SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks 2013; **6**:78–88
Published online 21 March 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/sec.534

RESEARCH ARTICLE

Toward early warning against Internet worms based on critical-sized networks

Emmanouil Magkos[1]*, Markos Avlonitis[1], Panayiotis Kotzanikolaou[2] and Michalis Stefanidakis[1]

[1] Department of Informatics, Ionian University, Plateia Tsirigoti 7, Kerkyra, 49100, Greece
[2] Department of Informatics, University of Piraeus, 80, Karaoli-Dimitriou, 18534, Piraeus, Greece

ABSTRACT

In this paper, we build on a recent worm propagation stochastic model, in which random effects during worm spreading were modeled by means of a stochastic differential equation. On the basis of this model, we introduce the notion of the *critical size* of a network, which is the least size of a network that needs to be monitored, in order to correctly project the behavior of a worm in substantially larger networks. We provide a method for the theoretical estimation of the critical size of a network in respect to a worm with specific characteristics. Our motivation is the requirement in real systems to balance the needs for accuracy (i.e., monitoring a network of a sufficient size in order to reduce false alarms) and performance (i.e., monitoring a small-scale network to reduce complexity). In addition, we run simulation experiments in order to experimentally validate our arguments. Finally, based on notion of critical-sized networks, we propose a logical framework for a distributed early warning system against unknown and fast-spreading worms. In the proposed framework, propagation parameters of an early detected worm are estimated in real time by studying a critical-sized network. In this way, security is enhanced as estimations generated by a critical-sized network may help large-scale networks to respond faster to new worm threats. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS
monitoring worm propagation; stochastic models; critical-size networks; early warning

***Correspondence**
Emmanouil Magkos, Department of Informatics, Ionian University, Plateia Tsirigoti 7, Kerkyra, 49100, Greece.
E-mail: emagos@ionio.gr

1. INTRODUCTION

Computer worms are autonomous programs that spread across a network by exploiting existing security vulnerabilities of interconnected computers. Scanning worms search for their targets by scanning target port(s) of other nodes in order to locate software applications with specific vulnerabilities. Worms can self-propagate and pollute a large portion of a network in a short period because of the relatively homogeneous software base and the high bandwidth connectivity in the Internet [1]. Depending on their strategy, scanning worms can also be seen as random, local preference, sequential, or topological scanning worms [2,3].

Whereas, nowadays, classical scanning worms represent a very small percentage of malware creation;[†] a new generation of advanced malware with self-propagating characteristics will soon constitute a real, severe threat to computer networks and critical infrastructures. For example, the recent Stuxnet worm [4,5], which attempted to seize control of industrial control systems, may be the first in a long line of highly selective, self-propagating malware [6,7] that are expected to emerge in the near future. Such new generation of smart worms may also be able to traverse nonInternet-connected systems: for example, the Stuxnet worm used local preference strategies and peer-to-peer (P2P) networking techniques to send instructions to infected machines that were not connected to the Internet. The threat of a future, advanced scanning strategy has also been studied in the literature, under the names of hitlist worms [3], routing worms [8,9] and importance scanning worms [10], permutation [11,3], or divide-conquer worms [9]. These are selective worms that spread faster by carefully selecting their victims instead of 'blindly' scanning the universe for possible targets [9,8,3]. Envisaged worms, such as the flash [3] or complete-scan [9] worms, could theoretically infect the entire vulnerable population within seconds.

[†] Compared to 10 years ago, individual, stand-alone malware (i.e., *worms* and *viruses*) do not rank as the number one threat in computer epidemics [54,4]. The main trend in malware propagation involves arbitrary code execution supported by Botnet infrastructures that control hundreds of thousands of hosts in order to generate high financial profits to their owners [55].

Worm (or in general, self-spreading malware) propagation is in fact a *stochastic process*, as random effects are present in real networks. This source of randomness takes root in the various nonuniform parameters that influence malware propagation [3,8,10,12–17]. Most of these parameters can be categorized as malware-related (e.g., scanning strategy, scan rate, IP random address selection, congestion), network-related (e.g., network bandwidth, traffic, topology), system-related (e.g., vulnerable hosts distribution, initially infected hosts), policy-related (e.g., network or host-level firewall policies, intrusion prevention, automatic quarantine), and human-related (e.g., removal tools, vulnerability patching, disconnecting or isolating hosts, blocking access to a service, operating system updating or restoring, training users, user awareness). Novel characteristics of current and future self-propagating malware, such as high stealthiness, polymorphism and context awareness, increase the inherent complexity of the propagation process. Sometimes, there are interdependencies among some of the above factors. For example, the worm's scan rate may be affected by the available bandwidth and the traffic created by the worm itself [12] or even by the delays in Domain Name System (DNS) replies [17]. Or, the distribution of the vulnerable hosts may depend on the security policies. Finally, the propagation of the next generation malware will be influenced by the specific characteristics of an underlying wireless communication infrastructure (e.g., wireless range, congestion, mobility of nodes), thus, increasing the overall complexity.

Mathematical models can help the security research community to understand the threat and study the propagation pattern during the lifetime of a worm [12,18,19]. For example, an analytical model can provide numerical solutions that explain the evolution of the worm's population, provide patterns for accurate prediction and damage assessment for future worm threats, and test new models for containment and disinfection of worms [3,18]. Recent research also suggests that, by analyzing a worm's behavior, we may have insights into effectively detecting and containing a fast-spreading worm [18,20]. For example, during the spread of a worm, a propagation model can be used to individuate and describe symptoms of worm activity, thus, providing useful data to an early detection system [20]. The extracted knowledge could also be used to trigger emergency response, for example, an automatic containment policy. A challenge for the propagation models is to take into account most of the above parameters, which affect the propagation rate of a worm.

