Journal of Theoretical Biology 418 (2017) 8–15

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

An epidemiological model of internet worms with hierarchical dispersal and spatial clustering of hosts

David E. Hiebeler[a,*], Andrew Audibert[a,b], Emma Strubell[c,d], Isaac J. Michaud[a,e]

[a] Department of Mathematics and Statistics, University of Maine, Orono, ME 04469, United States
[b] Bangor High School, Bangor, ME 04401, United States
[c] School of Computing and Information Science, University of Maine, Orono, ME 04469, United States
[d] College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA 01003, United States
[e] Department of Statistics, North Carolina State University, Raleigh, NC 27695, United States

ARTICLE INFO

Keywords:
Network models
Malicious software

ABSTRACT

Beginning in 2001, many instances of malicious software known as Internet worms have been using biological strategies such as hierarchical dispersal to seek out and spread to new susceptible hosts more efficiently. We measured the distribution of potentially susceptible hosts in the space of Internet addresses to determine their clustering. We have used the results to construct a full-size simulated Internet with 2^{32} hosts with mean and variance of susceptible hosts chosen to match our measurements at multiple spatial scales. Epidemiological simulations of outbreaks among the roughly 2.8×10^6 susceptible hosts on this full-sized network show that local preference scanning greatly increases the chances for an infected host to locate and infect other susceptible hosts by a factor of as much as several hundred. However, once deploying this strategy, the overall success of a worm is relatively insensitive to the details of its dispersal strategy over a wide range of parameters. In addition, although using localized interactions may allow malicious software to spread more rapidly or to more hosts on average, it can also lead to increased variability in infection levels among replicate simulations. Using such dispersal strategies may therefore be a high risk, high reward strategy for the authors of such software.

1. Introduction

Malicious software such as worms spreading through computer networks is a routine problem faced by all computer or network users and administrators. Attacks by such malware can cost billions of dollars per incident and spread around the world in a matter of hours or even minutes (Moore et al., 2002, 2003a; Albanese et al., 2004).

Since the emergence of Internet worms in the late 1980s, many attempts have been made to characterize their proliferation. Just a few years after Morris, the first computer worm, infected thousands of university and corporate machines (Eichin and Rochlis, 1989), IBM sponsored a comprehensive study on the dynamics and dangers of these Internet infections. One of the earliest applications of models from biological epidemiology to the spread of Internet worms (Kephart et al., 1993) found that the simplifying assumption of homogeneous contact used in many traditional epidemiological models greatly affected the dynamics of simulated outbreaks. Network topology is a major factor in the spread of worms, whether through peer-to-peer file sharing networks, email contacts, or the IP addressing system. Just as the spatial distribution of a population changes the dynamics of an epidemic (Duryea et al., 1999), so too can the topological distribution of computer hosts on a network have a major impact on the success of an Internet worm. Random graphs were used to simulate the small software sharing networks that served as vectors for early worm infections (Kephart and White, 1991, 1993), and small-world social networks have been used to model the spread of email worms (Moore and Newman, 2000; Zou et al., 2004). Connectivity of some social networks as well as the Internet as a whole has also been described using scale-free networks. Consequently, a great deal of work has been done regarding the spread of disease, both biological and computer-based, on scale-free networks (Pastor-Satorras and Vespignani, 2001, 2004; Albert and Barabási, 2002; Newman, 2002; Balthrop et al., 2004; Zou et al., 2004; Barthélemy et al., 2005; Doyle et al., 2005; Hwang et al., 2005). Others have looked at hierarchical (Wang et al., 2000; Grabowski and Kosiński, 2004) and random (Keeling, 2005) network models with clustering, but not in conjunction with the topology of the Internet Protocol version 4 (IPv4) Internet.

In the past 15 years, Internet worms have begun using biologically-inspired strategies for seeking out hosts to invade (Chen and Ji, 2005; Avlonitis et al., 2007). These tactics exploit inherent vulnerabilities in

* Corresponding author.
E-mail addresses: hiebeler@math.umaine.edu (D.E. Hiebeler).

http://dx.doi.org/10.1016/j.jtbi.2017.01.035
Received 19 August 2016; Received in revised form 20 December 2016; Accepted 19 January 2017
Available online 22 January 2017
0022-5193/

the structure of the Internet to spread more swiftly and efficiently than their predecessors (Zou et al., 2006).

In particular, a technique known as *local preference scanning* takes advantage of the tendency for machines that are nearby in IPv4 address space to make up a homogeneous network maintained by the same administrators and thus share the same software and vulnerabilities. To quickly infect clusters of susceptible hosts, local preference scanning worms choose new targets within the same local network ("subnet") as their infected host machine more often than they choose a machine completely at random. That is, they use localized dispersal when choosing new hosts to contact for attempted infection. This strategy for propagation typically results in persistent and widespread infections while entrenching extensive invasion within protected networks once a single host there has been compromised (Nazario, 2004; Pastor-Satorras and Vespignani, 2004). For example, one version of the Code Red worm infected more than 350,000 computers within 14 h, for a time infecting more than 2,000 new hosts per minute (Moore and Shannon, 2001).

Our study is based on the IPv4 addressing scheme, in which each IP address is a 32-bit number commonly expressed in the form $w.\ x.\ y.\ z$, where each of the four values w, x, y, and z is between 0 and 255 inclusive. (These values are often referred to as *octets*, as they are 8-bit values.) Although there are $256^4 \approx 4.3 \times 10^9$ possible IP addresses, many of them are reserved for various special purposes. The space of IP addresses can be thought of as a $256 \times 256 \times 256 \times 256$ four-dimensional lattice. Note that there is not a one-to-one correspondence between active IP addresses and computer hosts; a single computer host may have multiple IP addresses, and it is very common for a single public IP address to have many computer hosts "behind" it (such as multiple devices in a home sharing a network router with a single publicly-visible IP address).

In summer 2001, Code Red II, hereafter simply Code Red, was the first worm to successfully employ a local preference scanning strategy. Although this paper focuses on the specific strategy used by this worm, many subsequent worms have used similar methods to spread. Code Red's strategy was as follows (Moore et al., 2002). When a computer with IP address $w.\ x.\ y.\ z$ is infected and attempts to contact a new host to attempt infection:

1. With probability 0.375, it performs short-range dispersal, generating random IP addresses of the form $w.\ x.\ r_1.\ r_2$ where r_1 and r_2 are randomly chosen between 0 and 255. These addresses are referred to as being "in the same /16 subnet" as that of the infected host (they share the same first 16 bits with the source address).
2. With probability 0.5, the worm performs medium-range dispersal, generating random IP addresses of the form $w.\ r_1.\ r_2.\ r_3$, with r_1, r_2, and r_3 randomly chosen between 0 and 255. These addresses are referred to as being in the same /8 subnet as that of the infected host (they share the same first 8 bits with the source address).
3. With probability 0.125, it performs long-range dispersal, generating an IP address $r_1.\ r_2.\ r_3.\ r_4$ consisting of four randomly generated values, avoiding addresses of the form $127.\ x.\ y.\ z$ and $224.\ x.\ y.\ z$, corresponding to subnets reserved for loopback and multicast routing, respectively.

