Modeling Host-based Detection and Active Worm Containment

Frank Akujobi, Ioannis Lambadaris (fakujobi, ioannis@sce.carleton.ca)
Department of Systems and Computer Engineering
Carleton University, Ottawa, ON, Canada

Evangelos Kranakis (kranakis@scs.carleton.ca)
Department of Computer Science
Carleton University, Ottawa, ON, Canada

Technical Report: SCE-07-10

October 23, 2007

Abstract

Recent advancements in Internet worms propagation techniques has generated interest in the development of appropriate defense techniques against such worms. Modeling the behaviour of worm defense techniques to better understand and measure their defense capabilities is crucial to developing effective defenses. This paper presents a discrete-time model of our earlier proposed host-based worm detection and collaborative network containment defense technique, which we referred to as the Analytical Active Worm Containment (AAWC) model. The AAWC model captures the protection capability of the proposed technique by modeling the host population protected from fast spreading, scanning intrusion attack such as worms in a large scale network. Analysing the model alongside an existing discrete-time worm propagation model, we demonstrate quantitatively the effectiveness of our proposed detection and containment technique in defending against fast spreading scanning worms. Based on the host-based worm detection technique, we also develop a continuous-time probability model for worm detection interval which uniquely captures the relationship between worm scanning rate and the detection interval of the worm. Further, we investigate the introduction of immunization to our containment technique and show the resultant effect on a vulnerable population under attack using the developed model.

Keywords – Active worms, Intrusion, Detection, Containment, Immunization, Modeling.

1 Introduction

Fast spreading self-propagating malicious programs also known as active worms have been an enduring security threat on the Internet and large enterprise networks. The sophistication of their mode of attack and spread has been the subject of intense recent research and concern to both private and public sector institutions. Active worms propagate autonomously by infecting vulnerable computer systems which then become launching platforms for infecting other computer systems. Botnets and zombie networks created from such automated propagation of malicious code have also generated serious concern. With advanced scanning and propagation techniques, active worms have successfully spread across the Internet in a few seconds [1]. Malicious active worms which cause damage to systems or interfere with normal system operations have been known to account for huge financial losses on networks they traverse. The large-scale nature of their infestation and the severity of the havoc they inflict on vulnerable systems has motivated several efforts to understand their propagation mechanisms as well as contrive techniques for actively defending against their attacks. Active defense mechanisms take the battle to the worm [2] by automatically eliminating, isolating or patching infected systems as well as pro-actively protecting vulnerable uninfected systems from infection. Worm propagation has been likened to biological virus spreading and biological epidemiological models have been used to model their propagation [3][4]. The simple epidemic model (SI model) assumed that the vulnerable population size is constant and that no recovery or death of an infected host is possible. An infected system therefore will remain in an infected state perpetually. Kermack-McKendrick's SIR model [3][4] improved the SI model by considering that some infected host either recover from the infection or die with time. However, this model did not consider human countermeasures or active defense mechanisms which can reduce infection rates and isolate both infected and vulnerable systems. The two-factor model [5] modeled Internet worm propagation by considering human coutermeasures such as patching, physical removal of systems from the network and manual setup of router filters. This model did not account for active defenses which do not rely on human intervention. More analytical models which are not derived from epidemiology have also been developed. The Analytical Active Worm Propagation (AAWP) model [6] is a discrete time model that characterizes the propagation of worms that employ random scanning. In [6] the AAWP model was described as more accurate and realistic compared to the epidemiological model for the following reasons:

- In the AAWP model a host cannot infect other hosts before it is completely infected, but in the epidemiological model a host begins infecting other hosts even before it is completely infected. Therefore the observed propagation speed of the worm and number of infected hosts are different with the two models. The AAWP model uses a more accurate approach.

- The epidemiological model does not consider the time it takes to infect a host. Depending on characteristics such as the size of the worm and vulnerability the worm exploits, scanning worms take a varying amount of time to infect a vulnerable host. The AAWP model takes this time into consideration and [6] shows that the time to infect a host is an important factor in the spread of active worms.

- The AAWP model also takes into consideration the realistic case that a vulnerable host can be scanned by multiple copies of a worm at the same time. The epidemiological model ignores this case.

In this paper, we use the AAWP to model active worm propagation and we develop an Analytical Active Worm Containment (AAWC) model for modeling the effect of active worm defense based on our earlier proposed host based detection and network containment technique [7]. In [7] we proposed a distributed host-based intrusion detection and network-centric containment approach to defend against fast propagating worm attacks and used emulations on a live testbed to demonstrate the effectiveness of the proposed technique. This paper extends our work by presenting the AAWC model, a discrete time model for the proposed technique that shows the automated defense capabilities of the technique in a large scale network. The main contributions of this paper are:

- We developed a discrete-time Analytical Active Worm Containment (AAWC) model for modeling a host population protected from fast spreading, scanning intrusion attack in a large hierarchical network.

- Knowing that the detection interval of a worm attack is a major contributor to the containment time of the worm, we developed a continuous-time probability model that attempts to answer the question - What is the probability that the measured detection interval for a particular deployment of our defense technique is not greater than a certain desirable detection interval? We demonstrate how the scanning rate of a worm affects this probability. The model can be useful to network and security architects who deploy our proposed defense technique in large scale networks.

- We adapted the AAWP model [6] for scanning worms to capture the charateristics of the large scale hierarchical network topology used in developing the AAWC model.