Another challenge for computer epidemiology and security research is whether the monitoring of worm propagation within a small-scale network can accurately project its growth rate in networks of a larger scale (i.e., the Internet). We will call this process *worm projection*. The generalization of the results obtained by monitoring a small fraction of the Internet is subject to controversy [14,21]. On one hand, worm projection in small scales is difficult because of the heterogeneities of the various small-sized networks that comprise the Internet, as well as because of the non-uniformities of the worm's scanning strategies. As a result, under a pragmatic theoresis, worm projection cannot be successful within a single (uniform) or small-scale network, whereas a sufficiently large scale will diffuse the network heterogeneities and better describe the phenomenon. On the other hand, the cost and complexity of monitoring a very large network is high (e.g., the processing cost of realtime traffic analysers or for deep scanning of a large number of monitored network packets is not negligible—see, for example, the analysis in [22]). In [9], the performance of a detection system is related to the minimum size of a detection network needed to ensure that a worm is detected within a certain time. Intuitively, a small network size reduces the time for early detection but increases false alarms.

1.1. Our contribution

We build on a recent model [23] which describes the propagation of fast-spreading, random-scanning worms in the Internet, where random effects in worm-spreading velocity were modeled by means of a stochastic differential equation. In this paper, we validate the robustness of the model proposed in [23] and introduce the notion of the *critical size* of a network, which is the minimum size of a network that is sufficient to monitor, in order to accurately project the growth of such worms in larger networks. Furthermore, we provide a theoretical estimation of the critical size of a network, which is intrinsically linked to the characteristics of a specific worm. We run simulation experiments that validate our arguments. Then, we exploit this knowledge to design a framework for a distributed early warning system against fast-spreading malware. In our view, this means finding a network of critical size, over which monitoring worm propagation will accurately project the spread of the worm in larger networks. Our envisaged system involves a coalition of network domains of variable size and characteristics that cooperatively estimate the critical network size. We argue that networks of this scale can correctly estimate and disseminate, in a timely manner, the infection parameters for an early detected worm. In this way, security is enhanced as estimations of worm propagation generated by a critical-sized network may help larger networks respond faster to new worm threats.

2. RELATED WORK

Epidemiologic models for analyzing the spread of computer malware are not new [24]. Early attempts [12,3] that capture the strategy of random-scanning worms use the simple epidemic model [13] to study the initial part of worm spreading, where factors such as human countermeasures and congestions do not affect the worm propagation. In recent years, a number of deterministic models were designed to consider the parameters that affect the scanning worm propagation for random scanning (e.g., [12,3,18,25,16]), local preference (e.g., [14,19,15]), or other advanced scanning strategies (e.g., [9,8,19,17,10]). For example, the two-factor model in [12] takes into account the congestion caused by the worm scan packets, as well as the reactive (human) countermeasures that turn infected or susceptible nodes into a recovered state. Models that consider the preventive measures (e.g., antivirus and patch management [26]), the link bandwidth between systems [18,25,16], the network topology [27], the slow down caused by automatic treatment and containment measures [28,1,14,29,30], and the infection delay and user vigilance [31] have also been proposed in the literature.

Because of the observed randomness affecting worm propagation in the Internet (e.g., [13,1]), the so called *stochastic models* have also been emerged (e.g., [32,33]). These models, contrary to deterministic models that express a mean-field behavior, are based on the observation that worm propagation is an inherently random process. In the models of [32,33], randomness emerges because of the scanning strategy, whereas other sources of randomness, for example bandwidth limitation or network topology, are not covered. The above approaches propose discrete Markov models in order to predict propagation at early stages, introducing, as an appropriate variable, the amount of time for the next infection, as well as estimating its mean value and variance in order to construct robust detection protocols.

The problem of Internet modeling and analysis has been studied in various contexts [34] and is particularly challenging in monitoring and detecting worm propagation [14,16]. Normally, a detection architecture requires a large number of monitored networks to distinguish scanning worms from other activities. Staniford *et al.* [3] proposed the idea of a ‘Center for Disease Control’ that collects worm-related information from a very large number of monitors. The Honeynet Project [35], CAIDA [36], and the Internet Storm Center (DShield) [37] also use the same approach. CAIDA used one /8 network and two /16 networks to monitor the spread of the Code Red v2 worm [38]. Moore’s network telescopes [39] monitor a relatively large fraction of the IP space to project the spread of a worm over the global Internet. In the literature, there have also been attempts to use smaller scale architecture to monitor and early detect the propagation of a worm (e.g., [9,20,14,16]). In [20], for example, a set of distributed monitors offer observation data on a worm’s activities to a malware-warning center for early detection. In [14], a /8 network size is seen as sufficient to characterize and monitor the spread of scanning worms. In the distributed framework of [40], software agents are placed in ‘many’ local machines; they monitor for suspected executables and send CRC reports to a central server which decides whether executables behave as worms [40]. Most of the previous approaches require monitoring a large network. Furthermore, they are static, in the sense that the size of the network that will be monitored is predetermined.

Especially for fast-spreading worms, a challenge for security research is the high number of false alarms in anomaly-based intrusion detection systems [1,20,41]. The problem becomes worse given the threat of stealth worms and polymorphism where worms change behavior to evade detection and containment [42]. Furthermore, during the outbreak of a fast scanning worm, the difficulty of in-time human countermeasures has been pointed out in the literature [1,3]. In several circumstances, automatic containment measures [43,28], when paired with early detection, can slow down the worm infection [1,44]. Moore *et al.* [1] studied the challenges and effectiveness of automated containment for fast-spreading worms and presented a deployment scenario for distributed containment. In the system of [28], suspiciously behaving hosts are quarantined for a fixed time interval. In [30], early detection is incorporated with automatic containment that is based on local victim information. In [45], a host-based automatic containment system, destined for random or preferential worms, is based on the strategy of limiting the number of scans to unique IP addresses.