In April and May 2004, the Sasser worm utilized a similar dispersal strategy with a slightly different probability distribution. In August 2003, the Blaster worm used a combination of dispersal strategies. It would first use a variation of medium-range dispersal, and then after 20 connection attempts, it would switch to long-range dispersal. Since that time, worms generally use various types of hierarchical dispersal strategies such as the above.

To improve our understanding of how local preference scanning worms spread, we have performed scans of web servers on the Internet to determine the extent to which machines running certain server software and versions are clustered in IPv4 address space. We use that information to partition a simulated population of 2^{32} computer hosts into a hierarchy reflecting the observed structure imposed by IPv4 addressing, and apply state-based epidemiological models to simulate the dynamics of infected hosts among the roughly 2.8×10^6 susceptible hosts in the population.

Different types of worms have been classified based on the way they identify new hosts to infect (Staniford et al., 2002; Weaver, 2002) and their effectiveness assessed (Vogt, 2004; Zou et al., 2006). Worms using simple random scanning have been modeled a great deal due to their simplicity (Chen et al., 2003; Moore et al., 2003b), while less has been done to study other methods of dispersion such as preference scanning. Some have assumed a homogeneous distribution of susceptible hosts throughout the Internet to model local preference scanning worms (Chen et al., 2003), while others developed a model on a network made of smaller networks of varying sizes (Avlonitis et al., 2007), but with no correlation based on IP addresses alone. Because of their lack of an underlying network topology, plant epidemiological models that incorporate multiple dispersal distances (Filipe and Gibson, 1998; Filipe and Maule, 2004) are difficult to apply to preference scanning worms. Here we employ a hierarchical household/community model (Barbour, 1978; Dye and Hasibeder, 1986; Kotliar and Wiens, 1990; Daley and Gani, 1994; Hiebeler et al., 2011) with clustering of susceptible hosts as measured from scans of the Internet.

In this model, the population is completely connected, in the sense that any host can contact any other host. However, local preference scanning leads to variability in the rates of contact among different hosts depending on their location within the hierarchically structured population.

2. Model and measurements

2.1. The model

To apply the household epidemiological model to Internet worms, we partition a population of 2^{32} hosts into hierarchical groups reflecting the structure of IPv4 addressing. Various other topologies, such as physical network connectivity, links between web sites, or social connections among users' e-mail address books are also sometimes imposed by people or by worms and viruses as they spread (Balthrop et al., 2004; Alderson et al., 2005). In our model, we simulate the spread of worms over the IPv4 system, since it was the addressing scheme exploited by Code Red and many subsequent Internet worms, and is still the most prevalent addressing system used.

Although the space of IPv4 addresses may be thought of as a 4-D $256 \times 256 \times 256 \times 256$ integer lattice, Code Red and other worms using similar strategies do not separately utilize the final dimension of the lattice. The population may therefore be thought of as a collection of hosts in a 3-D $256 \times 256 \times 65536$ lattice, each characterized by address of the form $w.\ x.\ y$, where w and x are between 0 and 255, while y is between 0 and $65535 = 256^2 - 1$. We describe the population as containing 256 *neighborhoods* (representing /8 networks), each of which contains 256 *households* (representing /16 networks), each of which contains 65536 individual *hosts*. Here we make the simplifying assumption that a given IP address corresponds to a single potential computer host.

To explore the dynamics of worm proliferation, we use an SIR-V epidemiological model (Hiebeler et al., 2011) in which each host is in one of four states: Susceptible, Infected, Recovered/Removed, or Vaccinated. A susceptible host is one running software that can be exploited by the Internet worm in question; in the case of Code Red, this would be a vulnerable version of Microsoft's Internet Information Services (IIS). Infected hosts are those containing the worm of interest actively trying to infect others. Recovery corresponds to the worm being removed and the security hole patched. Here, once a host is recovered/resistant, it remains so. We also explored an SIRS-V model in which resistant machines slowly transition back to the susceptible

state, representing turnover of machines on the Internet as new (vulnerable) machines come on-line; qualitatively similar results were obtained from that model, and are not reported here. A vaccinated host corresponds to an IP address with a machine running software which cannot be infected by the worm in question (e.g., a different operating system), or which does not contain a machine. Both resistant and vaccinated hosts are immune to infection; the only difference is that a resistant host is one which was previously infected at some time.

In our model, each infected host contacts others and attempts infection at rate ϕ per unit time and recovers at rate μ, according to a Poisson process: the time until the next contact follows an exponential probability distribution with mean $1/\phi$, and similarly for the time until recovery. When contact is made, the target host is chosen at random from the entire Internet with probability α_0, from the same neighborhood (/8 network) as the infected attacker with probability α_1, and from the same household (/16 network) with probability α_2; these are referred to as long-, medium-, and short-distance dispersal respectively. The subscript i on α_i indicates how many leading octets of the originating host's IP address are preserved when generating a new address to contact. When a target is contacted, if it is susceptible, it immediately becomes infectious; otherwise, the contact has no effect.

Ordinary differential equations can be written to describe the proportions of hosts infected within given households, neighborhoods, or the entire population. Here we use deterministic differential equations, implicitly taking the mean over realizations of the stochastic process. Let I_{jk} be the proportion of hosts in household k, neighborhood j (i.e., hosts whose IP addresses begin with the octets $j.\ k$) that are infected, for $j, k \in \{0, 1, 2, \ldots, 255\}$. Similarly, let S_{jk}, R_{jk}, and V_{jk} be the proportions of susceptible, recovered, and vaccinated hosts. Let

$$I_{j\cdot} = \sum_{k=0}^{255} I_{jk}/256 \tag{1}$$

be the proportion of hosts within neighborhood j that are infected, computed by averaging among the 256 households within that neighborhood, and

$$I_{\cdot\cdot} = \sum_{j=0}^{255}\sum_{k=0}^{255} I_{jk}/65536 \tag{2}$$

be the proportion of hosts infected within the entire population, as averaged among all 65536 households. The quantities $S_{j\cdot}$, $S_{\cdot\cdot}$, $R_{j\cdot}$, $R_{\cdot\cdot}$, $V_{j\cdot}$, and $V_{\cdot\cdot}$ are similarly defined. For convenience, $U = S + I + R = 1 - V$ will refer to all unvaccinated hosts combined.

