Modeling and Analysis of Worm Interactions (War of the Worms)

Sapon Tanachaiwiwat
Ming Hsieh Department of Electrical Engineering
University of Southern California, CA
tanachai@usc.edu

Ahmed Helmy
Computer and Information Science and Engineering
University of Florida, FL
helmy@ufl.edu

Abstract—“War of the worms” is a war between opposing computer worms, creating complex worm interactions as well as detrimental impact on infrastructure. For example, in September 2003 the Welchia worms were launched to terminate the Blaster worms and patch the vulnerable hosts. In this paper, we try to answer the following questions: How can we explain the dynamic of such phenomena with a simple mathematical model? How can one worm win this war? How do other factors such as locality preference, bandwidth, worm replication size and reaction time affect the number of infected hosts? We propose a new Worm Interaction Model (based upon and extending beyond the epidemic model) focusing on random-scan worm interactions. We also propose a new set of metrics to quantify effectiveness of one worm terminating other worm. We validate our worm interaction model using extensive ns-2 simulations. This study provides the first work to characterize and investigate worm interactions of random-scan worms in multi-hop networks.

I. INTRODUCTION

Since the Morris worm incident in 1988, worms have been a major threat to Internet users. In addition, more and more worms carry destructive payload enabling them to perform distributed denial-of-service attacks, steal username/password or hijack victims’ files. Worms can be categorized as network worms and email worms. Network worms such as Slammer, Witty, and Code Red aggressively scan and infect vulnerable machines. Mass-mailing worms’ propagation such as Kamasutra, Love Bugs, and NetSky rely on social engineering techniques. Several worm propagation models have been proposed [7, 8, 9, 18, 20] to provide the insight into the dynamic of network worm propagations as well as effectiveness of available detection and response mechanisms. However, those worm propagation models have not considered the interaction among different worm types and, as we shall show, are inadequate to model war of the worms.

The war of the worms creates unprecedented dynamic and complex scenarios (fig.1) as well as detrimental impact on infrastructure.

Much of this work was performed at the University of Southern California with support from NSF awards: CAREER 0134650, ACQUIRE 0435505 and Intel.
accurately approximate these metrics with properly chosen network-delay factors.

- We derive the important parameter, *Epidemiological Threshold*, from worm’s scan rate ratio and initial infective ratio to quantify the degree of outbreak and guideline for effective worm containment.

We aim at using our proposed worm interaction model as the foundation of effective security response protocol design that deploys beneficial worms to terminate malicious worms [14]. We explain the necessary components of this protocol in Section VII.

Next we discuss the related work in Section II. We explain the basic epidemic model and related variables in Section III and then in Section IV we explain the basic definitions required for understanding the Worm Interaction Model. After that in Section V, we extend the basic epidemic model to build the Worm Interaction Model for one-sided interaction. In Section VI, we evaluate accuracy of our model as well as the effectiveness of worm termination based on proposed metrics. We conclude our work in Section VII.

II. RELATED WORK

In [2], the authors pointed out that traditional human intervened responses were too slow and the worm-terminating-worm scheme may be practical, since today’s defense (prevention, treatment, and containment) technology still cannot cope with such spreading speed [3, 9, 18]. Moore, Shannon, Voelker and Savage [9] evaluate the effectiveness of automated network-level worm containment based on IP-based blacklisting and signature-based filtering techniques. Interacting worms, however, do not need to filter the opposing worms, because they will remove the replications and vulnerabilities of the opposing worms from the vulnerable hosts. Our work aims to provide a framework explaining worm interactions relating to automated worm generation to counter on-going attacking worm.

In [2], the authors suggested modifying existing worms such as Code Red, Slammer and Blaster to terminate the original worm types. The modified code will retain portion of attacking method so it would choose and attack the same set of susceptible hosts. In this paper, we assume the existence of this technology. Other active defense such as automatic patching was also investigated in [17]; their work assumed the patch server and overlay network architecture. We provide the mathematical model that can explain the behavior of automatic-generated beneficial worm and automatic patch distribution using one-sided worm interaction. In [17] authors assumed patch blocking by worms after infection, and hence this scenario yields different type of interaction [14] from what we emphasize in this paper. Our work limits only to understanding and evaluating automated worm (with patch) generation but we do not touch upon details of vulnerabilities nor related software engineering techniques to generate patches or worms. Active defense using beneficial worms was also mathematically modeled in [12]; however, the author focused only on delay-limited worms with different type of interactions from ours and did not consider network-related factors.

