

Importance-Scanning Worm Using Vulnerable-Host Distribution

Zesheng Chen and Chuanyi Ji

School of Electrical & Computer Engineering

Georgia Institute of Technology, Atlanta, Georgia 30332

Email: {zchen, jic}@ece.gatech.edu

Abstract—Most Internet worms use random scanning. The distribution of vulnerable hosts on the Internet, however, is highly non-uniform over the IP-address space. This implies that random scanning wastes many scans on invulnerable addresses, and more virulent scanning schemes may take advantage of the non-uniformity of a vulnerable-host distribution. Questions then arise how attackers may make use of such information, and how virulent the resulting worm may be. These issues provide “worst-case scenarios” for defenders and “best-case scenarios” for attackers if the vulnerable-host distribution is available. This work develops such a scenario as the so-called *importance scanning*. Importance scanning results from Importance Sampling in statistics that scans IP-address space according to an empirical distribution of vulnerable hosts. An analytical model is developed to relate the infection rate of worms with the importance-scanning strategies. Experimental results based on parameters chosen from Code Red and Slammer worms show that an importance-scanning worm can spread much faster than both a random-scanning worm and a routing worm. Furthermore, a game-theory approach suggests that the best strategy for defenders is to scatter applications uniformly in the entire IP-address space.

Index Terms—Security, Worm propagation, Modeling, Game theory, Importance scanning

I. INTRODUCTION

As computers and communication networks become prevalent, the Internet is plagued by many worms [5], [6], [9]. Using self-propagating malicious codes, worms spread rapidly by infecting computer systems and disseminating themselves in an automated fashion using the infected nodes.

Most worms employ random scanning to select target IP addresses. Since the density of vulnerable hosts is low, a random scan hits a vulnerable machine with a small probability. Thus random scanning wastes many scans on invulnerable addresses. For example, Code Red infected a vulnerable population of 360,000 machines among 2^{32} IP addresses [16]. The probability for a random scan to hit a vulnerable target is thus only $\frac{360,000}{2^{32}} = 8.38 \times 10^{-5}$.

Future worms, however, are likely to employ more effective scanning strategies in identifying the targets. Hence it is important to study advanced scanning strategies, which can potentially be used to access worst-case scenarios. This work proposes a novel scanning method, referred to as *importance scanning*. Importance scanning is inspired by importance sampling in statistics [15], [4], [10]. The basic idea of importance sampling is to make rare events occur more frequently, and thus reduces the number of samples needed for accurately estimating the corresponding probability. Rare events for worm scanning correspond to hitting a target in a large population. Importance scanning thus allows attackers to focus on the most

relevant parts of an address space so that the probability of hitting vulnerable hosts can be made larger.

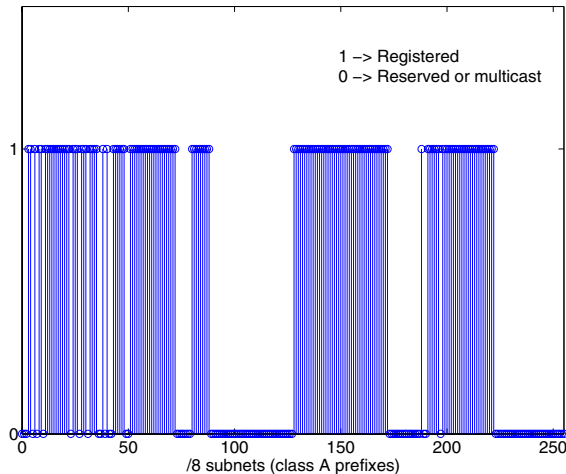
Importance scanning relies on a certain statistic of an underlying vulnerable-host distribution. An attacker can potentially obtain such information by querying a database of parties to the vulnerable protocol, stealthy scanning (partial) target address space, and/or searching records of old worms [11].

In view of the amount of information an attacker can obtain, random, flash [12], and routing [17] worms can be regarded as special cases of importance-scanning worms. In particular, a random worm has no information about vulnerable-host distribution and thus regards the distribution as uniform in IPv4 space. A flash worm acquires all knowledge and the target distribution is uniform only in the vulnerable-population space. A routing worm has knowledge from BGP routing tables about the space of existing hosts, and the corresponding distribution can be considered as uniform in the routing space.

