

Received September 23, 2019, accepted October 21, 2019, date of publication October 25, 2019, date of current version November 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2949606

# Detection of Infection Sources for Avian Influenza A(H7N9) in Live Poultry Transport Network During the Fifth Wave in China

XIN PEI<sup>1,2,3</sup>, ZHEN JIN<sup>1,2,3</sup>, WENYI ZHANG<sup>4</sup>, AND YONG WANG<sup>4</sup>

<sup>1</sup>Data Science and Technology, North University of China, Taiyuan 030012, China

<sup>2</sup>Complex System Research Center, Shanxi University, Taiyuan 030006, China

<sup>3</sup>Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University, Taiyuan 030006, China

<sup>4</sup>Chinese PLA Center for Disease Control and Prevention, Beijing 100071, China

Corresponding author: Zhen Jin (jinzhn@263.net)

This work was supported in part by the National Natural Science Foundation of China under Grant 11331009, Grant 61873154, Grant 11701528, and Grant 11501340, in part by the Graduate Students' Excellent Innovative Item of Shanxi Province under Grant 2017BY105, in part by the Shanxi Key Laboratory under Grant 201705D111006, in part by the National Key Research and Development Program of China under Grant 2016YFD0501500, and in part by the Shanxi Scientific and Technology Innovation Team under Grant 201705D15111172.

**ABSTRACT** Numerous studies have demonstrated that exposure to live poultry or live poultry markets is the significant risk factor for human infection with avian influenza A(H7N9). However, the specific live poultry markets that are major infection sources for A(H7N9) human cases have not been explored in detail. In this study, we extract data associated with poultry farms, live poultry markets and farmers' markets from Baidu Map using the JavaScript language and then construct the live poultry transport network. From this, we establish our A(H7N9) transmission model over the network based upon probabilistic discrete-time Markov chain. On the basis of the obtained network and model, we propose spatiotemporal backward detection and forward transmission algorithms to detect the most likely infection sources and to compute the first arrival times of the infection sources. Our simulations use these algorithms to identify the specific locations of the infection sources, the first arrival times of the infection sources and the most likely transmission map of the A(H7N9) virus along the live poultry transport network. The results reveal that, in addition to the hazards posed by the live poultry markets, backyard poultry also contributed to A(H7N9) human infections; this risk source was significant especially in the newly affected provinces, in the fifth wave of infection. In particular, by analyzing the temperature characteristics at a given location when the infection source arrived, we find that the risk of human infection with the influenza A(H7N9) virus was high under 9°C~19°C; moderate under 0°C~9°C or 19°C~25°C; and low for temperatures < 0°C or > 25°C. Our results suggest that strengthening the supervision of the live poultry market system and immunizing poultry at both live poultry markets and the backyard poultry operations under the high risk temperature band of 9°C~19°C, will be able to significantly contribute to the control of avian influenza A(H7N9) in the future.

**INDEX TERMS** Live poultry transport network, avian influenza A(H7N9), transmission model, detecting infection sources.

## I. INTRODUCTION

The novel avian influenza A(H7N9) virus emerged in 2013. It is a bird flu strain of the influenza virus A (avian influenza virus or bird flu virus) [1]. The avian influenza A(H7N9) virus is only transmitted between poultry or from poultry to human. Human can be infected through direct exposure to poultry, poultry secretions or excreta, inhalation of

The associate editor coordinating the review of this manuscript and approving it for publication was Usama Mir.

viral aerosols, and exposure to environments contaminated with the virus [2]–[4]. Human cases of A(H7N9) infection have occurred since 2013, during the annual winter-spring epidemics in mainland China [5], [6]. After peaking in 2013–2014, the human infection cohort in subsequent epidemics was generally smaller [7], but it sharply increased in the fifth epidemic wave in December 2016 [8]. This fifth epidemic wave (lasting from October 1st, 2016, to September 31st, 2017) was the most significant up until that point; 746 human cases were reported across 27 provinces

in mainland China. The A(H7N9) virus strains circulating among poultry had been classified as low pathogenicity avian influenza (LPAI) in the previous four epidemic waves in China [9], but evolved to be highly pathogenic in poultry in the fifth epidemic wave [10]. The earlier start date, larger epidemic size, wider epidemic range and higher pathogenicity in the fifth epidemic wave A(H7N9) prompted panic and aroused public concern.

Most of the confirmed A(H7N9) cases in human had an etiology involving a history of recent exposure to live poultry or potentially contaminated environments, especially the live poultry markets, where live poultry is sold [11], [12]. Wang *et al.* [13] demonstrated that contaminated of poultry-related environments, including live poultry markets and backyard poultry, are the two major sources for exposure to diseased poultry; they divided human cases into rural, urban and semiurban cases and discovered that in each type of human case, all had visited a live poultry market and/or backyard poultry, in differing proportions for each type of human cases. The closure of live poultry markets was highly effective in reducing the risk of human infection with A(H7N9) by reducing human exposure to poultry [14]. Furthermore, other studies also confirmed the significant role of the live poultry trade on the occurrence of human infection with A(H7N9) [15], [16]. However, the specific live poultry markets were the main infection sources of A(H7N9) cases in human and the characteristics of these sources have not yet been explored in detail.

Detecting the source of an infection for disease transmission depends on two key determinants: the network structure and the transmission model. A number of existing algorithms and mathematical theories can locate the source of diffusion in complex networks. For instance, Shah *et al.* [17]–[19] first presented a systematic study to find the virus source based upon rumor centrality by using maximum likelihood (ML) estimation in regular trees, general trees, and general graphs. The rumor centrality of a node is the number of permitted permutations of nodes that begin with that node and result in a virus graph. The rumor centrality measure is performed by selecting any node as the source node and calculating its rumor centrality for all nodes in the virus graph, referring to the node that maximizes the rumor center of the graph. Later, Luo *et al.* [20], [21] expanded this method to identify multiple infection sources in networks. In addition, the Bayesian inference algorithm based on dynamic message-passing equations via belief propagation (BP) was developed by Altarelli *et al.* [22] and Loco *et al.* [23]. The dynamical message passing (DMP) approach and discrete-time Markov chain approach have been widely used to study the spreading dynamics on networks [24], [25]. Furthermore, Antulov-Fantulin *et al.* [26] proposed using Monte-Carlo methods with soft-margin algorithms to detect the sources of epidemics, using snapshots of spreading patterns in static and temporal networks. The use of the NetSleuth algorithm, employing the minimum description length (MDL) principle, showed high accuracy in

the detection of seed nodes and the correct automatic identification of their numbers [27], [28]. Moreover, Shen *et al.* [29] developed a time-reversal backward-spreading algorithm to locate sources of infection, and tested this algorithm by employing epidemic spreading and consensus dynamics as typically dynamic processes and by applying it to the H1N1 pandemic in China. Thus, the detection of infection sources in networks is advancing in the field of infectious diseases.

In this paper, we extract data associated with poultry farms, live poultry markets and farmers' markets from Baidu Map by using the JavaScript language. We construct a live poultry transport network on the basis of the data obtained. Following this, we establish an A(H7N9) transmission model, based on probabilistic discrete-time Markov chain to describe the state transition for each of the type nodes in the network. According to the network constructed and model established, we propose spatiotemporal backward detection and forward transmission algorithms to detect the infection source of the epidemic. The simulation results from the algorithms are shown, which include the detected infection sources, first arrival time and maximum likelihood  $L(t, u)$  of infection sources, most likely spread map, and temperature characteristics in a given location at the arrival time of the A(H7N9) virus. In the final section of the paper, we present our major conclusions.

## II. DATA AND NETWORK CONSTRUCTION

Live poultry and their surrounding environments are significant risk factors for human infection with A(H7N9). According to the present situation in China, the susceptible sites associated with live poultry are classified into poultry farms, backyard poultry, live poultry markets and farmers' markets [30]–[32]. Poultry farms are engaged in hatching and growing poultry to produce meat and eggs, and are usually characterized by their large-scale and industrialization. When they have reached a sufficient size, live poultry from poultry farms are transported to live poultry markets for wholesale. Backyard poultry refers to the poultry raised in the backyards of farmers in rural areas, which are usually characterized by low levels of technology and the scarcity of biosecurity practices. After feeding is finished, backyard poultry are transported to the nearby live poultry markets or farmers' markets for sale. Live poultry markets are the secondary wholesale markets for live poultry, where the live poultry from poultry farms and backyard poultry are wholesaled to the farmers' markets. Farmers' markets are the fixed places, mainly used for retailing. In addition to retailing live poultry, farmers' markets also sell vegetables, fruits, aquatic products, and other agricultural products and foods. Farmers' markets are generally located in the community, to facilitate residents' purchases of agricultural products. Operators of these markets buy live poultry from live poultry markets that are an approximately one hour round-trip drive away, usually before the opening of the farmers' market each morning. It is worth

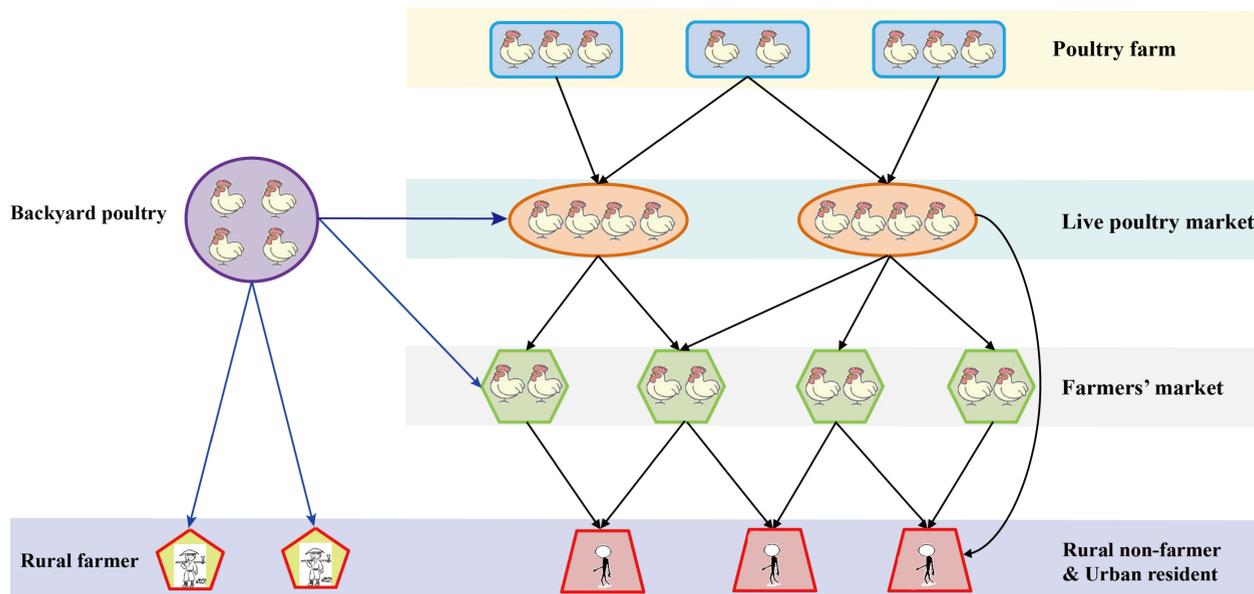


FIGURE 1. Transmission diagram of the A(H7N9) virus along live poultry transport route.

noting that there is no trading relationship between the same types of sites, because they are independent and competitive.

In the present study, the human infected with A(H7N9) virus are divided into three types [13]: urban residents, rural non-farmers and rural farmers. Urban residents and rural non-farmers may be infected by visiting a farmers' market or live poultry market, because they usually purchase live poultry or other agricultural products from nearby farmers' markets; the minority of people in this group go to nearby live poultry markets to purchase live poultry. Rural farmers may be infected through exposure to backyard poultry [33], [34].

The diagram presented in Fig 1 describes the transmission diagram of the A(H7N9) virus along the live poultry transport route. In the following sections, we collect data, construct network and establish transmission model according to this mode of live poultry transport and virus transmission.

**A. DATA COLLECTION**

1) DATA FOR A(H7N9) HUMAN CASES

The data on weekly reported human cases of avian influenza A(H7N9) used in the present study are obtained from the Center for Health Protection (<http://www.chp.gov.hk>) and cover the dates of October 1st, 2016 to May 17th, 2017. Individual case information includes geographic location, occupation, date of disease onset, date of reporting and the clinical condition of the environment at the time of disease reporting. Confirmed cases are all as defined by the World Health Organization criteria and the national authorities.

2) DATA FOR LIVE POULTRY TRANSPORT

Baidu Map ([map.baidu.com](http://map.baidu.com)) is the most popular online map service in China. It is similar to Google Map, comprises satellite images, street maps, street view and indoor

view perspectives, and functions as a route planner for travel by foot, car, or public transportation. We extract the data associated with poultry farms, live poultry markets and farmers' markets from each city in the Baidu Map database using JavaScript language. The search terms include "city name" and "poultry farm, chicken farm, and duckery"; "city name" and "live poultry or poultry trade, wholesale"; and "city name" and "farmers' market, agricultural market, vegetable market". The extracted data contain information about the names and locations of all sites. As a key additional data set, we obtain the opening and closing times of the live poultry markets in each province or city from government documents.

**B. DATA CLEANING AND PRE-PROCESSING**

1) DATA CLEANING

To ensure the accuracy of the data associated with live poultry, it is necessary to manually perform data cleaning to discard some extraneous data. For example, when we extract poultry farm information in a city using the search terms "city name" and "poultry farm, chicken farm, and duckery", livestock farms also appear in the search results, although they are not relevant to live poultry. We thus need to exclude data that are not relevant to live poultry from the obtained data set, such as livestock farms.

2) DATA PREPROCESSING

Using Eq.(1), we compute the distances  $\widehat{AB}$  between the following four types of sites based on their longitudes and latitudes:

- each A(H7N9) human case and each farmers' market;
- each A(H7N9) human case and each live poultry market;
- each farmers' market and each live poultry market;

- each live poultry market and each poultry farm.

$$\begin{aligned} \Gamma &= \sin(MLat.A) * \sin(MLat.B) \\ &\quad * \cos(MLon.A - MLon.B) \\ &\quad + \cos(MLat.A) * \cos(MLat.B), \\ \widehat{AB} &= R * \arccos(\Gamma) * \pi / 180. \end{aligned} \quad (1)$$

In Eq. (1),  $MLon.A$  ( $MLon.B$ ) is the longitude of site  $A$  ( $B$ ) and  $MLat.A$  ( $MLat.B$ ) is the latitude of site  $A$  ( $B$ ) in Baidu Map;  $\widehat{AB}$  is the spherical distance between sites  $A$  and  $B$ ;  $R$  is the equatorial radius.