The study in [7, 8, 20] investigated the propagation of the random scan network worm and local scan preference strategy which were used by Code Red II and Nimda. Extending beyond that study, our work emphasizes the effect of preferred-local-scanning strategy as well as underlying network characteristics such as reaction time, worm replication size, bandwidth on worm interactions.

The impact of router’s behavior and background traffic on the propagation dynamics of single worm type has been studied in [11]. This can drastically affect the dynamic of worm interactions. We discuss more on this issue in subsection VI.C.a

III. EPIDEMIC MODEL

From [3], the basic susceptible-infected-recovered (SIR) epidemic model which was developed for the study of biological infectious diseases has been used to explain the behavior of self-replicating network worms [7, 9, 18, 20].

In SIR model, vulnerable hosts fall in one of following states in sequence—susceptible, infected and recovered. Susceptible hosts have never had the disease and can catch it. Infected hosts have the disease and are contagious. Recovered hosts have already had the disease and are immune or cured.

Let \( N \) be the size of vulnerable population, \( I \) be the number of infected hosts at time \( t \), \( R \) be the recovered hosts at time \( t \), \( \beta \) be contact rate i.e. the rate of contact between hosts, \( \gamma \) be the removal or recovered rate and \( S \) which equals to \( N - I - R \) be the number of susceptible hosts at time \( t \). \( \beta \) and \( \gamma \) are assumed constant.

![Fig. 2. SIR Epidemic Models](image)

The fundamental nonlinear ordinary differential equations of SIR model are shown below,

\[
\frac{dS}{dt} = -\beta SI \tag{1}
\]

\[
\frac{dI}{dt} = \beta SI - \gamma I \tag{2}
\]

\[
\frac{dR}{dt} = \gamma I \tag{3}
\]

The transitions of states are shown in fig.2. An arrow represents a flow from one state to another. This model does
not consider the birth/death of population as well as the spatial distribution of susceptible hosts.

From (2), the epidemic is sustainable only if \( \frac{dI}{dt} > 0 \) which requires \( \frac{BS}{\gamma} > 1 \); such important ratio is called
Epidemiological Threshold,
\[
E_0 \equiv \frac{BS}{\gamma}. \tag{4}
\]

We would use SIR model as the foundation of our proposed Worm Interaction Model which we will discuss in Section V.

IV. DEFINITIONS

In previous section, we have discussed general definitions of SIR model. However we need additional elementary definitions and concepts that we use in Worm Interaction Model. We start by examining the important concept of the predator-prey relationships.

A. Predator-Prey Relationships

For every worm interaction type, there are two basic characters: Predator and Prey. The Predator, in our case the beneficial worm, is a worm that terminates and patches against another worm. The Prey, in our case the malicious worm, is a worm that is terminated or patched by another worm.

A predator can also be a prey at the same time for some other type of worm. Predator can vaccinate a susceptible node, i.e., infect the susceptible node (vaccinated nodes become predator-infected nodes) and apply a patch afterwards to prevent the nodes from prey infection. Manual vaccination, however, is performed by a user or an administrator by applying patches to susceptible nodes.

A termination refers to the removal of prey from infected nodes by predator; and such action causes prey-infected nodes to become predator-infected nodes. The removal by a user or an administrator, however, is referred to as manual removal.

We choose to use two generic types of interacting worms, A and B, as our basis throughout the paper. A and B can assume the role of predator or prey depending on the type of interactions.

B. Worm Life-cycle

Fig.3 illustrates the basic life cycle of predator and prey. Predator and prey search for susceptible hosts by using either TCP (such as Code Red and Code Red II) or UDP exhaustive scan (such as Slammer and Witty). Unlike UDP-scan worm, TCP-scan worms need to wait for its responses from valid destinations. The waiting time makes its scan rate much slower than the scan rate of UDP-scan worms.

Only predator (fig.3(a)) needs to check whether prey resides in the same host before it can terminate prey (fig.3(b)) and patch the host to prevent reinfection from prey (if patch available). Both types terminate itself after predefined timeout unless it has been terminated by opposing worm or manual removal process. For example, Welchia has embedded timeout which it will disable and terminate itself if the year from computer system’s date is 2004 [16]. The reason for self-termination is for reducing unnecessary workload and traffic on the host infected by Welchia. For simplicity, in our Worm Interaction Model (Section V), we assume that the patch is contained within predator’s payload and the time-out periods for both worms are indefinite. Our model also assumes that predator always detects prey if prey already infects the vulnerable host. Both predator and prey are subject to vaccination and manual removal.