- Using outputs from the AAWC model and the adapted AAWP model we analyzed the behaviour of our proposed collaborative containment technique in the presence of a simulated worm epidemic outbreak.

- Using the AAWC model we demonstrated quantitatively the containment capability of the proposed technique. We define containment of an infectious worm as the complete halting of further spread of the infectious worm to uninfected vulnerable hosts. Therefore, when a fast spreading worm is contained using the proposed containment

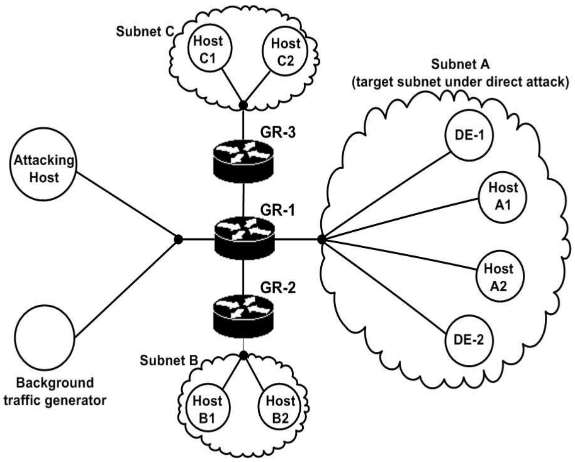

Figure 1: Experimental testbed used in earlier work [7].

technique significant portions of the vulnerable uninfected population are protected from infection.

Our work is unique in modeling a network-centric defense technique that is automatically triggered by distributed detection endpoints running anomaly-based intrusion detection software [7]. The work is also unique in combining outputs from an active worm propagation model and an active worm defense model to demonstrate the performance of the defense technique against a large scale active worm epidemic. Further, the probability model for detection interval developed in this paper uniquely captures the relationship between worm scanning rate and the detection interval of the worm.

The rest of this paper is organized as follows: Section II briefly explains our previous work on host-based intrusion detection and collaborative network-centric containment. In Section III, we develop two models - the discrete-time AAWC model used to show the protection capability of our collaborative containment technique and the continuous-time probability model for detection interval which captures the relationship between the scanning rate of a worm and its detection interval. In Section IV, we review the AAWP model and adapt it to our large-scale hierarchical network topology. In Section V, we use the AAWC model to demonstrate our technique's protection capabilities during a fast scanning worm attack modeled using the AAWP model. In addition, we show the effect of introducing immunization to the AAWC model. Section VI concludes the paper and points to future work.

2 Previous work on host-based detection and network containment of spreading worms

In [7] we proposed a distributed detection architecture that utilizes anomaly-based host intrusion detection (AHID) software running on detector end-points (DEs) located within logical cells as shown in the simple prototype in Fig. 1. During a fast propagating worm invasion, the DEs in the target cell detect the intrusion, alert the gateway router (GR) for the cell and send recorded logs of intrusion attempts to the GR. The GR performs an iterative statistical analysis on the received data to determine the attacking worm's intrusion traffic flow. The GR also runs our *reactive blocking protocol* [7] and implements a containment filter against the intrusion traffic according to the reactive blocking protocol. It then notifies participating peer GRs of the intrusion. The peer GRs in turn implement containment filters against the ingress intrusion traffic thus blocking the traffic from entering all cells existing on the peer GRs. The reactive blocking protocol determines whether or not a worm outbreak is prevalent and continues to spread containment notification to upstream peer routers if an outbreak is prevalent thus achieving collaborative containment. Experimenting on a live test-bed, the GR of the target cell successfully contained an emulated fast spreading worm automatically within 5 seconds of detection on any DE in the target cell. In comparison, recent simulations by Moore et al suggest that an effective worm containment should require a reaction time of well under 60 seconds [8]. Also, employing host-based anomaly detection in combination with statistical correlation of network heuristics, the technique proposed in [7] accurately detected fast spreading worms with zero false alerts since alerts were generated only when verifiable malicious intrusions occurred. Experimental results therefore indicated that our proposed technique can be a viable option for worm detection and rapid automated containment. In this paper, we advance our previous work by developing an analytical model that demonstrates quantitatively the protection capability of the technique. In the absence of a large experimental network, our model is used to simulate the effectiveness of the technique in a large-scale hierarchical network.

3 Modeling Host-based Detection and Active Worm Containment

In this section, we develop a discrete-time model for the behaviour of our proposed host-based detection and collaborative containment technique in a large hierarchical network (Fig. 2 and Fig. 3) which we call the Active Worm Containment (AAWC) model. We decided to use a hierarchical network topology in our analysis since the Internet and most well-designed large enterprise networks generally follow a hierarchical architecture [9][10]. In Fig. 2 and Fig. 3 the nodes represent network routers and the level L

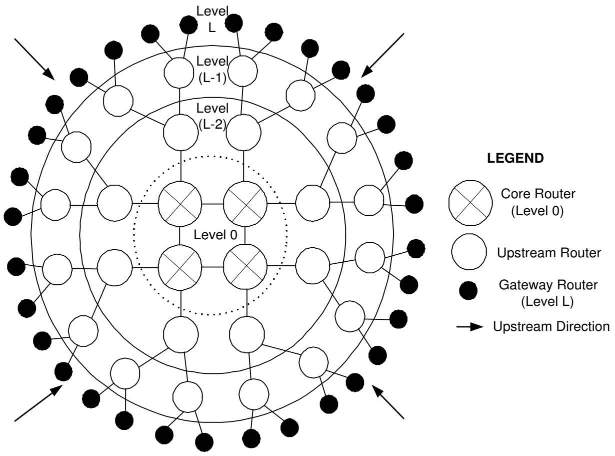

Figure 2: Typical large-scale hierarchical network topology.