### C. NETWORK CONSTRUCTION

Using the cleaned and preprocessed data, we construct the live poultry transport network as a directed and weighted graph  $G = (V, E)$ , where  $V$  denotes a set of nodes and  $E$  is the set of edges in the network. The weight of each edge is defined as the actual distance between the two linked nodes. The detailed network construction is described as follows:

**Nodes.** Nodes can be divided into four types: A(H7N9) infected human, farmers’ markets, live poultry markets (LPM) and poultry farms. The infected human nodes include three subtypes: urban residents, rural non-farmers and rural farmers. Backyard poultry (BP) are considered the “sun node” in every city, and are assumed to be only transported to live poultry markets, farmers’ markets, or to contact with rural farmers, according to the present situation in China.

**Edges.** Four directed edges are established, according to the distance between the different types of nodes as follows:

- Within 3 kilometers from farmers’ markets to infected urban residents or rural non-farmers;  
*Reason for the distance determination:* The Chinese population usually commutes 1 to 3 kilometers in their communities every day [35].
- Within 3 kilometers from live poultry market to infected urban residents or rural non-farmers;  
*Reason for the distance determination:* The Chinese population usually commutes 1 to 3 kilometers in their communities every day [35].
- Within 20 kilometers from live poultry markets to farmers’ markets;  
*Reason for the distance determination:* a) Investigations indicate that most respondents transport live poultry from live poultry markets to farmers’ markets over distances of approximately 10 kilometers [36]. b) According to the present situation in China, farmers’ market operators usually travel to purchase live poultry from live poultry markets that are an approximately one hour round-trip drive away (10-20 kilometers), prior to the opening of the farmers’ markets each morning. Thus, we determine 20 kilometers to be the maximum distance from live poultry markets to farmers’ markets.
- Within 100 kilometers from poultry farms to live poultry markets.  
*Reason for the distance determination:* a) In China, poultry farms are most often located in the suburbs or

TABLE 1. Descriptions of the variables used in the model.

Variable	Interpretation
$\partial j_{in}, \partial k_{in}, \partial m_{in}$	the set of neighbors directed to node $j, k, m$ in directed network
$\omega_{ij}, \omega_{jk}, \omega_{km}, \omega_{jm}$	the actual distance of the edge connecting node $i$ to $j, j$ to $k, k$ to $m, j$ to $m$
$\eta_{ij}(t)$	the spreading probability from poultry farm node $i$ to live poultry market node $j$ at time $t$
$\eta_{jk}(t)$	the spreading probability from live poultry market node $j$ to farmers’ market node $k$ at time $t$
$\eta_{km}(t)$	the spreading probability from farmers’ market node $k$ to rural non-farmer or urban resident node $m$ at time $t$
$\eta_{jm}(t)$	the spreading probability from live poultry market node $j$ to rural non-farmer or urban resident node $m$ at time $t$
$\lambda_{pl}(j, t)$	the probability of a susceptible live poultry market node $j$ being infected by its infected poultry farm neighbors at time $t$
$\lambda_{lf}(k, t)$	the probability of a susceptible farmers’ market node $k$ being infected by its infected live poultry market neighbors at time $t$
$\lambda_{fu}(m, t)$	the probability of a susceptible rural non-farmer or urban resident node $m$ being infected by its infected farmers’ market neighbors at time $t$
$\lambda_{lu}(m, t)$	the probability of a susceptible rural non-farmer or urban resident node $m$ being infected by its infected live poultry market neighbors at time $t$
$P_Y^b(*, t)$	the probability that the backyard poultry $*$ is in state $Y$ at time $t$ in the city, $Y \in \{S, C, I\}$
$P_Y^p(i, t)$	the probability that an arbitrary poultry farm node $i$ is in state $Y$ at time $t$ , $Y \in \{S, C, I\}$
$P_Y^l(j, t)$	the probability that an arbitrary live poultry market node $j$ is in state $Y$ at time $t$ , $Y \in \{S, C, I\}$
$P_Y^f(k, t)$	the probability that an arbitrary farmers’ market node $k$ is in state $Y$ at time $t$ , $Y \in \{S, C, I\}$
$P_Y^u(m, t)$	the probability that an arbitrary rural non-farmer or urban resident node $m$ is in state $Y$ at time $t$ , $Y \in \{S, C, I, R\}$
$P_Y^r(m', t)$	the probability that an arbitrary rural farmer node $m'$ is in state $Y$ at time $t$ , $Y \in \{S, C, I, R\}$

rural areas far from the live poultry markets. Investigation indicates that live poultry can be transported for sale much longer distances (in excess of 100 kilometers) [36]. b) We determine 100 kilometers as the maximum distance, by comparing the average degree and degree distributions, in the network within the different distances from poultry farms to live poultry markets in APPENDIX A.

Furthermore, the spatial transmission probability is proportional to the actual distance from the infectious node to the susceptible node, as shown in Eq.(2) in the next section. In summary, our creation of directed edges according to the distance between different types of nodes and the choice of distances is, thus scientific and reasonable.

**Time-varying.** Our network is temporally dynamic. Nodes and edges are deleted or added according to the closing and reopening of live poultry markets over time.

### III. TRANSMISSION MODEL

The transmission dynamics are considered for the A(H7N9) virus as a multi-group model. Fig 2 shows the state transition graph for the four types of nodes in this model. The main variables and parameters used are described in Table 1 and Table 2.

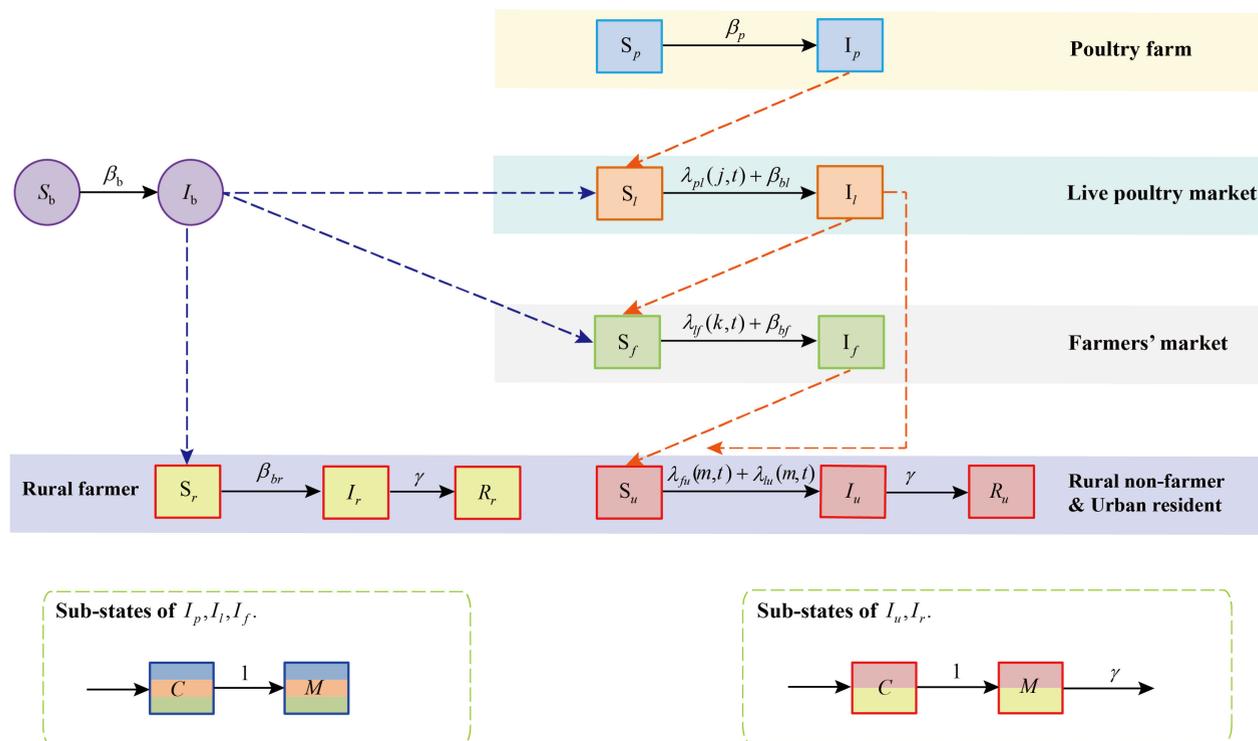


FIGURE 2. Node state transitions in the A(H7N9) transmission model.

TABLE 2. Descriptions and values of the parameters used in the model.

Parameter	Interpretation	Value (Source)	Range
$\gamma$	the recovery rate of infected human	0.053(CDC)	None
$\beta_{pl}$	the infection rate of infected poultry farm to susceptible live poultry market	0.23 (Assumption)	None
$\beta_{lf}$	the infection rate of infected live poultry market to susceptible farmer's market	0.36(LHS)	95%(0.3599,0.3602)
$\beta_{fu}$	the infection rate of infected farmer's market to susceptible rural non-farmer or urban resident	0.18(LHS)	95%(0.1799,0.1803)
$\beta_{lu}$	the infection rate of infected live poultry market to susceptible rural non-farmer or urban resident	0.46(LHS)	95%(0.4598,0.4602)
$\beta_{bl}$	the infection rate of infected backyard poultry to susceptible live poultry market	0.32(LHS)	95%(0.3198,0.3202)
$\beta_{bf}$	the infection rate of infected backyard poultry to susceptible farmer's market	0.32(LHS)	95%(0.3198,0.3202)
$\beta_{br}$	the infection rate of infected backyard poultry to susceptible rural farmer	0.26(LHS)	95%(0.2598,0.2602)

In our model, the nodes representing poultry farms, backyard poultry, live poultry markets and farmers' markets,

are classified as susceptible or infected. As such,  $S_p$  ( $S_b$ ) and  $I_p$  ( $I_b$ ) represent susceptible and infected poultry farms (backyard poultry), respectively.  $S_l$  ( $S_f$ ) and  $I_l$  ( $I_f$ ) describe susceptible and infected live poultry markets (farmers' markets), respectively. Contagion transfer was considered to occur only between different types of nodes, through the transport of live poultry. Human nodes were classified into three states as susceptible, infected or recovered. The avian influenza virus is only transmitted to human from backyard poultry, live poultry markets and farmers' markets; it cannot spread among human.  $S_r$ ,  $I_r$  and  $R_r$  represent susceptible, infected and recovered rural farmers, respectively.  $S_u$ ,  $I_u$  and  $R_u$  describe susceptible, infected and recovered rural non-farmers or urban residents, respectively.

The detailed state transition is described as follows:

- 1) The susceptible poultry farm ( $S_p$ ) can be infected with probability  $\beta_p$ , and then become an infected poultry farm ( $I_p$ );
- 2) The susceptible backyard poultry ( $S_b$ ) can be infected with probability  $\beta_b$ , and then become an infected backyard poultry ( $I_b$ );
- 3) The susceptible live poultry market ( $S_l$ ) can be infected by infected poultry farm neighbors ( $I_p$ ) or backyard poultry ( $I_b$ ) with probability  $\lambda_{pl}(j, t) + \beta_{bl}$ , and then become an infected live poultry market ( $I_l$ );
- 4) The susceptible farmer's market ( $S_f$ ) can be infected by infected live poultry market neighbors ( $I_l$ ) or backyard poultry ( $I_b$ ) with probability  $\lambda_{lf}(k, t) + \beta_{bf}$ , and then become an infected farmer's market ( $I_f$ );

- 5) The susceptible rural farmer ( $S_r$ ) can be infected by infected backyard poultry ( $I_b$ ) with probability  $\beta_{br}$ , and then recover with probability  $\gamma$ ;
- 6) The susceptible rural non-farmer or urban resident ( $S_u$ ) can be infected by infected farmers' market neighbors ( $I_f$ ) or live poultry market neighbors ( $I_l$ ) with probability  $\lambda_{fu}(m, t) + \lambda_{lu}(m, t)$ , and then recover with probability  $\gamma$ .

To describe states more precisely in different types of nodes at time  $t$ , we introduce two sub-states for infected nodes ( $I$ ): "contagious" ( $C$ ) and "maintained contagious" ( $M$ ) in Fig 2. Once an  $S$  node is infected, it first becomes  $C$  and then transits to  $M$ . The  $C$  state describes nodes that are newly infected. More specifically, a node that is  $C$  at time  $t$  means that this node is susceptible at time  $t - 1$  but becomes infected at time  $t$ . The  $M$  state describes nodes that are infected but not newly infected. More explicitly, once an  $S$  node is infected, it first becomes  $C$  at time  $t$  and then transits to  $M$  at time  $t + 1$ . The  $M$  node will stay infected until it recovers or is removed from the network [37].

According to the transmission dynamics of the A(H7N9) virus described above, we establish a transmission model based on a Markov chain to calculate the probabilities of each node in the various states. For simplicity, the state transition of a live poultry market is described here as an example, which can then be applied to other types of nodes. A susceptible live poultry market can be infected by infected poultry farm neighbors or backyard poultry in the city, it then becomes an infected live poultry market.  $\beta_{bl}$  is defined as the infection rate of infected backyard poultry nodes to a susceptible live poultry market node, and  $\eta_{ij}(t)$  is the spreading probability from a poultry farm node  $i$  to live poultry market node  $j$  at time  $t$ .  $\eta_{ij}(t)$  is a function of the actual distance of the link  $\omega_{ij}$  as follows:

$$\eta_{ij}(t) = \frac{\omega_{ij}^{-1}}{\sum_{i \in \partial j_{in}} \omega_{ij}^{-1}} \beta_{pl}. \quad (2)$$

Then, we can use Eq.(2) to calculate the probability  $\lambda_{pl}(j, t)$  of a susceptible live poultry market  $j$  being infected by its infected poultry farm neighbors at time  $t$ . Because the probability that node  $j$  is not infected means that none of its infected poultry farm neighbors spread the virus to it through their links, we can calculate the probability as follows:

$$\lambda_{pl}(j, t) = 1 - \prod_{i \in \partial j_{in}} [1 - \eta_{ij}(t) P_I^p(i, t - 1)]. \quad (3)$$

where, the product runs over all poultry farm neighbors of live poultry market node  $j$ .