3. A STOCHASTIC MODEL FOR WORM PROPAGATION

3.1. The stochastic differential model of Avlonitis *et al.* [23]

Let us assume a random-scanning worm that propagates over a network with N unique hosts (N interpreting the entire IP space scanned by the worm), where $N_s \leq N$ of these addresses could potentially become infected by the worm. In an arbitrary ensemble of hosts (i.e., a subnet) and for the simplest case (e.g., recovery and/or removal of hosts are not taken into account), the population N_s is split into infected and susceptible subpopulations, represented by *I(t)* and *S(t)*, respectively. The classical epidemic model can be expressed with the following ordinary differential equation.

$$\frac{dI(t)}{dt} = \frac{\beta}{N} S(t) I(t) \tag{1}$$

or

$$\frac{dI(t)}{dt} = f(I(t)) \tag{2}$$

where β is the constant scan rate and $f(I(t)) = \frac{\beta}{N}(N_s - \mathrm{I}(t))I(t)$ is the spreading ‘force’ over the scale of the entire network (i.e., the Internet).

A stochastic differential model for random-scanning worms was proposed in [23]. In [23], the scale over which the worm propagates was seen as a crucial factor, and it was argued that models that refer to different scales predefine the nature of the variables and parameters of the problem of worm propagation. Specifically, if a model tries to describe the behavior of the worm propagation in microscale (i.e., very few number of hosts), then a probabilistic model is the only choice, and $S(t), I(t)$ are interpreted as random variables (e.g., [32,33]). On the other hand, if a model is referred to the macroscopic behavior of a worm in the Internet (i.e., the macroscale), deterministic models are more appropriate, and $S(t), I(t)$ are interpreted as deterministic variables, (e.g., [24,12,3]). The link between these models is an approach that is able to describe worm propagation in the *mesoscale*, which is an appropriate scale for real-world monitoring systems.

Following the line of reasoning of Avlonitis *et al.* [23], in mesoscale, the population variables $S(t), I(t)$ are interpreted as stochastic variables. Moreover, the infection parameter β is also assumed stochastic: here, β models the total infection rate, incorporating randomness that is due to either choices/decisions made by the worm itself (e.g., scanning strategy, scan rate, random IP address selection) or changes in the network environment within which the worm evolves (e.g., bandwidth, congestion etc). β is fluctuating around a mean value $< \beta >$, where the corresponding noise is assumed to be the well-known limit of the white noise, that is,

$$\beta = < \beta > + \delta\beta \tag{3}$$

where

$$\delta\beta = \dot{w}, < \delta\beta > = 0 \tag{4}$$

and

$$< \delta\beta_i \cdot \delta\beta_j > = \sigma^2 \delta_{ij} \tag{5}$$

where σ^2 is the amplitude of the white noise $\dot{w}$.

Substituting the infection parameter β in the corresponding evolution equation (Equation (1)), a random fluctuating part δf of the spreading 'force' is obtained,

$$\frac{dI}{dt} = f_{<\beta>}(I) + \delta f \tag{6}$$

where

$$f_{<\beta>}(I) = \frac{< \beta >}{N}(N_s - I) \cdot I \tag{7}$$

and

$$\delta f = \frac{1}{N}(N_s - I) \cdot I \cdot \dot{w} \tag{8}$$

Equation (6) is the evolution equation describing the dynamics of worm propagation. It belongs to a general class of stochastic differential equations, being able to describe, with success, the evolution of dynamical systems in the mesoscale (e.g., [46–48]).

In this form, the stochastic differential model provides a quantitative estimate for the inherent randomness. Indeed, according to Equation (6), a measure of the resulting noise is

$$Q = \sqrt{< \delta f^2 >} \tag{9}$$

or [23]

$$Q(I) = \frac{\sigma}{N}(N_s - < I >) \cdot < I > \tag{10}$$

It is important to note that Equation (10) quantifies the inherent randomness of worm propagation emerged in real networks in terms of the variables and parameters of the problem.

3.2. Validation of Avlonitis *et al.* [23]—bridging the scales

It is emphasized that the stochastic equation proposed in Equation (6) is referred to a scale the size of which defines the mesoscale for the problem of worm propagation. We will validate the robustness of the proposed model by showing that it reproduces the classical epidemic model when averaging over the Internet is performed (up-scaling). Indeed, in order to find the propagation rate *dI/dt* in the macroscale, the stochastic differential equation (Equation (6)), must be averaged over the Internet, that is,

$$< dI/dt > = < \frac{< \beta >}{N}(N_s - I)I + Q(I)\dot{w} > \tag{11}$$

or

$$< dI/dt > = \frac{< \beta >}{N} < (N_s - I)I > \tag{12}$$

where, making use of Equations (4) and (5) and neglecting higher orders, the average value of the last term vanishes. The following evolution equation for the macroscale holds

$$d < I > /dt = \frac{< \beta >}{N}(N_s - < I >) < I > \tag{13}$$

As a result, when up-scaling is performed, the proposed stochastic differential equation coincides with the classical epidemic model, where the stochastic variable of the infected hosts in the mesoscale is replaced by the average number of infected hosts, for example, a deterministic variable in the macroscale. This result confirms the robustness of the model proposed in [23]. It also confirms the proposition that the nature of the variable of the infected population is determined from the scale over which the model is applied. We believe that the proposed formalism provides the missing link between the stochastic and deterministic propagation models found in literature, bridging together the different scales of observation.

At this point, the limitations of the adopted models emerge (as generally in nature or in artificial systems): when trying to describe systems in small scales, taking into account the inherent randomness, the exact values of the constitutive variables are replaced by the corresponding probabilities, whereas the description of systems in the macroscale gives exact results only for the average values of these variables.

4. ANALYTICAL ESTIMATION OF A CHARACTERISTIC SCALE

It turns out that the network scale, over which models are applied, is the key point for studying and predicting the worm behavior. Specifically, if robust monitoring is the goal, we need to be able to estimate the critical network size, which we define as the smallest size over which determinism not only wins but excels compared to randomness. In other words, monitoring worm propagation in a network of this size will accurately project the spread of the worm in larger networks.