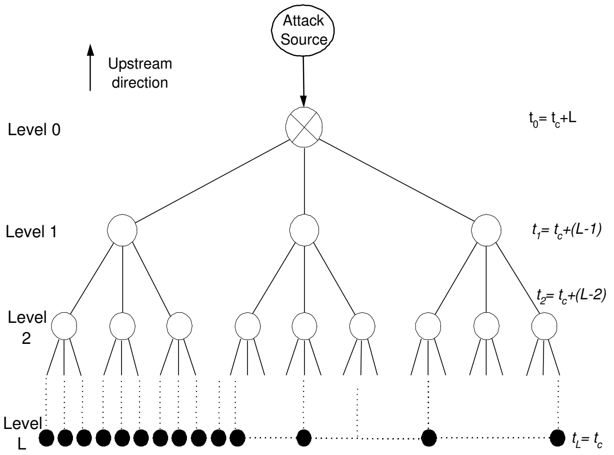

Figure 3: Hierarchical network topology used in AAWC modeling. $t_\delta = 1$.

nodes represent GRs which contain cells [1]. Nodes upstream of level L do not contain cells. When a GR implements a blocking filter against an intrusion traffic, the filter is applied to all cells contained in the GR hence protecting them from the intrusion traffic. Also, when an upstream router (UR) implements a blocking filter against an intrusion traffic, all cells contained in GRs downstream of that UR are protected from the intrusion traffic. Also, all hosts within the protected cells are considered to be contained hosts. While we are aware that our hierarchical network topology does not capture the exact topology of some production networks, it depicts the general topology of most large well-designed networks [9][10] and is sufficient to demonstrate the performance of our technique against large scale worm epidemics.

Table 1: Parameters for network topology and AAWC model

Notation	Explanation
L	number of hierarchical levels in network
y	number of nodes that connect to an upstream node
Q	number of hosts in each cell
s	number of Q-sized cells that exist on each GR
t_r	time intrusion traffic is released into the network
t_d	time DEs in the target cell detect the intrusion attempt
t_c	time GR for target cell implements a containment filter
t_δ	time interval for notification between nodes
N_i	total number of contained hosts after containment at level i

3.1 Discrete-time Modeling of Active Worm Containment

Using notations in Table I, we assume a threat model in which a single attack source releases malicious intrusion traffic at time t_r, targeted at hosts in the network (Fig. 3). In the hierarchical network model, level 0 represents the network core and level L represents the hierarchical location of GRs. Using our proposed intrusion detection and containment technique, we assume DEs within a target cell detect the attack at time t_d and the GR for the target cell implements a containment filter against the intrusion traffic at time t_c. In our hierarchical network model this initial contaiment action occurs at level L. After implementing a containment filter, we assume it takes a time interval of t_δ to notify the next upstream node of the containment action following our reactive blocking protocol. Using the reactive blocking protocol when a node is notified of a suspected attacker's profile [2] the node immediately blocks the malicious intrusion traffic and then monitors the suspected vulnerable target port for existence of worm activity. If the protocol determines that worm activity is not prevalent it disables the block

[1] Endpoints are logically located within a cell and a single GR typically contains multiple cells.

[2] In [1] we defined a *profile* as a 3-tuple consisting of *srcIP*, *dstport*, *proto*. *srcIP* is the source IP address in the IP header of packets captured by the DE, *dstport* is the target port and *proto* is the transport layer protocol used.

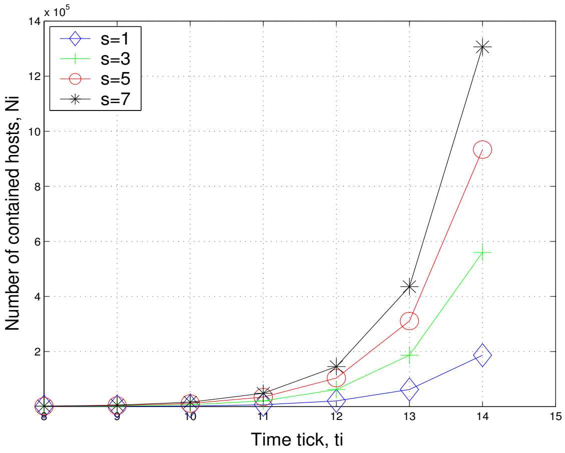

Figure 4: Effect of s on N_i. $y = 3, t_c = 8$.

thereby preventing a potential denial of service. If the protocol determines that worm activity is prevalent the block is maintained and the next upstream node is in turn notified of the suspected attacker's profile after a t_δ time interval. In the event of a worm outbreak, collaborative containment is achieved by rapidly spreading containment action up the hierarchy of the network towards the network core.