After this, we can then calculate the probability of an arbitrary live poultry market node  $j$  being susceptible at time  $t$  as follows:

$$P_S^l(j, t) = [1 - \lambda_{pl}(j, t) - \beta_{bl}] P_S^l(j, t - 1). \quad (4)$$

A susceptible live poultry market node first becomes contagious once it has been infected. We can calculate the

probability that an arbitrary live poultry market node  $j$  is contagious at time  $t$  as follows:

$$P_C^l(j, t) = [\lambda_{pl}(j, t) + \beta_{bl}] P_S^l(j, t - 1). \quad (5)$$

The infected state of a live poultry market node means that it is in either  $C$  or  $M$  state. In other words, the infected node is either a new infection at time  $t$  or the infected one before time  $t$ . Thus, the probability of an arbitrary live poultry market node  $j$  being infected at time  $t$ , is represented by:

$$P_I^l(j, t) = P_C^l(j, t) + P_M^l(j, t - 1). \quad (6)$$

All other terms in the model can be deduced in a similar manner, and the full model equations are given by the following.

$$\left\{ \begin{array}{l} P_S^b(*, t) = (1 - \beta_b) P_S^b(*, t - 1), \\ P_C^b(*, t) = \beta_b P_S^b(*, t - 1), \\ P_I^b(*, t) = P_C^b(*, t) + P_M^b(*, t - 1), \\ P_S^p(i, t) = (1 - \beta_p) P_S^p(i, t - 1), \\ P_C^p(i, t) = \beta_p P_S^p(i, t - 1), \\ P_I^p(i, t) = P_C^p(i, t) + P_M^p(i, t - 1), \\ P_S^l(j, t) = [1 - \lambda_{pl}(j, t) - \beta_{bl}] P_S^l(j, t - 1), \\ P_C^l(j, t) = [\lambda_{pl}(j, t) + \beta_{bl}] P_S^l(j, t - 1), \\ P_I^l(j, t) = P_C^l(j, t) + P_M^l(j, t - 1), \\ P_S^f(k, t) = [1 - \lambda_{lf}(k, t) - \beta_{bf}] P_S^f(k, t - 1), \\ P_C^f(k, t) = [\lambda_{lf}(k, t) + \beta_{bf}] P_S^f(k, t - 1), \\ P_I^f(k, t) = P_C^f(k, t) + P_M^f(k, t - 1), \\ P_S^u(m, t) = [1 - \lambda_{fu}(m, t) - \lambda_{lu}(m, t)] P_S^u(m, t - 1), \\ P_C^u(m, t) = [\lambda_{fu}(m, t) + \lambda_{lu}(m, t)] P_S^u(m, t - 1), \\ P_I^u(m, t) = P_C^u(m, t) + (1 - \gamma) P_I^u(m, t - 1), \\ P_R^u(m, t) = P_C^u(m, t - 1) + \gamma P_I^u(m, t - 1), \\ P_S^r(m', t) = [1 - \beta_{br}] P_S^r(m', t - 1), \\ P_C^r(m', t) = \beta_{br} P_S^r(m', t - 1), \\ P_I^r(m', t) = P_C^r(m', t) + (1 - \gamma) P_I^r(m', t - 1), \\ P_R^r(m', t) = P_C^r(m', t - 1) + \gamma P_I^r(m', t - 1). \end{array} \right. \quad (7a) \text{ to } (7f)$$

where the other terms  $\lambda_{lf}(k, t)$ ,  $\lambda_{fu}(m, t)$ , and  $\lambda_{lu}(m, t)$  are deduced in a manner similar to the above:

$$\begin{aligned} \eta_{jk}(t) &= \frac{\omega_{jk}^{-1}}{\sum_{j \in \partial k_{in}} \omega_{jk}^{-1}} \beta_{lf}, \\ \lambda_{lf}(k, t) &= 1 - \prod_{j \in \partial k_{in}} [1 - \eta_{jk}(t) P_I^l(j, t - 1)], \\ \eta_{km}(t) &= \frac{\omega_{km}^{-1}}{\sum_{k \in \partial m_{in}} \omega_{km}^{-1}} \beta_{fu}, \\ \lambda_{fu}(m, t) &= 1 - \prod_{k \in \partial m_{in}} [1 - \eta_{km}(t) P_I^f(k, t - 1)], \end{aligned}$$

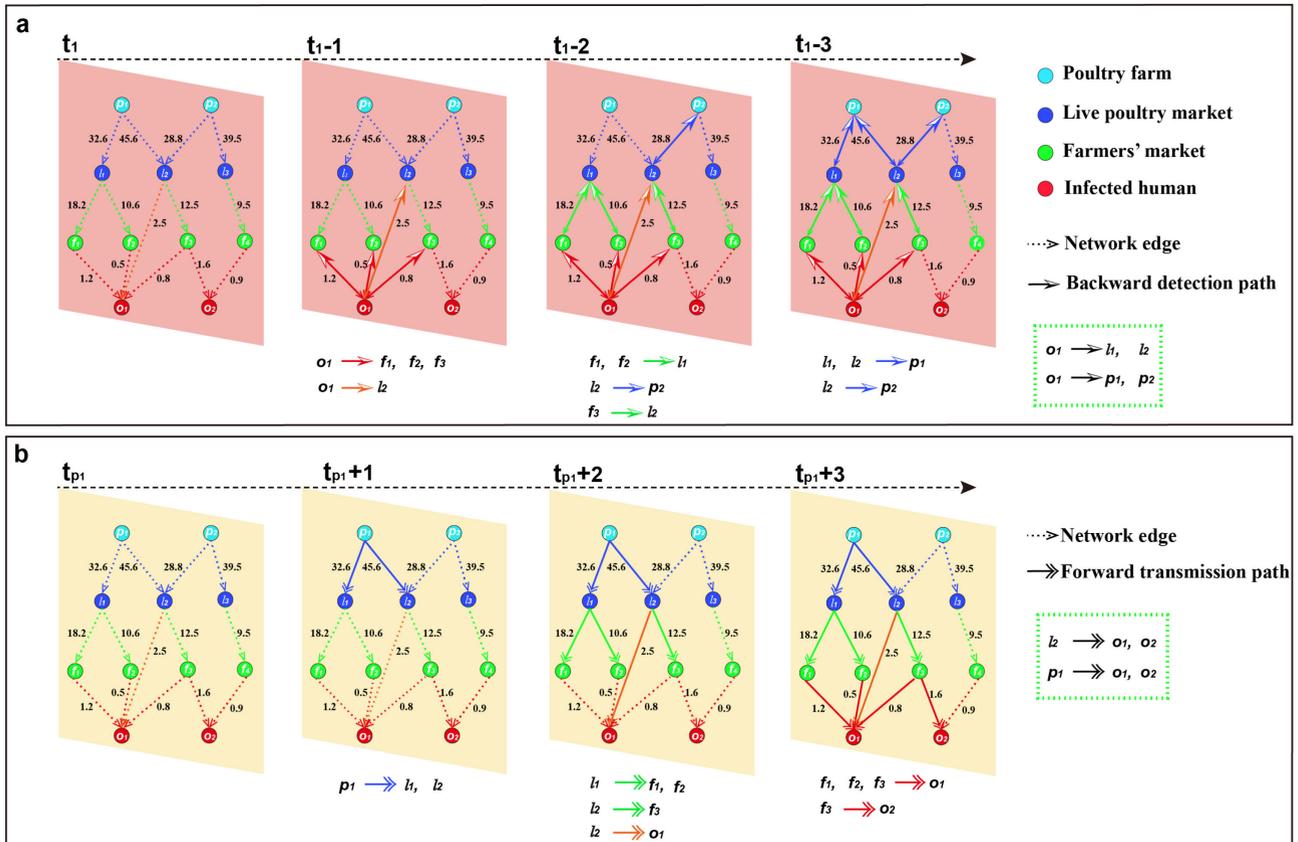


FIGURE 3. Flow chart of spatiotemporal backward detection and forward transmission algorithms in the directed and weighted network.

$$\eta_{jm}(t) = \frac{\omega_{jm}^{-1}}{\sum_{j \in \partial m_{in}} \omega_{jm}^{-1}} \beta_{lu},$$

$$\lambda_{lu}(m, t) = 1 - \prod_{j \in \partial m_{in}} [1 - \eta_{jm}(t) P_I^l(j, t-1)]. \quad (8)$$

IV. ALGORITHM

Based on the network constructed and on the model established, we present spatiotemporal backward detection and forward transmission algorithms to detect the most likely infection sources, and compute the first arrival time and maximum likelihood  $L(t, u)$  of the infection sources.

The rationale of the algorithms is the spread of the A(H7N9) virus along the reversed dynamic connections from each infected person node to exhaust all possible paths for spreading the virus, and then compute the probability of each detected node is an infection source using the Eq.(7). A simple example is provided next to describe the spatiotemporal backward detection and forward transmission algorithms presented in Fig 3. For the infected person node  $o_1$  at infection time  $t_1$ , we first reverse detect to derive the farmers' market nodes  $f_1, f_2$ , and  $f_3$  and the live poultry market node  $l_2$  at time  $t_1 - 1$ , and then obtain the live poultry market nodes  $l_1, l_2$  and the poultry farm node  $p_2$  at time  $t_1 - 2$ . Finally, we acquire the poultry farm nodes  $p_1, p_2$  at time  $t_1 - 3$

in Fig 3(A). This allow the identification of the most likely live poultry market or poultry farm infection source of  $o_1$  by computing the maximum likelihood value of the detected nodes  $l_1, l_2$  or  $p_1, p_2$ , respectively. In addition, poultry farm node  $p_1$  can spatiotemporal forward transmit the virus to both infected human nodes  $o_1$  and  $o_2$ , which is true for the live poultry market  $l_2$  in Fig 3(B). We can compute the first arrival time and maximum likelihood value of the detected infection sources  $p_1$  and  $l_2$  using a spatiotemporal forward transmission algorithm.

A. SPATIOTEMPORAL BACKWARD DETECTION

The spatiotemporal backward detection algorithm is inspired by the time-reversal backward spreading and reverse dissemination algorithm [29], [37], [38]. However, the spatiotemporal backward detection algorithm proposed here is more advanced than these two algorithms. Because our network dynamically changes over time, the probability that each node connects to its neighbor node varies with the spatial distance, and the number of infection sources is greater than one. A detailed description of the algorithm is shown in Table 3.

The live poultry transport network constructed here records not only the spatial location of the nodes and edges, but also the time at which each infected person becomes contagious. The real infection source is expected to transmit the virus both

TABLE 3. Spatiotemporal backward detection algorithm.

Algorithm 1: Spatiotemporal backward detection.	
1:	<b>Input:</b> Adjacency matrix of the network, a set of infected human nodes $O = \{o_1, o_2, \dots, o_n\}$ , a set of infection times of the infected human nodes $\{t_1, t_2, \dots, t_n\}$ , and a threshold $t_{max}$ .
2:	<b>Initialize:</b> $n$ sets of sources $U_m = \emptyset$ , a set of sources $U = \emptyset$ .
3:	<b>for:</b> ( $t$ starts from $t_1$ to $t_n$ ) <b>do</b>
4:	<b>for:</b> ( $o_m$ : starts from $o_1$ to $o_n$ that infection time is $t$ ) <b>do</b>
5:	<b>if:</b> ( $o_m$ has not been detected) <b>then</b> Start to detect the virus from infected person node $o_m$ separately and independently at time $t$ ; Add the detected nodes into the set $U_m$ , and mark the infected time as $t - t_s$ if the node is first detected.
6:	<b>end</b>
7:	<b>for:</b> ( $u$ : any node in the set $U_m$ ) <b>do</b>
8:	<b>if:</b> (node $u$ received virus from $o_m$ ) Compute the maximum likelihood $L(u, t)$ for infected person node $o_m$ ; Keep the first node with large maximum likelihoods in $U_m$ , and delete all the other nodes $u$ . Add set $U_m$ into the set $U$ .
9:	<b>end</b>
10:	<b>end</b>
11:	<b>end</b>
12:	<b>end</b>
13:	<b>Output:</b> $n$ sets of sources $U_m$ , a set of sources $U = \emptyset$ , and the maximum likelihood $L(u, t)$ that every detected node is infection source of $o_m$ .

spatially and temporally, therefore presenting a better match to the infected person node compared to the other nodes in the set of infection sources  $U$ . The core of the algorithm proposed here is the virus transmission model, as explained in the previous section, which provides an estimate of the probability that an arbitrary node is contagious ( $C$ ) at time  $t$ . For an arbitrary infected person node  $o_m$ , we use  $P_C(u, t_s|o_m)$  to denote the probability that an arbitrary node  $u$  is contagious ( $C$ ) after time  $t_s$ , starting with backward detection from the infected person node  $o_m$ , where  $t_s$  is the time span of the whole backward detection process. With the use of Bayes' rule, the probability that node  $u$  transmits the virus to  $o_m$  is proportional to the joint probability of node  $o_m$  reverse transmitting the virus to  $u$  at time  $t_m$ ,  $P(o_m|u) \sim P(u|o_m)$ .

$$P(u|o_m) = \prod_{o_m \in O} P_C(u, t_s|o_m). \tag{9}$$

Mathematically, the node with the maximum likelihood,  $L(u, t)$ , of being the most likely infection source is defined as follows:

$$L(u, t) = \ln \prod_{o_m \in O} P_C(u, t_s|o_m). \tag{10}$$

For computation convenience, we adopt a logarithmic function  $\ln(\cdot)$  in (10) to derive the maximum likelihood. Among all the nodes in  $U_m$ , we can estimate the most likely sources of node  $o_m$  infection by selecting the maximum value of  $L(u, t)$  as follows:

$$(u^*, t^*) = \arg \max_{u \in U_m} L(u, t). \tag{11}$$

TABLE 4. Spatiotemporal forward transmission algorithm.

Algorithm 2: Spatiotemporal forward transmission.	
1:	<b>Input:</b> Adjacency matrix of the network, a set of sources $U = \emptyset$ , a set of infected human nodes $O = \{o_1, o_2, \dots, o_n\}$ , a set of infection times of the infected human nodes $\{t_1, t_2, \dots, t_n\}$ , and a threshold $t_{max}$ .
2:	<b>Initialize:</b> $L^*(t, u) = \emptyset$ , $t^* = 0$ .
3:	<b>for:</b> ( $t$ starts from $t_1$ to $t_{max}$ ) <b>do</b>
4:	<b>for:</b> ( $u$ : any node in the set $U$ and the infection time $t_u$ ) <b>do</b>
5:	<b>if:</b> (We can transmit the virus to $o_m \in O$ if $o_m$ has not been transmit) <b>then</b> Compute the maximum likelihood $L(t, u)$ .
6:	<b>end</b>
7:	<b>end</b>
8:	<b>end</b>
9:	<b>Output:</b> A set of $L^*(t, u)$ , that the maximum likelihood and infection time of every infection source.

