

2008

## A Spatial SIRS Boolean Network Model for the Spread of H5N1 Avian Influenza Virus among Poultry Farms


Alexander Kasyanov

Leona Kirkland

Mihaela Teodora Matache

University of Nebraska at Omaha, [dvelcsov@unomaha.edu](mailto:dvelcsov@unomaha.edu)

Follow this and additional works at: <https://digitalcommons.unomaha.edu/mathfacproc>

 Part of the [Epidemiology Commons](#), [Mathematics Commons](#), [Other Public Health Commons](#), and the [Poultry or Avian Science Commons](#)

Please take our feedback survey at: [https://unomaha.az1.qualtrics.com/jfe/form/SV\\_8cchtFmpDyGfBLE](https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE)

---

### Recommended Citation

Kasyanov, Alexander; Kirkland, Leona; and Matache, Mihaela Teodora, "A Spatial SIRS Boolean Network Model for the Spread of H5N1 Avian Influenza Virus among Poultry Farms" (2008). *Mathematics Faculty Proceedings & Presentations*. 2.

<https://digitalcommons.unomaha.edu/mathfacproc/2>

This Conference Proceeding is brought to you for free and open access by the Department of Mathematics at DigitalCommons@UNO. It has been accepted for inclusion in Mathematics Faculty Proceedings & Presentations by an authorized administrator of DigitalCommons@UNO. For more information, please contact [unodigitalcommons@unomaha.edu](mailto:unodigitalcommons@unomaha.edu).

# A SPATIAL SIRS BOOLEAN NETWORK MODEL FOR THE SPREAD OF H5N1 AVIAN INFLUENZA VIRUS AMONG POULTRY FARMS

*Alexander Kasyanov<sup>1</sup>, Leona Kirkland<sup>2</sup>, and Mihaela Teodora Matache<sup>3</sup>*

<sup>1</sup>Laboratory for Avian Influenza and Poultry Disease Epidemiology, Federal Centre for Animal Health, 600901 Yur'evets, Vladimir, Russia,

<sup>2</sup>GR Exygnos, Inc., 6426 S 164th Ave., Omaha, NE 68135, USA,

<sup>3</sup>Department of Mathematics, University of Nebraska at Omaha, Omaha, NE 68182, USA  
kasjanov@arriah.ru, lkirkland@exygnos.org, dmatache@mail.unomaha.edu

## ABSTRACT

To predict the spread of Avian Influenza we propose a synchronous Susceptible-Infected-Recovered-Susceptible (SIRS) Boolean network of poultry farms, using probabilistic Boolean rules. Gravity models from transportation theory are used for the probability of infection of a node in one time step, taking into account farm sizes, distances between farms, and mean distance travelled by birds. Basic reproduction numbers are computed analytically and numerically. The dynamics of the network are analyzed and various statistics considered such as number of infected nodes or time until eradication of the epidemic. We conclude that mostly when large farms (eventually) become infected the epidemic is more encompassing, but for a farm that does not have a very large poultry population, the epidemic could be contained.

## 1. INTRODUCTION

The spread of Highly Pathogenic Avian Influenza (HPAI) H5N1 viruses across Asian and European countries has devastated domestic poultry industries. The development of strategies to moderate the spread of influenza among poultry flocks and humans is a top government priority. To investigate the spread of HPAI between poultry farms we propose a Susceptible-Infected-Recovered-Susceptible (SIRS) Boolean network model.

Various individual-based models have been successful in modelling real-world epidemics and understanding mechanisms of epidemic outbreaks [1]. The field of complex networks has now been recognized as an important line of study for epidemiology. For example, Barthélemy et.al. [2], or May and Lloyd [3], have published a number of papers on epidemics in scale-free networks. A large class of physical, biological, chemical networks have been modelled as Boolean networks in recent years (e.g. [4], [5], [6], [7]). General interest in Boolean networks and their applications started much earlier with publications such as the one by Kauffman [8], whose work on the self-organization and adaptation in complex systems has inspired many other research studies.

