An Integrated Approach to Detection of Fast and Slow Scanning Worms *

Frank Akujobi
Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Dr., Ottawa, ON K1S 5B6, Canada
fakujobi@sce.carleton.ca

Ioannis Lambadaris
Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Dr., Ottawa, ON K1S 5B6, Canada
ioannis@sce.carleton.ca

Evangelos Kranakis
Department of Computer Science
Carleton University
1125 Colonel By Dr., Ottawa, ON K1S 5B6, Canada
kranakis@scs.carleton.ca

ABSTRACT

The propagation speed of fast scanning worms and the stealthy nature of slow scanning worms present unique challenges to intrusion detection. Typically, techniques optimized for detection of fast scanning worms fail to detect slow scanning worms, and vice versa. In practice, there is interest in developing an integrated approach to detecting both classes of worms. In this paper, we propose and analyze a unique integrated detection approach capable of detecting and identifying traffic flow(s) responsible for simultaneous fast and slow scanning malicious worm attacks. The approach uses a combination of evidence from distributed host-based anomaly detectors, a self-adapting profiler and Bayesian inference from network heuristics to detect intrusion activity due to both fast and slow scanning worms. We assume that the extreme nature of fast scanning worm epidemics make them well suited for extreme value theory and use sample mean excess function to determine appropriate thresholds for detection of such worms. Random scanning worm behavior is considered in analyzing the stochastic time intervals that affect behavior of the detection technique. Based on the analysis, a probability model for worm detection interval using the detection scheme was developed. Simulations are used to validate our assumptions and analysis.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection—*Invasive software*

General Terms

Security, Intrusion Detection

Keywords

Worms, Anomaly detection, Bayesian inference, Detection interval, Probability model

*This research was partially funded by grants from Natural Sciences and Engineering Research Council (NSERC) and Mathematics of Information Technology and Complex Systems (MITACS).

ASIACCS '09, March 10 – 12, 2009, Sydney, NSW, Australia.

1. INTRODUCTION AND RELATED WORK

Worm detection techniques are typically designed based on some unique characteristic(s) of the worm to be detected. Fast scanning worms typically exhibit an abnormally high number of connections or traffic flows which are detectable in the network or on end hosts, while slow scanning worms propagate more stealthily, enabling them to blend with normal traffic patterns and evade intrusion detection systems (IDS) that depend on only anomalous network heuristics for detection. Most proposed techniques for detection of fast scanning worms [2] [22] [14] [4] are unable to detect slow scanning worms. Slow scanning worms are usually indistinguishable from normal traffic seen on the network, or seen by end host network connections and are therefore difficult to detect. However, both types of worms pose a serious threat to vulnerable systems and investigating detection systems capable of detecting and defending against both type of worms is relevant.

Some recent works have attempted to address the challenge of detecting both fast and slow worms. In [1], adaptively adjusting the detection threshold on end host detectors based on observed traffic was proposed as a way to detect both fast and slow worms. A supervised classifier predicts the time-varying distribution of outgoing traffic based on previous observations and this was used for the adjustment. In [16] a multi-resolution approach for worm detection was proposed to deal with the limitations of simple threshold-based detection methods. Using number of unique destinations contacted as a basis for anomaly detection, the multi-resolution approach used different thresholds during different time windows to detect attacks of different speeds. Faster scanning attacks were detected with smaller time windows while slower attacks were detected with larger time windows.

We point out that a common characteristic of most schemes proposed for worm detection is the use of connection counts, traffic rates and traffic trends [23] as the basis for anoma-

lous detection. This approach inherently carries a high rate of false positives and false negatives because worms are capable of propagating at rates and exhibiting trends similar to that of normal traffic flows and therefore can camouflage as normal traffic. Also, information about vulnerabilities and attempted exploits do not exist in the network layer, hence such techniques are unable to provide verifiable evidence of malicious intrusions. In fact, the assumption that malicious attacks necessarily cause anomalous activity in the network in terms of host or network traffic was recently challenged in [7].

Host-based Anomaly Intrusion Detection Systems (AIDS) which infer suspicious activity when a detector endpoint experiences an intrusion that attempts to alter a pre-defined standard state [1] of the endpoint have been more successful at detecting malicious worm intrusions irrespective of scanning behavior of worms. Typically, such attempts are in the form of anomalous system calls [11] or unauthorized intrusions which cause the host AIDS to trigger an alert. Recent work [6] and vendor implementations [19] have recorded success in using host AIDS for detecting unauthorized intrusions. Host AIDS are capable of leveraging large amounts of detailed context about applications and system behavior to effectively detect anomalous host behaviors [18]. The technique adopted in [6] shows that with properly instrumented detection software, host-based intrusion detection is effective and capable of eliminating false positives. Though host-based AIDS can successfully detect malicious intrusions on a host and therefore determine the attempted exploit, host-based AIDS on a single host alone is not capable of determining the actual traffic flow responsible for the intrusion. During multiple simultaneous attacks, determination of traffic flows responsible for the attacks become even more difficult, and at the same time crucial for rapid or automated defense.

In this paper, we propose an integrated detection system for fast and slow scanning worms which uses host-based AIDS in combination with Bayesian inference and a self-adapting profiler to achieve detection. Our underlying assumption is that intrusions that attempt to alter the standard state of a hardened endpoint is verifiable evidence of unauthorized or malicious activity. We also assume that fast scanning worm attacks are extreme events which exhibit an extremely high number of connection attempts on vulnerable targets. Such attacks have been known to cause very abnormal increase in observed network traffic patterns [10] [17] and therefore seem to be well suited for extreme value theory (EVT) [9] [15]. Hence, we use sample mean excess function [15] [9] to determine appropriate thresholds used for Bayesian inference during fast worm detection. The proposed detection technique is unique because it leverages evidence from distributed host-based AIDS about unauthorized intrusions, as well as correlation of network heuristics based on Bayesian inference and adaptive profiling to achieve *simultaneous detection* and *identification* of both fast and slow scanning intrusion traffic. Based on literature survey, the adaptive anomaly detection approach proposed in [1], which uses a traffic predictor to control a time-varying detection threshold for worm detection is closest to our work. Our work is however distinct from this work in the following ways.

[1] Pre-defined standard states of endpoints are typically determined by established security policies and standards.