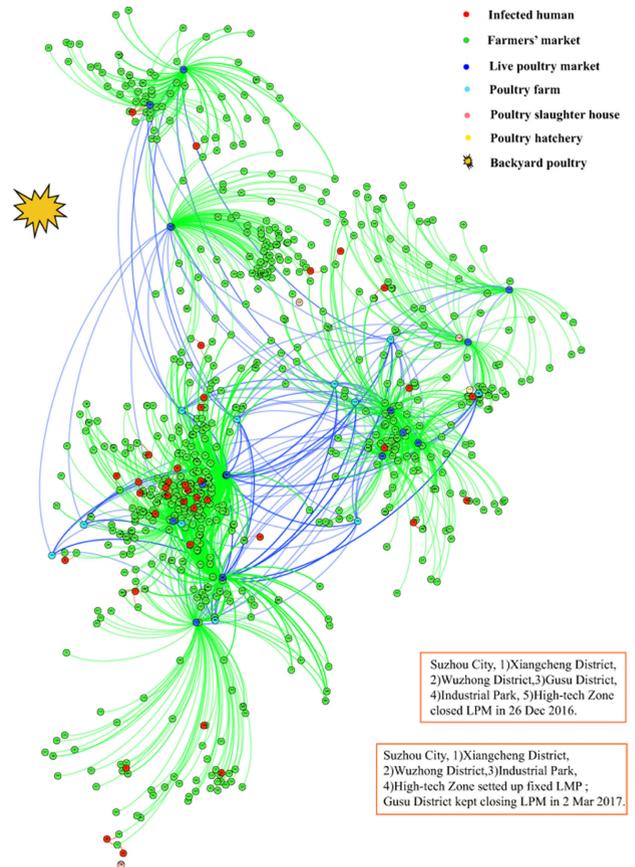
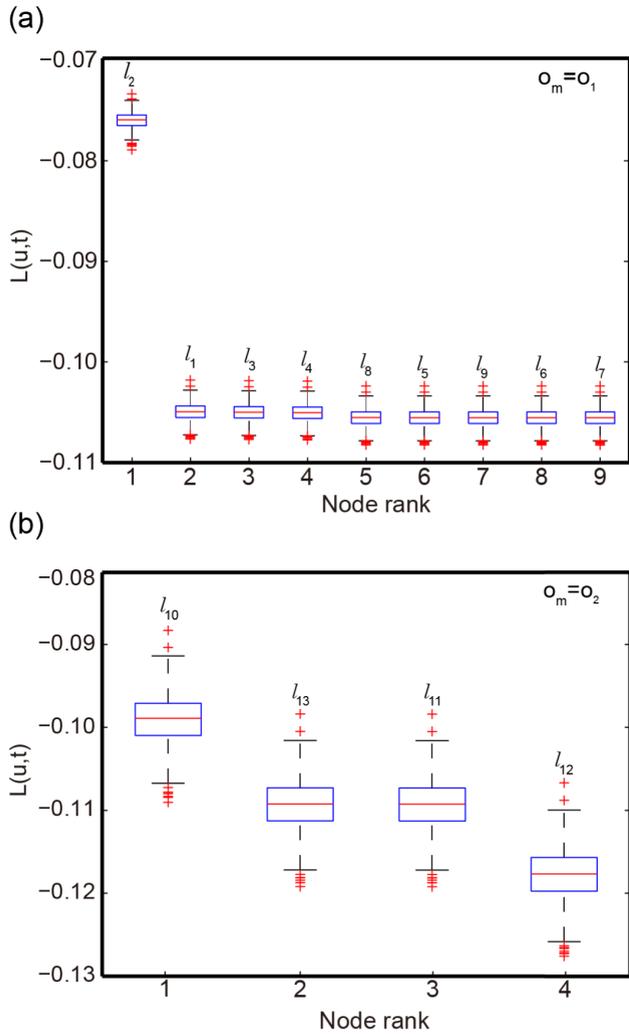


FIGURE 4. Live poultry transport network constructed for Suzhou City.

### B. SPATIOTEMPORAL FORWARD TRANSMISSION

Because a single infection source may infect more than one of its neighbors, it might be the source of multiple infected human nodes. For instance, the poultry farm node  $p_1$  can forward transmit the virus to both infected human nodes  $o_1$  and  $o_2$ , as can the live poultry market  $l_2$  in Fig 3(B). The spatiotemporal forward transmission algorithm is developed to compute the probability of node  $u$  being the infection source and the first arrival time of the virus. The detailed algorithm description is shown in Table 4.

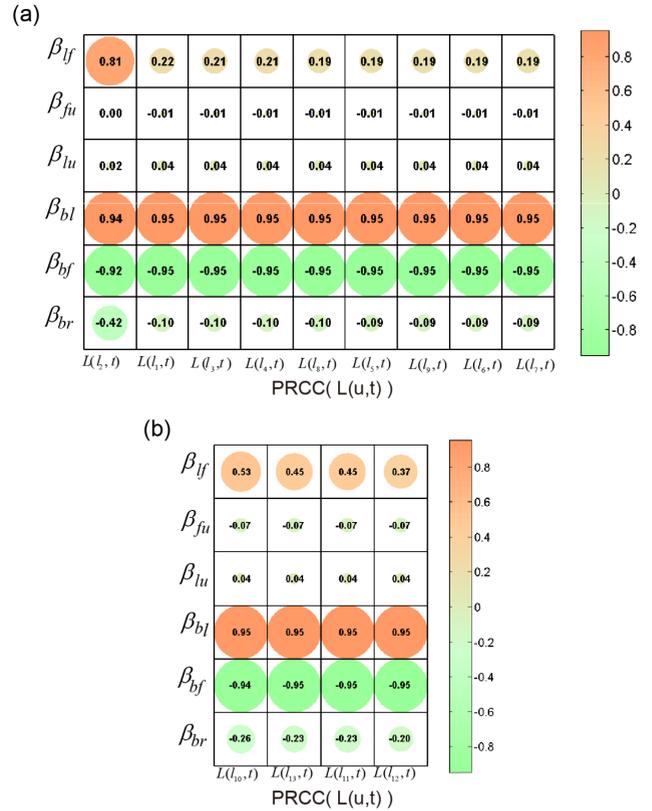


**FIGURE 5.** Box plots showing the sorting of  $L(u, t)$  for all detected susceptible live poultry market infection sources with infected human nodes  $o_1$  and  $o_2$  as examples. (a) The maximum likelihood  $L(u, t)$  of each detected live poultry market infection source of infected person  $o_1$ . The marks on boxes are the labels of live poultry market infection nodes in the network. (b) The maximum likelihood  $L(t, u)$  of each detected live poultry market infection source of infected person  $o_2$ .

$L(t, u)$  is used to denote the maximum likelihood of the virus spreading to infected human nodes at the infection source node  $u$  at time  $t$ , and  $P_C(o_m, t_s|u)$  to calculate probability of an arbitrary infected person node  $o_m$  being contagious ( $C$ ) after time  $t_s$ , starting spatiotemporal forward transmission from source node  $u$ , where  $t_s$  is the time span of the entire forward transmission process. For computing purposes, we adopt the logarithmic function  $\ln(\cdot)$  in (12) to derive the maximum likelihood.

$$L^*(t, u) = \ln \prod_{o_m \in O} P_C(o_m, t_s|u). \quad (12)$$

Furthermore, we also have an estimation of the infection scale  $I(t, u)$  as a byproduct, as in (13). Later, we can justify the effectiveness of the algorithm and maximum likelihood



**FIGURE 6.** Partial rank correlation coefficients (PRCCs) for parameters  $\beta_{if}, \beta_{fu}, \beta_{lu}, \beta_{bl}, \beta_{bf}, \beta_{br}$  with respect to the maximum likelihood  $L(u, t)$ . (a) The PRCCs for parameters  $\beta_{if}, \beta_{fu}, \beta_{lu}, \beta_{bl}, \beta_{bf}, \beta_{br}$  with respect to the maximum likelihood  $L(u, t)$  associated with all the detected suspected live poultry market infection sources ( $l_2, l_1, l_3, l_4, l_8, l_5, l_9, l_6, l_7$ ) of human case  $o_1$ . (b) The PRCCs for parameters  $\beta_{if}, \beta_{fu}, \beta_{lu}, \beta_{bl}, \beta_{bf}, \beta_{br}$  with respect to the maximum likelihood  $L(u, t)$  associated with all the detected suspect live poultry market infection sources ( $l_{10}, l_{13}, l_1, l_{12}$ ) of human case  $o_2$ .

method by examining the accuracy of  $I(t, u)$  as follows:

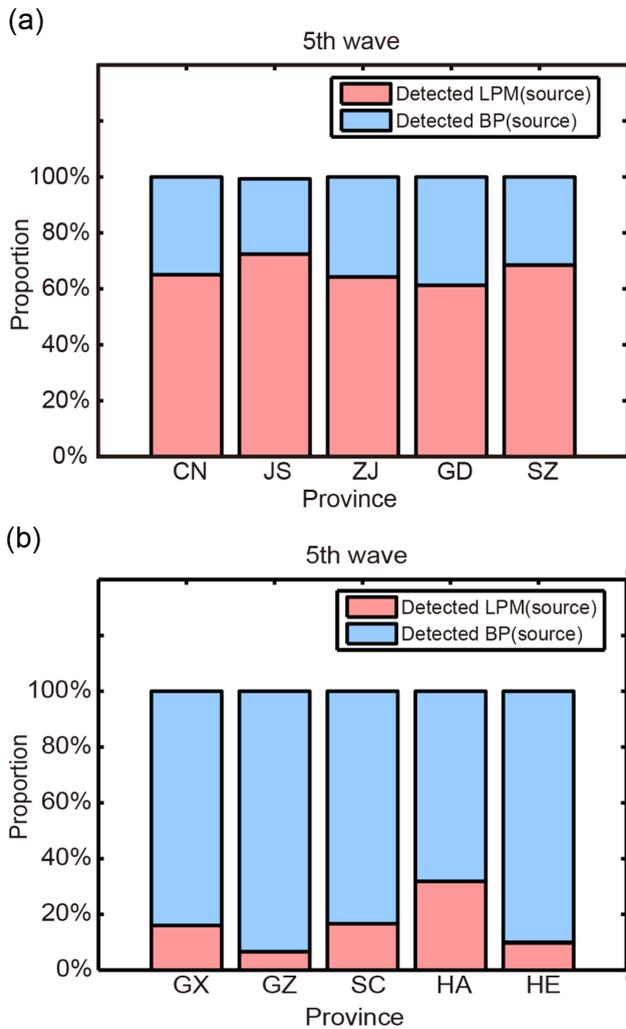
$$I(t, u) = \sum_{u \in U} \sum_{o_m \in O} P_I(o_m, t|u). \quad (13)$$

## V. RESULTS

In this section, we provide the simulation results from the above algorithms based on the constructed live poultry transport network and the established A(H7N9) transmission model. The results include detected infection sources, first arrival time and  $L(t, u)$  of the infection sources, and the most likely spread map of the A(H7N9) virus and the temperature characteristics of the arrival time.

### A. DETECTED INFECTION SOURCES

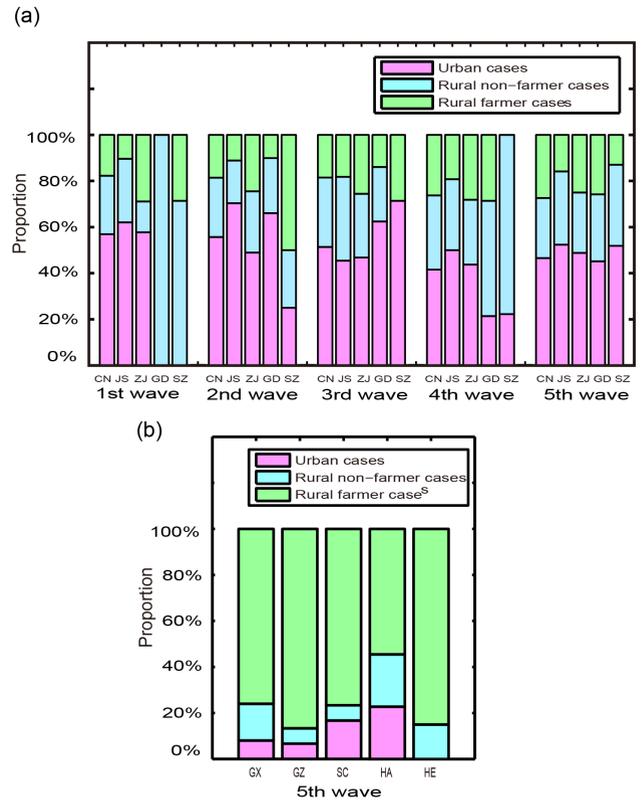
Take Suzhou City, where the A(H7N9) virus first occurred and the largest number of A(H7N9) human cases occurred in the 5th wave, as an example. Fig 4 shows our constructed live poultry transport network, and information about the closure or reopening of live poultry markets in Suzhou City. Fig 5 shows the rank of  $L(u, t)$  for all detected suspect live poultry market infection sources of human cases  $o_1$  and  $o_2$



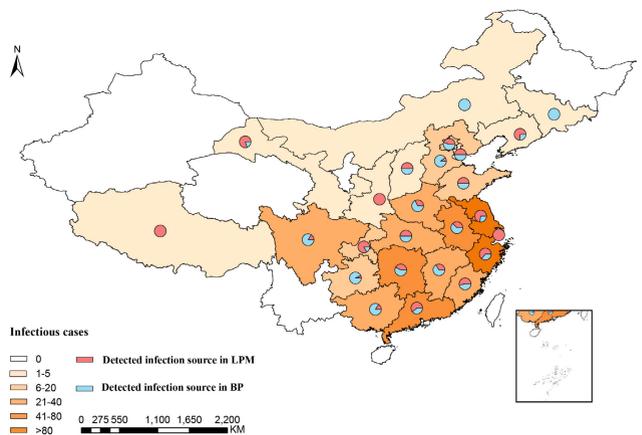
**FIGURE 7.** The proportion of infected human cases whose the most likely infection source is live poultry market or backyard poultry during the fifth wave epidemic. (a) Proportion of infected human cases that the most likely infection source is live poultry market or backyard poultry in Suzhou City (SZ), Jiangsu (JS), Zhejiang (ZJ), Guangdong Province (GD) and the whole mainland China (CN). (b) Proportion of infected human cases that the most likely infection source is live poultry market or backyard poultry in Guangxi (GX), Guizhou (GZ), Sichuan (SC), Henan (HA) and Hebei(HE) Province.

as examples. The parameters  $\beta_{lf}$ ,  $\beta_{fu}$ ,  $\beta_{lu}$ ,  $\beta_{bl}$ ,  $\beta_{bf}$ , and  $\beta_{br}$  in the transmission model Eq.(7), are sampled 1000 times within their respective ranges using a Latin hypercube sampling (LHS) method to optimize the sampling of parameter space. The results are presented as box plots depicting the median, lower and upper quartiles of  $L(u, t)$  from the 1000 simulations. For infected person node  $o_1$ , 9 live poultry market nodes are detected. Live poultry market node  $l_2$  is the most likely infection source of  $o_1$ , because its  $L(u, t)$  is far greater than that of the other 8 nodes in Fig 5(a). Similarly, the live poultry market node  $l_{10}$  is the most likely infection source of infected person node  $o_2$  in Fig 5(b). When the likelihood values  $L(u, t)$  of the first two nodes had approximately the same rank, they are both considered as the most likely infection sources.