To this end, by using the proposed model in Equation (6), a *characteristic* network size (i.e., a size over which determinism wins over randomness) is obtained by comparing the stochastic term with the first term in the second hand, that is, the classical epidemic term. The estimation is performed for the early stage of worm propagation where the saturation term in Equation (6) vanishes, for example,

$$\frac{da}{dt} =< \beta > a + \delta a, \tag{14}$$

with $a = I/N$ being the density of infected hosts and δa the corresponding fluctuation. Moreover, we assume a unit variation of the infected hosts. Then, the variation of the epidemic term is of the order $< \beta >$. This may be balanced against the fluctuating stochastic term, whose average over a characteristic scale L is of magnitude, $\left(\frac{<\delta a^2>}{L}\right)^{1/2}$. Equating the two terms, we obtain the following relation for the characteristic size, in units of the size of the network over which averaging is performed,

$$L_{char} = \frac{< \delta a^2 >}{< \beta >^2}. \tag{15}$$

Equation (15) is of crucial importance. To the best of our knowledge, it is the only exact relation estimating the characteristic network size. In Section 5.1, the above equation will be used to estimate the critical network size over which determinism not only wins but excels over randomness; that is, the classical epidemic models are accurate in predicting worm propagation.

For network monitoring and intrusion detection, these theoretical results can be very useful. Indeed, by monitoring a network of critical size, the growth of an unknown, fast-spreading worm may be correctly projected in a timely manner, meaning that a robust estimation of the infection rate and the expected damage associated with the worm could be given. In Section 6, we will discuss how this knowledge could be used to trigger mechanisms for early warning and emergency response against a fast-spreading worm.

5. SIMULATION AND VALIDATION

Next to the development of the theoretical model, we study the characteristics of worm infection spread via detailed discrete event simulation. In this section, we describe our simulator's setup and results.

The discrete event simulator code is written specifically for this simulation setup in the C language. The program uses the standard GNU C/C++ libraries and can be executed in any x86 architecture machine. An interconnected wired network with stable physical characteristics and error behavior is being modeled. All timing delays are being examined on top of this network; as such, these delays take into consideration possible transmission errors in underlying layers.

The code of the simulator models a fast User Datagram Protocol (UDP) scanning worm with a minimal payload packet. The employment of UDP enables an aggressive behavior of the worm without Transmission Control Protocol (TCP) handshaking delays. In this way, the scanning rate of an infected host is effectively limited only by the available bandwidth of its interconnecting network interface. In our setup, the average scanning rate is set to 1~probe per ms, which is also our base simulation timing unit. Following the theoretical model, the simulated worm is assumed to exhibit a uniform scanning strategy by targeting every node in the setup with equal probability. Our main goal is to study the early stage of rapid infection spread; consequently, each simulated computer node is modeled to be in one of two states, either susceptible or infected. That is, no recovery or immunization actions are taken during simulation execution.

During simulation time, we study an Internet portion of 256 C-class networks. Each network is treated as an independent LAN with a network backbone interconnecting all 256 LANs. Each LAN, internally, is assumed to have a total bandwidth of 100~Mbps[‡], with the same bandwidth available on egress nodes of every LAN toward the interconnecting backbone. Although arbitrary traffic is expected within and between LANs, the UDP worm-generated traffic is studied as the dominant factor of bandwidth limitation within each LAN. We assume 1% bandwidth overhead for each infected node in a LAN.

5.1. Estimation of the critical size

In order to validate the analytical results about the characteristic network size predicted by Equation (15), we apply the following procedure. For different subnet sizes of 150, 300, 450, 1200, and 2400 hosts, over which monitoring is performed for a given early time of worm propagation, the quantities $< \beta >$ and $< \delta a^2 >$ are estimated using the simulations outcomes (here, we use a sample of at least 20 subnets of each size). It is noted that, in order to

[‡] Within the limited complexity of our simulation setup, a larger bandwidth, that is 1~Gbps, was also simulated, showing a very small increase in the total infection rate.

estimate $<\beta>$, we fit simulation's monitoring data of infection propagation in time by means of an exponential function, because, at early times, the solution of Equation (14) follows an exponential law. Then, $<\beta>$ coincides with the exhibitor coefficient. Further, the estimation of $<\delta a^2>$ is straightforward. The results are depicted in Figure 1.

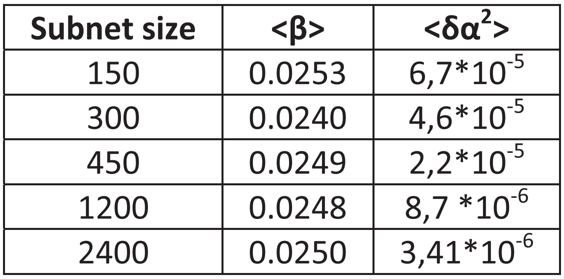
With substitution in Equation (15), it turns out that a characteristic network size of about 16 hosts is obtained. This means that, below this size, randomness always wins determinism, and the evolution of infected hosts proceeds in discrete jumps. Indeed, this is verified in Figure 2. On the other hand, above that scale, the effect of the deterministic term begins to become the dominant part for the evolution of infected hosts. It is our purpose to define the critical network size over which determinism not only wins but excels over randomness, and as a result, the evolution of infected hosts at this scale will coincide with the evolution on the Internet. To this end, in Figure 3, the variation of $<\delta a^2>$, with the subnet size over which monitoring is performed, is depicted. It can be seen that there is a power law dependency of $<\delta a^2>$ on the subnet size. Fitting the simulation's data, the following power law relation is estimated;

$$< \delta a^2 > (L) = \frac{0.02}{L^{1.1}}. \tag{16}$$

Without loss of generality, we assume that determinism excels over randomness when the ratio of Equation (15) is about 0.01. With the combination of Equations (15) and (16), it turns out that,

$$L_{crit} = \left(\frac{0.02}{0.01 < \beta>^2} \right)^{0.9} \tag{17}$$

or $L_{crit} = 1427$ hosts. The analytical estimated value of the critical network size is verified from the plot of infected hosts evolution in Figure 4. Indeed, for network sizes equal or greater than the critical size, the propagation of the infected hosts coincides with the evolution over the total population, whereas for networks of smaller sizes, a discrepancy is observed. As a result, the critical network size is correctly estimated from Equation (17).