First, the technique proposed in [1] is based on the assumption that a worm infection necessarily increases the outgoing connection rate of the infected host. While this may be true for fast scanning worms, slow scanning worms can propagate stealthily and may not cause an increase in the outgoing connection rate of the infected host. In comparison, our detection approach uses properly tuned host-based AIDS capable of providing verifiable evidence of unauthorized intrusions during a worm attack with very low false positives for detection, and correlates intrusion information received from independent detectors within a network cell only to "investigate" and probabilistically determine the traffic flow(s) responsible for the observed intrusion(s).

Second, the approach in [1] uses anomaly detectors that predict the time-varying distribution of outgoing traffic based on previous measurements and adjust the threshold settings for worm detection based on the predictions. Again, while fast propagating worms can be easily detected using this approach, slow scanning worms that exhibit connection rates similar to normal traffic or even less than normal traffic can evade this detection scheme. On the other hand, our proposed detection technique infers unauthorized intrusion when a detector endpoint experiences an intrusion attempt to alter a pre-defined standard state of the endpoint. In addition, we use Bayesian inference and a self-adapting profiler to distinguish network heuristics associated with slow scanning worms and therefore provide an integrated detection system capable of detecting simultaneous fast and slow scanning worm attacks.

The main contributions of this work are:

- We propose a unique integrated detection technique capable of detecting and identifying traffic flow(s) responsible for simultaneous fast and slow scanning malicious worm attacks. We use a combination of evidence from host-based anomaly detectors, a self-adapting profiler and Bayesian inference from network heuristics for detection.
- Based on the assumption that fast scanning worm attacks are extreme events that seem to be applicable to extreme value theory, we use sample mean excess function to determine appropriate thresholds for detection of such worms.
- Worm detection interval is modeled as a stochastic variable and we present an analysis of detection interval for both fast and slow scanning worms using the proposed detection technique.
- We then develop probability models for worm detection interval for both fast and slow scanning worms and validate the models using Markov's Inequality. Using the models, we show that the accuracy of detecting worms using our approach can be improved significantly by following certain network design principles. This can be useful to network and security architects deploying detection systems for worms.
- Experimenting on a live test-bed we evaluate the integrated detection technique and show that the results obtained concur with our analytical model and results.

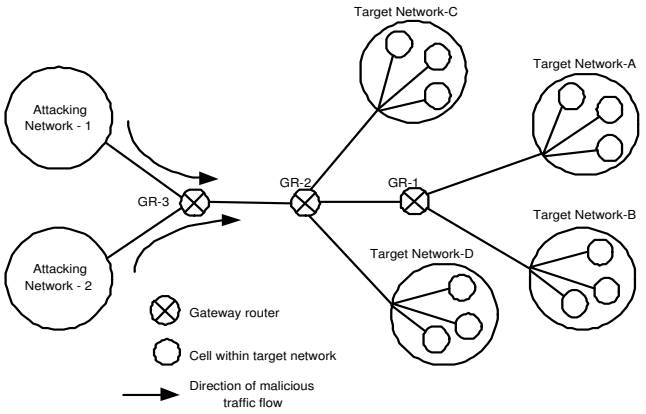


Figure 1: Typical worm attack on multiple networks

1.1 Outline

Section 2 and 3 present a description and analysis of the integrated detection technique. In section 4, we develop probability models for worm detection interval and validate the models using Markov's Inequality. Experimentation on a live test-bed with the proposed technique is presented in section 5. Section 6 concludes the paper and points to future work.

2. INTEGRATED DETECTION APPROACH

Fig. 1 depicts a typical attack scenario in which single or multiple attackers in Network-1 and Network-2 launch scanning worm attacks on several enterprise networks (Target Network-A, Target Network-B, Target Network-C, Target Network-D). Typically, well-designed enterprise networks are logically subdivided into cells or network zones as shown in Fig. 1. The detection scheme uses detector endpoints within distributed cells in a target network for detection of intrusion attempts and correlates captured intrusion information on the gateway router of the cells. The detection scheme consists of detection and correlation phases described in the next section.

2.1 Detection Phase

The scheme uses two instances of detector agents, fast worm detector agent (FDA) and slow worm detector agent (SDA) both running simultaneously on hardened detector endpoints (DEs) located within distributed cells in the network and responsible for capturing intrusion attempts targeted at the cells. The detector agents run similar host-based anomaly detection software configured to alert and capture intrusion data when anomalous system calls or unauthorized intrusions attempts are made. However, the FDA and SDA are used for detection of fast and slow scanning malicious worms respectively by capturing intrusion data during two different time intervals. The FDA captures intrusion data for a short interval while the SDA captures for a larger interval. Table 1 describes some of the detection algorithm parameters.

Table 1: Some Detection Algorithm Parameters

Notation	Explanation
SW_j	j^{th} slow worm detection window
FW_{ij}	i^{th} fast worm detection window within SW_j
t_s	duration of slow worm detection window
t_f	duration of fast worm detection window
U_j	set of profiles captured by the SDA during SW_j
X_{ij}	fast scanning worm profiles detected during FW_{ij}
Y_j	set of profiles forwarded to slow worm correlation engine

2.1.1 Fast Worm Detection

When an FDA running on a detector endpoint (DE) detects a malicious or unauthorized intrusion that attempts to alter the baseline configuration of the DE, the following occurs:

1. The FDA immediately sends a notification signal to other participating FDAs respectively within the cell. FDAs communicate only with other FDAs.

2. When the notification signal is received, the FDAs within the target cell start real-time recording of *profiles* for all network traffic originated from outside their cell and targeted at the DEs for a pre-set capture interval. We refer to the FDA capture interval as the *fast worm detection window* with duration t_f. We also define a *profile* as a 4-tuple consisting of *srcIP*, *dstport*, *proto*, *payload*. *srcIP* is the source IP address in the IP header of packets captured by the DE, *dstport* is the target port, *proto* is the transport layer protocol used and *payload* is the content of the payload of the IP packet. This *profile* format was chosen because it contains sufficient information to implement a traffic flow filter on most real-world routers. With deep packet inspection some routers are capable of performing intelligent content-based filtering [21].

3. At the end of the fast worm detection window, the FDAs on all DEs in the cell transfer their records to their upstream gateway router (GR) and continue monitoring the DEs for unauthorized intrusions.