The spread of HPAI among poultry farms has not yet been investigated in the context of Boolean networks. We

propose a new model in which each farm is considered a node of a Boolean network and can be in one of two states, "infected" or "not infected" by the disease. A node can become infected based on the number of other infected nodes in its neighborhood, their distance from the node under consideration (small distances allow for an easier spread of infection through wild bird or workers interaction of neighboring farms through the common market places), the size of the nodes (large farms have a bigger chance being infected through the synanthropic birds interaction and humans and equipment movement), and the distance travelled by birds in one time step. To define the probability of infection in one time step we use an approach similar to Xia et. al. [9] who have implemented a gravity model from transportation theory [10] to epidemiological coupling and dynamics using a transient force of infection exerted by infecteds in one location on susceptibles in a different location, proportional to the number of susceptibles and the number of infecteds, and inverse proportional to the distance between the locations. This is similar to Newton's gravitational law.

## 2. THE BOOLEAN NETWORK MODEL

In this section we describe the SIRS Boolean model. Consider a network with  $N$  nodes (farms). Each node  $c_n$  can take on two values 0 (not infected) or 1 (infected). The synchronous evolution of the nodes from time  $t$  to time  $t + 1$  is given by a Boolean rule which is considered the same for all nodes, but depends on varying parameters from one node to another. Initially all the nodes are considered susceptible (S). If a node is infected (I), it undergoes a period of cleaning and quarantine during which it could spread the disease to other nodes in its neighborhood; however the force of infection decreases with time, and the node recovers (R) completely eventually. After the quarantine the node becomes again susceptible (S), unless it goes out of business.

Let  $c_n(t)$  be the value of the node  $c_n$  at time  $t$ . Define the Boolean rule

$$c_n(t+1) = X(t) \cdot \chi_{\{0\}}(c_n(t)) + Y(t) \cdot \chi_{\{1\}}(c_n(t)) \quad (1)$$

where  $X(t)$  is a Bernoulli random variable  $X$  with param-

eter  $p_n(t)$  representing the probability that the susceptible node  $c_n$  becomes infected at time  $t$ , and  $Y(t) = 1$  if the node is infectious at time  $t + 1$ , and  $Y(t) = 0$  if the node is noninfectious at time  $t + 1$ . Here  $\chi_{\{a\}}(b) = 1$  if  $a = b$  and zero otherwise. If  $c_n(t)$  becomes 0 at time  $t$  during quarantine, then we set  $Y(t) = 0$  automatically until the end of quarantine. On the other hand, if a node goes out of business after an infection, then  $c_n = 0$  permanently. To define  $p_n(t)$  let  $B_n$  denote the size of the node  $c_n$ , that is  $B_n$  is the number of poultry at location  $c_n$ . Let  $\hat{c}_n$  denote the collection of all the farms in the neighborhood of node  $c_n$  (excluding the node itself). Then  $B(n) = \sum_{c_k \in \hat{c}_n} B_k$  is the total number of poultry in the neighborhood of node  $c_n$ . Let  $d_{nk}$  denote the physical distance between nodes  $c_n$  and  $c_k$ , with  $k \in \{1, 2, 3, \dots, N\}$ . We define the probability  $p_n(t)$  that the node  $c_n$  becomes infected at time  $t$  as follows:

$$p_n(t) = \sum_{c_k \in \hat{c}_n} c_k(t) (B_k/B(n))^\tau \frac{1}{1 + (d_{nk}/d_0)^\rho} f(t). \quad (2)$$

Here  $d_0$  represents the mean distance the infected wild birds are able to cover in one time step. The function  $f(t) \in [0, 1]$  is a random factor that accounts for a reduction of the probability of infection from the infectious node  $c_k$  while cleaning and disinfection take place. The factor  $B_k/B(n) = B_n B_k / \sum_{c_k \in \hat{c}_n} B_n B_k$  is a version of the size terms and  $1/(1 + (d_{nk}/d_0)^\rho)$  is a version of a distance kernel in the gravity model of [9]. Here  $\tau$  determines how the transient “emigration” probability scales with the donor population size, while  $\rho$  quantifies how attraction decays with distance.

In the next sections we analyze the actual network of farms and discuss the parameters of the model. Then we study the evolution of the disease in the network and we compute some related statistics.