*Uncertainty and sensitivity analyses.* The estimations of  $\beta_{lf}$ ,  $\beta_{fu}$ ,  $\beta_{lu}$ ,  $\beta_{bl}$ ,  $\beta_{bf}$ , and  $\beta_{br}$  are uncertain, so we use

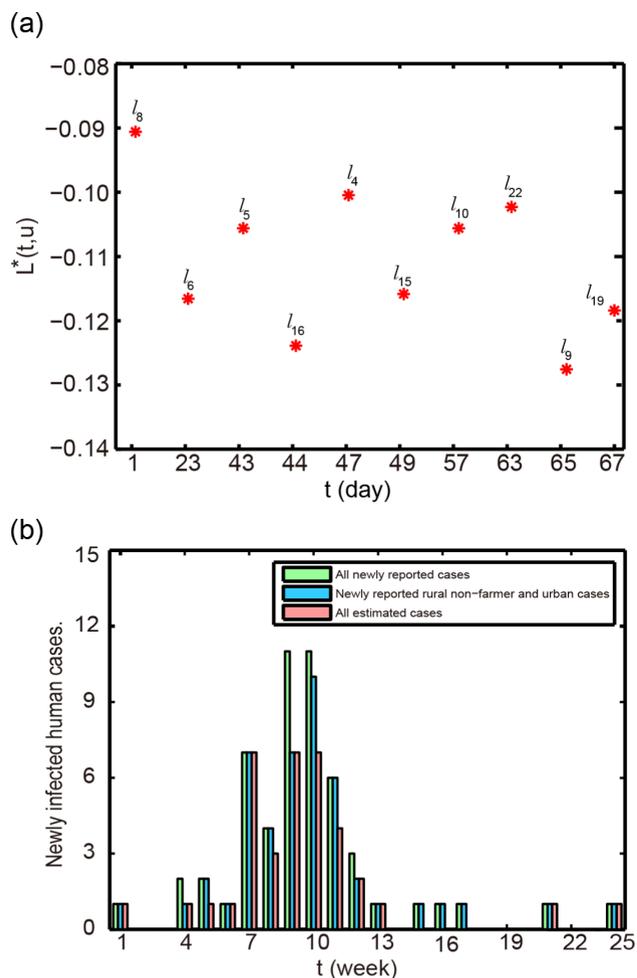


**FIGURE 8.** The proportion of infected urban residents, rural non-farmers and rural farmers. (a) The proportion of infected urban residents, rural non-farmers and rural farmers in the five waves epidemics in the whole mainland China (CN), Jiangsu (JS), Zhejiang (ZJ), Guangdong (GD) Province and Suzhou City (SZ). (b) The proportion of infected urban residents, rural non-farmer and rural farmer in the 5th wave epidemic in Guangxi (GX), Guizhou (GZ), Sichuan (SC), Henan (HA) and Hebei(HE) Province.



**FIGURE 9.** Geographic distribution of the proportion of human A(H7N9) cases whose detected infection source is live poultry market or backyard poultry during the 5th wave epidemic in mainland China.

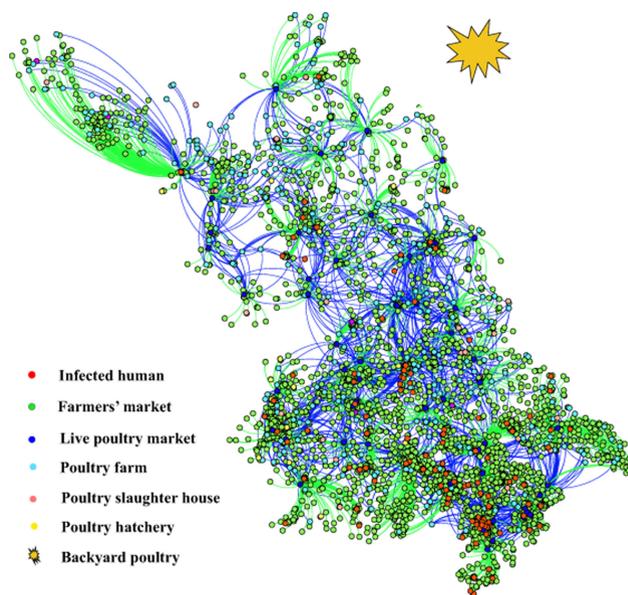
the Latin hypercube sampling method to sample these 6 parameters, which appear in the transmission model Eq.(7) and the derived expression  $L(u, t)$ . Latin hypercube sampling (LHS) is an extension of the Latin square sampling method. Uncertainty and sensitivity analyses based on LHS and partial rank correlation coefficients (PRCCs) have been applied to



**FIGURE 10.** (a) First arrival time and  $L(t, u)$  for each detected live poultry market infection source in Suzhou City. Numbers on the red stars are the labels of live poultry market nodes in the network; the time on the horizontal axis is unevenly divided for more convenient presentation. (b) The human infection scale of all newly reported cases, the newly reported rural non-farmer and urban cases, and the estimated cases.

disease transmission models in previous studies [39]–[42]. In our work, we perform uncertainty and sensitivity analyses for the parameters  $\beta_{lf}$ ,  $\beta_{fu}$ ,  $\beta_{lu}$ ,  $\beta_{bl}$ ,  $\beta_{bf}$ , and  $\beta_{br}$  in the model Eq.(7) using *LHS* with 1000 samples. The parameter values and their corresponding 95% confidence intervals are listed in Table 2, and a normal distribution function is used for all parameters [43]–[45]. As an example, the partial rank correlation coefficients (PRCCs) for the parameters with respect to the maximum likelihood  $L(u, t)$  associated with all detected suspected live poultry market infection sources for human cases  $o_1$  and  $o_2$  are shown in Fig 6. The results indicate that the PRCCs associated with the parameters  $\beta_{bl}$  and  $\beta_{bf}$  are the most influential in determining the magnitude of the aggregate  $L(u, t)$ , ( $|PRCC| > 0.5$ ), and are statistically significant ( $p < 0.05$ ).

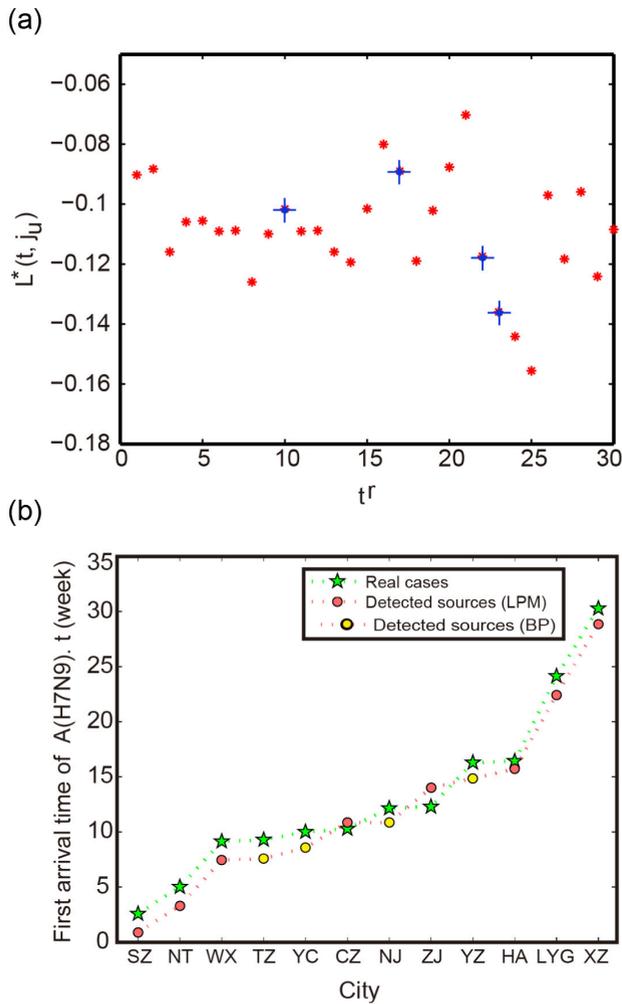
Using this approach, we analyze the infection sources of all infected human nodes in Suzhou City and found that approximately 68.5% of the infected human cases could be detected as having their infection source in the live poultry



**FIGURE 11.** The live poultry transport network constructed for Jiangsu Province.

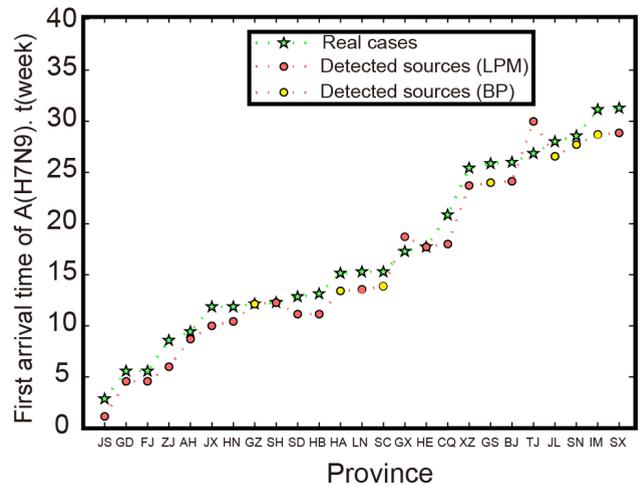
markets, and the infection source of the remaining 31.5% of infected human cases as a result of backyard poultry, as shown in Fig 7. Furthermore, we identify the infection source of all human cases in every province of mainland China. The proportion of infections originating in the live poultry market is also higher than that originating from backyard poultry in the provinces of Jiangsu, Zhejiang and Guangdong, as shown in Fig 7(a). The chi-square test is used to compare the proportion of infected human cases whose the most likely infection source is live poultry market or backyard poultry among different areas. There is no significant difference among the Jiangsu, Zhejiang and Guangdong Province ( $\chi^2 = 4.229, p = 0.376$ ), as shown in TABLE 6 (APPENDIX C). However, the source of more than 80% of infected human cases stems from backyard poultry in several provinces, including Guangxi, Guizhou, Sichuan, Henan and Hebei Provinces, and this proportion is 90% or higher in Guizhou and Hebei in Fig 7(b). Overall, the source of approximately 65.1% of human A(H7N9) infections in mainland China is the live poultry market while the remaining 34.9% resulted from backyard poultry. There is no significant difference among Guangxi, Guizhou, Sichuan, Henan and Hebei Provinces ( $\chi^2 = 5.282, p = 0.261$ ), as shown in TABLE 7 (APPENDIX C). The other live poultry transport networks in China, Zhejiang, Guangdong, Guangxi, Guizhou, Sichuan, Henan, and Hebei, are given in APPENDIX B.

As a result of the above discovery regarding infection sources, we compare the A(H7N9) human cases across the five epidemics. Fig 8(a) shows that the residences of infected human in mainland China gradually shifted from urban (urban residents) to rural areas (rural farmers and rural non-farmers) from the first to the fifth epidemics with a significant difference ( $\chi^2 = 21.207, p < 0.01$ ),



**FIGURE 12.** (a) First arrival time and  $L(t, u)$  of each detected live poultry market infection source in the whole Jiangsu Province.  $t^r$  on the horizontal axis represents the rank of the first arrival time. The blue cross mark denotes the live poultry market that are reported to be A(H7N9) positive. (b) First arrival time of infection sources in each infected city in Jiangsu Province. The sign on horizontal axis is the acronym of cities: Suzhou, Nantong, Wuxi, Taizhou, Yancheng, Changzhou, Nanjing, Zhenjiang, Yangzhou, Huai'an, Lianyungang, and Xuzhou. The cities are sorted by the first arrival time of infection source.

although the precise situation varied by province, as shown in TABLE 8 (APPENDIX C). This shift also explains the changes in the proportion of the sources is backyard poultry in Fig 8(a). In addition, for the newly infected provinces in the fifth wave epidemic, i.e. those that with no or fewer human cases in the previous four epidemics, such as Sichuan (the number of infected human in the 1st, 2nd, 3rd, 4th, and 5th wave epidemic is 0, 0, 0, 0, 30, respectively), Guangxi (0, 2, 0, 0, 25, respectively), Guizhou (0, 0, 1, 0, 15, respectively), Henan (4, 0, 0, 1, 22, respectively), and Hebei (1, 0, 0, 6, 20, respectively), the proportion of rural infected human cases is relatively high in the 5th wave epidemic, as shown in Fig 8(b). This reveals that backyard poultry is also an important risk factor for the spread of A(H7N9) virus, especially in the fifth wave, as shown in combination with Fig 8(b). Furthermore, we show the geographic



**FIGURE 13.** First arrival time of infection sources in each infected province of mainland China.

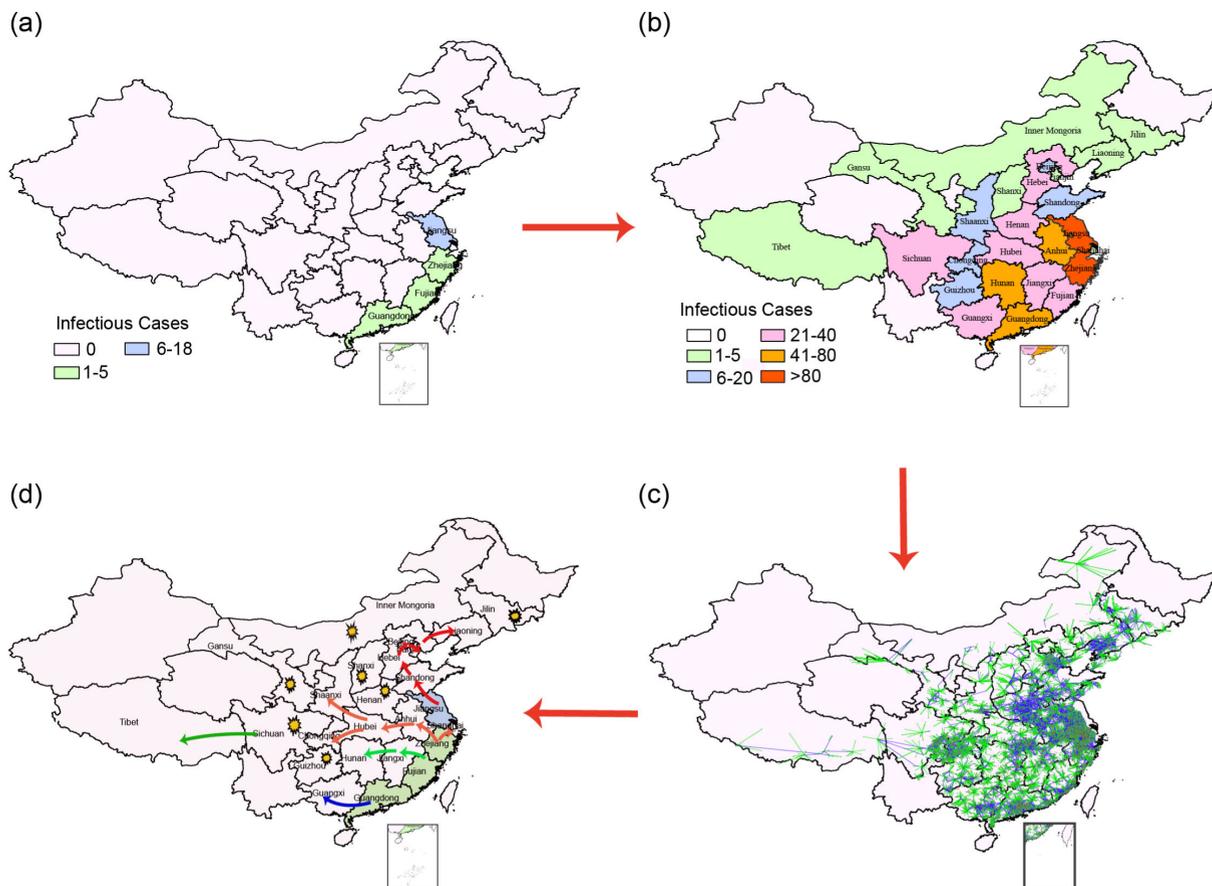
distribution of the proportion of human cases in which the detected infection source was live poultry markets and backyard poultry, in the 5th avian influenza A(H7N9) wave in mainland China, as shown in Fig 9. The difference in the proportion of human cases whose detected infection source is the live poultry market or backyard poultry across provinces was significant ( $\chi^2 = 121.206, p < 0.01$ ), as shown in TABLE 9 (APPENDIX C).