Using the hierarchical network topology (Fig. 3), it is assumed that the number of downstream nodes y that connect to an upstream node is the same for all upstream nodes and the number of hosts Q in each cell is also the same for all cells. The total number of contained hosts N_i as a result of a containment action carried out by a node at level i in an L-level network can be expressed as:

$$N_i = Qsy^{L-i} \tag{1}$$

The time of containment t_i at any level i in the hierarchical network is:

$$t_i = t_c + (L - i)t_\delta \tag{2}$$

Substituting,

$$N_i = Qsy^{\frac{t_i - t_c}{t_\delta}} \qquad 0 \le i \le L \tag{3}$$

Using equation 3 and assuming that t_δ is one time tick, Fig. 4, Fig. 5, Fig. 6 were generated to show the impact of variations in s, y and t_c on N_i. The figures show that our technique of applying containment filters on gateway routers and collaboratively spreading the containment action up the network hierarchy causes the contained population, N_i to increase exponentially over time if the worm attack persists. Fig. 4 and Fig. 5 show that a change in s and y respectively results in a directly proportional change in N_i. Increases in s or y result in an increase in the number of hosts contained by a single

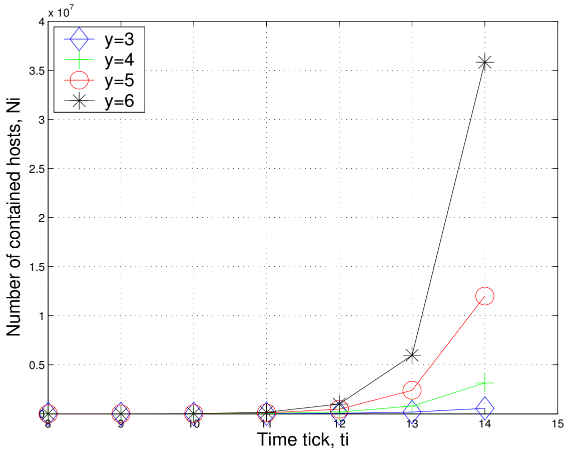

Figure 5: Effect of y on N_i. $s = 3, t_c = 8$.

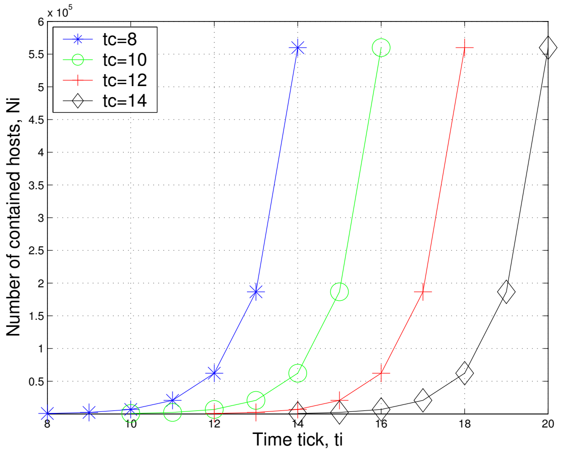

Figure 6: Effect of t_c on N_i. $y = 3, s = 3$.

containment filter on the gateway router. While an increase in s results in a proportional increase in N_i, increases in y result in an even greater increase in N_i. The tree-like structure of the hierarchical network topology is responsible for the greater changes in N_i caused by changes in y. Such tree-like structures are common on the Internet and large-scale networks [9][10]. Fig. 6 shows that increases in t_c results in a time-delay depicted as a time shift in the N_i curve. Such time delays in containment time on the gateway router can occur in practical systems due to an increase in detection interval of the worm which can also occur due to slowness of the detection system in detecting worm activity. As results in Section V show, such containment time delays result in infection of a greater number of vulnerable hosts before the worm is completely contained.

In the next section, we explain the relationship between worm detection interval and worm containment time. We also present the probability model for detection interval.

3.2 Continuous-time Probability Model for Detection Time

Using the AAWC model notations, the time t_c at which a gateway router (GR) for a target cell implements a containment filter can be expressed as

$$t_c = t_r + \alpha + \beta \tag{4}$$

where

- α is detection interval - time interval between t_r and t_d.
- β is containment interval - time interval between t_d and t_c.

From experimental observations in [7], the containment interval β was found to be dependent on our optimized router correlation algorithms and was relatively fixed for experiments carried out with the same version of router code. On the other hand, detection interval α was observed to vary with variations in scanning rate of the attacking worm. Variations in detection interval α was also observed to be the predominant cause of variations in the value of t_c. Based on the AAWC model, Fig. 6 and Fig. 10 show that the protection capability of the proposed detection and containment technique is significantly affected by the value of t_c and hence the value of detection interval α. In this section, we develop a probability model that attempts to address the question:

- What is the probability that the measured detection interval, T_{det} for a particular deployment of our defense technique is not greater that a certain desirable detection interval α? Mathematically, this is equivalent to $P(T_{det} \leq \alpha)$.

We model scanning of hosts in a target cell by a poisson process with an average rate of r h/s. Use of the poisson distribution to model scanning worm behaviour is not new. D.M. Nicol in [11] used a poisson distribution to model observed number of infection attempts due to scanning worm activity. Given a total of W hosts in the target cell comprising m detector endpoints (DEs) and $W-m$ non-detector endpoint hosts [3], we make assumptions that:

1. All hosts in the target network are vulnerable. Therefore, each scan results in an infection.

2. The worm's travel time from source to destination is negligible. This assumption is not unrealistic for fast propagating worms. Hence, the first successful infection of a vulnerable target host occurs at the time of worm release, t_r.