### 3. THE NETWORK OF FARMS

Information regarding the poultry farms are taken from the National Agriculture Statistics Service USDA ([www.nass.usda.gov](http://www.nass.usda.gov)) and topographic maps (1 : 100,000 Digital Raster Graphics; Conservation and Survey Division; School of Natural Resources; University of Nebraska-Lincoln). To simulate a network of farms we identify the geographical center of each county and compute the distances between these centers. We approximate each county by a square centered at the county center. In each square we apply a uniform geographical spread of the farms. The size of each farm is obtained as a random number from a Poisson distribution with mean equal to the average number of poultry per farm in each county. In Figure 1, we provide a network of 1198 poultry farms generated as above. This network is used further in the paper. We observe that Butler county accounts for about 63% and Polk county for about 32% of the poultry population of Nebraska.

We provide a boxplot for the node sizes in Figure 2 (a). The frequencies of the distances between nodes are in Figure 2 (b).

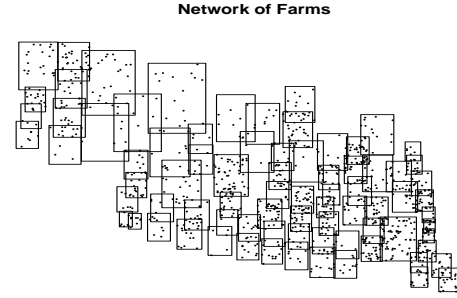


Figure 1. Geographical spread of the network.

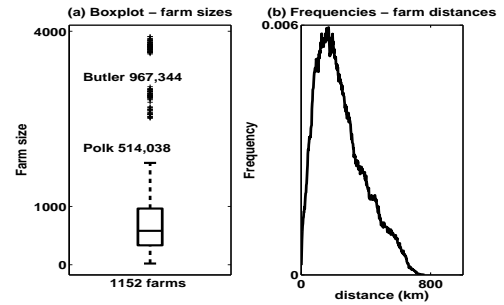


Figure 2. Boxplot of node sizes and distance frequencies. Most of the nodes have a rather small size, except farms in Butler and Polk counties, provided separately.

### 4. THE PARAMETERS OF THE MODEL

Recent studies have shown that the human influenza has an average time interval from infection of one individual to when their contacts are infected of about two days [11]. This number has been used by various authors in assessing the potential impact of a human pandemic of HPAI. At the same time two days is the minimum time needed to get preliminary results back from a diagnostic center for HPAI. Consequently we assume that the basic time step is two days, so the infections within a time step are secondary cases from infected nodes in the previous time step within a neighborhood. The basic neighborhood is considered a circle of radius  $R$  km centered at each node. Given an infected node, the probability that it will infect nodes outside its neighborhood is equal to zero. We will use mostly  $R = 100$  km, but the impact of the value of  $R$  will be considered in the analysis. The parameters  $\tau$  and  $\rho$  will be varied to understand the impact of how the transient “emigration” probability scales with the donor population size, and how attraction decays with distance. We do not possess real data to estimate these parameters. The parameter  $d_0$  is roughly estimated to 3.1 km from available information on home ranges for permanent resident birds and migrating birds of Nebraska. However, due to incomplete data, we believe that this number underestimates the true value of  $d_0$  and therefore we use values of  $d_0 \geq 10$  km.

The government quarantines an infected location for  $Q = 21$  time steps as specified in the USDA national response plan. During this process the disease can still spread to other locations due to migration of synanthropic

birds, rodents, humans and equipment movement, but the probability of infection decreases with time. To account for this, the random factor  $f(t)$  in formula 2 is set equal to 1 during the first time step after infection, and is subsequently given for all nodes by a Beta distribution  $\beta(1, h(T))$  where  $T$  is the number of time steps since the beginning of the quarantine, and  $h(T)$  is an increasing function of  $T$  ( $h(T) = T$  in simulations). We set  $c_n(t) = 0$  after 15 time steps of quarantine. After the quarantine the node re-enters the normal process if the location is repopulated. Small farms are assumed to have a 50% chance of going out of business versus repopulation.

Next we provide a formula for computing the basic reproduction numbers and generate simulations that allow us to understand the impact of a change in parameters on this quantity.

### 5. BASIC REPRODUCTION NUMBERS

Consider now the infection probability given by formula 2, used to compute the basic reproduction numbers, or the average amount of secondary infections generated by a primary infection. We assume that exactly one node, say  $c_K$ , is infected at time  $t = 0$ , that is  $c_K(0) = 1$ . We want to see what is the distribution of the number of infected nodes at time  $t = 1$ .