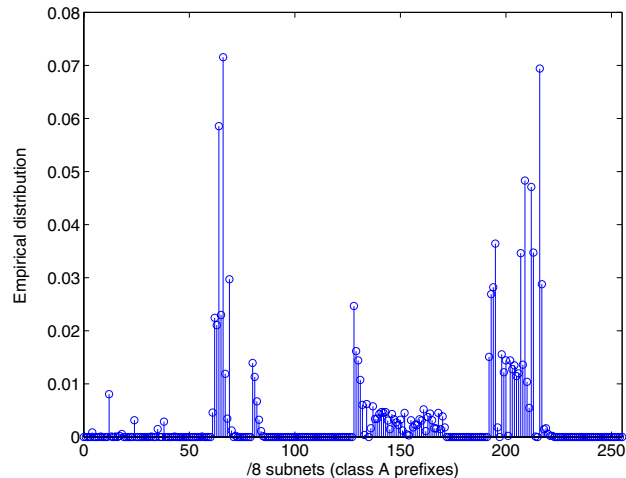
In this work, we assume that a probability distribution of vulnerable hosts is available/obtainable. We then intend to answer the following questions:

- How can an attacker design a fast importance-scanning worm by taking advantage of the knowledge of vulnerable-host distribution?
- How can we analyze quantitatively the relationship between the speed that worms can achieve and the knowledge that attackers can obtain?

To answer these questions, we focus on two quantities: the infection rate that characterizes how fast worms can spread at an early stage and the scanning strategy that is used to locate vulnerable hosts. We first derive a relationship between the infection rate and scanning strategies. We then model the spread of importance-scanning worms using the Analytical Active Worm Propagation (AAWP) model [1]. We derive the optimal scanning strategy, which maximizes the infection rate. The optimal strategy thus corresponds to the best-case scenario for attackers and the worst-case scenario for defenders. As the optimal strategy is difficult to achieve in reality, we derive a suboptimal scanning strategy as the approximation. To assess the virulence, we compare the importance scanning with the random and routable scanning. We take the empirical distribution of web servers as an example of vulnerable-host distribution. We show that an importance-scanning worm can spread nearly twice faster than a routing worm based on parameters chosen from real measurements. Moreover, we demonstrate, from a view of game theory, that defense mechanisms against importance-scanning worms require a uniform distribution of an application.



(a) IANA assignment on Jan. 27, 2005.



(b) Empirical distribution of web servers.

Fig. 1. Uneven distribution of vulnerable hosts.

The remainder of this paper is structured as follows. Section II provides the background on worm scanning methods, vulnerable-host distribution, and random worm propagation model. Section III characterizes the importance scanning strategy through an analytical model and experiments. We further discuss the defense strategy from the view of game theory in Section IV, and conclude the paper in Section V.

II. PRELIMINARIES

A. Scanning Methods

A worm spreads by employing distinct scanning mechanisms such as random, localized, and topological scanning [11]. *Random scanning* selects target IP addresses at random and is used by such famous worms as Code Red and Slammer. *Localized scanning* preferentially scans for hosts in the “local” address space and is used by Code Red II and Nimda worms. *Topological scanning* relies on the “address” information contained in the victim machine to locate new targets and is used by Morris worm.

Some advanced scanning methods have been developed in research community. For example, Weaver presented the “hitlist” idea [13] to speed up the spread of worms at the initial stage. There a list of potentially vulnerable machines is built up beforehand and targeted first when the worm is released. An extreme case for the hitlist-scanning worms is called *flash worms*, where all vulnerable machines are gathered into the list. The flash worms are considered to be the fastest possible worms [12], as every worm scan can hit a vulnerable host. One other method to improve the spread of worms is to reduce the scanning space. Attackers can potentially achieve this by using the information provided by BGP routing tables. This type of worm is called “routable-scanning worm” [14] or “routing worm” [17]. Zou et al. designed two types of routing worms [17]. One type is based on class A (x.0.0.0/8) address-allocations, and thus called “Class A routing worm”. Such worm can reduce the scanning space to 45.3% of the

entire IPv4 address space. The other type is based on BGP routing tables, and thus called “BGP routing worm”. Such worm can reduce the scanning space to only about 28.6% of the entire IPv4 address space. One other strategy that a worm can potentially employ is *DNS random scanning* [3], where a worm uses the DNS infrastructure to locate likely targets by guessing DNS names instead of IP addresses. Such a worm in an IPv6 Internet is shown to exhibit propagation speeds comparable to that of an IPv4 random-scanning worm.

Most of these advanced worms can propagate far faster than a traditional random-scanning worm. When these advanced worms are studied, however, the vulnerable hosts are assumed to be uniformly distributed in either the whole IPv4 address or the scanning space. Hence the information on a vulnerable-host distribution is not exploited by the worms.