C. Contact rate

As explained in subsection IV.B, each worm constantly scans the vulnerable hosts by issuing new worm replication to randomly chosen addresses. Let \( P_v \) be the probability of worm replication having a contact with a vulnerable host from the total address space \( \nu \) i.e. for IPv4 is \( 2^{32} \). We define a contact as a worm replication reaching a destined vulnerable host. Let \( P_e \) be the fraction of a vulnerable hosts reached by a worm replication,
\[
P_v \equiv \frac{N}{\nu}. \tag{5}
\]
where \( 0 \leq P_v \leq 1 \) and \( 0 \leq P_s \leq 1 \).

A worm replication can be significantly slowed down by network delay \( D \) including transmission delay, link delay, processing delay and queuing delay. Let \( \rho_A \) and \( \rho_B \) be the network-delay factor which attenuates contact rate of prey and predator.

Let \( v_A \) and \( v_B \) be the respective scanned address space of prey and predator, and let \( r_A \) and \( r_B \) be the respective scan rate of prey and predator where a scan rate is a frequency of a worm issuing its replication to chosen destinations. Thus the contact rate of prey \( \beta_A \) and the contact rate of predator \( \beta_B \) are

\[
\beta_A = \rho_A r_A v_A P_s = \frac{\rho_A r_A}{v_A} (7-a)
\]

\[
\beta_B = \rho_B r_B v_B P_s = \frac{\rho_B r_B}{v_B} (7-b)
\]

where \( 0 \leq \rho_A, \rho_B \leq 1 \).

Let \( D_A \) and \( D_B \) be respective network delay for prey and predator. We can derive \( \rho_A \) and \( \rho_B \) as follows:

\[
\rho_A = \frac{1/r_A}{1/r_A + D_A} = \left(1 + r_A D_A\right)^{-1} (7-c)
\]

\[
\rho_B = \frac{1/r_B}{1/r_B + D_B} = \left(1 + r_B D_B\right)^{-1} (7-d)
\]

The contact rate can be dynamic according to the network congestion (network delay), adaptive scan rate [14] and change of scanned address range.

Let \( e \) be number of targeted sub networks. For sub network \( j \), let \( h_{aj} \) and \( h_{bj} \) be the probability of network \( j \) being scanned for prey and predator, \( g_A \) and \( g_B \) be the worm replication size for prey and predator, \( q_{aj} \) and \( q_{bj} \) be average queue length of outgoing links for prey and predator, \( b_{aj} \) and \( b_{bj} \) be average bandwidth of outgoing links for prey and predator, \( c_{aj} \) and \( c_{bj} \) be average packet drop rate for prey and predator and \( u_{aj} \) and \( u_{bj} \) be average link delays for prey and predator. We can derive \( D_A \) and \( D_B \) as follows:

\[
D_A = \sum_{j=1}^{e} \left(h_{aj}(1-c_{aj})(u_{aj} + \frac{g_A(q_{aj} + 1)}{b_{aj}})\right) (7-e)
\]

\[
D_B = \sum_{j=1}^{e} \left(h_{bj}(1-c_{bj})(u_{bj} + \frac{g_B(q_{bj} + 1)}{b_{bj}})\right) (7-f)
\]

Hence, according to their strategies, \( D_A \) and \( D_B \) can be drastically different between prey and predator. Note that \( q_{aj}, q_{bj}, c_{aj} \) and \( c_{bj} \) are subject to background traffics also.

### D. Worm Interaction Ratios

To estimate how much relative characteristics of predator and prey impact on their propagations, we propose following worm interaction ratios: scan rate ratio, initial infective ratio. We further develop the concept of similarity and difference to gain insight into relationships between scan rate ratios, and between initial infective ratios. We also systematically derive minimum scan rate ratio and minimum initial infective ratio for effective termination in Section V.

#### a. Scan rate ratio

Scan rate ratio is the ratio of scan rates of one worm type to that of another worm type. Let \( X \) be a scan rate ratio of predator to prey,

\[
X = \frac{r_B}{r_A} (8)
\]

If we assume that \( D_A = D_B = D \approx 0 \) and \( v_A = v_B \) then contact rate ratio of predator to prey can be derived approximately as

\[
\frac{r_B}{r_A} \approx \frac{r_B (1 + r_A D)}{r_A (1 + r_B D)} = \frac{\beta_B}{\beta_A} . (9)
\]

\( X_i \) is similar to \( X_j \) only when \( \frac{r_B}{r_A} = \frac{r_B'}{r_A'} \), otherwise it is said to be different from \( X_j \) where \( i \) and \( j \) represent interacting pairs.