Subnet size	<β>	<δα^2>
150	0.0253	$6{,}7*10^{-5}$
300	0.0240	$4{,}6*10^{-5}$
450	0.0249	$2{,}2*10^{-5}$
1200	0.0248	$8{,}7*10^{-6}$
2400	0.0250	$3{,}41*10^{-6}$

Figure 1. Mean infection rate and fluctuation amplitude.

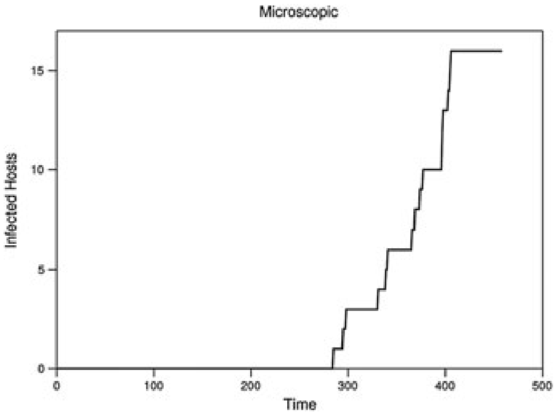


Figure 2. Infected hosts in microscale (size = 16 hosts).

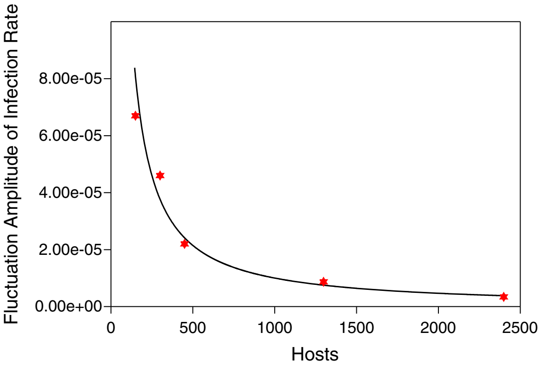


Figure 3. Fluctuation of infection rate versus network scale.

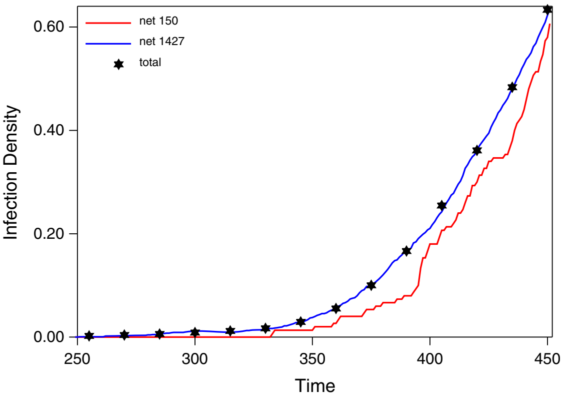


Figure 4. Worm evolution in critical/ noncritical sizes.

6. A FRAMEWORK FOR A DISTRIBUTED EARLY WARNING SYSTEM

The basic theoretical result of this paper is that by studying worm propagation within a network of proper scale that is, a critical size *CS* of subnetworks, it is possible to

DOI: 10.1002/sec

accurately project the behavior of a worm in larger networks. Worm projection basically involves collecting data and then estimating the infection rate and expected damage caused by the worm. Early projection results, paired with a well-established early warning policy, may lead to robust response strategies against fast-spreading, unknown worms.

One issue that needs to be addressed is that, when a new worm starts to propagate, we cannot know, a priori, whether the size of any given network is close to the critical size *CS*. In fact, worm projection introduces a *trade-off* between robustness and timely response. Specifically, for any given network of size $|N|$, there are three possibilities.

(1) Its size is far smaller than the critical size, $|N| < < CS$, which means that its projection is probably not accurate.
(2) Its size is far greater than the critical size, $|N| > > CS$, which means that its projection would be accurate, but it would also induce a time delay that is analogous to $|N|$. For very large scales, such a delay could be detrimental in the case of a fast-spreading worm where the time to respond is of paramount importance.
(3) Its size is close to the critical size, $|N| \simeq CS$, which, as shown in Section 5.1, will lead to an accurate estimation of the infection parameters within an optimally short time.

To this end, we envisage a distributed early warning system that, on the basis of the stochastic model of [23], will be able to estimate and disseminate, in a timely manner, the propagation parameters of a specific worm.

6.1. Assumptions and system model

We assume a hierarchical and distributed early warning system that is capable of incorporating various networks with minimum configuration effort. A high level description of the system is depicted in Figure 5. In our system model, there are *k* network domains, where each domain monitors *n* subnetworks. We assume that member domains are variably sized, that all different sizes sufficiently represent most regions of network scales, and that they have varying internal characteristics, (e.g., topology, bandwidth, traffic, installed operating systems, and applications). We also assume the existence of an early detection component (EDC) within each member domain. The EDC is able to detect the presence of a fast-spreading worm and is able to define the worm propagation model parameters. We will treat the EDC as a blackbox and assume that its functionality is administered within each network domain.

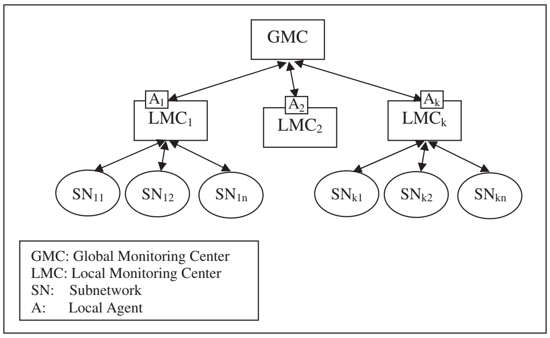


Figure 5. System model.

Each domain is locally monitored by one local monitoring center (LMC). A local agent *A* runs in each LMC and is programmed to act as a communication interface between the LMC and the root of the hierarchy, namely a global monitoring center (GMC). The operation of the local agent is practically the basic requirement in order to participate in the warning system. Finally, the GMC receives infection information from the LMCs and sends back warning information for an emergency response.

