spread”, *Preventive Veterinary Medicine*, Vol. 80 Nos 2-3, pp. 143-165.

Brockmann, D. and Helbing, D. (2013), “The hidden geometry of complex, network-driven contagion phenomena”, *Science*, Vol. 342 No. 6164, pp. 1337-1342.

Büttner, K., Krieter, J., Traulsen, A. and Traulsen, I. (2013), “Static network analysis of a pork supply chain in Northern Germany—characterisation of the potential spread of infectious diseases via animal movements”, *Preventive Veterinary Medicine*, Vol. 110 Nos 3-4, pp. 418-428.

Büttner, K., Krieter, J. and Traulsen, I. (2015), “Characterization of contact structures for the spread of infectious diseases in a pork supply chain in northern Germany by dynamic network analysis of yearly and monthly networks”, *Transboundary and Emerging Diseases*, Vol. 62 No. 2, pp. 188-199.

Büttner, K., Krieter, J., Traulsen, A. and Traulsen, I. (2016), “Epidemic spreading in an animal trade network–Comparison of distance-based and network-based control measures”, *Transboundary and Emerging Diseases*, Vol. 63 No. 1, pp. e122-e134.

Christley, R.M., Pinchbeck, G.L., Bowers, R.G., Clancy, D., French, N.P., Bennett, R. and Turner, J. (2005a), “Infection in social networks: using network analysis to identify high-risk individuals”, *American Journal of Epidemiology*, Vol. 162 No. 10, pp. 1024-1031.

Christley, R.M., Robinson, S.E., Lysons, R. and French, N.P. (2005b), “Network analysis of cattle movement in Great Britain”, in *Proceedings of a Meeting–Society for Veterinary Epidemiology and Preventive Medicine*, pp. 234-243.

Dubé, C., Ribble, C., Kelton, D. and McNab, B. (2008), “Comparing network analysis measures to determine potential epidemic size of highly contagious exotic diseases in fragmented monthly networks of dairy cattle movements in Ontario, Canada”, *Transboundary and Emerging Diseases*, Vol. 55 Nos 9-10, pp. 382-392.

Dubé, C., Ribble, C., Kelton, D. and McNab, B. (2009), “A review of network analysis terminology and its application to foot-and-mouth disease modelling and policy development”, *Transboundary and Emerging Diseases*, Vol. 56 No. 3, pp. 73-85.

Dubé, C., Ribble, C., Kelton, D. and McNab, B. (2011), “Estimating potential epidemic size following introduction of a long-incubation disease in scale-free connected networks of milking-cow movements in Ontario, Canada”, *Preventive Veterinary Medicine*, Vol. 99 Nos 2-4, pp. 102-111.

Freeman, L.C. (1978), “Centrality in social networks conceptual clarification”, *Social Networks*, Vol. 1 No. 3, pp. 215-239.

Han, J., Trienekens, J.H. and Omta, S.O. (2011), “Relationship and quality management in the Chinese pork supply chain”, *International Journal of Production Economics*, Vol. 134 No. 2, pp. 312-321.

Ji, C., de Felipe, I., Briz, J. and Trienekens, J.H. (2012), “An empirical study on governance structure choices in China’s pork supply chain”, *The International Food and Agribusiness Management Review*, Vol. 15, 1030-2016-82767, pp. 121-152.

Kao, R.R., Danon, L., Green, D.M. and Kiss, I.Z. (2006), “Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain”, *Proceedings of the Royal Society B: Biological Sciences*, Vol. 273 No. 1597, pp. 1999-2007.

Kim, Y., Choi, T.Y., Yan, T. and Dooley, K. (2011), "Structural investigation of supply networks: a social network analysis approach", *Journal of Operations Management*, Vol. 29 No. 3, pp. 194-211.

Kiss, I.Z., Green, D.M. and Kao, R.R. (2006), "The network of sheep movements within Great Britain: network properties and their implications for infectious disease spread", *Journal of The Royal Society Interface*, Vol. 3 No. 10, pp. 669-677.

Kukielka, E.A., Martinez-Lopez, B. and Beltrán-Alcrudo, D. (2017), "Modeling the live-pig trade network in Georgia: implications for disease prevention and control", *PloS One*, Vol. 12 No. 6, p. e0178904.

Lebl, K., Lentz, H.H., Pinior, B. and Selhorst, T. (2016), "Impact of network activity on the spread of infectious diseases through the German pig trade network", *Frontiers in veterinary science*, Vol. 3, p. 48.

LeBlanc, D.I., Villeneuve, S., Beni, L.H., Otten, A., Fazil, A., McKellar, R. and Delaquis, P. (2015), "A national produce supply chain database for food safety risk analysis", *Journal of Food Engineering*, Vol. 147, pp. 24-38.

Lee, K., Polson, D., Lowe, E., Main, R., Holtkamp, D. and Martínez-López, B. (2017), "Unraveling the contact patterns and network structure of pig shipments in the United States and its association with porcine reproductive and respiratory syndrome virus (PRRSV) outbreaks", *Preventive Veterinary Medicine*, Vol. 138, pp. 113-123.

Lentz, H., Kasper, M. and Selhorst, T. (2009), "Network analysis of the German cattle trade net– preliminary results", *Berliner und Münchener Tierärztliche Wochenschrift*, Vol. 122 Nos 5-6, pp. 193-198.

Lentz, H.H., Konschake, M., Teske, K., Kasper, M., Rother, B., Carmanns, R., Petersen, B., Conraths, F.J. and Selhorst, T. (2011), "Trade communities and their spatial patterns in the German pork production network", *Preventive Veterinary Medicine*, Vol. 98 Nos 2-3, pp. 176-181.