B. Distribution of Vulnerable Hosts

The distribution of vulnerable hosts in the Internet is not uniform. Figure 1(a) shows the assignment of the first byte of IPv4 address space by the Internet Assigned Number Authority (IANA) on January 27, 2005 [19]. There no hosts can exist in the reserved or multicast address ranges. More importantly, the distribution of vulnerable hosts is highly non-uniform over the IPv4 address space that is registered. Consider web servers as an example. Assume that a worm searches for public accessible web servers as targets. We can estimate the distribution of web servers as follows. We collected 13,866 IP addresses of web servers provided by the random URL generator from UROULETTE (<http://www.roulette.com/>) on January 24, 2005. These addresses were then used to form an *empirical distribution*, where

$$P_e(i) = \frac{\text{number of addresses with the first byte equal to } i}{\text{total number of collected addresses}}, \quad (1)$$

where $i = 0, 1, \dots, 255$. The results are shown in Figure 1(b) and match IANA assignment on IPv4. It is observed

that the distribution of web servers is far from uniform in the address space that is routable. Statistical analysis of *network telescope* observations also shows that the victims of Code Red and Slammer worms have a highly non-uniform geographical distribution [5], [6]. Moreover, DShield [18] data indicate that the distributions of vulnerable hosts among prefixes follow a *power law* [8].

C. Worm Propagation Model for Random-Scanning Worms

We now review worm propagation model as a preparation for relating the rate of worm spread with the distribution of vulnerable hosts. A simple model, known as *susceptible* \rightarrow *infected* (SI) model, has been used to model the spread of random-scanning worms in various earlier works [11], [17], [3]. It assumes that each host has only two states: susceptible or infected. Once infected, a host remains infected.

In this paper, we adopt a discrete-time SI model, as importance scanning (sampling) is usually performed in discrete time [15]. In particular, we use Analytical Active Worm Propagation (AAWP) model, which was developed by Chen et al. in [1]. In the AAWP model, the spread of random-scanning worms is characterized as following:

$$I_{t+1} = I_t + (N - I_t)[1 - (1 - \frac{1}{\Omega})^{sI_t}], \quad (2)$$

where I_t is the number of infected hosts at time t ($t \geq 0$); N is the number of vulnerable hosts; s is the scanning rate of the worm; and Ω is the scanning space. At $t = 0$, I_0 hosts are infected, representing the number of hosts on the hitlist.

When a worm begins to spread, $I_t \ll N$ and $sI_t \ll \Omega$. The AAWP model can be approximated by

$$I_{t+1} \approx I_t + N \cdot \frac{sI_t}{\Omega} = (1 + \frac{sN}{\Omega})I_t = (1 + \alpha)I_t, \quad (3)$$

where $\alpha = \frac{sN}{\Omega}$, called the infection rate [17]. Infection rate represents the average number of vulnerable hosts that can be infected per unit time by one infected host during the early stage of worm propagation. Based on Equation (3), $I_t \approx (1 + \alpha)^t I_0$, i.e., the number of infected hosts increases exponentially. Therefore, to speed up the spread of worms during the early stage, attackers should design effective scanning methods to increase the infection rate. For instance, a traditional random worm scans the whole IPv4 address space and thus $\Omega = 2^{32}$. The infection rate of this worm is $\alpha_0 = \frac{sN}{2^{32}}$. Comparatively, a BGP and a Class-A routing worm can achieve faster infection rates with the same scanning rate and the number of targets: $\alpha_1 = \frac{sN}{0.286 \times 2^{32}} = 3.5\alpha_0$ and $\alpha_2 = \frac{sN}{0.453 \times 2^{32}} = 2.2\alpha_0$.

III. IMPORTANCE SCANNING

In this section, we first introduce the concept of importance scanning by relating the infection rate with the vulnerable-host distribution. We then model the spread of importance-scanning worms. Finally, we compare the importance scanning with the random and routable scanning through experiments.

A. Infection Rate

Let $I(A_n)$ denote the vulnerability of an address A_n .

$$I(A_n) = \begin{cases} 1, & \text{if address } A_n \text{ is vulnerable to a worm;} \\ 0, & \text{otherwise.} \end{cases}$$

Thus, $\sum_{A_n} I(A_n) = N$. Let $p(A_n)$ denote the true vulnerable-host distribution, i.e., the probability that $I(A_n) = 1$.

$$p(A_n) = \frac{I(A_n)}{N} = \begin{cases} \frac{1}{N}, & \text{if } I(A_n) = 1; \\ 0, & \text{if } I(A_n) = 0. \end{cases}$$

Let $p^*(A_n)$ denote the probability that the worm scans the address A_n . $p^*(A_n)$ can be a uniform distribution as in random-scanning worms or a non-uniform biasing distribution as in flash worms. $p^*(A_n)$ is chosen by an attacker. The choice of the scanning distribution $p^*(A_n)$ is essential to the effectiveness of importance scanning. As we shall see, $p^*(A_n)$ depends on the true probability distribution $p(A_n)$.