Due to the underlying poisson distribution the interval between infections is an exponential random variable with mean $\frac{1}{r}$ and the measured detection interval, T_{det} is the sum of inter-infection times until all DEs in the target cell are scanned. Therefore, T_{det} is also an exponential random variable. According to our proposed host detection algorithm, all DEs in the target cell must have records of the suspected attacker's profile to eliminate false alerts and therefore must be scanned by the worm for detection to be achieved. If X is the number of scanned non-detector endpoints in the target cell before all m detector endpoints are successfully scanned, then the value of X can range from 0 to $W-m$. X is therefore a uniformly distributed random variable $X \sim U(0, W-m)$.

The cumulative distribution function (cdf) of T_{det} which we represent as $F_{T_{det}}$ can be expressed as:

$$F_{T_{det}} = P(T_{det} \leq \alpha) = 1 - e^{-\frac{\alpha r}{m+X}} \qquad \alpha \geq 0 \tag{5}$$

It can be shown using total probability that:

$$P[YinA] = \int_{-\infty}^{+\infty} P[YinA|X = x) f_X(x) dx \tag{6}$$

where $f_X(x)$ is the probability distribution function (pdf) of X. Similarly, $P(T_{det} \leq \alpha)$ can be expressed as:

$$P(T_{det} \leq \alpha) = \int_{0}^{W-m} P(T_{det} \leq \alpha|X = x) f_X(x) dx \tag{7}$$

where $f_X(x) = \frac{1}{W-m}$. Solving,

$$P(T_{det} \leq \alpha) = \frac{1}{W-m} \int_{0}^{W-m} 1 - e^{-\frac{\alpha r}{m+x}} dx \tag{8}$$

Using numerical integration, Fig. 7, Fig. 8 were generated from equation 8. They depict the effects of r and $\frac{m}{W}$ respectively on $P(T_{det} \leq \alpha)$. Fig. 7 shows

[3] We do not differentiate between different non-detector endpoints based on hardware or software configuration.

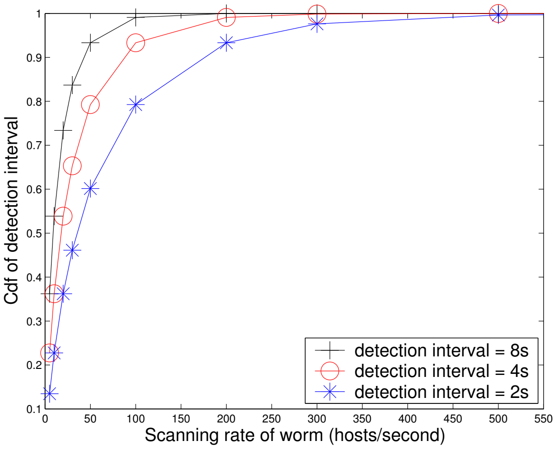

Figure 7: Effect of worm scanning rate, r on cdf of α. $W = 254, m = 3$.

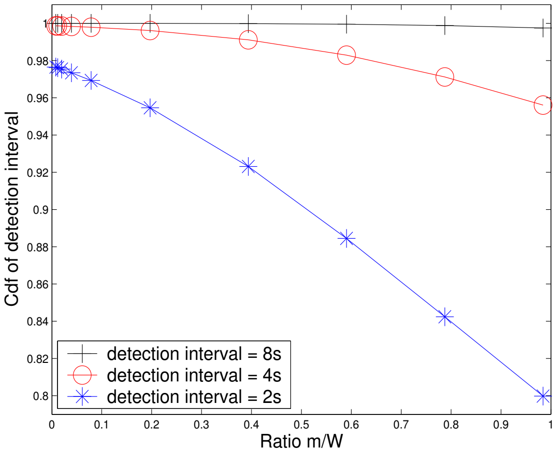

Figure 8: Effect of ratio $\frac{m}{W}$ on cdf of α. $W = 254, r = 300$.

that the probability that a measured detection interval will not be greater than a chosen detection interval value increases exponentially with increase in scanning rate of the worm. Hence, *faster spreading worms are more likely to be detected within shorter intervals*. In addition, for a given deployment of our defense technique and using Fig. 7, it is feasible to determine with high probability an upper bound for detection interval for a particular scanning worm rate. For example, Fig. 7 shows that for a deployment with $W = 254$ and $m = 3$, $P(T_{det} \leq \alpha) > 95\%$ for $\alpha = 2$ and $r = 300$. This implies that there is over 95% probability that the measured detection interval would be less than 2 seconds for a worm scanning at a rate of 300 hosts/second. Fig. 8 shows that the probability that a measured detection interval will not be greater than a chosen detection interval value is not significantly affected by variations in the number of detector endpoints in the target cell. Variations in $\frac{m}{W}$ from 0 to 1 only caused an 18% change in $P(T_{det} \leq \alpha)$ for $\alpha = 2$. For $\alpha = 4$, the change in $P(T_{det} \leq \alpha)$ was an insignificant 4.5%.

Table 2: Parameters for AAWP Model

Notation	Explanation
M	total number of vulnerable machines
H	size of entire population scanned by worm
r	scanning rate (the average number of machines scanned by an infected machine per unit time)
p	patching rate (the rate at which an infected or vulnerable machine becomes invulnerable)
z	size of hitlist (the number of infected machines at the beginning of the spread of active worms)
d	death rate (the rate at which an infection is detected on a machine and eliminated without patching)
n_i	number of infected machines at time tick i
m_i	number of vulnerable machines at time tick i

4 Modeling Propagation of Active Worms

In order to quantify the protection capability of the proposed technique using AAWC model, it is important to characterize the worm spread. In this section, we briefly review the AAWP model for worm propagation and then adapt the model to our hierarchical network topology.