2.1.2 Slow Worm Detection

The SDAs perform continuous real-time capturing of *profiles* of all network traffic originated from outside their cell and targeted at the DEs in epochs of interval t_s which we refer to as the *slow worm detection window*. During a slow worm detection window, if an SDA running on a DE detects a malicious or unauthorized intrusion that attempts to alter the DE's baseline configuration it captures the nature of the attempted alteration and continues real-time recording of incoming traffic profiles. The capture reveals useful information about a possible exploit and vulnerability on hosts in the cell. At the end of a slow worm detection window, the SDAs on all DEs in the cell transfer their records to their upstream GR and immediately start the next epoch of recording. Unlike the FDAs, the SDAs do not wait for an alert before capturing intrusion data. Intrusion data is captured in periodic slow worm detection windows of duration t_s.

DEs are dedicated to the function of detecting malicious intrusions. They do not initiate communication with any

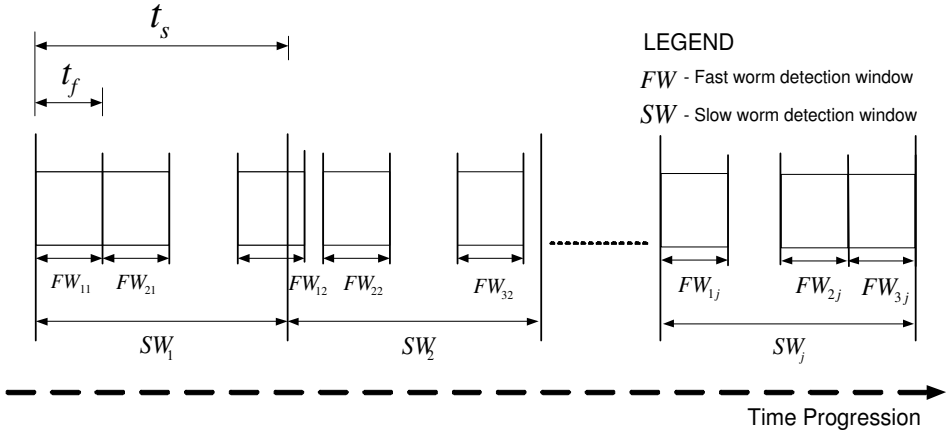


Figure 2: **Series of epochs showing fast and slow worm detection windows**

host outside their cell nor do they participate in normal traffic transactions. Traffic hits from outside their cell, recorded by FDAs and SDAs must therefore be as a result of verifiable unauthorized intrusions or benign network scans.

2.1.3 Detection Windows

We refer to an epoch that spans a capture interval as a *detection window*. *Fast worm detection window* refers to an epoch of duration t_f started as a result of an intrusion attempt detected by an FDA. The *slow worm detection window* refers to a periodic epoch of duration t_s which runs continuously on each SDA. Fig. 2 shows a snapshot of a series of epochs during which the FDA and SDA carry out real-time recording of network traffic profiles. Typically, $t_s > t_f$, hence during multiple simultaneous fast and slow scanning worm attacks there could be multiple fast worm detection windows within a single slow worm detection window as depicted in Fig. 2. At the end of a fast worm detection window, all profiles recorded by an FDA running on a DE in the cell are transferred to the GR for correlation. Thereafter, the FDA continues to monitor for future intrusion attempts. The SDAs wait until the end of the slow worm detection window before transfering captured records to the GR. The next slow worm detection window is started immediately after the transfer.

2.2 Correlation Phase

A process running on the upstream gateway router (GR) monitors the transfer of records from the FDAs and SDAs on DEs in the target cell. The GR runs two correlation engines, fast worm correlation engine (FCE) which executes a fast worm correlation algorithm (FCA) and a slow worm correlation engine (SCE) which executes a slow worm correlation algorithm (SCA). Both FCA and SCA use Bayesian inference to probabilistically determine the most likely profile(s) of the unauthorized intrusion(s). Records from the FDAs are forwarded to the FCE (Fig. 4).

2.2.1 Fast Worm Correlation Algorithm (FCA)

The subroutines that comprise the FCA are:

Fast worm bayesian modeling.

Let B_i be the event that hits from traffic profile i were recorded by an FDA in the target cell. Also, let N_{ij} be the number of hits belonging to traffic profile i recorded by the FDA on the j^{th} DE. If there are m DEs in the target cell and y profiles recorded by FDAs on DEs in the target cell, the a priori probability $P(B_i)$ that event B_i occurred is expressed as:

$$P(B_i) = \frac{\sum_{j=1}^{m} N_{ij}}{\sum_{i=1}^{y}\sum_{j=1}^{m} N_{ij}} \quad (1)$$

Let A be the event that an unauthorized intrusion due to a fast scanning worm is in progress. Host-based AIDS on the FDAs provide verifiable evidence of this event. We model the observation of profile i by the FDA on the j^{th} DE using a Bernoulli random variable, I_{ij}. When the j^{th} FDA experiences a hit as a result of traffic profile i, $I_{ij} = 1$, otherwise $I_{ij} = 0$. We expect that during a fast scanning worm attack, all FDAs in the target cell will experience worm scan hits. Therefore, the likelihood function $P(A|B_i)$ is the probability that a profile i is observed on an FDA, and is expressed as:

$$P(A|B_i) = \frac{\sum_{j=1}^{m} I_{ij}}{m} \quad (2)$$

If profile i is observed on all FDAs in the target cell, then $P(A|B_i) = 1$.

The a posteriori probability, $P(B_i|A)$ of event B_i is a measure of how responsible profile i is for the observed unauthorized intrusion and using Bayes theorem it can be expressed

as:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{y} P(A|B_i)P(B_i)} \quad (3)$$

The following *conditions* are executed to identify a suspicious profile associated with the fast scanning worm.

```
If (y = 1), then P(B_i|A) = 1,
   {
    select profile i = 1;
    trigger automated containment;
   }
If (y > 1), then P(B_i|A) < 1,
   {
    transition to Threshold-based selection
    subroutine;
   }
If (y = 0),
   {
    abort and wait to be activated again;
   }
```

Fast worm threshold-based selection.