For example, \( X_1 \simeq 1:2 \) is similar to \( X_2 \simeq 2:4 \) but the first ratio has \( r_A = 1/\text{sec}, r_B = 2/\text{sec} \) and the latter has \( r_A = 2/\text{sec}, r_B = 4/\text{sec} \). To differentiate between \( X_1 \) and \( X_2 \), we use scan-rate-ratio multiplicative factor \( k_i \), from above example we have \( k_i = 1.0 \) for \( X_1 \) and \( k_i = 2.0 \) for \( X_2 \). We use \( X = 1:1 \) as the absolute reference. The importance of this concept is shown in Section V.

#### b. Initial Infective ratio

Initial infective ratio is the ratio of infective of one worm type to that of another worm type at initial release time of both worms. Let \( Y \) be an initial infective ratio of predator to prey,

\[
Y = \frac{I_B(0)}{I_A(0)} (10)
\]

where \( I_A(0) \) and \( I_B(0) \) = number of initial infectives of prey and predator respectively at their released times.

This ratio is only valid when there is no difference in launching time of prey and predator. We can also consider the ratio of infected hosts for any \( t \); however, we shall consider the delay of launching the opposing worm i.e. reaction time in the next subsection.

Again \( Y_i \) is similar to \( Y_j \) only when \( \frac{I_B(0)}{I_A(0)} = \frac{I_B(0)}{I_A(0)} \), otherwise it is said to be different from \( Y_j \).
We assume that timeout period in the model is indefinite for both prey and predator. Let \( \gamma_S \) be the manual vaccination rate and let \( \gamma_A \) and \( \gamma_B \) be the removal rate for prey and predator. Hence, the susceptible rate is

\[
\frac{dS}{dt} = -\beta_A S A - \beta_B S B - \gamma_S S . \tag{11}
\]

Since prey relies on susceptible hosts to expand its population, the increase of prey infection rate is determined by the contacts of the susceptible hosts and prey infectives. The decrease of prey infection rate is determined by manual removal and prey termination caused by the contacts of prey infectives and predator infectives whose contacts are initiated by predator. Hence the prey infection rate is

\[
\frac{dI_A}{dt} = \beta_A S A - \beta_B I_B B - \gamma_A I_A . \tag{12}
\]

Because predator can terminate prey as well as vaccinate susceptible hosts, the increase of predator infection rate is determined by the contacts of predator with either the susceptible hosts or prey infectives whose contacts are initiated by predator. The decrease of predator infection is, however, determined only by manual removal. Thus

\[
\frac{dI_B}{dt} = \beta_B S B + \beta_B I_B B - \gamma_B I_B . \tag{13}
\]

The removal rate is simply the sum of vaccination of susceptible hosts, manual removal of both prey and predator.

\[
\frac{dR}{dt} = \gamma_S S + \gamma_A I_A + \gamma_B I_B . \tag{14}
\]

From (12), the epidemiological threshold for prey is

\[
E_A = \frac{\beta_A S A}{\beta_B I_B B + \gamma_A I_A} = \frac{\beta_A S}{\beta_B I_B B + \gamma_A} . \tag{15}
\]

If we want prey to be contained by predator i.e. \( E_A < 1 \) at \( t=0 \), and we assume \( \gamma_A = \gamma_B = 0 \) (because of user’s unawareness of worms’ presence at the early stage of propagation) and \( \beta_A, \beta_B, I_A(0), I_B(0) > 0 \), we requires that the minimum scan rate ratio to be

\[
\frac{\beta_B}{\beta_A} = X > \frac{S(0)}{I_B(0)} = \frac{S(0)}{YI_A(0)} . \tag{16}
\]

or minimum infective ratio to be
\[ Y > \frac{S(0)}{XI(0)}. \] (17)

From (13), the epidemiological threshold for predator is
\[ E_B = \frac{\beta_B S I_B + \beta_B I_B I_B}{\gamma_B I_B} = \frac{\beta_B(S + I_B)}{\gamma_B}. \] (18)

With the same assumption as for (16), predator can always spread with \( \gamma_B = 0 \) (\( E_B = \infty \)). Note that after predator or prey terminate themselves because of the time out (as shown in Section IV.B), \( I_a \) and \( I_A \) become zero.