4.1 Review of the AAWP Model

Active worms often propagate through random scanning and the Analytical Active Worm Propagation (AAWP) model [6] was chosen to model worm propagation in our analysis because it more accurately captures the behavior of random scanning worms. In addition, it is a discrete time model similar to our AAWC model. The AAWP model shows that the number of newly infected hosts in each time tick as a result of a random scanning worm attack is determined by parameters such as the size of the total population that the worm scans, the total number of vulnerable hosts in the population, the scanning rate of the worm, the patching rate, the death rate, and the time it takes for the worm to complete infection on a vulnerable host.

Using parameters in Table II, the model assumes that a worm randomly scans the entire population, H and requires one time tick to infect a vulnerable host. Therefore, the probability that a host is hit by one scan is $\frac{1}{H}$. If at time tick $i = 0$ there are $n_0 = z$ infected hosts and m_0 vulnerable hosts then the effective initial scanning rate will be $n_0 r$ and there will be $(m_i - n_i)\left[1 - (1 - \frac{1}{H})^{n_i r}\right]$ newly infected hosts on the next time tick.

It was shown in [6] that with death rate d and patching rate p, the total number of infected hosts n_{i+1} on the next time tick can be expressed as

$$n_{i+1} = n_i + (m_i - n_i)\left[1 - (1 - \tfrac{1}{H})^{n_i r}\right] - (d + p)n_i.$$

Also, the total number of vulnerable hosts (including infected ones) reduce by a factor of $(1 - p)$ after every time tick. Hence, $m_{i+1} = (1 - p)m_i$ and $m_i = (1 - p)^i m_0 = (1 - p)^i M$.

Therefore,

$$n_{i+1} = (1 - d - p)n_i + \left[(1 - p)^i M - n_i\right]\left[1 - (1 - \frac{1}{H})^{n_i r}\right] \tag{9}$$

where $i < 0$, $n_0 = z$ and $m_0 = M$. According to [6], the recursion stops when there are no more vulnerable hosts left or when the worm can no longer increase the total number of infected hosts.

4.2 AAWP Model in a Hierarchical Network

We use the hierarchical network topology (Fig. 3) to depict a large scale network and model worm propagation by adapting parameters of the AAWP model to suit the hierarchical network topology. We make the following assumptions in adapting the AAWP model:

1. The entire host population are vulnerable to the worm attack but there are no infected hosts in the network prior to the worm attack.
2. A single attacker is the only initial scanning source with a scanning rate of r hosts per second (h/s). Hence, $n_0 = z = 1$. After a successful infection, an infected host initiates scans on other hosts in the network similar to the attacker. This simulates active worm spreading.
3. Infection on infected hosts are eliminated only by patching. Hence death rate $d = 0$.
4. For simplicity we assume that one time tick is equivalent to one second.

The total number of hosts in the network (Fig. 3), $H = Qsy^L$. Applying the AAWP model (equation 9) and using notations in Table I and II, the total number of infected hosts can be expressed as:

$$n_{i+1} = (1 - p)n_i + \left[(1 - p)^i M - n_i\right]\left[1 - (1 - \frac{1}{Qsy^L})^{n_i r}\right] \tag{10}$$

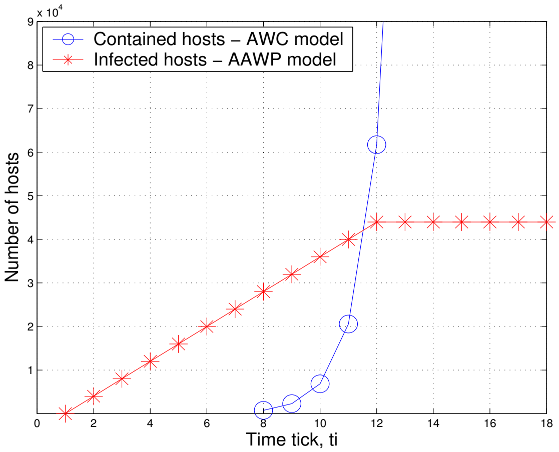

Figure 9: Results with AAWP and AAWC models. $y = 3, s = 3, t_c = 8, r = 4000, L = 10, p = 0$.

5 Effectiveness of the AAWC Model

We define effectiveness of the AAWC model in terms of the maximum number of hosts that the active worm successfully infects before further spread of the worm is completely contained. The smaller the maximum number of infected hosts before containment, the more efficient the simulated model.

In our analysis, we assume that the attacking worm is a fast spreading scanning worm and the adapted AAWP model (equation 10) is used to model the number of infected hosts in the network. The AAWC model (equation 3) is used to model the number of contained hosts as a result of our detection and collaborative defense technique. We also assume that the spread of the worm is successfully contained when the number of contained hosts exceed the number of infected hosts in the network. Fig. 9 shows simulation results using both models. While the AAWP model shows proportional growth in infected population before detection and containment of the worm, the AAWC model shows an even greater growth in the number of contained hosts after time t_c. Using our technique, a perimeter protecting the contained hosts is created on a gateway router or upstream router after a containment action is taken thus preventing further direct scans from the attacker. For a scanning worm attack, the worm spread is stopped when the number of contained hosts exceed the number of infected hosts, thus preventing further increase in the number of infected hosts (see Fig. 9). For a single worm attack scanning a 10-level network with a total vulnerable population of 44.4 million [4] hosts, at a rate of $4,000$scans/second [5], Fig. 9 shows that the total number of hosts infected before complete containment is about $44,000$ hosts, 0.1% of the total vulnerable population. The remaining 99.9% were protected from the attack. The result also shows that complete containment was achieved within 12 seconds after release of the worm. As an example, the

[4] Computed using equation 1, with $i = 0$ and $L = 10$.