Based on our assumption that fast scanning worm intrusions are well suited to extreme value theory, we use sample mean excess function to determine a threshold, Φ_f used for identifying profiles with a posteriori probabilities associated with fast scanning worm intrusion. The following *if loop* is carried out for all y profiles observed on the DEs.

```
If (P(B_i|A) > Φ_f),
   {
    then select profile i;
    trigger automated containment;
   }
If (P(B_i|A) ≤ Φ_f),
   {
    then do not select profile i;
   }
```

It is expected in the *Correlation phase* that during a fast scanning malicious worm or distributed denial of service (DDoS) attack, there is a very high probability that FDAs located in a cell under attack will observe early intrusion attempts and therefore record the *profile(s)* of the intrusion traffic. Previous work reveals that several recent scanning worms such as Code RED II [10] and Nimda [5] preferentially target other hosts from IP address ranges closer to the vulnerable target host (i.e. in the same /24 or /16 network). Intrusion attempts as a result of flash worm activity which does not exhibit such scanning worm patterns will also be captured by the FDAs if the DEs run the same vulnerable software as hosts within the cell they are located and therefore cannot be differentiated from the vulnerable hosts. The assumption will fail only if an attacker has prior knowledge of the DEs and instruments a worm that selectively avoids intrusion attempts on the DEs. A hitlist worm is an example of such a worm. Also, worm infections that require the vulnerable host to first initiate outgoing connections would not be detectable using our technique since DEs do not initiate communication with any host outside their cell. However, the vast majority of worms seen in the wild are scanning worms which lack precise knowledge of hosts and ports on the target network that are currently active [8], and therefore can be detected using our proposed technique.

2.2.2 Brief discussion of threshold determination

Suppose that $X_1, X_2, ..., X_n$ is a sequence of independent and identically distributed random variables from an unknown distribution function $F(x)$ and let u be a predetermined threshold for the random variables. The exceedance of X over u given that X exceeds the threshold can be expressed as:

$$Y = [X - u | X \geq u]$$

Let F_u denote that conditional distribution of the exceedance $Y = X - u$ given that X exceeds the threshold. Hence,

$$F_u(y) = \frac{F(y+u) - F(u)}{1 - F(u)}$$

According to an extreme value theorem by Picklands [13] and Balkema & de Haan [3], the distribution of the exceedances converges in distribution to the generalized Pareto distribution, $G_{\xi,\sigma}(y)$ provided a high enough threshold u is chosen. This offers an opportunity for appropriate selection of thresholds when extreme events such as fast scanning worm invasions occur.

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - (1 + \frac{\xi}{\sigma}y)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp(-\frac{y}{\sigma}) & \text{if } \xi = 0 \end{cases}$$

where $0 \leq y \leq (x_F - u)$ if $\xi \geq 0$, and $0 \leq y \leq -\frac{\sigma}{\xi}$ if $\xi < 0$. x_F denotes the rightmost point of the distribution function $F(x)$.

The peak over threshold (POT) method [15] [9] is one approach to selecting a threshold in extreme value statistics. It offers a graphical tool, the mean excess function $e(u)$ which is used to determine appropriate thresholds based on theoretical knowledge that the mean excess function for a generalized Pareto distribution is a straight line with a positive gradient. Hence,

$$e(u) = \frac{\sigma + \xi u}{1 - \xi} \qquad \sigma + \xi u > 0$$

The sample mean excess function $e_n(u)$ is an empirical estimate of the mean excess function and is defined as:

$$e_n(u) = \frac{\sum_{i=1}^{n} \max(0, X_i - u)}{\sum_{i=1}^{n} 1_{X_{i>u}}}$$

where $1_{X_{i>u}}$ is the indicator function with value 1 if $X > u$ and 0 otherwise.
If the sample mean excess plot is approximately linear with positive gradient above a certain threshold value u, then it is an indication that the exceedances follows a generalized Pareto distribution with positive shape parameter, and as a consequence the threshold u was appropriately chosen.

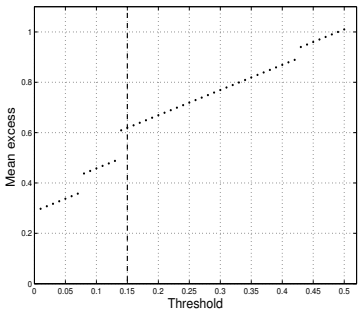


Figure 3: Sample mean excess plot

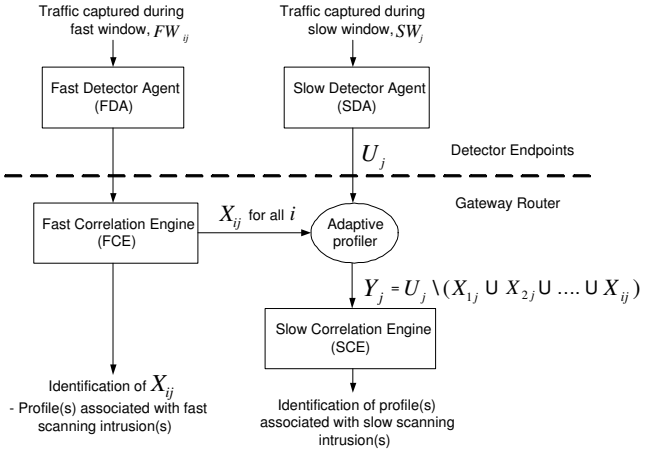


Figure 4: Flow diagram of integrated detection of fast and slow scanning intrusions.

Fig. 3 shows a sample mean excess plot for $P(B_i|A)$ generated using simulations of fast and slow worm attacks. The sample mean excess plot is linear with a positive gradient over a wide region. The vertical line, $u = 0.15$ marks our chosen threshold since there is evidence that the plot is linear with a positive gradient beyond the chosen threshold. This method of determining an appropriate threshold was similarly applied in our experimentation in section 5.1.