New infections may arise via long-, medium-, or short-distance dispersal. A differential equation describing the rates of these three types of newly arising infections along with the rate of recovery of infected hosts for a particular household is

$$\frac{dI_{jk}}{dt} = \phi\{\alpha_0 I_{\cdot\cdot} S_{jk} + \alpha_1 I_{j\cdot} S_{jk} + \alpha_2 I_{jk} S_{jk}\} - \mu I_{jk}. \tag{3}$$

This differential equation may be averaged across all households to obtain a differential equation describing the dynamics of the infection level among the entire population

$$\frac{dI_{\cdot\cdot}}{dt} = \phi\left\{\alpha_0 I_{\cdot\cdot} S_{\cdot\cdot} + \alpha_1 \sum_{j=0}^{255} I_{j\cdot} S_{j\cdot}/256 + \alpha_2 \sum_{j=0}^{255}\sum_{k=0}^{255} I_{jk} S_{jk}/65536\right\} - \mu I_{\cdot\cdot}. \tag{4}$$

The quantities I_{jk}, $I_{j\cdot}$, and $I_{\cdot\cdot}$ indicate the prevalence of infection at various spatial scales within the network. We also quantify the clustering of infected (or unvaccinated) hosts at two spatial scales, as follows. Following earlier work (Hiebeler, 2006), let $Q_{g,a|b}$ be the probability that if one randomly selects a host in state b from the population, then a second host selected at random within the same group g is in state a, where a and b are one of S, I, R, U, V. The "group" may refer to a neighborhood or household. For example, the neighborhood-level clustering of unvaccinated hosts is given by

$$Q_{n,U|U} = \frac{\sum_{j=0}^{255} (U_{j\cdot})^2}{256 U_{\cdot\cdot}},$$

while the household-level clustering of unvaccinated hosts is

$$Q_{h,U|U} = \frac{\sum_{j=0}^{255}\sum_{k=0}^{255} (U_{jk})^2}{65536 U_{\cdot\cdot}}.$$

The variance of the level of unvaccinated hosts among groups g (neighborhoods or households) may be expressed in terms of the global level and the corresponding clustering parameter, $\mathrm{Var}_g[U] = U_{\cdot\cdot}(Q_{g,U|U} - U_{\cdot\cdot})$. The minimum variance of 0 occurs when the local clustering is the same as the global density, $Q_{g,U|U} = U_{\cdot\cdot}$ (all groups have the same proportion $U_{\cdot\cdot}$ of unvaccinated hosts), while the maximal variance of $U_{\cdot\cdot}(1 - U_{\cdot\cdot})$ occurs when $Q_{g,U|U} = 1$ (fraction $U_{\cdot\cdot}$ of groups are completely unvaccinated, while the rest are completely vaccinated).

The first-generation reproduction ratio $\mathcal{R}_0$ can be calculated for the above model as a function of the dispersal parameters,

$$\mathcal{R}_0 = (\phi/\mu)(\alpha_0 U_{\cdot\cdot} + \alpha_1 Q_{n,U|U} + \alpha_2 Q_{h,U|U}). \tag{5}$$

This quantity gives the expected number of secondary infections caused by a single randomly-chosen unvaccinated host (before it recovers) which has been made infectious in an otherwise-healthy population, taking into account the prevalence and clustering of vaccination. Note that the number of tertiary infections per secondary infection and so on will tend to increase over subsequent generations; details also depend on where the initial infection is released (Hiebeler et al., 2011).

2.2. Measurement

Over the past decade, we have scanned random samples of hosts on the Internet to estimate the prevalence ($U_{\cdot\cdot}$) and clustering ($Q_{n,U|U}$ and $Q_{h,U|U}$) of potentially susceptible hosts. For simplicity in our study, we assume that all servers running the currently-dominant version of IIS are susceptible, although in practice, worms such as Code Red were only able to infect hosts running particular unpatched versions of IIS.

Measurements were taken during January 14–16, 2002, January 15–19, 2005, and January 7–11, 2010. Each time, we generated approximately 10^6 unique random IP addresses, excluding the reserved addresses 127.0.0.0/8, 224.0.0.0/8, 172.16.0.0/12, 10.0.0.0/8, 192.168.0.0/16, and addresses for which the final (fourth) octet is 0 or 255. Note that Code Red only avoided addresses of the first two forms. To keep their code simple and small (which allows them to spread to a new host more quickly and with a smaller chance of their network traffic being detected), many worms use fairly simple strategies for excluding reserved addresses and simply live with the slightly increased failure rate when seeking new hosts to infect. For each address, an attempt was made to contact a web server at that address on HTTP port 80. If contact was established, a GET request was made for the "robots.txt" file, any response was recorded, and the connection closed. We allotted a timeout period of two minutes for the initial contact, and thirty seconds for a response after sending the GET request. When a web server responds to a request, it also provides information about what software and version the server is running; this information was recorded for each web server we found. Although easily spoofed, we assume the number of servers reporting incorrect or no information is relatively small. The most common version of IIS was 4.0 in 2002, 5.0 in 2005, and 6.0 in 2010.

To measure clustering, during each scan, a set of 249 IP addresses of computers running the then-dominant version of IIS was chosen from the computers which responded during the measurement of the random sample above (this number was chosen because in the initial

Table 1
Results of scans to determine prevalence of all web servers and their clustering at various scales. The total numbers of IP addresses scanned were: 992,116 (2002); 992,214 (2005); 992,047 (2010). "Times increase" indicates how much more likely one is to find a web server of any kind when performing a medium-, short-, or very short-distance jump (preserving 1, 2, or 3 leading octets of the IP address respectively) when starting from a web server running the then-dominant version of Microsoft IIS. (The version of IIS on the hosts that served as the starting points for these subsequent localized scans is referred to as the "Source" version.).

			Medium jumps (same /8)		Short jumps (same /16)		Very-short jumps (same /24)	
Year	"Source" version	Servers found out of approx. 10^6 scanned $U_{..}$	Servers found out of 2.49×10^5 $Q_{n,U\|U}$	Times increase $Q_{n,U\|U}/U_{..}$	Servers found out of 2.49×10^5 $Q_{h,U\|U}$	Times increase $Q_{h,U\|U}/U_{..}$	Servers found	Times increase
2002	IIS 4.0	3629=0.3658%	–	–	21769=8.743%	23.90×	–	–
2005	IIS 5.0	5218= 0.5259%	9587=3.850%	7.32×	28522=11.45%	21.77×	1420/ 3735=38.02%	72.29×
2010	IIS 6.0	11964= 1.206%	8231=3.306%	2.74×	33463=13.44%	11.14×	1568/ 4980=31.49%	26.11×

scan in 2002, we found 249 machines running the dominant version of IIS). The version of IIS running on the hosts at these 249 IP addresses is referred to as the "Source" version in Table 1. For each address within the set of 249 initial IP addresses, 1000 new distinct addresses were generated using short-range (same /16 network) dispersal, generating a new sample of 2.49×10^5 addresses which were measured in the same manner as above. Another sample of 2.49×10^5 addresses were then generated using medium-range (same /8) dispersal. In 2002, only short dispersal was used for the second phase of measurements; medium jumps were not performed. In 2005, we additionally scanned 15 addresses in the same /24 network for each of the 249 machines to look at "very short" range dispersal; in 2010, we scanned 20 such additional addresses per host.

Table 1 shows the results of the scans, indicating all hosts that we found running web servers of any kind. Table 2 shows results indicating only hosts running the dominant version of IIS. In 2010, of the 10^6 randomly selected IP addresses scanned, we found 1.196% running web servers, 7.564% of which were some version of IIS. Out of all the machines we scanned, 0.0655% were running IIS version 6.0. From these measurements, we projected that out of the 2^{32} IPv4 addresses, approximately 5.138×10^7 belonged to machines running web servers, 3.887×10^6 of which were running IIS, and 2.813×10^6 IIS 6.0, the most popular version of IIS at the time. As predicted, hosts with more leading octets in common with a host known to have a web server were more likely to also be running a web server with similar software. We used the results from the 2010 scan within our epidemiological simulations, assuming the worst-case scenario that all hosts running IIS 6.0 are unvaccinated (and therefore initially susceptible), namely $U_{..} = 6.55 \times 10^{-4}$, $Q_{n,U|U} = 2.892 \times 10^{-3}$, and $Q_{h,U|U} = 4.539 \times 10^{-2}$.