[5] Slammer exhibited an initial scanning rate of 4000scans/second [12].

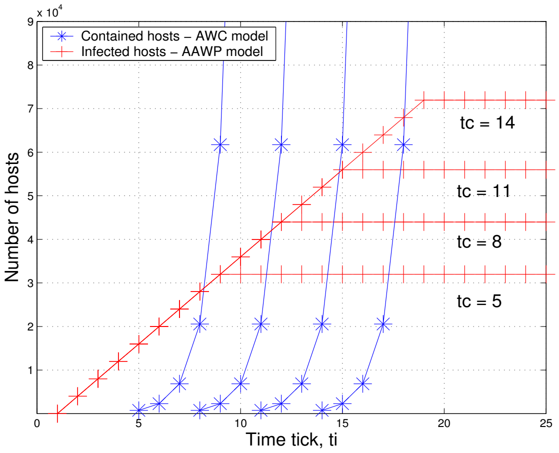

Figure 10: Effect of t_c on infected population. $y = 3, s = 3, r = 4000, L = 10, p = 0$.

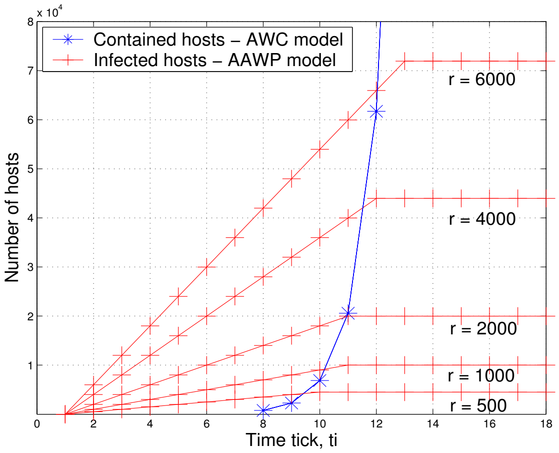

Figure 11: Effect of worm scanning rate, r on infected population. $y = 3, s = 3, t_c = 8, r = 4000, L = 10, p = 0$.

Slammer worm infected more than 90% of vulnerable hosts on the Internet within 10 minutes [12]. Projecting from the simulation results in Fig. 9, our automated detection and collaborative containment technique has the capability of containing a single Slammer-like worm attack within seconds of detection, thereby protecting about 99.9% of vulnerable hosts. Using the models we further investigated the impact of varying time t_c and worm scanning rate. Fig. 10 shows that the total number of hosts infected before the worm is contained as well as the time taken to completely contain the worm spread increases proportionately with an increase in t_c. Equation 4 shows that both detection and containment intervals are contributors to containment time t_c. It is therefore crucial for any viable worm defense mechanism to minimise worm detection and containment intervals. Fig. 11 shows that increasing the worm scanning rate increases total number of hosts

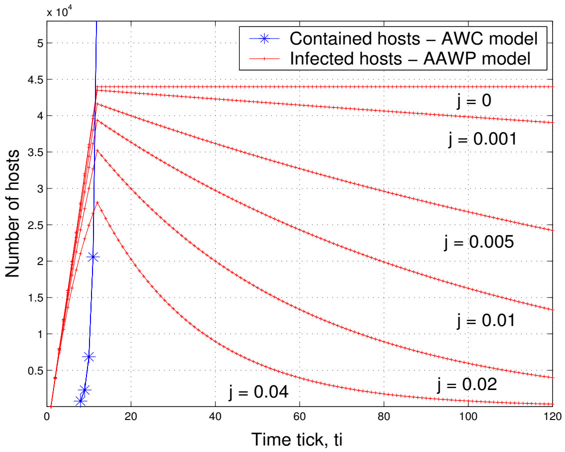

Figure 12: Effect of patching rate, p on Infected population. $y = 3, s = 3, t_c = 8, r = 4000, L = 10$.

infected before the worm is contained. There is however an insignificant increase in the time taken to completely contain the worm spread.

Worm defense using our automated detection and collaborative network containment technique can effectively and quickly stop further direct worm scans but does not address the infectious state of hosts infected before containment of the worm. For complete eradication of infection we study the effect of introducing immunization by patching to the network containment defense mechanism in the next section.