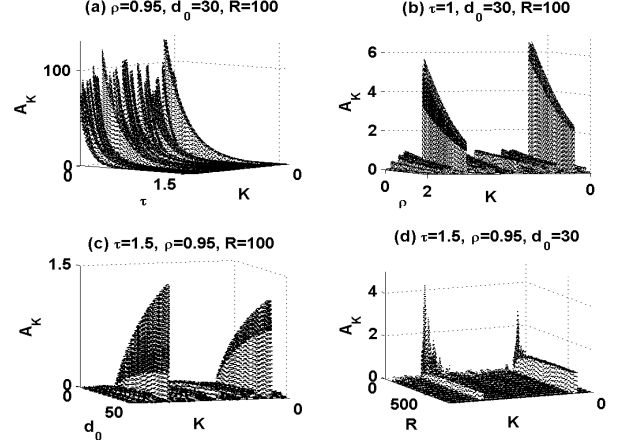
Let  $c_n$  be a node in the neighborhood of node  $c_K$ . Then  $p_n = (B_K/B(n))^\tau / (1 + (d_{nK}/d_0)^\rho)$ . This is the probability that  $c_n(1) = 1$  given that  $c_K(0) = 1$  and  $c_k(0) = 0$ , for all nodes  $c_k, k \neq K$ . Thus, at time  $t = 1$  the number  $m$  of nodes that are turned ON can vary from 0 to the number  $M_K$  of nodes in the neighborhood of node  $c_K$  (not including the node  $c_K$ ). So if  $p_1, p_2, \dots, p_{M_K}$  are the probabilities corresponding to the  $M_K$  nodes of the neighborhood of node  $c_K$ , then the probability  $q_K(m)$  that exactly  $m$  nodes are infected at time  $t = 1$  is given by  $q_K(m) = \sum p_{i_1} p_{i_2} \dots p_{i_m} \prod_j (1 - p_j)$ , where the sum is over all possible combinations of  $m$  nodes out of  $M_K$ ,  $1 \leq i_1 < i_2 < \dots < i_m \leq M_K$ , and  $j = 1, 2, \dots, M_K, j \neq i_l, l = 1, \dots, m$ . Thus the random variable giving the number of nodes that are infected at time  $t = 1$  is:  $P(m \text{ infected nodes}) = q_K(m), m = 0, 1, \dots, M_K$ . Then  $\sum_{m=0}^{M_K} q_K(m) = 1$  by the following result.

**Remark 2.** For any integer  $k > 0$  and any real numbers  $a_1, a_2, \dots, a_k$ , we have that  $\sum_{l=0}^k \sum a_{i_1} a_{i_2} \dots a_{i_l} \cdot \prod_j (1 - a_j) = 1$ , where the sum is over  $1 \leq i_1 < i_2 < \dots < i_l \leq k$ , and  $j = 1, 2, \dots, k, j \neq i_n, n = 1, 2, \dots, l$ . We make the convention that  $l = 0$  means that there is only one term in the inside sum and all the factors of this term are of the type  $(1 - a_j)$ .

**Proof:** The proof is by induction on  $k$ . Let  $S_k$  denote the sum in Remark 2. Clearly  $S_1 = a_1 + (1 - a_1) = 1$ . Also  $S_{k+1} = S_k \cdot a_{k+1} + S_k \cdot (1 - a_{k+1}) = S_k$ .  $\diamond$

Then the average number of infected nodes given only one infected node at time  $t = 0$ ,  $c_K(0) = 1$ , is  $A_K = \sum_{m=0}^{M_K} m \cdot q_K(m)$ .

**Remark 3.** For any integer  $k > 0$  and any real numbers  $a_1, a_2, \dots, a_k$ , we have that  $\sum_{l=0}^k l \sum a_{i_1} a_{i_2} \dots a_{i_l}$ .