2.2.3 Adaptive Profiler

Let FW_{ij} be the i^{th} fast worm detection window within the j^{th} slow worm detection window SW_j (see Table 1 and Fig. 2). We use X_{ij} to model a set with elements corresponding to profile(s) identified as associated with fast scanning worm intrusions(s) by the FCE from records captured during FW_{ij} (Fig. 4). We also use U_j to model the set with elements corresponding to all profiles captured by the SDAs in the cell during the slow worm detection window SW_j. For each slow worm detection window, the adaptive profiler tags traffic profiles determined to belong to fast scanning intrusions and periodically adapts the input into the slow worm correlation engine (SCE) by filtering out fast scanning intrusion profiles. This ensures that only profiles that have not been previously selected by the FCE as associated with fast scanning worms are forwarded to the SCE (Fig. 4). If Y_j is the set with elements corresponding to profiles forwarded to the SCE, then Y_j is expressed as:

$$Y_j = U_j \setminus (X_{1j} \cup X_{2j} \cup ..... \cup X_{ij}) \quad (4)$$

This profiler algorithm ensures that for every slow worm detection window, SW_j, the corresponding Y_j is adapted with outputs, X_{ij} from the FCE. At the end of a slow worm detection window, only profiles that are not deemed to belong to fast scanning worms by the FCE are forwarded to the SCE for slow worm detection and identification. This mechanism provides a capability for detection and identification of traffic flow(s) responsible for simultaneous fast and slow worm attacks. The SCE runs the slow worm correlation algorithm (SCA) described in the next section.

2.2.4 Slow Worm Correlation Algorithm (SCA)

The SCE runs the slow worm correlation algorithm (SCA) on Y_j. The subroutines that comprise the SCA are:

Slow worm bayesian modeling.

Let $S_i : S_i \in Y_j$ be the event that hits from traffic profile i were recorded by an SDA in the target cell. Also, let M_{ij} be the number of hits belonging to traffic profile i recorded by the SDA on the j^{th} DE. If there are m DEs in the target cell and n profiles recorded by SDAs on DEs in the target cell, the a priori probability $P(S_i)$ that event S_i occurred is expressed as:

$$P(S_i) = \frac{\sum_{j=1}^{m} M_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} M_{ij}} \quad (5)$$

Let H be the event that an unauthorized intrusion due to a slow scanning worm is in progress. Host-based AIDS on the SDAs provide verifiable evidence of this event. We refer to SDAs that experience an intrusion attempt and therefore have evidence of an intrusion attempt as witness slow detector agents (WSDA). Assuming there are x WSDAs out of m SDAs in the target cell, we model the observation of profile i by the WSDA on the j^{th} DE using a Bernoulli random variable, L_{ij}. When the WSDA on the j^{th} DE has record of a hit as result of traffic profile i, $L_{ij} = 1$, otherwise $L_{ij} = 0$. We expect that during a slow scanning worm attack, all WSDAs in the target cell will have records of the malicious worm scan hits. Therefore, the likelihood function $P(H|S_i)$ is the probability that a profile i is observed on a WSDA, and is expressed as:

$$P(H|S_i) = \frac{\sum_{j=1}^{x} L_{ij}}{x} \quad (6)$$

If profile i is observed on all WSDAs in the target cell then $P(H|S_i) = 1$ and if profile i is not observed on any WSDA, $P(H|S_i) = 0$.

The a posteriori probability, $P(S_i|H)$ of event S_i is a measure of how responsible profile i is for the observed unauthorized intrusion and using Bayes theorem it can be expressed

as:

$$P(S_i|H) = \frac{P(H|S_i)P(S_i)}{\sum_{i=1}^{n} P(H|S_i)P(S_i)} \qquad i \leq n \tag{7}$$

The following *conditions* are executed to identify a suspicious profile associated with the slow scanning worm.

```
If (n = 1), then P(S_i|H) = 1,
   {
    select profile i = 1;
    trigger automated containment;
   }
If (n > 1), then 0 < P(S_i|H) < 1,
   {
    select profile i in sequence ordered by
    P(S_i|H); trigger automated containment;
   }
If (n = 0),
   {
    abort and wait to be activated again;
   }
```

For a single slow worm attack, $n = 1$ and the GR identifies a single traffic profile $i = 1$. For multiple simultaneous slow worm attacks the GR identifies all traffic profiles in a sequence ordered by the value of $P(S_i|H)$. Hence, profile u is identified before profile v if,

$$P(S_u|H) > P(S_v|H) \quad \forall(u, v) \tag{8}$$

It is expected in the Correlation phase that during a slow scanning malicious worm attack, all WSDAs in the target cell will have a record of the malicious worm scan hits. However, the slow scanning worm will be sucessfully detected if at least one of the SDAs in the target cell experiences a malicious intrusion attempt due to the worm.

3. ANALYSIS OF DETECTION

In this section we present a stochastic analysis of the detection interval for fast and slow random scanning worms. Stochastic variance in such intervals have been known to impact analysis of worm behavior [12].

3.1 Fast Worm Detection Interval, t_{fd}

Detection interval for fast worms, t_{fd} is the interval between the time a worm scan first hits a target cell and the time the worm is successfully detected by the FDAs in the target cell. It comprises the total inter-infection interval, t_{fv} and the total time to infect, t_{infect}.

$$t_{fd} = t_{fv} + t_{infect}$$

3.1.1 Total inter-infection interval, t_{fv}

Inter-infection interval for fast worms is the time interval between successive hits experienced by hosts in a target cell as a result of a particular fast scanning worm. The total inter-infection interval, t_{fv} is the sum of inter-infection intervals until all FDAs in the target cell have experienced worm scan hits. For successful detection of worm activity all FDAs in the target cell must have records of the worm's traffic profile and therefore must be scanned by the worm for detection to be achieved. We model scanning of hosts in the target cell by a Poisson process with an average rate of r hosts/second (h/s). Use of Poisson distribution to model scanning worm behavior is not new [12]. The inter-infection interval between hosts is an exponential random variable [2] with mean $\frac{1}{r}$ and the total inter-infection interval, t_{fv} is the sum of inter-infection intervals until all DEs in the target cell are scanned. Therefore, t_{fv} is also an exponential random variable. Let us assume that there are a total of W hosts in the target cell comprising m DEs and $W - m$ non-detector endpoint hosts and that the hosts are scanned only once in a single worm attack instance.
If G is the number of scanned non-detector endpoints in the target cell before all m DEs are successfully scanned, then G is a uniformly distributed random variable $G \sim U(0, W-m)$. t_{fv}, being the sum of inter-infection intervals until all DEs in the target cell are scanned is an exponential random variable with mean $\frac{m+G}{r}$. Hence,

$$E[t_{fv}|G] = \frac{m+G}{r} \qquad G \sim U(0, W-m)$$

3.1.2 Total time to infect, t_{infect}

This is the time interval it takes to scan and infect a vulnerable host in the target cell. This time is largely dependent on the nature of the intrusion attack and the vulnerability being exploited on the endpoint. For analysis we assume that:

1. All hosts in the target network are vulnerable and t_{infect} is uniform for all vulnerable hosts. Therefore, each scan results in an infection.
2. t_{infect} is negligible for virulent worms.