From the above, we can see that, in addition to live poultry markets, backyard poultry is an important infection source of human A(H7N9) infections, and the risk presented by backyard poultry was high, especially in the fifth wave epidemic in newly infected provinces. This may be due to the accumulation of live poultry in rural areas after the closing of the live poultry markets in some cities or provinces. Therefore, live poultry market closure was effective for controlling the human risk of infection with avian influenza A(H7N9) virus, but cleaning and transforming live poultry markets, changing the trading mode, and immunizing live poultry would have more substantial effects than just closing live poultry markets.

### B. FIRST ARRIVAL TIME AND THE MAXIMUM LIKELIHOOD VALUE OF INFECTION SOURCES

It is known that a single infection source may infect multiple infected persons. For example, the live poultry market node  $l_2$  is detected as the infection source of person node  $o_1$  at time  $t_1 - 2$ , as shown in Fig 3, but also transmit virus to infected person node  $o_2$  at time  $t_2 - 2$ . There are many similar examples, and thus, we compute the first arrival time of each detected infection source.

**Note.** Fig 10(a) shows the first arrival time and maximum likelihood  $L(t, u)$  of the detected live poultry market infection sources in Suzhou City, where the total number of live poultry markets is 22, 10 of these markets are detected as the most likely infection sources. Whereas the deviation between the estimated cases and all newly reported cases is



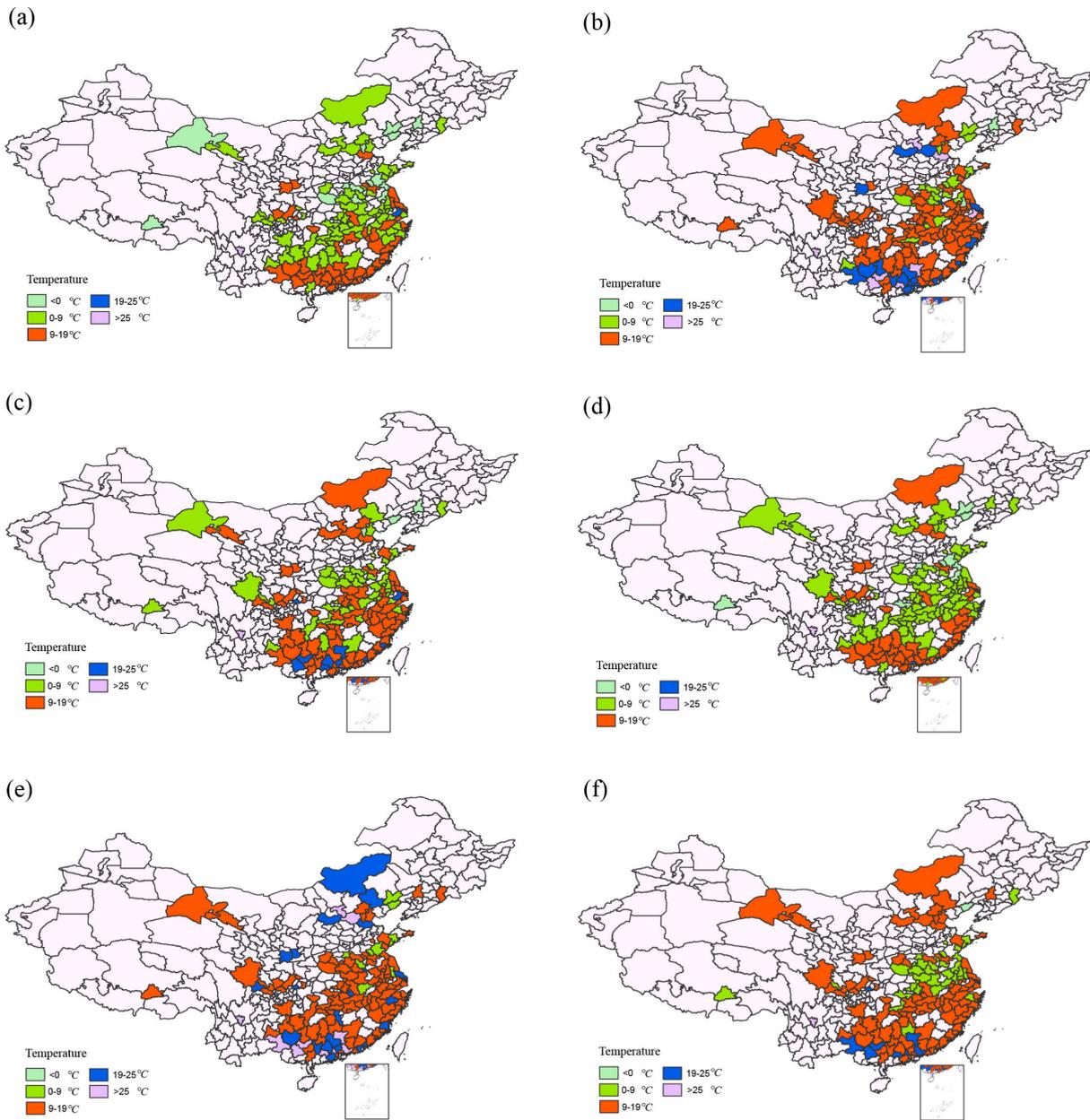
**FIGURE 14.** Inference of the most likely spread map of the A(H7N9) virus in mainland China. (a) The geographic distribution of A(H7N9) human cases in the beginning of the 5th wave. (b) The geographic distribution of A(H7N9) human cases by May 17th, 2017. (c) The constructed live poultry transport network of infected cities in mainland China. (d) The map of the most likely spread map of the A(H7N9) virus. The provinces marked with yellow star, represent that the infection sources for their first human cases were backyard poultry.

relatively large, the deviation between the estimated cases and the sum of reported rural non-farmer and urban cases is small, as shown in Fig 10(b). This is because the virus can only be transmitted to urban residents and rural non-farmers from the live poultry market infection source using the forward transmission algorithm. However, backyard poultry are the infection sources of rural farmers and some non-farmers. The small deviation between the estimated cases and the sum of reported rural non-farmer and urban cases demonstrates that our algorithm is valid. The large deviation between the estimated cases and all newly reported cases reveals the high risk of backyard poultry spreading the A(H7N9) virus.

**City.** Our network displayed connections between adjacent cities in the same province, especially in the urban fringes. Therefore, we expand from a single city to the whole province, to analyze the characteristics of A(H7N9) transmission among cities. Fig 11 shows our constructed live poultry transport network for Jiangsu Province. Fig 12(a) shows the first arrival time and maximum likelihood  $L(t, u)$  of the detected live poultry market infection sources in the whole of Jiangsu Province, where there are 75 live poultry markets, 30 of which are detected as the most likely infection sources.

In addition, four of these markets were reported as A(H7N9) positive markets [46]. We must take into account that not every live poultry market was tested for A(H7N9) virus on any given day, and therefore we are using limited data to illustrate the live poultry markets in which the virus was detected. Cities are sorted according to the first arrival time of the infection source, as shown in Fig 12(b), which shows that the A(H7N9) virus spread from southeast to northwest China through the live poultry trade. In Changzhou, Zhenjiang, and Huaian City, the infection sources for the first human cases were the live poultry market (LPM) in their adjacent cities, and the first arrival time of the infection source was later than human cases in their own cities. In Taizhou, Yancheng, Nanjing, Yangzhou City, the virus was first found in backyard poultry (BP).

**Province.** Furthermore, we expand the research object from a single province to all of mainland China to analyze the characteristics of A(H7N9) transmission among infected provinces through the live poultry transport network. In Fig 13, the first arrival time of infection sources to each province are displayed. The infection sources of the first human cases were the live poultry markets (LPM) in



**FIGURE 15.** The geographic distribution of weekly temperature on the first and peak arrival time of infection sources. (a) Weekly minimum temperature on first arrival time; (b) Weekly temperature on first arrival time; (c) Weekly average temperature on first arrival time; (d) Weekly minimum temperature during the peak arrival time; (e) Weekly maximum temperature during the peak arrival time; (f) Weekly average temperature during the peak arrival time.

the neighboring provinces of some provinces. The A(H7N9) virus was first detected in backyard poultry (BP) in Guizhou, Henan, Sichuan, Gansu, Jilin, Shaanxi Province and Inner Mongolia Autonomous Region.

**C. THE MOST LIKELY TRANSMISSION MAP**

Human A(H7N9) influenza cases first occurred in Jiangsu, Guangdong, Fujian and Zhejiang Provinces, at the beginning of the 5th wave epidemic, as shown in Fig 14(a). By May 17th, 2017, the epidemic had spread across 25 provinces in mainland China, as shown in Fig 14(b).

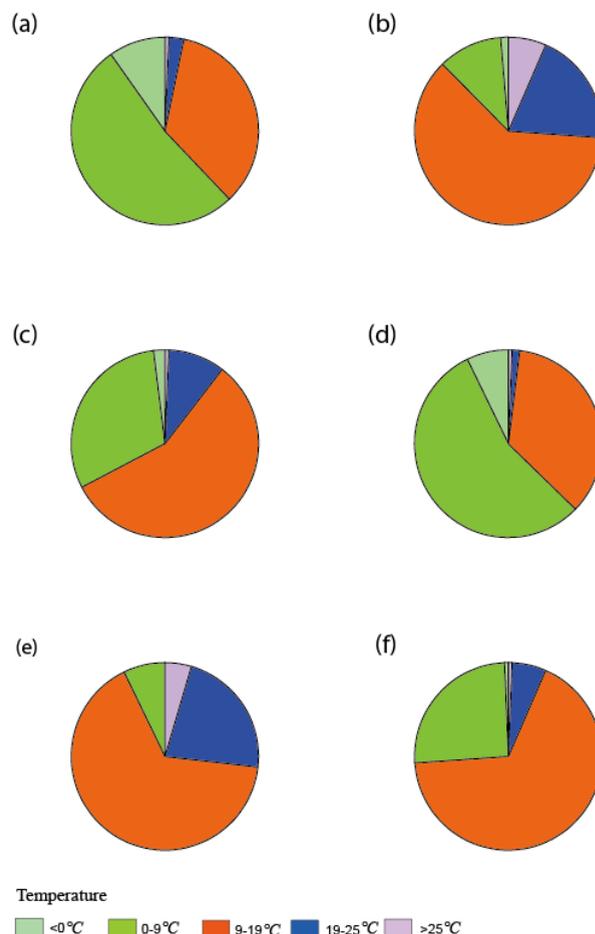
Fig 14(c) shows the live poultry transport network constructed for infected cities in mainland China. Based on this network, we infer the most likely spread map of A(H7N9) virus using Algorithm 1 and Algorithm 2, as shown in Fig 14(d). The A(H7N9) virus spread westward along four paths (i.e. the red, pink, green and blue paths) from the Yangtze and Pearl River Deltas, Zhejiang and Jiangsu Provinces and from there, spread to more provinces. In the 7 provinces marked with a yellow star (i.e. Guizhou, Henan, Sichuan, Gansu, Jilin, and Shaanxi Provinces and Inner Mongolia Autonomous Region), A(H7N9) virus was first found in the backyard poultry kept within the area.

### D. TEMPERATURE CHARACTERISTICS

Previous studies have demonstrated that temperature contributed significantly to the occurrence of human infection with the influenza A(H7N9) virus [47]–[49]. In this subsection, we analyze the temperature characteristics on the first and peak arrival times of the infection sources detected in each city. Fig 15 shows the geographic distribution of weekly temperature on the first and peak arrival time of infection sources. The weekly temperatures (minimum, maximum and average temperature) during the study period were obtained from the Weather Underground (<http://www.wunderground.com>). Fig 16 shows the proportion of infected cities in which the weekly temperature is concentrated in different temperature ranges on the first and peak arrival times of infection sources. At the first arrival time of the infection sources, the minimum temperature in 52.3% of infected cities was concentrated in the range of 0°C~9°C, the proportions of infected cities with minimum temperatures of < 0°C, 9°C~19°C, 19°C~25°C and > 25°C were 9.8%, 34.6%, 2.6%, and 0.7%, respectively, as shown in Fig 16 (a). The maximum temperature in 61.4% of infected cities was concentrated in the range of 9°C~19°C, the proportion of infected cities with maximum temperatures of < 0°C, 0°C~9°C, 19°C~25°C and > 25°C were 1.3%, 11.1%, 19.6%, and 6.6%, respectively, as shown in Fig 16 (b). The average temperature in 56.9% of infected cities was concentrated in the range of 9°C~19°C, the proportion of infected cities with average temperatures of < 0°C, 0°C~9°C, 19°C~25°C and > 25°C are 2.0%, 30.7%, 9.8%, and 0.6%, respectively, as shown in Fig 16 (c). During the peak arrival time of infection sources, the temperature presented the same characteristics, although the proportions varied among temperature ranges. In summary, the risk of human infection with the influenza A(H7N9) virus was high when the temperature is 9°C~19°C, moderate when the temperature was 0°C~9°C or 19°C~25°C, and low when the temperature was < 0°C or > 25°C.