6.1.1. The local agent.

The local agent A_i runs on the LMC_i and can be programmed locally, by the domain administrator, in order to enforce a particular worm detection strategy. This is achieved by enabling the agent to configure and manage the local EDC component in order to collect infection data from all the subnetworks it monitors. Specifically, the agent collects data from the local EDC, and when enough data are available, it runs the propagation model in order to estimate the infection parameters, that is, an estimation of the fluctuation δa_i^2 and of the worm infection rate β_i.

6.1.2. The role of the LMC.

The LMC receives, through the local agent it controls, information (when this is available) concerning the infection parameters as well as the size of each subnetwork it monitors. Then, the LMC uses a secure communication channel in order to send the infection information to the *GMC*. This information includes the size $|LMC_i|$ of the domain and estimations for the worm infection rate β_i and the fluctuation δa_i^2. Thus,

$$LMC_i \rightarrow GMC : \left[|LMC_i|, \beta_i, \delta a_i^2\right] \qquad (18)$$

As an extension, in addition to the worm propagation estimations, each LMC could also send to the GMC (a selection of) mitigation policies and techniques that the domain is enforcing in response to the threat of the specific worm.

6.1.3. The role of the GMC.

When a worm starts to propagate, the GMC receives, from the LMCs, information including the size of the domains and estimations of the infection parameters. In time, the GMC will be able to reproduce the power law of Equation (16) and to estimate the critical size *CS* from Equation (17). As soon as a robust estimation of the critical size *CS* is available, the GMC will directly disseminate the correct infection parameters, that is, the parameters that were estimated by any LMC whose size is the closest to the *CS*.

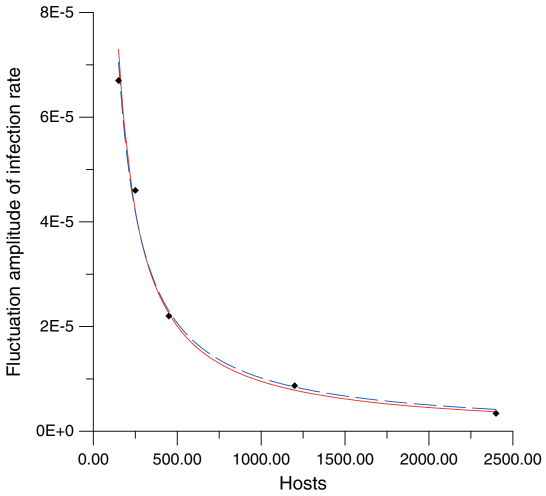


Figure 6. Critical size: convergence of the estimations.

It is noted that although LMCs with small sizes will start sending their data at earlier times, the reproduction of the power law of Equation (16) and, as a result, the initial estimations of *CS* via Equation (17) will gradually converge to the correct CS for the specific worm, as more data from LMCs with larger sizes will arrive. In any case, convergence takes place within the same order of magnitude for the corresponding correct *CS* value. Indeed, this is depicted in Figure 6 where we reproduced the power law for the two cases: (i) where all five sizes of the LMCs (i.e., 150, 300, 450, 1200, and 2400 hosts) were used, as in Figure 3 (with the solid line); and (ii) where only the first three, small-sized LMCs were included (with the dashed line). It is evident that the convergence to the correct *CS* values proceeds with a very small error.

As an extension, the GMC could also disseminate mitigation information back to the LMCs. In this way, early mitigation controls could be implemented by the LMCs. For example, according to the assessment of the collected mitigation policies, the LMCs may properly adapt their network or host-level firewall policies, intrusion prevention systems, automatic quarantine policies, or even manually disconnect or isolate particular types of hosts or services.

6.2. Deployment and implementation issues

Our approach does not compete but is meant to complement current systems for early detection and automatic containment. For example, the proposed framework could be integrated with any system that is based on worm propagation models for early detection (e.g., [20]) or with systems that describe mechanisms for automatic containment (e.g., [28]).

As early detection has a cost [20,22], especially in large-sized networks, in real systems, it is desirable that the EDC is initiated when there is a nonnegligible evidence that a worm is ‘at the gates’. To this end, we suggest the deployment of an anomaly-based detection system that will collect preliminary data from default locations, preanalyze the network traffic, and make a decision on whether the examined traffic overcomes a threshold and hides a potential scanning worm’s behavior, for example, the number of scans to unused IP addresses or to a specific TCP/UDP port by ingress or egress traffic. As in [20], a recursive filtering algorithm (e.g., a Kalman filter) could be activated at the EDC in order to correctly estimate the necessary model parameters (the average worm scan rate and the number of infected hosts). We scope away deployment and implementation issues for such a system.

Typically, the local agent will run as an extra logic into the subnet routers or at the edge of the domains as a firewall logic. We also refer the reader to deployment scenarios studied in [1].

Finally, any suitable early warning system should fulfill the following requirements:

1. *Scalability*. The warning system should be scalable in order to monitor a large number of networks of variable size.
2. *Extensibility and flexibility*. The insertion or deletion of a network should not affect other participating networks.
3. *Technology independence*. The participation of networks should not depend on specific network technologies or systems. Worm detection and mitigation/containment technologies should be extensible to adapt in different underlying technologies such as networking and OS technologies.
4. *Computationally efficient*. The system should be able to efficiently estimate, in a relatively short time, the worm propagation parameters.
5. *Mitigation capabilities*. The system should be able to collect information from various networks regarding the mitigation policies and techniques used, as well as their effect on the worm mitigation.
6. *Warning and mitigation efficiency*. The system should be able to disseminate worm mitigation policies and techniques almost automatically in all the involved entities.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we build upon the stochastic approach of [23], which is used in order to study random-scanning worms in the Internet. The stochastic model of [23] was validated and extended by giving a theoretical estimation of the critical size of a network that needs to be monitored in order to project the behavior of a worm in larger networks. On the basis of the notion of critical-sized networks, which was also validated by experimental simulations, we described a framework for a distributed early warning system, where a coalition of network domains of variable size and characteristics cooperatively estimate and disseminate the infection parameters for an early detected worm instance. We expect that estimations generated by a network of critical size will

DOI: 10.1002/sec

help larger networks to respond in a more timely manner to new worm threats. In this way, the resilience of networks against fast-spreading malware can be improved. In a future study, we intend to elaborate on the design and implementation of a distributed intrusion prevention system against fast-spreading Internet worms.