For fast scanning worms, detection interval t_{fd} can therefore be expressed as:

$$t_{fd} = t_{fv} \tag{9}$$

Then,

$$E[t_{fd}] = E[E[t_{fv}|G]]$$

$$\overline{t_{fd}} = \frac{W+m}{2r} \tag{10}$$

Similarly, since t_{fv} is the sum of independent inter-infection intervals, $Var(t_{fv})$ can be expressed as

$$Var(t_{fv}|G) = \frac{m+G}{r^2}$$

Using total variance,

$$Var(t_d) = Var(t_{fv}) = E[Var(t_{fv}|G)] + Var(E[t_{fv}|G])$$

$$E[Var(t_{fv}|G)] = E[\frac{m+G}{r^2}] = \frac{W+m}{2r^2}$$

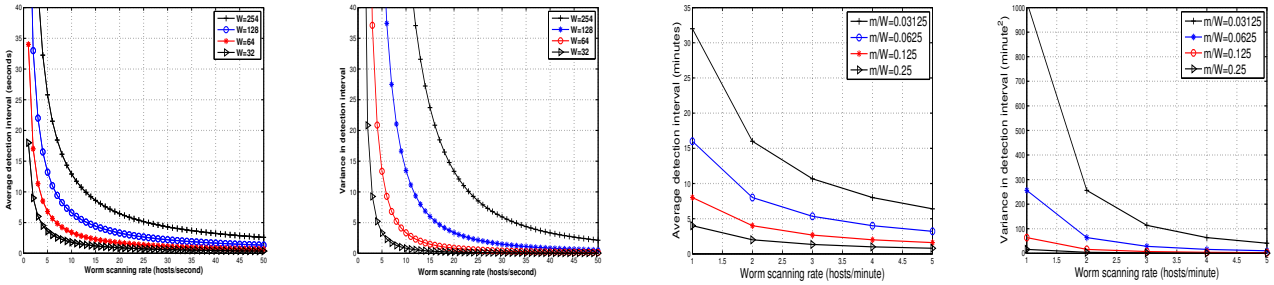
Also,

$$Var(E[t_{fv}|G]) = \frac{1}{12r^2}(W^2 + m^2 - 2Wm)$$

Hence

$$Var(t_{fd}) = \frac{1}{12r^2}(W^2 + m^2 - 2Wm + 6m + 6W) \tag{11}$$

[2]The inter-infection intervals are independent and identically distributed (i.i.d).



(a) Average fast worm detection interval, $\overline{t_d}$.

(b) Variance in fast worm detection interval.

(c) Average slow worm detection interval, $\overline{t_{sd}}$. $W = 128$.

(d) Variance in slow worm detection interval. $W = 128$

Figure 5: Average detection intervals and variance in detection intervals. $m = 4$

Using Matlab simulations, Fig. 5(a) and Fig. 5(b) were generated from (10) and (11). The figures show that average detection interval $\overline{t_{fd}}$, as well as variance in t_{fd} decrease progressively with increase in worm scanning rate. The target cell size, W was varied to investigate impact of cell size on detection, and Fig. 5(a) and Fig. 5(b) show that mean and variance of t_{fd} increases progressively with increase in W. Smaller cell sizes will therefore result in faster host-based detection and containment of fast worm attacks.

3.2 Slow Worm Detection Interval, t_{sd}

Detection interval for slow worms, t_{sd} is the interval between the time a worm scan first hits a target cell and the time the worm is successfully detected by at least one DE in the target cell. It comprises the total inter-infection interval, t_{sv} and the total time to infect, t_{infect}.

$$t_{sd} = t_{sv} + t_{infect}$$

3.2.1 *Total inter-infection interval,* t_{sv}

Inter-infection interval for slow worms is the time interval between successive hits experienced by hosts in a target cell as a result of a particular slow scanning worm. The total inter-infection interval, t_{sv} is the sum of inter-infection intervals until at least one DE in the target cell experiences a worm scan hit. For successful detection of slow worm activity, at least one SDA in the target cell must have record of the worm's traffic profile and therefore must be scanned by the worm for detection to be achieved. We model scanning of hosts in the target cell by a Poisson process with an average rate of r hosts/minute (h/m). The inter-infection interval between hosts is an exponential random variable with mean $\frac{1}{r}$ and the total inter-infection interval, t_{sv} is the sum of inter-infection intervals until at least one DE in the target cell is scanned. Therefore, t_{sv} is also an exponential random variable. Let us assume that there are a total of W hosts in the target cell comprising m DEs, and that the hosts are scanned only once in a single worm attack instance. Hence, the probability of randomly scanning a DE in a target cell is $\frac{m}{W}$.
We model the number of hosts scanned until the first DE is scanned as a geometric random variable Z. t_{sv}, being the sum of inter-infection intervals until the first DE in the target cell is scanned is an exponential random variable with mean $\frac{Z}{r}$. Hence,

$$E[t_{sv}|Z] = \frac{Z}{r}$$

3.2.2 *Total time to infect,* t_{infect}

Based on the same assumptions made in section 3.1.2, t_{infect} is considered negligible for slow scanning virulent worms. Detection interval t_{sd} can therefore be expressed as:

$$t_{sd} = t_{sv} \tag{12}$$

Then,

$$E[t_{sd}] = E[E[t_{sv}|Z]]$$

$$\overline{t_{sd}} = \frac{W}{m * r} \tag{13}$$

Similarly, since t_v is the sum of independent inter-infection intervals, $Var(t_v)$ can be expressed as