**Figure 3.** Plot of the average number of infected nodes in one time step,  $A_K$ , versus  $K$ , the index of the initial infected node. This is done in four different scenarios corresponding to the variation of one of the parameters ( $\tau$ ,  $\rho$ ,  $d_0$ ,  $R$ ) while keeping the other ones fixed as mentioned in the titles of the subplots. Observe that  $A_K$  is decreasing as a function of  $\tau$ ,  $\rho$  or  $R$ , and increasing as a function of  $d_0$ . The two peaks correspond to Butler and Polk counties. The values of  $A_K$  are impacted most dramatically by changes in  $\tau$ . When  $R$  increases there is an approximate threshold value beyond which the neighborhood size makes no difference since far away farms will not be infected.

$\prod_j (1 - a_j) = a_1 + a_2 + \dots + a_k$ , where the sum is over  $1 \leq i_1 < i_2 < \dots < i_l \leq k$ , and  $j = 1, 2, \dots, k, j \neq i_n, n = 1, 2, \dots, l$ .

**Proof:** The proof is by induction on  $k$ . Let  $S_k$  denote the sum in Remark 3. Observe that  $S_1 = 0 \cdot (1 - a_1) + 1 \cdot a_1 = a_1$ . Also  $S_{k+1} = S_k \cdot (1 - a_{k+1}) + S_k \cdot a_{k+1} + a_{k+1} \cdot \sum_{l=0}^k \sum a_{i_1} a_{i_2} \dots a_{i_l} \prod_j (1 - a_j)$  where the second sum is over  $1 \leq i_1 < i_2 < \dots < i_l \leq k$ , and  $j = 1, 2, \dots, k, j \neq i_n, n = 1, 2, \dots, l$ . Thus,  $S_{k+1} = S_k + a_{k+1} \cdot 1 = a_1 + a_2 + \dots + a_{k+1}$ .  $\diamond$

So the average number of nodes infected at time  $t = 1$  or the basic reproduction numbers given  $c_K(0) = 1$  are

$$A_K = p_1 + p_2 + \dots + p_{M_K} \quad K = 1, 2, \dots, N. \quad (3)$$

We graph  $A_K$  versus  $K$  and one other parameter ( $\tau$ ,  $\rho$ ,  $d_0$ , and  $R$  respectively) in Figure 3. A modification of the fixed parameters mentioned in the titles of the plots does not change the shape of the graphs, only the values of  $A_K$ . For example, when  $\tau$  is varied, an increase in the fixed  $\rho$  generates overall smaller values of  $A_K$  due to the fact that the distance kernel in formula 2 decreases.

Now we can focus on one parameter combination and analyze the average number of infected nodes by time steps and time until eradication of the epidemic.

### 6. NETWORK EVOLUTION AND SOME STATISTICS

We set the parameters as follows:  $\tau = 1.5, \rho = 0.95, d_0 = 30$  km, and  $R = 100$  km which yields an average of 195 farms per neighborhood. In the next graph we list the nodes horizontally (in the alphabetic order of the counties) and represent the infected ones by dots. We iterate formula 2 exactly 50 time steps. In Figure 4 we start with one infected node in Butler county and we plot dots for

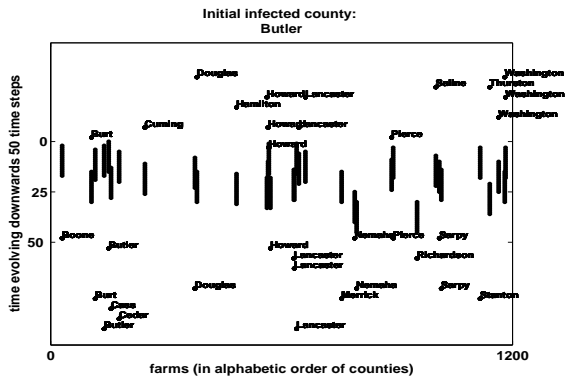


Figure 4. Sample spread of the infection starting with one infected node in Butler county. The infected nodes are listed: Butler, Howard, Lancaster, and Pierce counties, followed by the rest of the listed counties at various times.

all the nodes infected at each time step listing their names. Note that Butler is followed by Howard, Lancaster, Pierce, and then the rest of the listed counties at various times. A total of 34 nodes are infected and the infection is contained during the 50 time steps.