Lu, G. and Shang, G. (2017), "Impact of supply base structural complexity on financial performance: roles of visible and not-so-visible characteristics", *Journal of Operations Management*, Vol. 53, pp. 23-44.

Lu, X., Horn, A.L., Su, J. and Jiang, J. (2019), "A universal measure for network traceability", *Omega*, Vol. 87, pp. 191-204.

Martínez-López, B., Perez, A.M. and Sánchez-Vizcaíno, J.M. (2009), "Combined application of social network and cluster detection analyses for temporal-spatial characterization of animal movements in Salamanca, Spain", *Preventive Veterinary Medicine*, Vol. 91 No. 1, pp. 29-38.

Ministry of Commerce of the People's Republic of China (2017), "The natioanl important product traceability system in China", available at: https://zycpzs.mofcom.gov.cn/ (accessed 12 June 2017).

Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G. and Calistri, P. (2009), "Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread", *Preventive Veterinary Medicine*, Vol. 92 No. 4, pp. 341-350.

Newman, M.E. (2002), "Assortative mixing in networks", *Physical Review Letters*, Vol. 89 No. 20, p. 208701.

Nöremark, M., Håkansson, N., Lewerin, S.S., Lindberg, A. and Jonsson, A. (2011), "Network analysis of cattle and pig movements in Sweden: measures relevant for disease control and risk based surveillance", *Preventive Veterinary Medicine*, Vol. 99 Nos 2-4, pp. 78-90.

Osadchiy, N., Gaur, V. and Seshadri, S. (2016), "Systematic risk in supply chain networks", *Management Science*, Vol. 62 No. 6, pp. 1755-1777.

Perera, S., Perera, H.N. and Kasthurirathna, D. (2017), "Structural characteristics of complex supply chain networks", *2017 Moratuwa Engineering Research Conference (MERCon)*, IEEE, pp. 135-140.

Rautureau, S., Dufour, B. and Durand, B. (2012), "Structural vulnerability of the French swine industry trade network to the spread of infectious diseases", *Animal*, Vol. 6 No. 7, pp. 1152-1162.

Relun, A., Grosbois, V., Sánchez-Vizcaíno, J.M., Alexandrov, T., Feliziani, F., Waret-Szkuta, A., Molia, S., Etter, E.M.C. and Martínez-López, B. (2016), "Spatial and functional organization of pig trade in different European production systems: implications for disease prevention and control", *Frontiers in Veterinary Science*, Vol. 3, p. 4.

Robinson, S.E., Everett, M.G. and Christley, R.M. (2007), "Recent network evolution increases the potential for large epidemics in the British cattle population", *Journal of The Royal Society Interface*, Vol. 4 No. 15, pp. 669-674.

Salines, M., Andraud, M. and Rose, N. (2017), "Pig movements in France: designing network models fitting the transmission route of pathogens", *PloS One*, Vol. 12 No. 10, p. e0185858.

Shao, Y., Li, M., Zhang, W., Ji, Y. and Hayes, D.J. (2018), "World's largest pork producer in crisis: China's african swine fever outbreak", *Agricultural Policy Review*, Vol. 2018 No. 3, p. 1.

The Lancet (2012), "Food safety in China: a long way to go", *The Lancet*, Vol. 380, p. 75, available at: https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(12)61157-X.pdf (accessed 14 July 2012).

Vernon, M.C. and Keeling, M.J. (2012), "Impact of regulatory perturbations to disease spread through cattle movements in Great Britain", *Preventive Veterinary Medicine*, Vol. 105 Nos 1-2, pp. 110-117.

Volkova, V.V., Howey, R., Savill, N.J. and Woolhouse, M.E.J. (2010), "Potential for transmission of infections in networks of cattle farms", *Epidemics*, Vol. 2 No. 3, pp. 116-122.

Wang, T., Sun, Y. and Qiu, H.J. (2018), "African swine fever: an unprecedented disaster and challenge to China", *Infectious diseases of poverty*, Vol. 7 No. 1, p. 111.

Wang, H., Xia, T. and Guan, Z. (2019), "Market power and food safety in the China pork industry", *Agribusiness*, Vol. 35 No. 1, pp. 97-113.

Wasserman, S. and Faust, K. (1994), *Social Network Analysis: Methods and Applications*, Vol. 8, Cambridge University Press, Cambridge.

Webb, C.R. (2005), "Farm animal networks: unraveling the contact structure of the British sheep population", *Preventive Veterinary Medicine*, Vol. 68 No. 1, pp. 3-17.

Webb, C.R. (2006), "Investigating the potential spread of infectious diseases of sheep via agricultural shows in Great Britain", *Epidemiology and Infection*, Vol. 134 No. 1, pp. 31-40.

Wu, L., Wang, H., Zhu, D., Hu, W. and Wang, S. (2016), "Chinese consumers' willingness to pay for pork traceability information—the case of Wuxi", *Agricultural Economics*, Vol. 47 No. 1, pp. 71-79.

Yu, J., Chen, R., Han, J., Lu, H., Trienekens, J.H. and Omta, S.O. (2013), "The impact of supply chain integration on firm performance in the pork processing industry in China", *Chinese Management Studies*, Vol. 7 No. 2, pp. 230-252.

Zhou, X., Li, Y., Wang, Y., Edwards, J., Guo, F., Clements, A.C., Huang, B. and Magalhaes, R.J.S. (2015), "The role of live poultry movement and live bird market biosecurity in the epidemiology of influenza A (H7N9): a cross-sectional observational study in four eastern China provinces", *Journal of Infection*, Vol. 71 No. 4, pp. 470-479.

Corresponding author
Xin Lu can be contacted at: xin.lu@flowminder.org