3. Simulation

3.1. Generating simulated internets

We first generate a 3-D *simulated Internet*, a population of $256^4 \approx 4.29 \times 10^9$ hosts distributed as a set of 256 neighborhoods, each containing 256 households, each with 65536 hosts. This is done by adapting a similar method for generating 2-D spatially structured heterogeneous lattices (Thomson and Ellner, 2003). The previous work on lattices generated distributions of suitable and unsuitable habitat, as follows. First, the desired proportion of sites in the lattice are marked as suitable. Then, a pair of randomly-chosen sites (one suitable, one unsuitable) is chosen. If swapping the states of the two sites will move the measured clustering in the lattice closer to the desired clustering, the swap is performed. The process is then repeated until the desired clustering is reached, or a maximum number of iterations has been exceeded. This method is modified for our hierachically-structured population, with "suitable habitat" now corresponding to "unvaccinated hosts," as described below.

First, the population is initialized so that each household has the same proportion $U_{..} = 6.55 \times 10^{-4}$ of unvaccinated hosts as estimated from our scans. We then cluster the population distribution, first among neighborhoods as follows, and subsequently among households:

1. Choose two neighborhoods at random, where the probability of choosing a neighborhood is proportional to the number of unvaccinated hosts within that neighborhood. Neighborhoods that are either completely unvaccinated or vaccinated are excluded from this choice.
2. Choose a random household within each of those two neighborhoods, where the probability of choosing a household is proportional to the number of unvaccinated hosts within that household.
3. Convert one vaccinated host to an unvaccinated host in the household whose neighborhood was less vaccinated, and convert one unvaccinated host to a vaccinated one in the other household. This increases the neighborhood-level clustering $Q_{n,U|U}$ while maintaining the overall prevalence $U_{..}$ of unvaccinated hosts.
4. Repeat until $Q_{n,U|U}$ has reached or exceeded the desired level as measured from our scans, 2.892×10^{-3}.

This is then similarly performed at the household level, where hosts from two households within the same neighborhood are randomly chosen and adjusted, which maintains the neighborhood clustering

Table 2
Results of scans to determine prevalence of web servers running the dominant version of IIS and their clustering at various scales. Results are as in Table 1, but only servers running the then-dominant version of Microsoft IIS, and therefore potentially vulnerable to a worm targeting that server software, are counted.

			Medium jumps (same /8)		Short jumps (same /16)		Very-short jumps (same /24)	
Year	Server version	Servers found out of approx. 10^6 scanned $U_{..}$	Servers found out of 2.49×10^5 $Q_{n,U|U}$	Times increase $Q_{n,U|U}/U_{..}$	Servers found out of 2.49×10^5 $Q_{h,U|U}$	Times increase $Q_{h,U|U}/U_{..}$	Servers found	Times increase
2002	IIS 4.0	249= 0.0251%	–	–	8747=3.513%	140.0×	–	–
2005	IIS 5.0	558=0.0562%	1691=0.6791%	12.08×	12920=5.189%	92.33×	1122/3735=30.04%	534.5×
2010	IIS 6.0	655=0.0660%	720=0.2892%	4.382×	11301=4.539%	68.77×	1062/4980=21.33%	323.2×

$Q_{n,U\setminus U}$ and prevalence $U_{..}$ but increases the household clustering $Q_{h,U\setminus U}$. This is repeated until $Q_{h,U\setminus U}$ reaches or exceeds the desired value from our measurements, 4.539×10^{-2}.

The above procedure generates a simulated Internet with 2^{32} hosts, with the desired mean and variance of susceptible hosts at the household and neighborhood scales; note however, that third and higher moments of the distribution of susceptible hosts are not specified (and were not measured in our scans). Our simulated Internet has a total of 2,813,203 unvaccinated hosts that are eligible to be actively involved in the epidemiological simulations.

3.2. Worm simulation

When simulating spatial epidemiological models, it is fairly typical to simulate all attempted infection events, i.e., contacts between infected hosts and others, including those events where attempts to infect non-susceptible hosts fail. Because the proportion of unvaccinated hosts is extremely low in our model (≈0.06%) and the total population size so large (2^{32}), this straightforward simulation methodology is too inefficient for practical use here. We therefore used the differential Eqs. (3)–(4) to construct an improved simulation algorithm which only simulates successful infections; failed infection attempts are filtered out of the simulation.

Population-wide rates of successful infections arising from long-, medium-, and short-distance contacts are denoted by r_0, r_1, and r_2, respectively, while the total rate of recoveries is given by r_μ. These rates, described in Table 3, neglect the rates of failed infection attempts, where an infected host contacts another host which is not susceptible. Let $r_T = r_0 + r_1 + r_2 + r_\mu$ be the total rate of all successful events in the SIR-V simulation. The simulation proceeds as follows:

1. The time until the next event is chosen from an exponential distribution with mean $1/r_T$.
2. The type of the next event is chosen at random; the probability that the next event will be of a given type is the ratio of the rate of that event-type to r_T. For example, the probability that the next event is a successful short-distance infection is r_2/r_T.
3. The location of the event is chosen.
   (a) A recovery event occurs in neighborhood j, household k with probability $\mu I_{jk}/r_\mu$.
   (b) A long-distance infection occurs in neighborhood j, household k with probability $S_{jk}/(256^2 S_{..})$.
   (c) A medium-distance infection occurs in neighborhood j with probability $I_{j\cdot}S_{j\cdot}/\sum_{m=0}^{255} I_{m\cdot}S_{m\cdot}$; once the neighborhood has been chosen, the infection occurs specifically within household k in neighborhood j with probability $S_{jk}/(256 S_{j\cdot})$.
   (d) A short-distance infection occurs in neighborhood j with probability $\sum_{\ell=0}^{255} I_{j\ell}S_{j\ell}/\sum_{m=0}^{255}\sum_{k=0}^{255} I_{mk}S_{mk}$; once the neighborhood has been chosen, the infection occurs specifically within household k with probability $I_{jk}S_{jk}/\sum_{\ell=0}^{255} I_{j\ell}S_{j\ell}$.

Table 3
The four event types within the epidemiological model and their population-wide rates. A contact between an infectious host and a non-susceptible host produces no change in the state of the system and is therefore a "wasted" or failed event. By using the probability of infections succeeding with contacts at various scales, the rates of only successful infections may be computed.

Event type	Attempted rate	Successful rate
Long infection	$256^4\phi\alpha_0 I_{..}$	$r_0 := 256^4\phi\alpha_0 I_{..}S_{..}$
Medium infection	$256^4\phi\alpha_1 I_{..}$	$r_1 := 256^3\phi\alpha_1 \sum_{j=0}^{255} I_{j\cdot}S_{j\cdot}$
Short infection	$256^4\phi\alpha_2 I_{..}$	$r_2 := 256^2\phi\alpha_2 \sum_{j=0}^{255}\sum_{k=0}^{255} I_{jk}S_{jk}$
Recovery	$256^4 I_{..}\mu$	$r_\mu := 256^4 I_{..}\mu$

The simulation begins with a single infectious host. However, only infection events occur (i.e., no recoveries) until a specified number of initially infected hosts, N_{I_0}, is achieved. In a real worm outbreak, the individual responsible for first releasing the worm would likely similarly ensure (perhaps via repeated introductions) that the worm had infiltrated some minimum number of hosts before leaving it to proliferate unsupervised. It is also likely that fewer recovery events occur at such an early stage in the epidemic, before anti-malware software has been updated to identify and remove the new threat. The simulation runs until there are no more infectious hosts remaining.