To see the importance of scan rate ratio and initial host ratio, we plot numerical solutions from our aggressive one-sided interaction model using four sets of variables in this model: (1) similar scan rate ratios with a fixed initial host ratio where \( k_1 = 1, k_2 = 2, \ldots, k_5 = 5 \). (fig. 5(b)), (2) similar initial host ratios with a fixed scan rate ratio where \( l_1 = 1, l_2 = 2, \ldots, l_5 = 5 \). (fig. 5(d)), (3) different scan rate ratios with a fixed initial host ratio (fig. 5(a)), and (4) different initial host ratios with a fixed scan rate ratio (fig. 5(c)).

This can be observed from fig.5 that the change of predator scan rate causes multiplicative change in prey infection rate while the change of predator initial infected host ratio only causes additive changes in prey infection rate. In other words, change of predator scan rate is repetitively applied to prey infection rate while change of initial predator infected hosts is only applied once to prey infection rate (at initial release). Note that extremely large \( X_i \) can cause adverse effect on the predator contact rate because of excessive network congestion (with large \( D_B \) and hence low \( \beta_B \)) and possibly a router outag[e [11].

Other important finding here in fig.5 (b) is that similar scan rate ratios with different \( k_i \) in fixed population yields equal maximum of prey infectives; however, times that take to reach the exact number of infectives between similar scan rate ratios decrease proportionally with the increase of \( k_i \). To be specific, we shall explain an example from fig.5 (b) where \( X_i = 1:1, X_2 = 2:2, \ldots, X_5 = 5:5 \). We can observe that the time required to reach maximum of prey infectives for \( X_i \) is 1.3185 seconds which is 0.2 (or \( k_i/k_j \)) of time required to reach maximum of prey infectives for \( X_j \) which is 6.5927 seconds. This observation is also applied to any other \( X_i \). This also hints that automated generating worm that preserve the characteristic of original worm i.e. same scan rate will optimally limit the maximum prey infectives to 20% of vulnerable populations no matter what original scan rate is.

Furthermore, this observation also applies to sets of similar initial infective ratio for equal maximum of prey infectives as shown in fig.5 (d). However, in this figure, we keep the ratio of susceptible hosts to initial predator infectives to initial prey infectives similar, e.g. \( X=1:1 \) with \( S=998, Y=2:2 \) with \( S=1996, \ldots, Y=5:5 \) with \( S=4990 \). We can observe that all infective fractions of prey for different \( l_i \) overlapping exactly at the same positions for all \( t \). This means that the number of total vulnerable hosts does not affect the relative fraction of infections as long as the ratio of susceptible hosts to predator infectives to prey infectives are similar.

To understand the unique characteristics of the worm interactions better, in fig.6, we compare the instantaneous prey infectives of the SIR model using 6 different removal rates with those of the aggressive one-sided interaction model (\( \gamma = 0.004, 0.002, 0.04, 0.02, 0.4 \) and \( \beta = 0.004 \)). For the aggressive one-sided interaction model (\( \beta_A = \beta_B = 0.004 \)), we assume that manual removal is negligible (\( \gamma_A = \gamma_B = 0 \)). We can see much shaper drop of instantaneous prey infectives in aggressive one-sided interaction model than those of SIR model. Those chosen removal rates in the SIR model are constant multipliers with the prey infectives; while the aggressive one-sided interaction model has an increasing multiplier which is the product of the predator infectives and the predator contact rate.
Earlier we make implicit assumption that predator is launched at the same time as prey is (reaction time = 0). However, the realistic reaction time which is the time after prey is launched can be significant based on how predator is generated i.e. manually or automatically. Automated predator generation should be several orders faster than the manual predator generation is. Hence reaction time depends significantly on how soon prey is detected and how early predator is generated.

For non-zero reaction time, when predator has not been generated yet at time \( t \), the prey infection rate \( \frac{dI_A}{dt} = \beta_A SI_A \) (we assume that manual removal rate is negligible during this period). After a first predator generated at time \( t+\Delta t \) (given reaction time = \( \Delta t \)) then \( \frac{dI_A}{dt} = \beta_A SI_A - \beta_B I_A \). Beyond this point and before the timeout of both prey and predator, the prey infection rate should simply follow (12). The impact of reaction time is investigated in depth in VI.C.a.

VI. SIMULATION RESULTS

To validate our Worm Interaction Model, we investigate network worm propagation using ns-2 simulation [10]. Instead of using high-level simulation, we choose ns-2 because we want to investigate the impact of network congestion caused by worm propagations on the worm interactions.

Our main goal is to verify the accuracy of our mathematical model and have better understanding of worm interaction in a rich set of environments. We choose the Slammer-like worm characteristic as basic behavior of a worm in this simulation. Slammer is chosen because, despite its simplicity, it still holds the world record of fastest-spread worm yet [6].