$$Var(t_{sv}|Z) = \frac{Z}{r^2}$$

Using total variance,

$$Var(t_{sd}) = Var(t_{sv}) = E[Var(t_{sv}|Z)] + Var(E[t_{sv}|Z])$$

$$E[Var(t_{sv}|Z)] = E\left[\frac{Z}{r^2}\right] = \frac{W}{r^2 m}$$

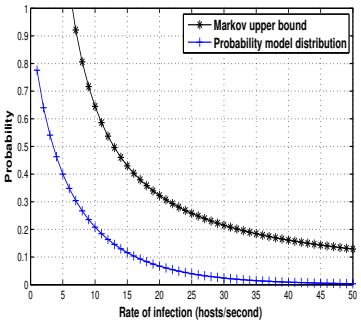
Also,

$$Var(E[t_{sv}|Z]) = Var\left(\frac{Z}{r}\right) = \frac{W(W-m)}{m^2 r^2}$$

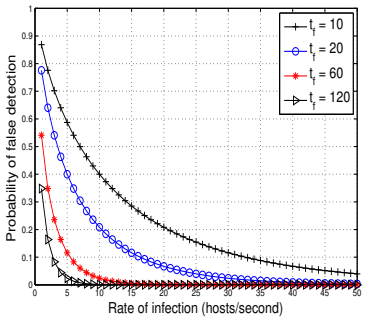
Hence

$$Var(t_{sd}) = \frac{W}{r^2 m}\left(1 + \frac{W-m}{m}\right) \tag{14}$$

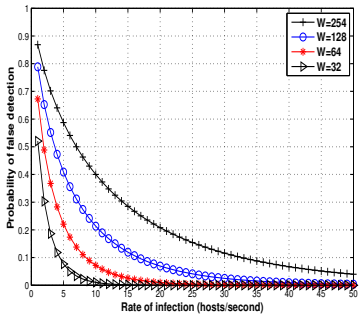
Using Matlab simulations, Fig. 5(c) and Fig. 5(d) were generated from (13) and (14). The figures show that average detection interval $\overline{t_{sd}}$, as well as variance in t_{sd} decrease progressively with increase in worm scanning rate. The ratio m/W was varied to investigate the impact of number of detectors and cell size on distributed cellular detection. Fig. 5(c) and Fig. 5(d) show that mean and variance of t_{sd}



(a) Probability model distribution vs. Markov upper bound. $t = 20$.

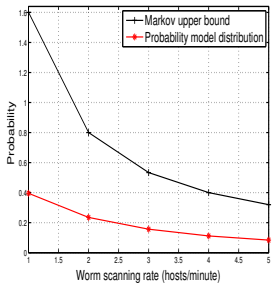


(b) Effect of t_f on false detection probability. $W = 254, m = 4$.

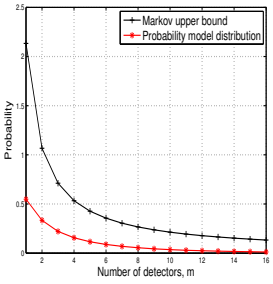


(c) Effect of W on false detection probability. $t_f = 10, m = 4$.

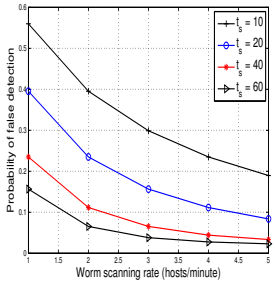
Figure 6: Simulations with probability model for *fast worm* detection interval.



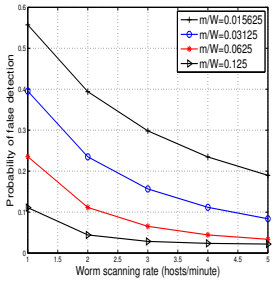
(a) Effect of scanning worm rate on Probability model distribution vs. Markov's upper bound. $W = 128, m = 4, t = 20$.



(b) Effect of m/W ratio on Probability model distribution vs. Markov's upper bound. $W = 128, t = 20, r = 3$.



(c) Effect of t_s on false detection probability. $W = 128, m = 4$.



(d) Effect of m/W ratio on false detection probability. $t_s = 20, m = 4$.

Figure 7: Simulations with probability model for *slow worm* detection interval.

increases progressively with decrease in m/W. Smaller cell sizes or deployment of more detectors will therefore result in faster host-based detection of slow scanning worm attacks.

4. PROBABILITY MODEL FOR DETECTION INTERVAL

Scanning rates vary with different types of worms and in practice there is interest in predicting responsiveness and accuracy of worm defense systems. In this section, we present a probability model for detection interval and use Markov's Inequality to validate the model.

4.1 Fast Worm

Using assumptions made in section 3.1.1, if $G \sim U(0, W - m)$ is the number of scanned non-detector endpoints in the target cell before all m DEs are successfully scanned, then the cumulative distribution function (cdf) of t_{fd} can be expressed as a function of G, $P_{(G)}(t_{fd} \leq t)$:

$$P_{(G)}(t_{fd} \leq t) = 1 - e^{-\frac{rt}{m+G}} \qquad t \geq 0$$

It can be shown using total probability that:

$$P(t_{fd} \leq t) = \int_0^{W-m} P_{(G)}(t_{fd} \leq t | G = g) f_G(g) dg$$

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

$$P(t_{fd} \leq t) = \frac{1}{W-m} \int_0^{W-m} 1 - e^{-\frac{rt}{m+g}} dg \qquad (15)$$

This probability model computes the probability, $P(t_{fd} \leq t)$ that t is an upper bound for detection interval t_{fd}. According to Markov's Inequality,

$$P(t_{fd} \geq t) \leq \frac{E[t_{fd}]}{t}$$

Since,

$$P(t_{fd} > t) \leq P(t_{fd} \geq t) \leq \frac{E[t_{fd}]}{t}$$

then,

$$\{1 - P(t_{fd} \leq t)\} \leq \frac{E[t_{fd}]}{t} \qquad (16)$$

Fig.6(a) was generated using $\{1-P(t_{fd} \leq t)\}$ and $\frac{E[t_{fd}]}{t}$ and it shows that the developed probability distribution (15) for t_{fd} satisfies (16) irrespective of worm scanning rates.