4. Results

Simulations were performed for a variety of dispersal parameters α_0, α_1, and α_2, by letting each parameter vary between 0 and 1 in steps of 1/32, subject to the constraint $\alpha_0 + \alpha_1 + \alpha_2 = 1$. There are 561 such parameter combinations; we ran 100 replicated simulations for each. We used an initial infection size of $N_{I_0} = 100$ infectious hosts. Other parameters used were contact rates of ϕ= 100 and 1850, and recovery rate $\mu = 1$ which can be assumed without loss of generality by rescaling the units of time in the model, as only the ratio ϕ/μ affects qualitative dynamics. The reproduction ratio $\mathcal{R}_0$ from Eq. (5) is minimized when $\alpha_0 = 1$, i.e. pure long-distance dispersal (random scanning) is used. In that case, the critical value $\mathcal{R}_0 = 1$ needed for the infection to avoid immediately decreasing to extinction occurs when $\phi = 1/U_{..} \approx 1527$. We chose our two values of ϕ to be slightly above and well below that value. In practice, the recovery rate μ is likely time-dependent, most likely increasing as both public awareness of the malware and the number of anti-malware software tools targeting it grow; this would be equivalent to decreasing ϕ over time in our model. Our study showed that results varied gradually as ϕ varied over a fairly wide range of values (including several more values explored but not reported here), with no abrubt changes in results.

Three quantities were recorded for each simulation: initial growth rate, final proportion $R_{..}/U_{..}$ of unvaccinated hosts affected by the infection, and household-scale relative clustering $Q_{h,I\setminus I}/I_{..}$ of infected hosts at the time of peak infection, i.e. when $I_{..}$ achieves its maximum. The initial growth rate was measured as the slope of an interpolated least-squares fitted line for the logarithm of the infection level over an initial period of 10 time units, which closer exploration and inspection of simulation results indicated was fairly indicative of initial behavior before density-dependent effects tend to become significant.

Fig. 1 displays the initial growth rate as a function of the dispersal parameters. The location of Code Red's dispersal parameters are highlighted within the figure. For comparison, the instantaneous initial growth rate may be calculated theoretically as follows. Using (1), (2), and (4), note that when the infection is at low density then $U_{jk} \approx S_{jk}$. Given that an infectious host which is randomly located from among all unvaccinated hosts will be in neighborhood j and household k with probability $U_{jk}/(65536 U_{..})$, the instantaneous initial growth rate is

$$\phi(\alpha_0 U_{..} + \alpha_1 Q_{n,U\setminus U} + \alpha_2 Q_{h,U\setminus U}) - \mu. \quad (6)$$

Because $Q_{h,U\setminus U}$ is much larger than $Q_{n,U\setminus U}$ and $U_{..}$ as shown in Table 2, this quantity increases as one approaches $\alpha_2 = 1$, i.e. entirely short-distance dispersal. When measured over an initial time interval rather than instantaneously at t=0, saturation of local networks can begin to play a role; this happens more quickly with larger values of ϕ and when dispersal is more localized.

The final outbreak size $R_{..}/U_{..}$ is shown in Fig. 2; this indicates what proportion of unvaccinated hosts are infected over the entire course of the outbreak. This quantity quickly grows as α_0 and α_2 move away from zero, showing that a little bit of localized contacts and random scanning working together go a long way. Slightly higher infection levels can be reached by including more localized dispersal than Code Red used, but

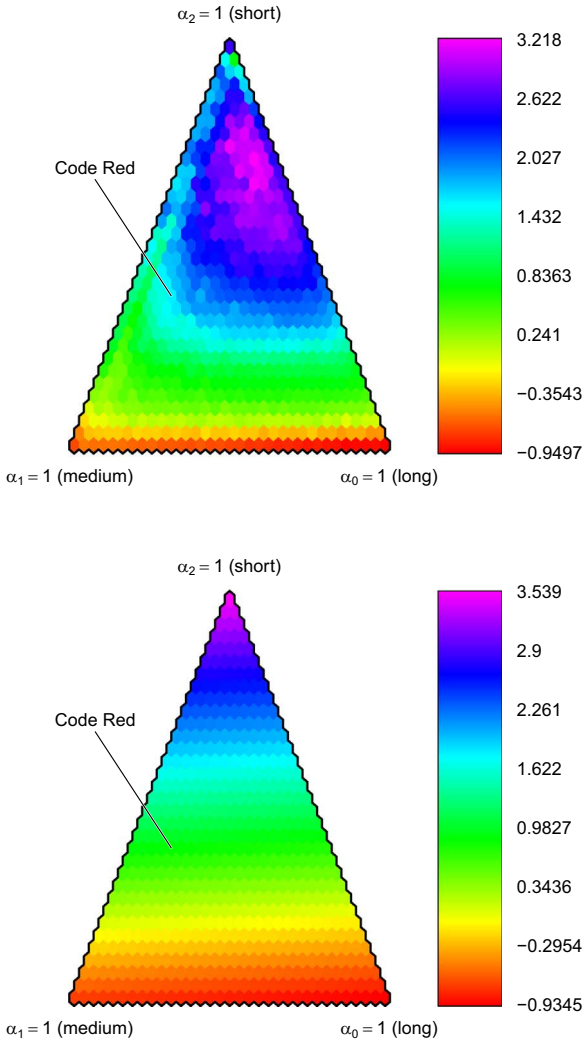

Fig. 1. The initial growth rate for the SIR-V model with $\phi = 100$ is shown as a function of the proportions of contacts which are short, medium, and long distance. Those proportions each varied from 0 to 1 in steps of 1/32, subject to the constraint that they sum to one. Results here are displayed using barycentric coordinates. The location of parameters for Code Red ($\alpha_0 = 0.125$, $\alpha_1 = 0.5$, $\alpha_2 = 0.375$) is highlighted. Top: measurements from stochastic simulations; each hexagon within the triangle displays the mean of 100 replicated simulations. Bottom: theoretical instantaneous growth rate from Eq. (6).