Future malware is also expected to take advantage of the ubiquity of wireless networking technologies (e.g., bluetooth and WiFi), which smartphones and other portable computational devices are equipped with [49,50]. Given that such devices may not always be Internet-connected, fast-spreading malware will also need to exhibit local preference behavior and employ P2P networking techniques for self-propagation within wireless broadcast range. For example, the notion of a *mobile malnet* (i.e., a botnet created from computational 802.11 devices) [51,52] has already been proposed in the context of involuntary location tracking [52]. Malware attacks directed at smartphones will also exploit the relatively homogeneous software base of current smartphones. Furthermore, given the pervasiveness of geolocation technologies in modern smartphones, future malware propagation strategies may also be based on context awareness [53]. Our future work direction will be motivated by the need to treat random and scale effects in the problem of self-propagating malware in wireless networking environments.

REFERENCES

1. Moore D, Shannon C, Voelker G, Savage S. Internet quarantine: requirements for containing self-propagating code. *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2003*., IEEE, 2003; 1901–1910.
2. Weaver N, Paxson V, Staniford S, Cunningham R. A taxonomy of computer worms. *WORM'03: Proceedings of the 2003 ACM workshop on Rapid malcode*, ACM: New York, NY, USA, 2003; 11–18.
3. Staniford S, Paxson V, Weaver N. How to own the internet in your spare time. *Proceedings of the 11th USENIX Security Symposium*, USENIX Association: Berkeley, CA, USA, 2002; 149–167.
4. Cisco. Cisco 2010 Annual Security Report June 2010. URL http://www.cisco.com.
5. Symantec Security Response. W32.Stuxnet Dossier Version 1.4 (February 2011) February 2011. URL http://www.symantec.com.
6. Brenner JF. Why isn't cyberspace more secure? *Communications of the ACM* 2010; **53**: 33–35.
7. Krebs B. Stuxnet'' worm far more sophisticated than previously thought. *Krebs on Security*; URL http://krebsonsecurity.com.
8. Zou CC, Towsley D, Gong W, Cai S. Advanced routing worm and its security challenges. *Simulation* 2006; **82**(1): 75–85.
9. Wu J, Vangala S, Gao L, Kwiat K. An efficient architecture and algorithm for detecting worms with various scan techniques. *NDSS"04: Proceedings of the 11th Annual Network and Distributed System Security Symposium*, 2004.
10. Chen Z, Ji C. Measuring network-aware worm spreading ability. *26th IEEE International Conference on Computer Communications, INFOCOM 2007*, IEEE, 2007; 116–124.
11. Weaver N. Warhol worms: The potential for very fast Internet Plagues, http://www.iwar.org.uk/comsec/resources/worms/warhol-worm.htm 2001.
12. Zou CC, Gong W, Towsley D. Code red worm propagation modeling and analysis. *CCS'02: Proceedings of the 9th ACM conference on Computer and communications security*, ACM: New York, NY, USA, 2002; 138–147.
13. Zou C, Gong W, Towsley D, Gao L. The monitoring and early detection of internet worms. *ACM Transactions on Networking* 2005; **13**(5): 961–974.
14. Chen Z, Gao L, Kwiat K. Modeling the spread of active worms. 22nd Annual Joint Conference of the IEEE Computer and Communications INFOCOM 2003, IEEE, 2003; 1890–1900.
15. Chen Z, Chen C, Ji C. Understanding localized-scanning worms. *Proceedings of the 26th IEEE International Performance Computing and Communications Conference, IPCCC 2007, April 11–13, 2007, New Orleans, Louisiana, USA*., IEEE Computer Society, 2007; 186–193.
16. Weaver N, Hamadeh I, Kesidis G, Paxson V. Preliminary results using scale-down to explore worm dynamics. *WORM'04: Proceedings of the 2004 ACM workshop on Rapid malcode*, ACM: New York, NY, USA, 2004; 65–72.
17. Kamra A, Feng H, Misra V, Keromytis A. The effect of DNS delays on worm propagation in an IPv6 Internet. *Proceedings of IEEE INFOCOM*, 2005.
18. Serazzi G, Zanero S. Computer virus propagation models. In *MASCOTS Tutorials, Lecture Notes in Computer Science*, Vol. 2965, Springer: Berlin, 2003; 26–50.
19. Zou CC, Towsley D, Gong W. On the performance of internet worm scanning strategies. *Performance Evaluation* 2006; **63**(7): 700–723.
20. Zou C, Gao L, Gong W, Towsley D. Monitoring and early warning for Internet worms. *Proceedings of the 10th ACM conference on Computer and communications security, ACM CCS 2003*, ACM: New York, NY, USA, 2003; 190–199.
21. Pouget F, Dacier M, Pham V. Understanding threats: a prerequisite to enhance survivability of computing systems. *International Journal of Critical Infrastructures (IJCIS)* 2008; **4**(1): 153–171.