We simulate prey (A) and predator (B) which may have different scan rates and initial infectives. Even we consider the manual removal earlier in the model; we do not model human interaction because we only want to focus on worm interaction.

We simulate 1000 vulnerable hosts in following topologies:

1. Star-shaped topology: we want to test our model with network having small number of hops (1-2 hops) that has moderate constant bandwidth (512 kbps) and constant delay (1ms) between hosts.

2. Transit-stub topology: this two-level topology will help us test our model with bottleneck network having large number of hops (1-4 hops) that has moderate bandwidth (512 kbps access-link, 10 Mbps local network) and delay between hosts (average 1 ms). There are 10 local networks; each local network has 100 hosts with one of them acting as a router. One AS can have one or two local networks. The links between routers in this topology are generated by BRITE Internet topology generator [1].

Each worm use UDP scan to transfer worm replication to random chosen vulnerable hosts of the network. The default packet size is 404 bytes which similar to Slammer worm. The results are based on the average of at least 10 simulation runs.

A. Model Accuracy

In this section, we compare the simulation results with our proposed model. We use the \( X = 2:2 \) for every model evaluation. We assume \( P_v = 1.0 \) (the perfect probability of worm replication having a contact with a vulnerable host from the total address space \( \nu \) i.e. the vulnerable host address range is exactly the same as scan address range) for both prey and predator. This assumption makes the worm propagate much faster than the real worm propagation in the Internet. We will investigate the scenario that \( P_v \) is much less than 1.0 in subsection VI.C.

Fig.7 shows that our numerical solution of our aggressive one-sided interaction mathematical model with \( \rho_a=0.65 \) and \( \rho_b=0.65 \) closely matches with simulation results, especially to the results of worm interaction in transit stub topology. Although, from our simulation, using equation 7(c) and 7(d), our estimation for \( \rho_a \) and \( \rho_b \) of transit stub topology around 0.64 based on \( D_a \) and \( D_b = 0.284 \) second derived from 90% of scanning traffics going outside the local network and average queue size is 49 for outbound link.

Hence, this figure suggests that different topologies can yield different interaction characteristics. The reason that simulated propagation in transit stub topology is slower than that of star topology (estimated from simulation, \( \rho_A=0.98 \) and \( \rho_B=0.98 \), \( D_A \) and \( D_B = 0.0084 \) second) because of higher average number of packets in queues and higher average number of hops; even local network bandwidth in the transit stub topology is much higher than bandwidth of star topology but, based on address range, local network traffic only accounts for 10% of all traffic. When compared with simulation results in the transit-stub topology, our model estimation for prey infectives are almost perfect (<1% error) for aggressive one-sided interaction.

Such congestion of outbound scan can be significantly reduced if prey and predator scan local-network addresses more than outside-network addresses [10]. To see the advantage of choosing local network preference effectively, we will investigate the effect of local network preference on worm interaction in subsection VI.C.

![Fig. 7. Comparison between mathematical model and simulation results from ns-2 star, and transit-stub topologies](image-url)
B. Effectiveness of Termination

The effectiveness of termination is defined as a measure of how efficient predator can terminate prey in aggressive one-sided interaction. In other words, how much damage does prey causes to the network after predator is launched to terminate prey.

We quantify such effectiveness as two important characteristics of prey:
- Total infectives \( T \): the number of hosts ever infected by prey including prey infectives that have been removed. Total infectives of prey can be derived from
  \[
  T = \int_{t=0}^{\infty} \beta_0 S I_d dt .
  \]  

  We have mentioned maximum infectives earlier but \( T \) is the real damage while the maximum infectives is only the maximum of instantaneous number of prey infectives at that time. From (12), we can see that \( T \geq \max \) maximum infectives.
- Individual life span \( L \): the time between the start of infection and the end of infection i.e. infectious period for individual replication of prey caused by prey termination. Individual life span is derived from weighted average of instantaneous life span that weighs on number of terminated node \( w_i \) by predator at that time instant.
  From (13), we can derive the instantaneous life span at time \( t \) of each terminated prey from the inverse of product of predator contact rate and predator infectives at time \( t \), i.e. \( \frac{1}{(\beta_0 I_B)} \). The sum of product of that term and number of terminated prey gives the total life span for all prey infectives, and hence we divide that term by total infectives to get average individual life span.
  \[
  L = \frac{1}{T} \sum_{t=0}^{\infty} W_i (\frac{1}{(\beta_0 I_B)}),
  \]  

  where \( W_i = (\Delta \beta_0 I_A I_B) - \Delta I_A \)

  We will investigate the effect of deployment scenarios on these metrics as well as verify the accuracy of the model with metrics observed from the simulation next.