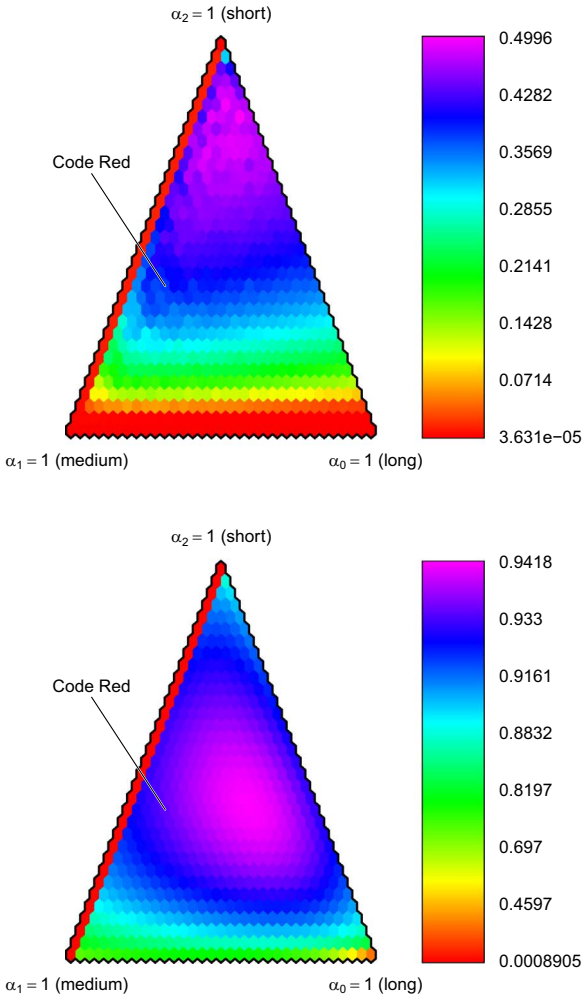

Fig. 2. The rescaled final resistance level (i.e. $R_{..}/U_{..}$, the proportion of unvaccinated hosts that were affected during the entire course of the outbreak) is shown. Top: $\phi = 100$; Bottom: $\phi = 1850$ (note logarithmic scale for color values in bottom image).

the results are not very sensitive to the exact values of α_0, α_1, and α_2 over much of their range.

Fig. 3 displays the household-scale relative clustering of infectious hosts, which indicates how much more likely it is that a second host chosen at random from within the same household as a randomly-chosen infected host will also be infected, as compared with a second host chosen from the entire population. With only short dispersal ($\alpha_2 = 1$), household-scale relative clustering $Q_{h,I|I}/I_{..}$ is approximately 65536 because the infection is contained within one household. With only long-distance dispersal ($\alpha_0 = 1$), we obtain $Q_{h,I|I}/I_{..} \approx Q_{h,U|U}/U_{..} \approx 69$ (from Table 2) for sufficiently large values of ϕ. For smaller values of ϕ, stochastic fluctuations in the model cause the clustering to deviate substantially from this (≈ 697 in Fig. 3).

Figs. 1–3 display sample means among 100 replicate simulations for each combination of parameter values; Fig. 4 displays the standard errors for the sample means of the final outbreak size with $\phi = 100$ from Fig. 2. If there is no long-distance dispersal, the infection never spreads very far; if there is too much, the infection spreads relatively equally among replicated simulations. Variability tends to be highest when there is a small amount of long-distance dispersal. Along the strip of parameters slightly inside the top-left edge of the triangle, variability tends to increase

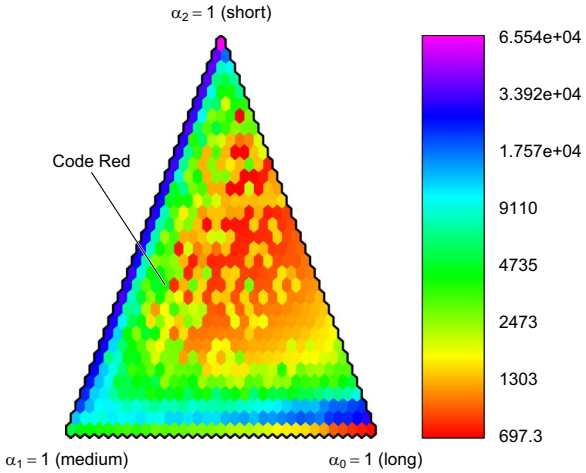

Fig. 3. The relative clustering (logarithmically scaled) of infectious hosts at the household scale $Q_{h,I|I}/I_{..}$ when the system is at its peak infection level, with $\phi = 100$.

as one moves up toward a higher proportion of short-range interactions with less medium-range contact. Localized dispersal is therefore a relatively higher-risk strategy with potentially higher rewards for a worm seeking to maximize its number of hosts affected. Similar results (not shown) were seen in the variability of the initial growth rate.

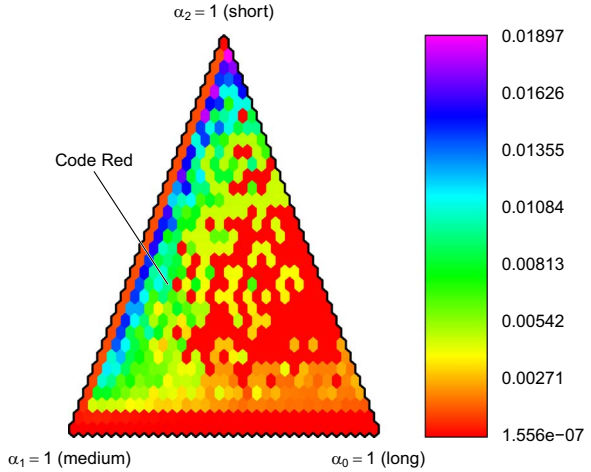

Fig. 4. The standard errors of the final resistance level with $\phi = 100$ whose means are shown in the top of Fig. 2.

5. Discussion

Population models on heterogeneous landscapes in both lattices and community-structured models have shown that spatial clustering of suitable habitat or susceptible hosts strongly impacts the spread of populations or epidemics (Hiebeler, 2004; Hiebeler and Criner, 2007; Hiebeler et al., 2011; Liao et al., 2013a, 2013b). One extreme case of clustered immunity in biological systems is with childhood vaccines, where some communities may have 18 times as many unvaccinated individuals as the population at large (McNeil Jr, 2002). Our measurements show very strong clustering of software platforms among web servers in IPv4 address space, far greater than the clustering of habitat typically observed in biological systems.

Because of the clustering of unvaccinated hosts, localized movement is much more likely to find new hosts to infect (several hundred times more likely, using very-short jumps within the same /24 network). However, this strategy also has several disadvantages. Localized spreading can lead to greater local saturation of the infection, eventually causing infected hosts to waste their time contacting already-infected hosts nearby; this is a likely cause of the increased variability seen among replicated simulations with localized dispersal as seen in Fig. 4. To avoid this problem, some worms, such as Zotob, have begun using adaptive strategies whereby the values of α_0, α_1, and α_2 are modified based on the success or failure of infection attempts so far. Rapid spreading within a local area is also more likely to cause detection. A user walking into a lab with a single malfunctioning machine is likely to simply use another machine; if the entire lab is exhibiting problems, it's more likely to be reported. A local cluster of infected machines will also produce network traffic which is more likely to be detected by administrators than a single infected machine.

To be most successful, some proportion of infection attempts must take place over larger spatial scales in order to colonize new regions of Internet address space, along with local dispersal to exploit those newly colonized regions. Fig. 2 shows that once α_2 increases somewhat modestly, the proportion of the population affected by an outbreak can be quite large relative to its maximum possible value. For example, with $\phi = 1850$, the overall maximum final proportion of unvaccinated hosts affected (R_∞/U_∞) was 0.94 when $\alpha_0 = 3/8$, $\alpha_1 = 7/32$, and $\alpha_2 = 13/32$. With a relatively modest value of $\alpha_2 = 1/16$ (with $\alpha_0 > 0$ to avoid the worm being completely trapped within the region it starts in), the *minimum* final R_∞/U_∞ was approximately 0.86. This is in comparison with a value of only 0.33 with pure long-distance contact.