### VI. CONCLUSION

In conclusion, this study first constructed a live poultry transport network using data extracted from Baidu Map using the JavaScript language. Then, an A(H7N9) transmission model of a network based on probabilistic discrete-time Markov chains was established. Significantly, we introduced spatiotemporal backward detection and forward transmission algorithms to detect the infection source, compute the first arrival time of A(H7N9) virus, and analyze their characteristics. By detecting the infection source of each human case, we discovered that, in addition to the live poultry markets, backyard poultry were also a very significant source of infection for A(H7N9) human cases. The risk from backyard poultry infection was high, which was most especially seen in the 5th wave of avian influence in newly infected provinces. Therefore, we believe that live poultry market closures were effective in controlling the risk to human from the avian



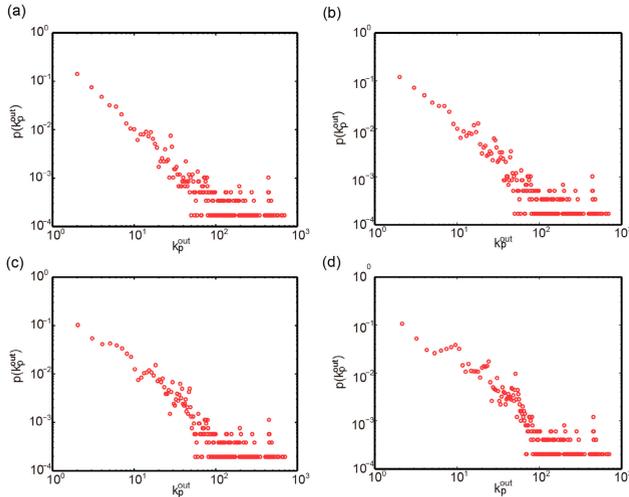
**FIGURE 16.** The proportion of infected cities in which the weekly temperature is < 0°C, 0°C~9°C, 9°C~19°C, 19°C~25°C or > 25°C for (a) Weekly minimum temperature on first arrival time; (b) Weekly maximum temperature on first maximum arrival time; (c) Weekly average temperature during the peak arrival time; (d) Weekly minimum temperature during the peak arrival time; (e) Weekly maximum temperature during the peak arrival time; (f) Weekly average temperature during the peak arrival time.

influenza A(H7N9) virus infection. However, beyond simply closing the live poultry markets, there would be a more significant effect from cleaning and transforming the live poultry markets, changing the trading mode and immunizing live poultry. Meanwhile, we drew the most likely transmission map of the A(H7N9) virus along the live poultry transport network. The result shows that the A(H7N9) virus spread westward along the Yangtze River Delta and Pearl River Delta, and then the virus was spread to more provinces mainly from the Yangtze River Delta. Furthermore, we analyzed the temperature characteristics of A(H7N9) virus transmission according to the arrival time we obtained for detected infection sources in each city. When the temperature was 9°C~19°C, the risk of human infection with the influenza A(H7N9) virus was high; 0°C~9°C and 19°C~25°C, represented moderate risk and the temperature range below 0°C or above 25°C, indicated low infection risk.

These results provide meaningful suggestions for the prevention and control of A(H7N9) avian influenza infection.

**TABLE 5.** Average out-degree of poultry farms and average in-degree of live poultry markets in the network for different distances of poultry farms to live poultry markets.

Distance	80km	100km	120km	150km
Average out-degree of poultry farm $\langle k_p^{out} \rangle$	6.25	8.67	11.46	15.98
Average in-degree of live poultry market $\langle k_l^{in} \rangle$	29.36	40.96	54.16	75.50

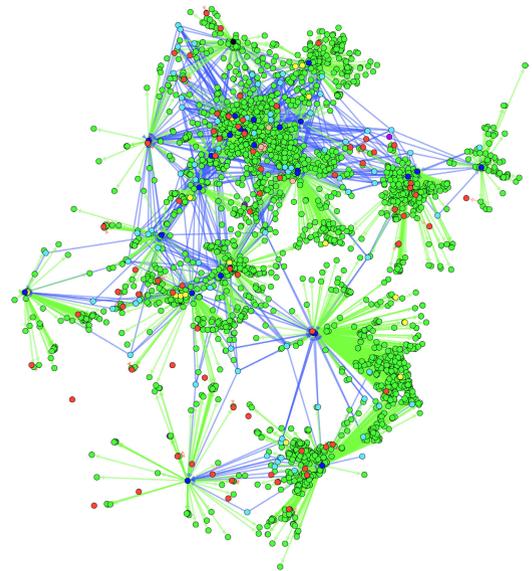


**FIGURE 17.** The out-degree distribution of poultry farm nodes, in the network within different distances from poultry farm to live poultry market. (a) Within 80 kilometers from poultry farm to live poultry market. (b) Within 100 kilometers from poultry farm to live poultry market. (c) Within 120 kilometers from poultry farm to live poultry market. (d) Within 150 kilometers from poultry farm to live poultry market.

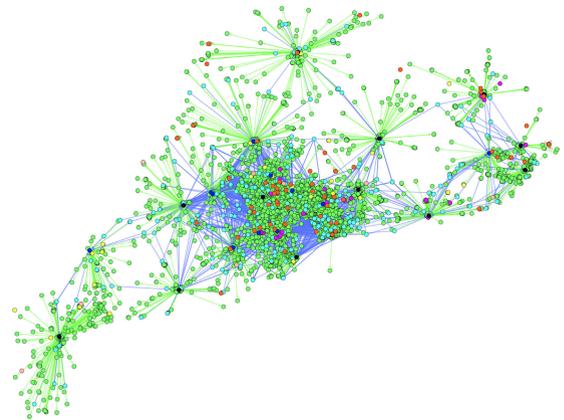
However, some deficiencies exist in our study. For instance, we constructed edges to the networks according to the spherical distance between different nodes, ignoring the real live poultry transport paths along the actual roads, due to the difficulty of acquiring data. Moreover, we did not take into account the change in the infection rate with spatiotemporal changes in temperature, and humidity. Research into a more perfectly representative transmission model on a constructed real network is a difficult challenge, requiring more sophisticated data mining technology and a more comprehensive concept of the necessary mathematical modeling. Our work may shed some light on future studies into infectious diseases within real and complex networks.

**APPENDIX A  
THE REASON B) FOR THE DISTANCE DETERMINATION  
FROM POULTRY FARM TO LIVE POULTRY MARKET**

We determined the maximum distance between poultry farms and live poultry market by comparing the average degree and degree distributions, in the network within different distances (80km, 100km, 120km, 150km) from poultry farm to live poultry market. In the directed network  $G = (V, E)$  with  $N$  nodes, where  $V$  is the set of nodes and  $E$  is the set of edges. Each node (individual)  $v$  has a certain out-degree  $k_v^{out}$  and



**FIGURE 18.** The constructed live poultry transport network of Zhejiang Province.



**FIGURE 19.** The constructed live poultry transport network of Guangdong Province.

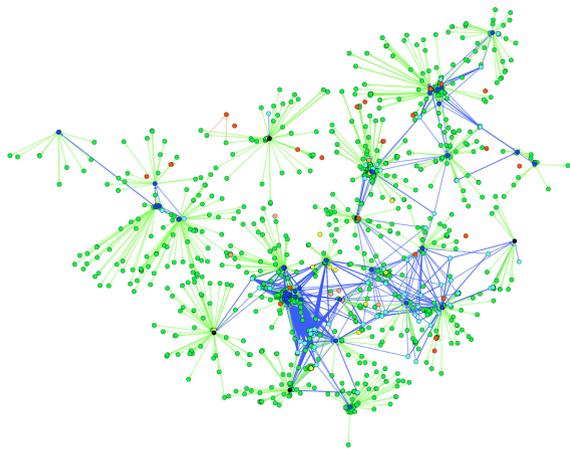
in-degree  $k_v^{in}$ .  $k_v^{out}$  represents the number of edges connected from node  $v$  to other nodes,  $k_i^{in}$  represents the number of edges connected from other nodes to node  $i$  [50]. The average out-degree  $\langle k^{out} \rangle$  and average in-degree  $\langle k^{in} \rangle$  are denoted as

$$\langle k^{out} \rangle = \frac{\sum_{v \in V} k_v^{out}}{N}, \quad \langle k^{in} \rangle = \frac{\sum_{v \in V} k_v^{in}}{N}. \quad (14)$$

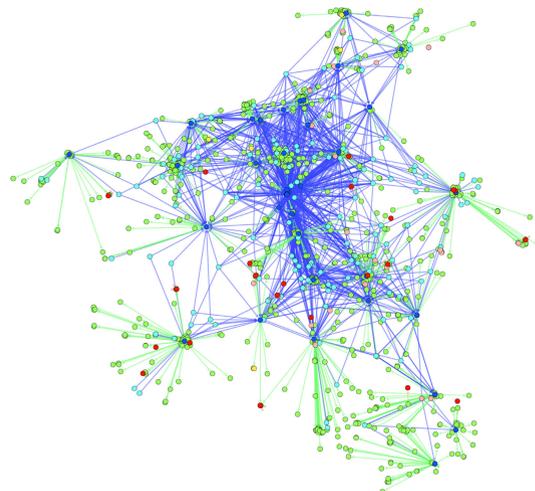
In our study, we are concerned about the average out-degree of poultry farm  $\langle k_p^{out} \rangle$  and the average in-degree of live poultry market  $\langle k_l^{in} \rangle$ . That is:

$$\langle k_p^{out} \rangle = \frac{\sum_{i \in V_p} k_i^{out}}{N_p}, \quad \langle k_l^{in} \rangle = \frac{\sum_{j \in V_l} k_j^{in}}{N_l}. \quad (15)$$

where  $V_p$  is the set of poultry farm nodes,  $N_p$  is the number of poultry farm nodes,  $V_l$  is the set of live poultry market nodes,  $N_l$  is the number of live poultry market nodes.



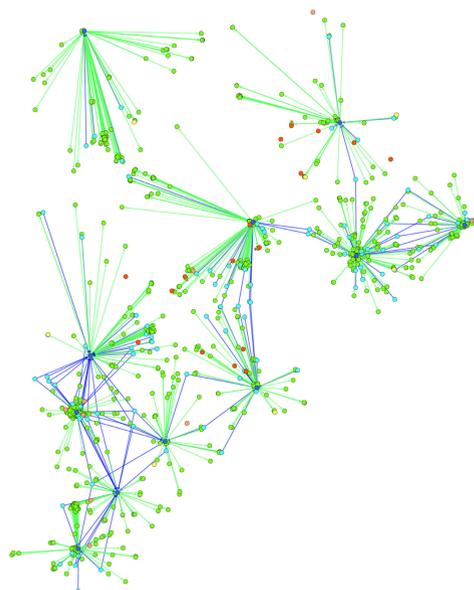
**FIGURE 20.** The constructed live poultry transport network of Guangxi Province.



**FIGURE 23.** The constructed live poultry transport network of Henan Province.



**FIGURE 21.** The constructed live poultry transport network of Guizhou Province.



**FIGURE 24.** The constructed live poultry transport network of Hebei Province.



**FIGURE 22.** The constructed live poultry transport network of Sichuan Province.

According to Eq.(15), we calculated the average out-degree of the poultry farms and average in-degree of the

live poultry market, in the network within different distances (80km, 100km, 120km, 150km) from poultry farm to live poultry market in Table 5. We found that the number of average live poultry market neighbors of poultry farms  $\langle k_p^{out} \rangle$  and the poultry farm neighbors of live poultry market  $\langle k_l^{in} \rangle$  were too dense, when the distance between the poultry farm and live poultry market was more than 100 kilometers. This is more consistent with the present situation in China, where the distance between poultry farm and live poultry market is less than 100 kilometers.

In addition, we verified the reasonableness of distance selection by comparing the degree distribution of the network within different distances.  $p(k_p^{out})$  is the out-degree distribution of poultry farm nodes in the network, which represents the probability that out-degree is  $k$  of a arbitrarily poultry

**TABLE 6. Characteristics of most likely infection sources of H7N9 human cases in different areas.**

Characteristic	Mainland China (n = 672)	Jiangsu (n = 144)	Zhejiang (n = 84)	Guangdong (n = 62)	Suzhou (n = 54)	$\chi^2$	p
Live poultry market n (% versus backyard poultry)	437 (65.1%)	105 (72.3%)	54 (64.3%)	38 (61.3%)	37 (68.5%)	4.229	0.376

**TABLE 7. Characteristics of most likely infection sources of H7N9 human cases in different areas.**

Characteristic	Guangxi (n = 25)	Guizhou (n = 15)	Sichuan (n = 30)	Henan (n = 22)	Hebei (n = 20)	$\chi^2$	p
Live poultry market n (% versus backyard poultry)	4 (16.0%)	1 (6.7%)	5 (16.7%)	7 (31.8%)	2 (10.0%)	5.282	0.26

**TABLE 8. Characteristics of residences of H7N9 human cases among five waves.**

	Characteristic	1st wave	2nd wave	3rd wave	4th wave	5th wave	$\chi^2$	p
Mainland China	Urban cases	74(56.9%)	171(55.7%)	111(51.4%)	49(41.5%)	313(46.6%)	21.207	0.007
	Rural farmer cases	33(25.4%)	79(25.7%)	65(30.1%)	38(32.2%)	175(26.0%)		
Jiangsu	Urban cases	18(62.1%)	19(70.4%)	10(45.4%)	13(50.0%)	313(52.4%)	5.229	0.733
	Rural farmer cases	8(27.6%)	5(18.5%)	8(36.4%)	8(30.8%)	46(31.7%)		
Zhejiang	Urban cases	26(57.8%)	46(48.9%)	22(46.8%)	14(43.8%)	41(48.8%)	4.108	0.847
	Rural farmer cases	6(13.3%)	25(26.6%)	13(27.7%)	9(28.1%)	22(26.2%)		
Guangdong	Urban cases	0(0%)	72(66.0%)	45(62.5%)	3(21.4%)	28(45.2%)	20.565	0.008
	Rural farmer cases	1(100%)	26(23.9%)	17(23.6%)	7(50.0%)	18(29.0%)		
Suzhou	Urban cases	0(0%)	1(25.0%)	5(71.4%)	2(22.2%)	28(51.9%)	20.646	0.008
	Rural farmer cases	5(71.4%)	1(25.0%)	0(0%)	7(77.8%)	19(35.1%)		

**TABLE 9. Characteristics of most likely infection sources of H7N9 human cases among different provinces.**

Province	Live poultry market n (% versus backyard poultry)
Shanghai	5 (100%)
Jiangsu	105 (72.9%)
Zhejiang	54 (64.3%)
Anhui	22 (37.9%)
Jiangxi	13 (34.2%)
Hubei	14 (50.0%)
Hunan	25 (42.4%)
Guangdong	38 (61.3%)
Fujian	17 (53.1%)
Guangxi	4 (16.0%)
Guizhou	1 (6.7%)
Sichuan	5 (31.8%)
Henan	7 (83.3%)
Chongqing	5 (50.0%)
Shandong	7 (10.0%)
Hebei	2 (75.0%)
Liaoning	3 (100%)
Jilin	0 (0%)
Tianjin	0 (0%)
Shaanxi	3 (100%)
Gansu	4 (80.0%)
Beijing	4 (44.4%)
Shanxi	1 (50.0%)
Tibet	3 (100%)
Inner Mongolia	0 (0%)
$\chi^2$	121.206
p	<0.001

farm node [51]. That is:

$$p(k_p^{out}) = \frac{N_p(k)}{N_p} \tag{16}$$

where  $N_p(k)$  is the number of poultry farm nodes whose out-degree is  $k$ .

The simulation results show that the out-degree distribution of poultry farm nodes are all power law distributions [52], in the network within different distances from poultry farm to live poultry market in Fig 17.

**APPENDIX B  
THE CONSTRUCTED LIVE POULTRY TRANSPORT NETWORKS**

See Figs. 18–24.

**APPENDIX C  
STATISTICAL ANALYSIS**

See Tables 6–9.