The quantity  $I(t) = \sum_{n=1}^N c_n(t)$  is the number of infected nodes at time  $t$ . We generate frequency plots of  $I(t)$  for  $t = 1, 2, \dots, 60$ , starting with one infected county. For example the first graph of Figure 5 corresponds to the infection of Butler. We observe that small and large values of  $t$  correspond to mostly small values of  $I(t)$ , while for medium values of  $t$  there are higher frequencies of larger values of  $I(t)$ . The epidemic may not be contained. For smaller counties, the plots are concentrated around small values of  $I(t)$  for all  $t$ .

Now consider the time until the eradication of the disease starting with one infection, averaged over multiple sample evolutions. The results are in the second graph of Figure 5. The two peaks are for Butler and Polk counties. The overall network average is about 18 time steps.

We note that the infection of small counties has little impact on the network, unless they are close enough to one of the bigger nodes. When the spread of the disease is more encompassing, the bigger nodes are infected and spread the disease to other nodes faster and throughout a wider area. However, for a medium node the disease could be contained rather fast. On the other hand, it could be that even small nodes spread the disease to bigger nodes and produce an outbreak. However, for most cases the infection spreads to only a few or no other nodes.

## 7. REFERENCES

- [1] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42(4), pp. 599–653, 2000.
- [2] M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, "Dynamical patterns of epidemic outbreaks in complex heterogeneous networks," *Journal of Theoretical Biology*, vol. 235, pp. 275–288, 2005.
- [3] M. E. J. May and A. L. Lloyd, "Infection dynamics on scale-free networks," *Physical Review E*, vol. 64, pp. 066112, 2001.
- [4] R. Albert and A.-L. Barabasi, "Dynamics of complex systems: scaling laws for the period of boolean networks," *Physical Review Letters*, vol. 84(24), pp. 5660–5663, 2000.
- [5] M. T. Matache and J. Heidel, "Asynchronous random boolean network model based on elementary cellular automata rule 126," *Physical Review E*, vol. 71, pp. 026232, 2005.
- [6] C. Goodrich and M. T. Matache, "The stabilizing effect of noise on the dynamics of a boolean network," *Physica A*, vol. 379, pp. 334–356, 2007.
- [7] H. Lahdesmaki, S. Hautaniemi, I. Shmulevich, and O. Yli-Harja, "Relationships between probabilistic boolean networks and dynamic bayesian networks as models of gene regulatory networks," *Signal Processing*, vol. 86(4), pp. 814–834, 2006.
- [8] S. A. Kauffman, *The origins of order*, Oxford University Press, Oxford, 1993.
- [9] Y. Xia, O. N. Bjørnstad, and B. T. Grenfell, "Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics," *The American Naturalist*, vol. 164, pp. 267–281, 2004.
- [10] S. Erlander and N. F. Stewart, *The gravity model in transportation analysis: theory and extensions*, International Science, Netherlands, 1990.
- [11] N. M. Ferguson, D. A. T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, and D. S. Burke, "Strategies for containing an emerging influenza pandemic in southeast asia," *Nature articles*, vol. 437, pp. 209–214, 2005.

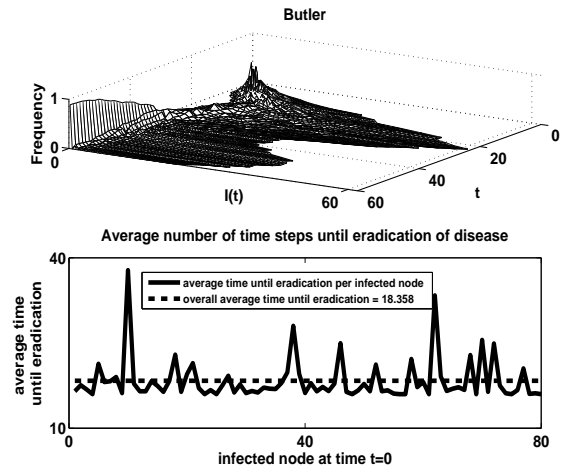


Figure 5. Frequency of  $I(t)$ ,  $t = 1, 2, \dots, 60$  when the initial infected nodes are in Butler county. For small and large values of  $t$  less nodes are infected as expected. Around  $t = 20$  which is the time around which the first infections are cured larger values of  $I(t)$  are more frequent. The second graph represents the average number of time steps for eradication of the network infection versus the initial infected node. The overall average is represented by a horizontal line.