Let R be the number of hosts that can be infected per unit time by one infected host during the early stage of worm propagation. R can be expressed as

$$R = \sum_{n=1}^s I(A_n), \quad (4)$$

where s is the scanning rate as the number of scans that an infected host sends per unit time. Therefore, the infection rate is given by

$$\begin{aligned} \alpha &= E_*[R] = \sum_{n=1}^s E_*[I(A_n)] = \sum_{n=1}^s \sum_{A_n} I(A_n)p^*(A_n) \\ &= N \sum_{n=1}^s \sum_{A_n} p(A_n)p^*(A_n) = sN \sum_{A_n} p(A_n)p^*(A_n), \end{aligned} \quad (5)$$

where $E_*[\cdot]$ denotes that the expectation is taken with respect to the scanning distribution $p^*(A_n)$. It is noted that

$$\alpha \leq \sum_{n=1}^s \sum_{A_n} p^*(A_n) = s, \quad (6)$$

for any $p^*(A_n)$.

Hence the infection rate is strongly influenced by the choice of the scanning distribution $p^*(A_n)$. A choice of $p^*(A_n)$ determines a scanning strategy; and a good choice, in view of an attacker, should maximize the infection rate α . Two special cases have been observed on ‘‘choosing’’ $p^*(A_n)$. The first case is the random-scanning worms, where $p^*(A_n) = \frac{1}{2^{32}}$. Thus, $\alpha = \frac{sN}{2^{32}} = \alpha_0$. The second case is the flash worms, where $p^*(A_n) = p(A_n)$. In this case, $p^*(A_n)$ obtains the optimal scanning strategy $p_{opt}^*(A_n)$, which leads to $\max_{p^*(A_n)} \{\alpha\} = s$. This means that every scan from the worm would hit a vulnerable host.

One interpretation of $p_{opt}^*(A_n)$ suggests that a good scanning strategy of a worm is to concentrate the scans in areas that are more likely to find a vulnerable host. The vulnerable-host probability distribution $p(A_n)$, however, cannot be obtained without probing the entire IP address space or getting a complete database of parties to the vulnerable protocol. Therefore, attackers may not acquire the entire knowledge of $p(A_n)$. However, partial knowledge can be obtained, e.g., aggregation on the subspaces of IP addresses.

B. Group Distributions

We consider such partial information as the marginal of $p(A_n)$, referred to as *group distributions*. The group distributions capture the statistics of groups of addresses, rather than individual addresses. The vulnerable-host probability distribution in groups is essentially the marginal of the true distribution $p(A_n)$. Many methods exist to form such groups of addresses. For example, IP addresses can be grouped by using the conventional 4-byte description. In [15], this approach is applied to study importance sampling of the size of the Internet. Here, we extract relevant groups in a more general setting by defining the networks. In particular, we regard a *network* as a group of IP addresses that can be identified by such diverse methods as either the first byte of IP addresses (/8 subnets) or IP prefixes in CIDR.

We assume that the Internet is composed of m networks. Let D_i ($i = 1, 2, \dots, m$) denote the set of addresses in network i , which has Ω_i ($\Omega_i \geq 0$) addresses. Thus, $\sum_{i=1}^m \Omega_i = \Omega = 2^{32}$. We define group distribution $p_g(i)$ ($i = 1, 2, \dots, m$) as the proportion of vulnerable hosts in network i , i.e.,

$$p_g(i) = \frac{N_i}{N} = \sum_{A_n \in D_i} p(A_n), \quad (7)$$

where N_i is the population of vulnerable hosts in network i .

The partition of networks reflects the knowledge that attackers can obtain. For example, in one extreme case of random-scanning worms, $m = 1$ and $\Omega_1 = 2^{32}$. In the other extreme case of flash worms, $m = 2^{32}$ and $\Omega_i = 1$ ($i = 1, 2, \dots, 2^{32}$). Another choice of partitioning networks is based on the first byte of IP addresses (/8 subnets), where $m = 2^8$ and $\Omega_i = 2^{24}$ ($i = 1, 2, \dots, 2^8$). The amount of knowledge collected in the latter case is between the cases of random worms and flash worms.