IPv6 addressing will increase the size of the Internet address space to 2^{128} possible addresses, 2^{96} times the size of IPv4. One often-used example to illustrate the magnitude of this larger address space is the following: assume that the world population is 8×10^9 people, and that there are roughly 10^{14} cells per human body. There would still be more than 10^{14} IPv6 addresses available per cell in every human body. Put another way, suppose each of the 4.3×10^9 possible IPv4 addresses has a single computer host; if we were to place these hosts within the IPv6 address space and release a new worm capable of scanning 10^9 hosts per second (roughly 300×10^6 times faster than Code Red), it would still take roughly 3×10^{12} years to find a host. The distribution of hosts in the address space will therefore be much too sparse for long-distance scanning to have any reasonable hope of success; even preferential random scanning will only be viable if hosts are extremely clustered in IPv6 address space. Worms spreading in IPv6 address space will need to carry information about where they should seek new susceptible hosts, at the expense of making such worms larger and slower to copy. It is likely that IPv4 will remain in widespread use for quite some time, and consequently worm creators will continue to favor the simple and effective scanning methods there over the more elaborate and ambitious techniques that would be necessary to infect machines over IPv6 (Albanese et al., 2004; Bellovin et al., 2006; Chen and Ji, 2007). However, the continued rapid growth of mobile devices such as smartphones with IPv6 addresses may eventually become a lucrative population for malicious software to target.

Our measurements show that the proportion of IP addresses with web servers increased more than threefold over eight years as the Internet grew. Interestingly, the clustering of servers at the household level $Q_{h,U|U}$ did not follow a similar trend, while clustering at the neighborhood level $Q_{n,U|U}$ actually decreased from 2005 to 2010. The fractions of all web servers running the dominant version of IIS were 6.86%, 10.7%, and 5.48% respectively in 2002, 2005, and 2010. The clustering of servers around these hosts running IIS reflects the pattern of deployment of new versions of IIS rolling out, which may vary among different versions. Additional scans have been performed more recently than 2010. However, the methodology used for those scans was quite different, and detailed results of those scans are not included here. But to give one point of comparison with our historical scans, in September 2015, 1.988% of approximately 5.65×10^6 measured IP addresses were running a web server of any kind, with 1.629× and 10.05× increases when using medium and short jumps. Version 7.5 was the most prevalent version of IIS at that time; it was found on 0.0610% of scanned hosts, with 2.796× and 150.5× increases at the medium and short scales.

Although many worms propagate by exploiting vulnerabilities in other software, measuring web servers is a relatively unobtrusive way to obtain a sample which is likely to be qualitatively representative of the distribution of hosts in the Internet. Performing more detailed scans of a given host is more likely to set off intrusion detection systems, causing additional worry and work for the administrators of those machines; we have attempted to make our scanning process minimally disruptive. Also, even if hosts which were actually susceptible to a particular worm were randomly distributed among all hosts running, say, IIS 6.0, such machines would still be highly clustered in the IP address space because of the strong clustering of IIS 6.0 hosts themselves. In fact, unpatched hosts are likely to be even more clustered within IIS 6.0 servers, because machines with similar IP addresses are more likely to be maintained by the same group of people exercising a common set of security practices. The spread of worms could be slowed by randomizing the locations of hosts in IP address space, i.e., by randomly assigning IP addresses to hosts regardless of their location or the organization they belong to. This would make the distribution of susceptible hosts less clustered, i.e., more uniform, among different regions in IP address space. Previous epidemiological models have shown that varying the clustering of unvaccinated hosts among communities within a population can change the proportion of hosts affected by an order of magnitude or more (Hiebeler et al., 2011). Unfortunately, randomizing the assignment of IP addresses would also make it more difficult to administer their allocations among different organizations.

Because the spread of a worm is not uniform across the space of IP addresses, strategies for placing monitors to detect the spread of malicious software may also be designed to take advantage of this heterogeneity. In particular, it would be advantageous to focus monitoring and detection efforts in those regions of the Internet that contain more susceptible hosts, as worms are both more likely to appear in those regions, and also able to spread more quickly (and do more damage) once there.

Future work with these models will explore the effects of worms such as Zotob using adaptive strategies to adjust dispersal parameters over time, and allowing for a time-varying recovery rate μ. Initial measurements taken during the spread of Code Red I showed a contact rate orders of magnitude larger than those studied here, but the rate surely decreased as awareness of the worm spread. The recovery rate may in fact be tied to prevalence of infections at local and global spatial scales, as media coverage of an outbreak will affect people's tendency to employ protective measures.

In conclusion, we have measured the prevalence and clustering of web servers in IPv4 address space over the past decade, showing a growing number of web servers and quantifying their clustering. Assuming a worst-case scenario that all web servers running IIS 6.0 are vulnerable, we constructed an artificial network with the same mean and variance of susceptible machines as seen in our measurements. Simulations of a full-size IPv4 address space (2^{32} hosts) with approximately 2.8×10^6 susceptible hosts showed that the local preference scanning strategies used by Code Red II and later worms greatly increases the ability of infected hosts to successfully locate additional susceptible hosts to infect. Highly localized scanning can however lead to greater variability in the success of a worm, and may increase chances of detection. We hope this paper serves as an introduction to quantify some of the well-known clustering of hosts in IPv4 address space, and that improved measurements and simulations may guide placement of monitoring systems to increase chances of early detection of intrusions.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant no. DMS-0746603 to D.E.H.

References

Albanese, D.J., Wiacek, M.J., Salter, C.M., Six, J.A., 2004. The case for using layered defenses to stop worms. Network Architecture and Applications Division of the Systems and Network Attack Center (SNAC). Information Assurance Directorate, U. S. National Security Agency; http://www.nsa.gov/snac/support/WORMPAPER.pdf).

Albert, R., Barabási, A.-L., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74 (1), 47–97.

Alderson, D., Li, L., Willinger, W., Doyle, J.C., 2005. Understanding Internet topology: principles, models, and validation. IEEE/ACM Trans. Netw. 13 (6), 1205–1218.

Avlonitis, M., Magkos, E., Stefanidakis, M., Chrissikopoulos, V., 2007. A spatial stochastic model for worm propagation: scale effects. J. Comput. Virol. 3 (2), 87–92.

Balthrop, J., Forrest, S., Newman, M., Williamson, M.M., 2004. Technological networks and the spread of computer viruses. Science 304 (April), 527–529.

Barbour, A.D., 1978. Macdonald's model and the transmission of bilharzia. Trans. R. Soc. Trop. Med. Hyg. 72 (1), 6–15.

Barthélemy, M., Barrat, A., Pastor-Satorras, R., Vespignani, A., 2005. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. J. Theor. Biol. 235, 275–288.

Bellovin, S., Cheswick, B., Keromytis, A., 2006. Worm propagation strategies in an IPv6 Internet. Login 31 (1), 70–76.

Chen, Z., Ji, C., 2005. Spatial-temporal modeling of malware propagation in networks. IEEE Trans. Neural Netw. 16 (5), 1291–1303.