22. Gamer T, Scholler M, Bless R. A granularity-adaptive system for in-network attack detection. *Proceedings of the IEEE/IST Workshop on Monitoring, Attack Detection and Mitigation*, 2006; 47–50.
23. Avlonitis M, Magkos E, Stefanidakis M, Chrissikopoulos V. A novel stochastic approach for modeling random scanning worms. *13th Panhellenic Conference on Informatics—PCI 2009, 12–14 Sep. 2009, Corfu, Greece.*, IEEE Computer Society, 2009; 176–179.
24. Kephart JO, White SR. Directed-graph epidemiological models of computer viruses. *IEEE Symposium on Security and Privacy*, 1991; 343–361.
25. Kesidis G, Hamadeh I, Jin Y, Jiwasurat S, Vojnović M. A model of the spread of randomly scanning internet worms that saturate access links. *ACM Transactions on Modeling and Computer Simulation* 2008; **18**(2): 1–14.
26. Faghani M, Saidi H, Ataei M. Effects of security solutions on worm propagation. *International Symposium on Telecommunications, IST 2008*, IEEE, 2008; 25–29.
27. Ganesh A, Massoulie L, Towsley D. The effect of network topology on the spread of epidemics. *24th Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2005.*, IEEE, 2005; 1455–1466.
28. Zou CC, Gong W, Towsley D. Worm propagation modeling and analysis under dynamic quarantine defense. *WORM'03: Proceedings of the 2003 ACM workshop on Rapid malcode*, ACM: New York, NY, USA, 2003; 51–60.
29. Avlonitis M, Magkos E, Stefanidakis M, Chrissikopoulos V. Treating scalability and modelling human countermeasures against local preference worms via gradient models. *Journal in Computer Virology* 2009; **5**(4): 357–371.
30. Gu G, Sharif M, Qin X, Dagon D, Lee W, Riley G. Worm detection, early warning and response based on local victim information. *IEEE 20th Annual Computer Security Application Conference, IEEE Computer Society*, 2004.
31. Wang Y, Wang C. Modeling the effects of timing parameters on virus propagation. *WORM'03: Proceedings of the 2003 ACM workshop on Rapid malcode*, ACM: New York, NY, USA, 2003; 61–66.
32. Nicol DM. The impact of stochastic variance on worm propagation and detection. *WORM'06: Proceedings of the 4th ACM workshop on Recurring malcode*, ACM: New York, NY, USA, 2006; 57–64.
33. Rohloff KR, Bacşar T. Deterministic and stochastic models for the detection of random constant scanning worms. *ACM Transactions on Modeling and Computer Simulation* 2008; **18**(2): 1–24.
34. Floyd S, Paxson V. Difficulties in simulating the Internet. *IEEE/ACM Transactions on Networking (TON)* 2001; **9**(4):403.
35. Project TH. Know your enemy 2010. URL http://honeynet.org.
36. CAIDA. The cooperative association for internet data analysis 2010. URL http://www.caida.org/home/.
37. ISC. Internet storm center 2010. URL http://isc.sans.org/.
38. Moore D, Shannon C. The spread of the code-red worm (crv2). *Cooperative Association for Internet Data Analysis (CAIDA): analysis: security: code red),[online] 30 July 30 2001*, http://www. caida.org/analysis/security/codered/coderedv2_analysis. xml *(Accessed* 2005; **3**.
39. Moore D. Network telescopes: observing small or distant security events. *Proceedings of the 11th USENIX security symposium*, 2002.
40. Rozenberg B, Gudes E, Elovici Y. A distributed framework for the detection of new worm-related malware. *Proceedings of the 1st European Conference on Intelligence and Security Informatics*, Springer, 2008; 190.
41. Axelson S. Intrusion detection systems: a survey and taxonomy. *Technical Report Technical Report 99–15*, Department of Computer Engineering, Chalmers University of Technology, Goteborg, Sweden 2000.
42. Cavallaro L, Lanzi A, Mayer L, Monga M. LISABETH: automated content-based signature generator for zero-day polymorphic worms. *Proceedings of the fourth international workshop on Software engineering for secure systems*, ACM: New York, NY, USA, 2008; 41–48.
43. Wang C, Knight JC, Elder MC. On computer viral infection and the effect of immunization. *ACSAC'00: Proceedings of the 16th Annual Computer Security Applications Conference*, IEEE Computer Society: Washington, DC, USA, 2000; 246.
44. Wang J, Liu Y, Tian D, Wei D. Internet worm early detection and response mechanism. *The Journal of China Universities of Posts and Telecommunications* 2007; **14**(3): 79–84.
45. Sellke S, Shroff N, Bagchi S. Modeling and automated containment of worms. *IEEE Transactions on Dependable and Secure Computing* 2008; **5**(2): 71–86.
46. Horsthemke W, Lefever R. Noise-induced Transitions. *Springer Series in Synergetics* 2004; 15.
47. Haken H. *Synergetics: Introduction and Advanced Topics*. Springer: Berlin, 2004.
48. Avlonitis M, Zaiser M, Aifantis EC. Some exactly solvable models for the statistical evolution of

internal variables during plastic deformation. *Probabilistic Engineering Mechanics* 2000; **15**: 131–138.

49. Hu H, Myers S, Colizza V, Vespignani A. WiFi networks and malware epidemiology. *Proceedings of the National Academy of Sciences* 2009; **106**(5): 1318.
50. Wang P, González M, Hidalgo C, Barabási A. Understanding the spreading patterns of mobile phone viruses. *Science* 2009; **324**(5930):1071.
51. Traynor P, Butler K, Enck W, McDaniel P, Borders K. Malnets: large-scale malicious networks via compromised wireless access points. *Security and Communication Networks* 2010; **3**(2–3): 102–113.
52. Husted N, Myers S. Mobile location tracking in metro areas: malnets and others. *Proceedings of the 17th ACM conference on Computer and communications security*, ACM, 2010; 85–96.
53. Schlegel R, Zhang K, Zhou X, Intwala M, Kapadia A, Wang X. Soundcomber: a stealthy and context-aware sound trojan for smartphones. *Proceedings of the 18th Annual Network and Distributed System Security Symposium (NDSS)*, 2011; 17–33.
54. PandaLabs. Quarterly report PandaLabs (July-September 2010) June 2010. URL http://prensa.pandasecurity.com.
55. Franklin J, Paxson V, Perrig A, Savage S. An inquiry into the nature and causes of the wealth of internet miscreants. *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, 2007.

DOI: 10.1002/sec