Recall that the goal of importance scanning is to maximize the infection rate. From Equation (5), we have the infection rate

$$\alpha = sN \sum_{i=1}^m \sum_{A_n \in D_i} p(A_n) p^*(A_n). \quad (8)$$

Refer the location of an address A_n that is in network i as the *interface*, denoted by b ($b = 0, 1, \dots, \Omega_i - 1$). Let $p_i(b)$ denote the true probability of finding a vulnerable host with the interface equal to b given that the host is in network i , i.e., $p_i(b) = \frac{I(A_n)}{N_i}$. Similarly, define *group scanning distribution* $p_g^*(i)$ as the probability of scanning network i , and *interface scanning distribution* $p_i^*(b)$ as the probability of scanning interface b given that a scan hits network i for the scanning distribution $p^*(A_n)$. We can obtain that

$$p(A_n) = p_g(i) \cdot p_i(b) \quad (9)$$

$$p^*(A_n) = p_g^*(i) \cdot p_i^*(b), \quad (10)$$

where A_n is in the network i with interface equal to b . From

Equations (9) and (10), the infection rate becomes

$$\begin{aligned} \alpha &= sN \sum_{i=1}^m \sum_{b=0}^{\Omega_i-1} p_g(i) p_i(b) p_g^*(i) p_i^*(b) \\ &= sN \sum_{i=1}^m \left[p_g(i) p_g^*(i) \sum_{b=0}^{\Omega_i-1} p_i(b) p_i^*(b) \right]. \end{aligned} \quad (11)$$

We assume that attackers can only obtain the information about group distribution $p_g(i)$ and cannot acquire the knowledge about interface distribution $p_i(b)$ further. Therefore, if a scan hits the network i , the Ω_i hosts in this network will be targeted by that scan with the same likelihood, i.e., $p_i^*(b) = \frac{1}{\Omega_i}$. Hence Equation (11) yields

$$\alpha = sN \sum_{i=1}^m \frac{p_g(i) p_g^*(i)}{\Omega_i}. \quad (12)$$

Equation (12) provides the relationships among the infection rate, the group distribution, and the group scanning distribution. It is noted that $\alpha = sN \sum_{i=1}^m v_i p_g^*(i) \leq sN \sum_{i=1}^m \max_k \{v_k\} p_g^*(i) = sN \max_k \{v_k\}$, where $v_i = \frac{p_g(i)}{\Omega_i}$, referred to as the *vulnerable-host density*. The equality holds when $p_g^*(j) = 1$, $j = \arg \max_k \{v_k\}$; or 0, otherwise. This means that the optimal importance scanning of a worm is to scan only the network with the largest vulnerable-host density.

C. Worm Propagation Model for Importance-Scanning Worms

We now model the spread dynamics of importance-scanning worms based on the information of group distribution.

At time t ($t \geq 0$), let $I_{t,i}$ denote the average number of infected hosts in network i , and thus the total number of infected hosts $I_t = \sum_{i=1}^m I_{t,i}$. The rate at which network i is scanned is $sI_t p_g^*(i)$. As importance scanning worm employs random scanning *within* each network, on the next time tick, the number of infected hosts in network i can be derived by the AAWP model, which is

$$I_{t+1,i} = I_{t,i} + (N_i - I_{t,i}) [1 - (1 - \frac{1}{\Omega_i})^{sI_t p_g^*(i)}], \quad (13)$$

where $i = 1, 2, \dots, m$ and $t \geq 0$. $I_{0,i}$ is the number of initially-infected hosts in network i . We assume that $\frac{1}{\Omega_i} \ll 1$, which yields

$$I_{t+1,i} \approx I_{t,i} + sI_t \frac{(N_i - I_{t,i}) p_g^*(i)}{\Omega_i}. \quad (14)$$

Summing over $i = 1, 2, \dots, m$ on both sides, we can obtain

$$I_{t+1} = I_t + sI_t \sum_{i=1}^m \left(\frac{N_i - I_{t,i}}{\Omega_i} \right) p_g^*(i) \quad (15)$$

$$\leq I_t + sI_t \sum_{i=1}^m \max_k \left\{ \frac{N_k - I_{t,k}}{\Omega_k} \right\} p_g^*(i) \quad (16)$$

$$= [1 + s \cdot \max_k \left\{ \frac{N_k - I_{t,k}}{\Omega_k} \right\}] I_t. \quad (17)$$

The equality holds when

$$p_g^*(j) = \begin{cases} 1, & j = \arg \max_k \left\{ \frac{N_k - I_{t,k}}{\Omega_k} \right\}; \\ 0, & \text{otherwise.} \end{cases}$$

When $t = 0$, $N_i \gg I_{t,i}$ and then $\max_k \left\{ \frac{N_k - I_{t,k}}{\Omega_k} \right\} \approx N \max_k \{v_k\}$, which leads to $\alpha = sN \max_k \{v_k\}$.