**REFERENCES**

- [1] World Health Organization. *Health Topics: Influenza*. Accessed: Aug. 5, 2019. [Online]. Available: <http://www.who.int/influenza/human-animal-interface/influenza-h7n9/en/>
- [2] R. E. Kahn and J. A. Richt, “The novel H7N9 influenza a virus: Its present impact and indeterminate future,” *Vector-Borne Zoonotic Diseases*, vol. 13, no. 6, pp. 347–348, 2013.
- [3] Y. Qin, P. W. Horby, T. K. Tsang, E. Chen, L. Gao, J. Ou, T. H. Nguyen, T. N. Duong, V. Gasimov, L. Feng, and P. Wu, “Differences in the epidemiology of human cases of avian influenza A(H7N9) and A(H5N1) viruses infection,” *Clin. Infectious Diseases*, vol. 61, no. 4, pp. 563–571, 2015.
- [4] L. Zhou, E. Chen, C. Bao, N. Xiang, J. Wu, S. Wu, J. Shi, X. Wang, Y. Zheng, Y. Zhang, R. Ren, C. M. Greene, F. Havers, A. D. Iuliano, Y. Song, C. Li, T. Chen, Y. Wang, D. Li, D. Ni, Y. Zhang, Z. Feng, T. M. Uyeki, and Q. Li, “Clusters of human infection and human-to-human transmission of avian influenza A(H7N9) virus, 2013–2017,” *Emerg. Infectious Diseases*, vol. 24, no. 2, pp. 397–400, 2018.
- [5] R. Gao, B. Cao, Y. Hu, Z. Feng, D. Wang, W. Hu, and X. Xu, “Human infection with a novel avian-origin influenza A (H7N9) virus,” *New England J. Med.*, vol. 368, no. 20, pp. 1888–1897, 2013.
- [6] J. Zhang, Z. Jin, G.-Q. Sun, X.-D. Sun, Y.-M. Wang, and B. Huang, “Determination of original infection source of H7N9 avian influenza by dynamical model,” *Sci. Rep.*, vol. 4, May 2014, Art. no. 4846.
- [7] N. Xiang et al., “Assessing change in avian influenza A(H7N9) virus infections during the fourth epidemic—China, September 2015–August 2016,” *MMWR. Morbidity Mortality Weekly Rep.*, vol. 65, no. 49, pp. 1390–1394, 2016.
- [8] L. Zhou, R. Ren, L. Yang, C. Bao, J. Wu, D. Wang, C. Li, N. Xiang, Y. Wang, D. Li, H. Sui, Y. Shu, Z. Feng, Q. Li, and D. Ni, “Sudden increase in human infection with avian influenza A(H7N9) virus in China, September–December 2016,” *Western Pacific Surveill. Response J.*, vol. 8, no. 1, pp. 6–14, 2017.
- [9] M. C. Chan, R. W. Chan, L. L. Chan, C. K. Mok, K. P. Hui, J. H. Fong, K. P. Tao, L. L. Poon, J. M. Nicholls, P. Y. Guan, and J. M. Peiris, “Tropism and innate host responses of a novel avian influenza A H7N9 virus: An analysis of *ex-vivo* and *in-vitro* cultures of the human respiratory tract,” *Lancet Respiratory Med.*, vol. 1, no. 7, pp. 534–542, 2013.
- [10] S. Su, M. Gu, D. Liu, J. Cui, G. F. Gao, J. Zhou, and X. Liu, “Epidemiology, evolution, and pathogenesis of H7N9 influenza viruses in five epidemic waves since 2013 in China,” *Trends Microbiol.*, vol. 25, no. 9, pp. 713–728, 2017.
- [11] B. J. Cowling et al., “Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: A population-based study of laboratory-confirmed cases,” *Lancet*, vol. 382, no. 9887, pp. 129–137, 2013.
- [12] P. Wu, L. Wang, B. J. Cowling, J. Yu, V. J. Fang, F. Li, L. Zeng, J. T. Wu, Z. Li, G. M. Leung, and H. Yu, “Live poultry exposure and public response to influenza A(H7N9) in urban and rural China during two epidemic waves in 2013–2014,” *PLoS One*, vol. 10, no. 9, 2015, Art. no. e0137831.
- [13] X. Wang et al., “Epidemiology of avian influenza A H7N9 virus in human beings across five epidemics in mainland China, 2013–17: An epidemiological study of laboratory-confirmed case series,” *Lancet Infectious Diseases*, vol. 17, no. 8, pp. 822–832, 2017.

- [14] H. Yu, J. T. Wu, B. J. Cowling, Q. Liao, V. J. Fang, S. Zhou, P. Wu, H. Zhou, E. H. Y. Lau, D. Guo, M. Y. Ni, Z. Peng, L. Feng, H. Jiang, H. Luo, Q. Li, Z. Feng, Y. Wang, W. Yang, and G. M. Leung, "Effect of closure of live poultry markets on poultry-to-person transmission of avian influenza A H7N9 virus: An ecological study," *Lancet*, vol. 383, no. 9916, pp. 541–548, 2014.
- [15] C. Chen, S. Lu, P. Du, H. Wang, W. Yu, H. Song, and J. Xu, "Silent geographical spread of the H7N9 virus by online knowledge analysis of the live bird trade with a distributed focused crawler," *Emerg. Microbes Infections*, vol. 2, no. 1, pp. 1–7, 2013.
- [16] X. Zhou, Y. Li, Y. Wang, J. Edwards, F. Guo, A. C. A. Clements, B. Huang, and R. J. S. Magalhaes, "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," *J. Infection*, vol. 71, no. 4, pp. 470–479, 2015.
- [17] D. Shah and T. Zaman, "Detecting sources of computer viruses in networks: Theory and experiment," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 38, no. 1, pp. 203–214, 2010.
- [18] D. Shah and T. Zaman, "Rumors in a network: Who's the culprit?" *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5163–5181, Aug. 2011.
- [19] D. Shah and T. Zaman, "Rumor centrality: A universal source detector," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 40, no. 1, pp. 199–210, 2012.
- [20] W. Luo and W. P. Tay, "Identifying multiple infection sources in a network," in *Proc. Conf. Rec. 46th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Nov. 2012, pp. 1483–1489.
- [21] W. Luo, W. P. Tay, and M. Leng, "Identifying infection sources and regions in large networks," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2850–2865, Jun. 2013.
- [22] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Lage-Castellanos, and R. Zecchina, "Bayesian inference of epidemics on networks via belief propagation," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 112, no. 11, 2014, Art. no. 118701.
- [23] A. Y. Lokhov, M. Mézard, H. Ohta, and L. Zdeborová, "Inferring the origin of an epidemic with a dynamic message-passing algorithm," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 1, 2014, Art. no. 012801.
- [24] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, "Discrete-time Markov chain approach to contact-based disease spreading in complex networks," *Europhys. Lett.*, vol. 89, no. 3, p. 38009, 2010.
- [25] W. Wang, M. Tang, H. E. Stanley, and L. A. Braunstein, "Unification of theoretical approaches for epidemic spreading on complex networks," *Rep. Prog. Phys.*, vol. 80, no. 3, 2017, Art. no. 036603.
- [26] N. Antulov-Fantulin, A. Lančić, T. Šmuc, H. Štefancić, and M. Šikić, "Identification of patient zero in static and temporal networks: Robustness and limitations," *Phys. Rev. Lett.*, vol. 114, no. 24, 2014, Art. no. 248701.
- [27] B. A. Prakash, J. Vreeken, and C. Faloutsos, "Efficiently spotting the starting points of an epidemic in a large graph," *Knowl. Inf. Syst.*, vol. 38, no. 1, pp. 35–59, 2014.
- [28] B. A. Prakash, J. Vreeken, and C. Faloutsos, "Spotting culprits in epidemics: How many and which ones?" in *Proc. ICDM*, Brussels, Belgium, Dec. 2012, pp. 11–20.
- [29] Z. Shen, S. Cao, W.-X. Wang, Z. Di, and H. E. Stanley, "Locating the source of diffusion in complex networks by time-reversal backward spreading," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat.*, vol. 93, no. 3, p. 032301, 2016.
- [30] Y. Chen et al., "Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: Clinical analysis and characterisation of viral genome," *Lancet*, vol. 381, no. 9881, pp. 1916–1925, 2013.
- [31] C.-J. Bao, L.-B. Cui, M.-H. Zhou, L. Hong, G. F. Gao, and H. Wang, "Live-animal markets and influenza A (H7N9) virus infection," *New England J. Med.*, vol. 368, no. 24, pp. 2337–2339, 2013.
- [32] J. Zhang, W. Jing, W. Zhang, and Z. Jin, "Avian influenza A (H7N9) model based on poultry transport network in China," *Comput. Math. Methods*, vol. 2018, Nov. 2018, Art. no. 7383170.
- [33] Z. Peng, P. Wu, L. Ge, R. Fielding, X. Cheng, W. Su, M. Ye, Y. Shi, Q. Liao, H. Zhou, L. Zhou, L. Li, J. Wu, S. Zhang, Z. Yu, X. Wu, H. Ma, J. Lu, B. J. Cowling, and H. Yu, "Rural villagers and urban residents exposure to poultry in China," *PLoS One*, vol. 9, no. 4, 2014, Art. no. e95430.
- [34] X.-X. Wang, W. Cheng, Z. Yu, S.-L. Liu, H.-Y. Mao, and E.-F. Chen, "Risk factors for avian influenza virus in backyard poultry flocks and environments in Zhejiang Province, China: A cross-sectional study," *Infectious Diseases Poverty*, vol. 7, no. 1, 2018, Art. no. 65.
- [35] E. Ge, R. Zhang, D. Li, X. Wei, X. Wang, and P.-C. Lai, "Estimating risks of inapparent avian exposure for human infection: Avian influenza virus A (H7N9) in Zhejiang Province, China," *Sci. Rep.*, vol. 7, Jan. 2017, Art. no. 40016.
- [36] J. Kurscheid, M. Stevenson, P. A. Durr, J.-A. L. M. L. Toribio, S. Kurscheid, I. G. A. A. Ambarawati, M. Abdurrahman, and S. Fenwick, "Social network analysis of the movement of poultry to and from live bird markets in Bali and Lombok, Indonesia," *Transboundary Emerg. Diseases*, vol. 64, no. 6, pp. 2023–2033, 2017.
- [37] J. Jiang, S. Wen, S. Yu, Y. Xiang, and W. Zhou, "Rumor source identification in social networks with time-varying topology," *IEEE Trans. Depend. Sec. Comput.*, vol. 15, no. 1, pp. 166–179, Jan./Feb. 2016.
- [38] L. Cheng, B. V. Dongen, and W. Van Der Aalst, "Scalable discovery of hybrid process models in a cloud computing environment," *IEEE Trans. Services Comput.*, to be published.
- [39] S. M. Blower and H. Dowlatabadi, "Sensitivity and uncertainty analysis of complex models of disease transmission: An HIV model, as an example," *Int. Stat. Rev.*, vol. 62, no. 2, pp. 229–243, 1994.
- [40] S. M. Blower, A. R. Mclean, T. C. Porco, P. M. Small, P. C. Hopewell, M. A. Sanchez, and A. R. Moss, "The intrinsic transmission dynamics of tuberculosis epidemics," *Nature Med.*, vol. 1, no. 8, pp. 815–821, 1995.
- [41] M. A. Sanchez and S. M. Blower, "Uncertainty and sensitivity analysis of the basic reproductive rate: Tuberculosis as an example," *Amer. J. Epidemiol.*, vol. 145, no. 12, pp. 1127–1137, 1997.
- [42] M. Li, G. Sun, J. Zhang, Z. Jin, X. Sun, Y. Wang, B. Huang, and Y. Zheng, "Transmission dynamics and control for a brucellosis model in Hinggan League of Inner Mongolia, China," *Math. Biosci. Eng.*, vol. 11, no. 5, pp. 1115–1137, 2014.
- [43] K. J. Sharkey, R. G. Bowers, K. L. Morgan, S. E. Robinson, and R. M. Christley, "Epidemiological consequences of an incursion of highly pathogenic H5N1 avian influenza into the British poultry flock," *Proc. Roy. Soc. B, Biol. Sci.*, vol. 275, no. 1630, pp. 19–28, 2007.
- [44] S. Tang, Y. Xiao, L. Yuan, R. A. Cheke, and J. Wu, "Campus quarantine (Fengxiao) for curbing emergent infectious diseases: Lessons from mitigating A/H1N1 in Xi'an, China," *J. Theor. Biol.*, vol. 295, pp. 47–58, Feb. 2012.
- [45] Y. Xiao, S. Tang, and J. Wu, "Media impact switching surface during an infectious disease outbreak," *Sci. Rep.*, vol. 5, Jan. 2015, Art. no. 7838.
- [46] *Ministry of Agriculture and Rural Affairs of the People's Republic of China*. Accessed: Aug. 5, 2019. [Online]. Available: <http://www.moa.gov.cn/>
- [47] L.-Q. Fang, X.-L. Li, K. Liu, Y.-J. Li, H.-W. Yao, S. Liang, Y. Yang, Z.-J. Feng, G. C. Gray, and W.-C. Cao, "Mapping spread and risk of avian influenza A (H7N9) in China," *Sci. Rep.*, vol. 3, no. 7468, 2013, Art. no. 2722.
- [48] Y. Zhang, C. Feng, C. Ma, P. Yang, S. Tang, A. Laud, W. Sun, and Q. Wang, "The impact of temperature and humidity measures on influenza A (H7N9) outbreaks—Evidence from China," *Int. J. Infectious Diseases*, vol. 30, pp. 122–124, Jan. 2015.
- [49] X.-L. Li, Y. Yang, Y. Sun, W.-J. Chen, R.-X. Sun, K. Liu, M.-J. Ma, S. Liang, H.-W. Yao, G. C. Gray, L.-Q. Fang, and W.-C. Cao, "Risk distribution of human infections with avian influenza H7N9 and H5N1 virus in China," *Sci. Rep.*, vol. 5, no. 1, 2016, Art. no. 18610.
- [50] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, p. 47, Jan. 2002.
- [51] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the worldwide Web," *Nature*, vol. 401, no. 6749, pp. 130–131, 1999.
- [52] A.-L. Barabási, "Scale-free networks: A decade and beyond," *Science*, vol. 325, no. 5939, pp. 412–413, Jul. 2009.



**XIN PEI** is currently pursuing the Ph.D. degree in data science and technology with the North University of China. She is also a Visiting Student with the Complex System Research Center and Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University. She is working on the study of data-driven dynamics of epidemic spreading in networks. Her main research interests include mathematical biology, epidemiology, complex networks, and data analytic.



**ZHEN JIN** is currently a Professor with the Complex Systems Research Center, and the Director with the Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University, China. He is also a part-time Professor with the North University of China. He has more than 150 publications in various international journals. His main research interests include mathematical biology, epidemiology, complex networks, nonlinear dynamics, and data analytic.



**YONG WANG** is currently a Professor with the Chinese PLA Center for Disease Control and Prevention. He has more than 100 publications in various journals. His main research interests include infectious diseases epidemiology and field epidemiology.

...



**WENYI ZHANG** is currently an Associate Professor with the Chinese PLA Center for Disease Control and Prevention. He has more than 80 publications in various international journals. His main research interests include the surveillance and control of infectious diseases, the ecology of vector-borne diseases and zoonosis, and spatial epidemiology.