Chen, Z., Ji, C., 2007. Measuring Network-Aware Worm Spreading Ability. Vol. 26 of IEEE International Conference on Computer Communications (INFOCOM 2007) May, pp. 116–124.

Chen, Z., Gao, L., Kwiat, K., 2003. Modeling the spread of active worms. In: Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03), San Francisco, California, USA, March 2003. URL ⟨http://www.ieee-infocom.org/2003/papers/46_03.PDF⟩

Daley, D., Gani, J., 1994. A deterministic general epidemic model in a stratified population. In: Kelly, F. (Ed.), Probability, Statistics, and Optimisation. John Wiley & Sons, Ltd., New York, NY, Ch. 8, pp. 117–132.

Doyle, J.C., Alderson, D.L., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R., Willinger, W., 2005. The robust yet fragile nature of the Internet. Proc. Natl. Acad. Sci. USA 102 (41), 14497–14502.

Duryea, M., Caraco, T., Gardner, G., Maniatty, W., Szymanski, B.K., 1999. Population dispersion and equilibrium infection frequency in a spatial epidemic. Physica D 132, 511–519.

Dye, C., Hasibeder, G., 1986. Population dynamics of mosquito-borne disease: effects of flies which bite some people more frequently than others. Trans. R. Soc. Trop. Med. Hyg. 80, 69–77.

Eichin, M.W., Rochlis, J.A., 1989. With Microscope and Tweezers: An Analysis of the Internet Virus of November 1988. In: Proceedings of the 1989 IEEE Symposium of Research in Security and Privacy. Oakland, CA, May, pp. 326–343.

Filipe, J., Gibson, G., 1998. Studying and approximating spatio-temporal models for epidemic spread and control. Philos. Trans. R. Soc. Lond. B 353, 2153–2162.

Filipe, J., Maule, M., 2004. Effects of dispersal mechanisms on spatio-temporal development of epidemics. J. Theor. Biol. 226, 125–141.

Grabowski, A., Kosiński, R., 2004. Epidemic spreading in a hierarchical social network. Phys. Rev. E 70, 031908.

Hiebeler, D.E., Criner, A.K., 2007. Partially mixed household epidemiological model with clustered resistant individuals. Phys. Rev. E 75, 022901.

Hiebeler, D.E., Michaud, I.J., Ackerman, H.H., Iosevich, S.R., Robinson, A., 2011. Multigeneration reproduction ratios and the effects of clustered unvaccinated individuals on epidemic outbreak. Bull. Math. Biol. 73 (12), 3047–3070.

Hiebeler, D.E., 2004. Competition between near and far dispersers in spatially structured habitats. Theor. Popul. Biol. 66 (3), 205–218.

Hiebeler, D.E., 2006. Moment equations and dynamics of a household SIS epidemiological model. Bull. Math. Biol. 68 (6), 1315–1333.

Hwang, D.-U., Boccaletti, S., Moreno, Y., López-Ruiz, R., 2005. Thresholds for epidemic outbreaks in finite scale-free networks. Math. Biosci. Eng. 2 (April (2)), 317–327.

Keeling, M., 2005. The implications of network structure for epidemic dynamics. Theor. Popul. Biol. 67, 1–8.

Kephart, J.O., White, S.R., 1991. Directed-graph epidemiological models of computer viruses. In: Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy. Oakland, CA, pp. 343–359.

Kephart, J.O., White, S.R., 1993. Measuring and modeling computer virus prevalence. IEEE Symposium on Research in Security and Privacy. Oakland, CA, May, pp. 2–15.

Kephart, J.O., Chess, D.M., White, S.R., 1993. Computers and epidemiology. IEEE Spectr., 20–26.

Kotliar, N.B., Wiens, J.A., 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. Oikos 59, 253–260.

Liao, J., Li, Z., Hiebeler, D.E., El-Bana, M., Deckmyn, G., Nijs, I., 2013a. Modelling plant population size and extinction thresholds from habitat loss and habitat fragmentation: effects of neighbouring competition and dispersal strategy. Ecol. Model. 268, 9–17.

Liao, J., Li, Z., Hiebeler, D.E., Iwasa, Y., Bogaert, J., Nijs, I., 2013b. Species persistence in landscapes with spatial variation in habitat quality: a pair approximation model. J. Theor. Biol. 335, 22–30.

McNeil Jr., D. G., 2002. When parents say no to child vaccinations. The New York Times. Nov. URL ⟨http://www.nytimes.com/2002/11/30/us/when-parents-say-no-to-child-vaccinations.html⟩

Moore, C., Newman, M., 2000. Epidemics and percolation in small-world networks. Phys. Rev. E 61 (5), 5678–5682.

Moore, D., Shannon, C., 2001. The spread of the Code-Red worm (CRv2). The Cooperative Association for Internet Data Analysis. URL ⟨http://www.caida.org/research/security/code-red/coderedv2_analysis.xml⟩

Moore, D., Shannon, C., Brown, J., 2002. Code-Red: A case study on the spread and victims of an Internet worm. In: Proceedings of the Second ACM SIGCOMM Workshop on Internet Measurement. New York, NY, pp. 273–284.

Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, S., Weaver, N., 2003a. Inside the Slammer worm. IEEE Secur. Priv. 1 (4), 33–39.

Moore, D., Shannon, C., Voelker, G., Savage, S., 2003b. Internet Quarantine: Requirements for Containing Self-Propagating Code. In: Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03). San Francisco, CA, pp. 1901–1910.

Nazario, J., 2004. Defense and Detection Strategies Against Internet Worms. Artech House, Boston.

Newman, M., 2002. The spread of epidemic disease on networks. Phys. Rev. E 66, 016128.

Pastor-Satorras, R., Vespignani, A., 2001. Epidemic dynamics and endemic states in complex networks. Phys. Rev. E 63, 066117.

Pastor-Satorras, R., Vespignani, A., 2004. Evolution and Structure of the Internet. Cambridge University Press, Cambridge, United Kingdom.

Staniford, S., Paxson, V., Weaver, N., 2002. How to 0wn the Internet in Your Spare Time. In: Proceedings of the 11th USENIX Security Symposium. Berkeley, CA, pp. 149–167.

Thomson, N.A., Ellner, S.P., 2003. Pair-edge approximation for heterogeneous lattice population models. Theor. Popul. Biol. 64, 271–280.

Vogt, T., 2004. Simulating and Optimising Worm Propagation Algorithms, February. URL ⟨http://web.lemuria.org/security/WormPropagation.pdf⟩

Wang, C., Knight, J., Elder, M., 2000. On computer viral infection and the effect of immunization. Vol. 16 of IEEE Computer Security Applications Conference (ACSAC'00). New Orleans, LA, pp. 246–256.

Weaver, N., 2002. Potential Strategies for High Speed Active Worms: A Worst Case Analysis, March. URL ⟨www.cgisecurity.com/lib/worms.pdf⟩

Zou, C.C., Towsley, D., Gong, W., October 2004. Email Worm Modeling and Defense. In: Proceedings of the 13th International Conference on Computer Communications and Networks (ICCCN'04). Chicago, IL, pp. 409–414.

Zou, C.C., Towsley, D., Gong, W., 2006. On the performance of Internet worm scanning strategies. Perform. Eval. 63 (7), 700–723.