The above derivation results in the optimal importance-scanning strategy, which maximizes the infection rate.

Optimal importance-scanning:

- 1) At each time step t , the worms first find out the subnet that has the largest value of left vulnerable-host density, i.e., $j = \arg \max_k \left\{ \frac{N_k - I_{t,k}}{\Omega_k} \right\}$.
- 2) Then all infected hosts concentrate on scanning this subnet. That is, $p_g^*(j) = 1$ and $p_g^*(i) = 0$, $\forall i \neq j$.

This optimal importance scanning, however, is not realistic. First, N may not be known in advance, and thus N_i is unknown. Secondly, the subnet that has the largest value of left vulnerable-host density changes with time, and therefore the optimal assignment of $p_g^*(i)$ is relevant to time. Even though N is given, it requires each infected host to know $I_{t,i}$, which leads to a lot of information exchange between infected hosts. Nevertheless, the optimal importance scanning provides the best scenario of importance scanning and the baseline for a suboptimal selection of $p_g^*(i)$.

A simple strategy for suboptimal importance scanning is to assume $p_g^*(i) = \frac{p_g(i)/\Omega_i}{\sum_{j=1}^m p_g(j)/\Omega_j}$. That is, the probability of worm scanning network i is proportional to the vulnerable-host density for this network. If $\Omega_1 = \Omega_2 = \dots = \Omega_m$, then $p_g^*(i) = p_g(i)$. For this scanning strategy, the Equation (13) becomes

$$I_{t+1,i} = I_{t,i} + (N_i - I_{t,i}) \left[1 - \left(1 - \frac{1}{\Omega_i} \right)^{s I_t \frac{p_g(i)/\Omega_i}{\sum_{j=1}^m p_g(j)/\Omega_j}} \right]. \quad (18)$$

A suboptimal importance-scanning worm:

- 1) Before a worm is released, the attackers first obtain the group distribution of vulnerable hosts $p_g(i)$, and then encode the group scanning distribution $p_g^*(i) = \frac{p_g(i)/\Omega_i}{\sum_{j=1}^m p_g(j)/\Omega_j}$ in the worm code.
- 2) At each time step t , the worm scans the network i with the probability $p_g^*(i)$.

D. Experiments

How much does the importance scanning help a worm in speeding up the propagation? We employ the model in Equation (13) to study the propagation due to the optimal and suboptimal importance-scanning strategies, and compare them with routing and random worms.

To implement the model in Equation (13), we need to obtain the group distribution $p_g(i)$. Here we use the web-server distribution as an example of vulnerable-host distribution. In other words, we assume that worms attack the vulnerable hosts that have the same group distribution as that of the web server. The empirical distribution $p_e(i)$ (defined by Equation (1)) is expected to accurately reflect the relative distribution of the number of web servers as a function of the first byte values. Hence, we assume

$$p_g(i) = p_e(i). \quad (19)$$

Similar approach is also applied in [15].

The parameters we use in importance-scanning worms are comparable to those in Code Red and Slammer worms for

evaluating the propagation. Code Red worm has a vulnerable population $N = 360,000$ and a scanning rate $s = 358$ per minute [16]. We then refer such importance-scanning worm as the importance-scanning (IS) Code Red. Figure 2(a) shows how the propagation due to importance-scanning Code Red compares with those of random and BGP routing Code Red, assuming a hitlist of 10 (i.e., $I_0 = 10$). The importance-scanning Code Red greatly increase their spreading speed by using the information on the vulnerable-host distribution. The optimal importance-scanning Code Red is able to infect 80% vulnerable hosts in as few as 38 minutes, while the BGP routing Code Red needs 113 minutes. The suboptimal importance-scanning Code Red spreads slower than the optimal one, but only uses 60 minutes to infect the same number of hosts.