C. Deployment Scenarios

We simulate the aggressive one-sided worm interaction in different deployment scenarios which we shall discuss next. We deploy two worms in transit-stub topology as shown earlier. In addition, now we have 3 outbound bandwidths for the same transit-stub topology: 512 kbps, 1 Mbps and 2 Mbps.

In subsection VI.A we have shown that our model shows good approximation if we can estimate network-delay factor properly.

Here, we focus on the imperfection caused by deployment scenarios that affect the aggressive one-sided interaction. Effect of following factors on our proposed metrics \( T \) and \( L \) are investigated. We still assume \( P_v = 1.0 \) (the vulnerable host address range is exactly the same as scan address range) for prey and predator in subsection VI.C.a and VI.C.b.

a. Reaction time

The reaction time is the time required before launching predator by automated worm-generation or programmer. Automated worm [2] or patch generation [9] should take much less time compared to manual process. Our model assumes that predator starts scanning immediately as soon as it vaccinate susceptible hosts or terminate prey; hence the delay of patch applying will not be considered in our model. Furthermore, we also assume that patch takes effect instantly without need of rebooting the machine and both predator and prey have the same scanning rate.

As shown in fig. 8 (a), prey individual life span grows exponentially with reaction time. In addition, the lower outbound bandwidth causes drops in contact rate of both prey and predator. Hence the slow down of worm interaction causes prey to survive longer.

![Fig. 8. Effect of reaction time on (a) prey individual life span and (b) total infectives](image)

In fig. 8 (b), the reaction time causes significant increase of total infections only in low reaction time range. We know that three outbound bandwidth links have the same maximum infective (because of same scan rate ratios) implying equal of total prey infectives.

Furthermore because the total infectives is the non-decreasing function of time, for slower reaction time, such number can be extremely high, but since it cannot grow beyond the number of vulnerable hosts and hence, it causes the gradually drop in the growth rate of total infectives and finally saturate at level of population size.

Surprisingly, even reaction time is zero with equal scan rate \( X=1.0 \); the prey’s total infectives can be as high as 35% of total population (if \( X=2.0 \), prey’s total infectives is only 2% of total population) and individual life span is 1.5 second. To contain the prey within 50% total
infectives, the predator’s reaction time needs to be less than 0.5 second.

From the model with reaction time = 0, using $\rho_A = 0.65$ and $\rho_B = 0.65$, our estimations in worm interaction model are only off by 4% for simulated total infected hosts and 9% for the simulated infectious period when varying reaction times. We can observe that our model has better approximation for total infectives in the slow reaction time range. We expect the closer estimation if we consider the drop of contact rate modeling caused by congested outbound link [19]. Note that with large reaction time, prey’s random scan can cause the router outage [11]. Thus it may prevent predator from reaching majority of hosts and significantly slow down the predator infection.

b. Worm replication size

This factor is a transmission overhead reflecting the efficiency of coding and compression technique that automatic generation or programmer uses. Some examples of worm replication sizes and compression mechanisms can be found [13, 16]. We assume that the predator’s reaction time is zero for this part of experiment.

The increase of worm replication size in fig. 9 (a) causes the linear increase of prey’s individual life span; the effect is clearly seen in highly congested access-linked networks i.e. 512 kbps link (50% individual life span increase with 200% increase of predator replication size). The increase of replication size slightly increases total infectives (fig. 9(b)) even in the highly congested access-linked networks. It suggests that the delay caused by any size of packet slows down on both prey replication and predator replication.

From the model, if patch size is 1212 bytes in 512kbps access link, with $\rho_A = 0.6$ and $\rho_B = 0.575$, our estimations in worm interaction model are only off by 8% for simulated total infected hosts and 9% for the simulated infectious period when varying worm replication size.

c. Local preference

By focusing on scanning the hosts in the same subnet addresses, worm can avoid scanning invalid addresses and reducing the packet drops caused by bottleneck of outbound access of local network to the Internet. Since majority of addresses in IPv4 are not fully utilized, appropriate local preference must be an important factor. Many network worms already try to minimize the probability of scanning invalid addresses by using this technique [8, 16]. In particular, the local preference has significant impact on the probability of worm replication having a contact with a vulnerable host, based on local preference as follows.

$$P_c = \sum_i f_i p_i$$

where $f_i$ is the fraction of strategy $i$ being used, $p_i$ is the probability of worm replication having a contact with a vulnerable host address from the chosen address space $\nu_i$ of strategy $i$ which has $n_i$ vulnerable hosts in such address range. For example, $\nu_i$ of worm preferring vulnerable hosts in the same /16 address has address size = $2^{16}$ instead of $2^{32}$, furthermore, if in that address range, there are only $2^{14}$ vulnerable hosts then $p_i = 0.25$.