5.1 AAWC Model with Immunization

Immunization by quickly deploying patches on infected hosts has been proposed as an effective defense strategy for worms [13][14][15]. Results in Fig. 9, Fig. 10 and Fig. 11 reveal that for very fast spreading worms, network containment is capable of containing worm spread within few seconds of propagation thus protecting a significant portion of the vulnerable population from infection. Fig. 11 also show that with increase in scanning rate of the worm the total number of hosts infected before worm containment also increases. However, network containment does not recover already infected hosts. We therefore investigate introducing immunization to our network containment defense approach. In this analysis, the effect of patching is introduced by varying the patching rate, p in the AAWP model (equation 10). With a non-zero patching rate, after complete containment of the worm the number of infected hosts does not remain constant as in Fig. 9, Fig. 10 and Fig. 11. Rather, the number of infected hosts n_i at time t_i after containment decrease by the patching rate on every subsequent time tick. Using equation 10, results show that the total number of hosts infected before the worm is contained was reduced from $44,000$ (Fig. 9) to $28,000$ (Fig. 12), a 36.4% reduction, by introducing immunization at a patching rate of 0.04 hosts per second. However, the time taken to completely contain the worm spread does not change significantly. This hybrid approach of combining network containment and immunization also ensures that hosts infected before the worm spread is contained are recovered and do not become launching platforms for future worm attack.

6 Conclusion and future work

In this paper, we presented a discrete-time model for our earlier proposed host-based detection and collaborative network containment technique which we referred to as the Analytical Active Worm Containment (AAWC) model. In order to investigate the protection capabilities of the proposed detection and containment technique, we used the well known Analytical Active Worm Propagation (AAWP) model and the AAWC model to simulate worm propagation and our proposed technique respectively. The results showed that our detection and containment technique is capable of automatically and rapidly containing a fast spreading scanning worm thus protecting a significant proportion of vulnerable hosts in a large network in the event of a scanning worm outbreak.

We also investigated the introduction of immunization to our containment technique and studied the effects on a vulnerable population under attack. We observed that while collaborative network containment of worms can halt further worm spread within a short interval, it does not recover hosts that were successfully infected before the containment. Introducing immunizaton by patching to our containment technique not only resulted in recovery of infected hosts, it also reduced the number of hosts that were successfully infected before the containment.

A study of the impact of detection interval of a scanning worm on the protection capability of our technique was carried out and results showed that the detection interval of a worm is a major contributor to its containment time which is a determinant of the number of hosts protected by our containment technique. We then developed a probability model for detection interval which revealed the direct relationship between the scanning rate of a worm and its detection interval. Our results showed that faster spreading worms are more likely to be detected within shorter intervals. The probability model can be useful to network and security architects who deploy our proposed defense technique in large scale networks.

For future work, we intend to extend our host-based detection and network containment approach and model to defending against DDoS attacks. While the AAWC model presented in this paper was our first attempt at modeling our detection and containment technique, we intend to investigate stochastic modeling of the technique as part of our future work.

References

[1] S. Staniford, V. Paxson, and N. Weaver. How to Own the Internet in Your Spare Time. In in Proc. of the USENIX Security Symposium, 2002. http://www.iciri.org/vern/papers/cdc-usenixsec02/index.html.

[2] D. Nicol, M. Liljenstam. Models of Active Worm Defense. In Proceedings of the Measurement, Modeling and Analysis of the Internet (IMA Workshop '04), Urbana-Champaign, Illinois, January 2004.

[3] J.C. Frauenthal. Mathematical Modeling in Epidemiology. New York: Springer-Verlag, 1980.

[4] J. O. Kephart and S. R. White. Measuring and Modeling Computer Virus Prevalence. In Proceedings of the IEEE Symposimum on Security and Privacy, 1993.

[5] C. C. Zou, W. Gong, D. Towsley. Code Red Worm Propagation Modeling and Analysis, Proc. of the 9th ACM Conference on Computer and Communication Security (CCS02),Washington DC Nov. 2002.

[6] Z. Chen, L. Gao, and K. Kwiat. Modeling the Spread of Active Worms. In INFOCOM 2003, 2003.

[7] F. Akujobi, I. Lambadaris, E. Kranakis. Endpoint-driven Intrusion Detection and Containment in Enterprise Networks. In Proceedings of the IEEE Military Communications Conference (MILCOM) 2007, Orlando, Florida, October 2007.

[8] D. Moore, C. Shannon, G. M. Voelker, and S. Savage. Internet Quarantine: Requirements for Containing Self-Propagating Code. In IEEE INFOCOM, 2003.

[9] D. Alderson, L. Li, W. Willinger, and J. C. Doyle, Understanding Internet topology: Principles, models, and validation, in IEEE/ACM Transactions on Networking, 13(6):12051218, 2005.

[10] L. Li, D. Alderson, W. Willinger, and J. Doyle. A first-principles approach to understanding the internets router-level topology. In Proc. ACM SIGCOMM, 2004.

[11] D. Nicol, The impact of stochastic variance on worm propagation and detection. In WORM 06: Proceedings of the 2006 ACM Workshop on Rapid Malcode. Fairfax, VA, USA.

[12] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver. Inside the Slammer Worm. IEEE Magazine of Security and Privacy, pages 3339, July/August 2003.

[13] Z. Chen, L. Gao, and C. Ji. On Effectiveness of Defense Systems against Active Worms. Technical Report, 2003.

[14] C. C. Zou, D. Towsley, and W. Gong. Email worm modeling and defense. In 13th International Conference on Computer Communications and Networks, pages 409414, October 2004.

[15] K. G. Anagnostakis, M. B. Greenwald, S. Ioannidis, A. D. Keromytis, and D. Li. A Cooperative Immunization System for an Untrusting Internet. In Proceedings of the 11th IEEE International Conference on Networking (ICON), pages 403–408, Sept./Oct. 2003.