It is noted that when most of the vulnerable hosts are infected, the spread of suboptimal importance-scanning Code Red slows down. This is because the suboptimal strategy uses the same group scanning distribution all the time. As the infected hosts become saturated, the network that initially has more vulnerable hosts has actually a fewer vulnerable machines. To overcome this, the suboptimal importance scanning can choose to switch to routable scanning when there are a few vulnerable hosts left. Figure 2(b) shows results for the same experiments, assuming a hitlist of 13,866, which is the number of web servers collected from UROULETTE. The suboptimal importance scanning switches to Class A routable scanning when 80% vulnerable hosts are infected. Compared with the propagation of a BGP routing worm, the importance-scanning worms spread out faster before the victim population becomes saturated.

Figure 2(c) shows the propagation comparison among an optimal importance-scanning Slammer, a suboptimal importance-scanning Slammer, a Class A routing Slammer, and a random Slammer, with the parameters $N = 75,000$, $s = 4,000$ /second, and $I_0 = 10$ as in [3]. The suboptimal importance-scanning Slammer can propagate nearly twice faster than the Class A routing Slammer.

In regard of storage requirement for group distribution information, each $p_g(i)$ needs 4 bytes and each /8 prefix needs 1 byte. Therefore, the total number of bytes added is $5 \times 256 = 1280$. We can reduce this payload by removing the entries with $p_g(i) = 0$, where $i = 0, 1, \dots, 255$. Since there are only 102 entries with non-zero $p_g(i)$ according to the empirical distribution $p_e(i)$ in Figure 1(b), the table can be stored in a $102 \times 5 = 510$ bytes payload. Hence, the scanning rate of importance-scanning worms will not decrease much.

IV. GAME THEORY FOR ATTACKERS AND DEFENDERS

When an application is introduced to the Internet, defenders can choose how to deploy this application in networks. That is, the group distribution $p_g(i)$ can be controlled by the defenders. Then there would be a game between attackers and defenders. The attackers attempt to maximize infection speed (characterized by infection rate α) by choosing the optimal group scanning distribution $p_g^*(i)$, while the defenders endeavor to minimize worm propagation speed by customizing the group distribution $p_g(i)$. Thus, a *minmax* strategy for the

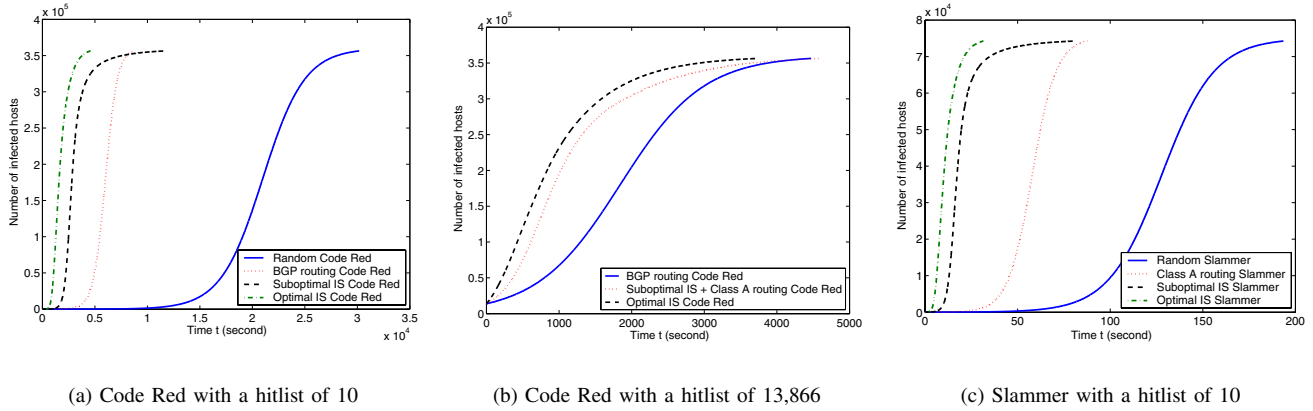


Fig. 2. Code Red and Slammer worms propagation comparisons.

defenders and a *maximin* strategy for the attackers lead to optimal strategies and value in Equation (12):

$$\alpha_{opt} = \min_{p_g(i)} \max_{p_g^*(i)} \{\alpha\} = \max_{p_g^*(i)} \min_{p_g(i)} \{\alpha\}. \quad (20)$$

This is a classical two-person zero-sum game [7]. Since the optimal choice of $p_g^*(i)$ is

$$p_g^*(j) = \begin{cases} 1, & j = \arg \max_k \{v_k\}; \\ 0, & \text{otherwise,} \end{cases}$$

which leads to $\alpha = sN \max_k \{v_k\}$, $p_g(i)$ needs to minimize $\max_k \{v_k\}$. This yields $v_1 = v_2 = \dots = v_m = \frac{1}{\Omega}$ and $p_g(i) = \frac{\Omega_i}{\Omega}$. That is, the defenders should deploy the application uniformly in the entire IP-address space.