In our simulation, each worm uses 2 strategies and the valid addresses only occupy 10% of its scanned address range (earlier we assume that all the scanned addresses are valid). Two strategies are assigned as following:

1. (1) random-scan ($p_1 = 0.999$ with $f_1$)
2. (2) preferred-local-scan ($p_2 = 0.99$ with $f_2$).

We assume that all of reaction time is zero and the worm replication size of both worms is 404 bytes.

In Fig.10 (a) and (b), we vary $f_1$ and $f_2$ of both worms to see how this factor affects the interaction between two worms. Given prey local preference = 0 (with purely random scanning strategy), predator can terminate prey effectively according to individual life span if its local preference is between 0.6 and 0.8.

![Fig. 9. Effect of worm replication size on (a) prey individual life span and (b) total infectives](image)

(a) (b)

![Fig. 10. Effect of local preference on (a) prey individual life span and (b) on total infectives](image)

(a) (b)

The total infectives of the prey can be limited to as low as 2.2% of the total population with predator local preference = 0.6. However if prey local preference is 0.5, predator can merely, at best, contain prey at 80% of the total population.

Hence predator should always utilize the local preference appropriately since majority of worms in the Internet exploit local preference scanning as well.

Our estimations in worm interaction model are off by 36% for simulated total infected hosts and 33% for the simulated individual life span when varying local preference. We expect these errors occur because we have not incorporated the reduction of scanned address space caused...
by excessive predator’s local scanning strategy. Predator can also further enhance the contact rate by using the knowledge of domain such as the currently used address space [8] and the topology of networks [7]. This is subject to future investigation.

VII. SUMMARY AND FUTURE WORK

Based on our worm interaction study, we find that worm interaction causes drastic change in the worm propagation model. Such interaction cannot be explained by earlier works based on the epidemic model even when the removal process is used. We mathematically model and explain aggressive one-sided interaction. In addition, we propose the important metrics to quantify the effectiveness of the worm termination which can also be used as a guideline for other security responses. The predator reaction time reduction and the effective local preference utilization are the key factors to contain the number of prey total infectives.

Our new worm interaction model is validated through extensive simulations. We find that scan rate ratio has much more impact on worm propagation pattern than initial infected host ratio does. With similar scan rate ratios, it always results in the same maximum prey infectives. Our model shows promising accuracy in estimating individual life span and total infectives for different scenarios.

Our worm interaction models can be used as a major component in designing an effective protocol of controlled worm deployment to counter ongoing worm attacks. For example, if there are 50,000 vulnerable hosts, and \( p_1 = 1E-2, \ p_2 = 1E-6, \ f_1 = f_2 = 0.5 \) and prey scan rate is 100/sec with same network-delay as our simulated transit-stub networks with reaction time = 30 seconds, to contain prey total infectives within 70%, 50% or 20% of total vulnerable hosts, predator requires scan rate at least 500/sec, 1000/sec, or 1500/sec, respectively.

Using our model can avoid deploying underestimated predator’s scan rate which causes predator to lose to prey (high total infectives and long individual life span of prey), or overestimated predator’s scan rate which causes excessive congested networks deteriorating both regular traffic and predator contact rate. Our model, however, does not incorporate the router misbehavior when experiencing excessive scan from prey as well as the background traffic which affects network delay factor. We also have not modeled other in-place security defenses.

The basic architecture supporting our approach should at least contain prey detection, predator generation, predator updates, inter-predator communication protocol and prey/predator network access control. Prey detection is the first necessary component to activate the predator generation. Hence prey detection needs to have low false negative. Host-level [13] and network-level detection [8] can be tightly integrated. After predator is launched, predator infectives may collaborate among others to reduce overlapped scanning address space using inter-predator protocol. Distributed non-overlapped sequential scan can be deployed if knowledge of AS is known to predator. Other infrastructure-based defenses such as [3, 9] must also be aware of and adapt to predator’s characteristics. Our future work will focus on such architecture and protocol design and evaluation to support this new security paradigm.

REFERENCES

[1] BRITE: Boston Representative Internet Topology Generator