V. CONCLUSIONS

In order to defend effectively against Internet worms, we need to study potential scanning techniques that attackers may employ. In this paper, we present a new scanning method called *importance scanning*, which can use vulnerable-host distribution information to increase worm propagation speed. This scanning strategy can be combined with other scanning methods such as hitlist scanning. It is noted that when the naming distribution information is exploited, the importance scanning can also be applied to DNS worms [3], which is worth further investigating. Moreover, when IPv4 is updated to IPv6, an importance-scanning worm will not be slowed down much if vulnerable hosts are still distributed in a clustered fashion. A game theory approach suggests that the best strategy for defenders is to evenly distribute the applications in the entire address space.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their valuable comments. Support from NSF ECS 0300605 is gratefully acknowledged.

REFERENCES

- [1] Z. Chen, L. Gao, and K. Kwiat, "Modeling the Spread of Active Worms," in *Proc. of INFOCOM 2003*, San Francisco, April, 2003.
- [2] D.J. Daley and J. Gani, "Epidemic Modelling: An Introduction," *Cambridge University Press*, 1999.
- [3] H. Feng, A. Kamra, V. Misra, and A. D. Keromytis, "The Effect of DNS Delays on Worm Propagation in an IPv6 Internet," in *Proc. of INFOCOM 2005*, March 2005.
- [4] P. Heidelberger, "Fast Simulation of Rare Events in Queueing and Reliability Models," *ACM Transactions on Modeling and Computer Simulation*, vol.5, no.1 pp. 43-85, Jan. 1995.
- [5] D. Moore, C. Shannon, and J. Brown, "Code-Red: A Case Study on the Spread and Victims of an Internet Worm," *ACM SIGCOMM/USENIX Internet Measurement Workshop*, Marseille, France, Nov 2002.
- [6] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the Slammer Worm," *IEEE Security and Privacy*, 1(4):33-39, July 2003.
- [7] G. Owen, "Game Theory," *Academic Press*, New York.
- [8] M. Abu Rajab, F. Monrose, and A. Terzis, "On the Effectiveness of Distributed Worm Monitoring," to appear in *Usenix Security 2005*.
- [9] C. Shannon and D. Moore, "The Spread of the Witty Worm," *IEEE Security and Privacy*, vol. 2 No 4, Jul-Aug 2004, pp. 46-50, Aug 2004.
- [10] P.J. Smith, M. Shafi, and H. Gao, "Quick Simulation: A Review of Importance Sampling Techniques in Communications Systems," *IEEE Journal on Selected Areas in Communications*, vol.15, pp.597-613, May 1997.
- [11] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in Your Spare Time," in *Proc. of the 11th USENIX Security Symposium (Security '02)*, 2002.
- [12] S. Staniford, D. Moore, V. Paxson, and N. Weaver, "The Top Speed of Flash Worms," in *Proc. ACM CCS WORM*, October 2004.
- [13] N. Weaver, "Warhol Worms, the Potential for Very Fast Internet Plagues," <http://www.cs.berkeley.edu/~nweaver/warhol.html>.
- [14] J. Wu, S. Vangala, L. Gao, and K. Kwiat, "An Effective Architecture and Algorithm for Detecting Worms with Various Scan Techniques," in *Network and Distributed System Security Symposium*, 2004.
- [15] S. Xing and B.-P. Paris, "Measuring the Size of the Internet via Importance Sampling," *IEEE Journal on Selected Areas in Communications*, 21(6), pages 922-933, August 2003.
- [16] C. C. Zou, L. Gao, W. Gong, and D. Towsley, "Monitoring and Early Warning for Internet Worms," *10th ACM Conference on Computer and Communication Security (CCS'03)*, Oct. 27-31, Washington DC, USA, 2003.
- [17] C. C. Zou, D. Towsley, W. Gong, and S. Cai, "Routing Worm: A Fast, Selective Attack Worm based on IP Address Information," *19th ACM/IEEE/SCS Workshop on Principles of Advanced and Distributed Simulation (PADS'05)*, June 1-3, Monterey, USA.
- [18] Distributed Intrusion Detection System (DShield), <http://www.dshield.org/>.
- [19] Internet Protocol V4 Address Space, <http://www.iana.org/assignments/ipv4-address-space>.