4.2 Slow Worm

Using assumptions made in section 3.2.1, if Z is the number of hosts scanned until the first DE is scanned in the target cell, then the cumulative distribution function (cdf) of t_{sd} can be expressed as a function of Z, $P_{(Z)}(t_{sd} \leq t)$:

$$P_{(Z)}(t_{sd} \leq t) = 1 - e^{-\frac{rt}{Z}} \qquad t \geq 0$$

It can be shown using total probability that:

$$P(t_{sd} \leq t) = \sum_{z=1}^{W-m+1} P_{(Z)}(t_{sd} \leq t | Z = z) P_Z(z)$$

where $P_Z(z) = \frac{m}{W}(\frac{W-m}{W})^{z-1}$. Solving,

$$P(t_{sd} \leq t) = \sum_{z=1}^{W-m+1} (1 - e^{-\frac{rt}{z}}) \frac{m}{W} \left(\frac{W-m}{W}\right)^{z-1} \qquad (17)$$

This probability model computes the probability, $P(t_{sd} \leq t)$ that t is an upper bound for detection interval t_{sd}. According to Markov's Inequality,

$$\{1 - P(t_{sd} \leq t)\} \leq \frac{E[t_{sd}]}{t} \qquad (18)$$

Fig.7(a) and Fig.7(b) were generated using $\{1 - P(t_{sd} \leq t)\}$ and $\frac{E[t_{sd}]}{t}$ and they show that the developed probability distribution (17) for t_{sd} satisfies (18) irrespective of worm scanning rate and m/W ratio.

4.3 False Detection

False detection occurs when either a false negative or a false positive occurs. With the proposed detection scheme, a false negative would occur if the host AIDS on the DEs fail to detect occurrence of a malicious intrusion activity. Scenarios in which this can happen include:

1. If the host-based AIDS running on the DEs is not tuned to detect anomalies caused by the unauthorized intrusion worm activity. This scenario can be avoided by ensuring proper tuning.

2. If for fast scanning worms, the worm takes more time to scan all the DEs in the target cell than the FDA capture interval, t_f. Note that for fast scanning worms, the time it takes to scan all DEs in the target cell is the total inter-infection interval t_{fv} (section 3.1.1), which is equivalent to the detection interval, t_{fd} (9). Using (15), the probability that this false detection scenario occurs can be expressed as:

$$P(t_{fd} > t_f) = 1 - P(t_{fd} \leq t_f)$$
$$= 1 - \frac{1}{W-m} \int_0^{W-m} 1 - e^{-\frac{rt_f}{m+g}} dg \qquad (19)$$

   Fig. 6(b) shows that the probability of false detection decreases progressively with increase in worm scanning rate or with increase in t_f. Fig. 6(c) also shows that this probability reduces with smaller cell sizes.

3. If for slow scanning worms, the worm takes more time to scan at least one DE in the target cell than the SDA capture interval, t_s configured on the DEs. For slow scanning worms, the time it takes to scan at least one DE in the target cell is the total inter-infection interval t_{sv} (section 3.2.1), which is equivalent to the detection interval, t_{sd} (12). Using (17), the probability that this false detection scenario occurs can be expressed as:

$$1 - \sum_{z=1}^{W-m+1} (1 - e^{-\frac{rt_s}{z}}) \frac{m}{W} \left(\frac{W-m}{W}\right)^{z-1} \qquad (20)$$

   Fig. 7(c) shows that the probability of false detection decreases progressively with increase in worm scanning rate or with increase in t_s. Fig. 7(d) also shows that this probability reduces with increase in m/W. Therefore, increasing number of detectors or reducing the cell size reduces false detection probability.

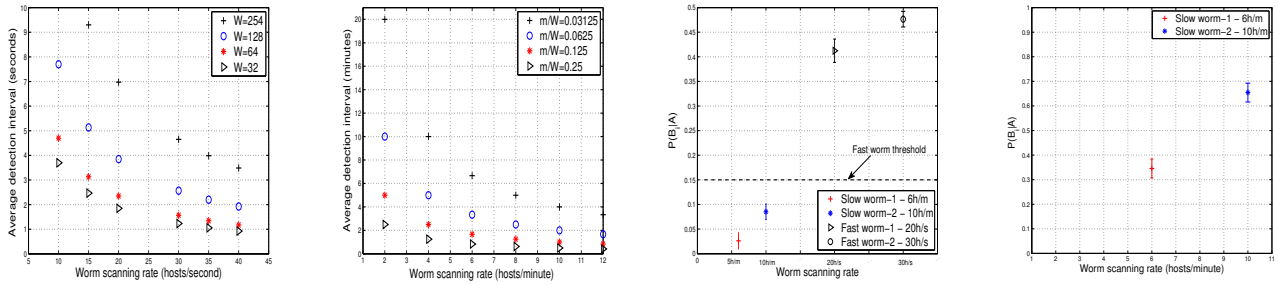
5. EXPERIMENTATION

To evaluate the functionality and performance of the proposed detection scheme on a live testbed, we emulated self propagating slow worm attacks using a modified *blaster worm* source code. To emulate multiple malicious attacks the source code was used to instrument two worms that exploited two different vulnerabilities. The first, *worm-1* was instrumented to create a directory named */root/infected-1* on the target host and copy a file named *malicious-1* into that directory over TCP port 888. The second, *worm-2* was instrumented to create a directory named */root/infected-2* on the target host and copy a file named *malicious-2* into that directory over UDP port 999. Fig. 1 shows the network topology of our live test-bed. Hosts in network-1 and network-2 were used to launch *worm-1* and *worm-2* random attacks respectively on hosts in the target networks (network A, network B, network C, and network D). We used OpenVZ virtualization [3] [20] to create the required vulnerable host population in the target networks. Up to 254 virtual hosts were created on individual Linux workstations running OpenVZ kernel-2.6.22 to emulate a class C network population in each target network.

For our test, the host-based AIDS running on the fast worm detector agent (FDA) and the slow worm detector agent (SDA) were emulated using different instances of snort-based IDS that constantly monitored the directory structure and content of the DE, and generated an alert when a file named *malicious-1* or *malicious-2* was found in a directory named */root/infected-1* or */root/infected-2* respectively on the DE. In our implementation, the snort-based IDS was used for real-time recording on the FDA and SDA [4]. t_f and t_s were set to 10 seconds and 25 minutes on the FDA and SDA respectively. For effectiveness, the malicious attacks randomly scanned hosts in one target network before selecting another target network. Worm scanning rates were varied between 10hosts/second (h/s) to 40h/s and from

[3]OpenVZ is an operating system-level virtualization technology based on the Linux kernel and operating system.

[4]Note that our emulation of host-based detection with snort-based alerts and real-time logging was only used to demonstrate the behavior of the proposed detection technique. Other host-based AIDS software such as Thirdbrigade host AIDS, Cisco Security Agent and Tripwire host AIDS can be used for detection in enterprise deployments.



(a) Average fast worm detection interval, $\overline{t_{fd}}$. $\Phi_f = 0.15$.

(b) Average slow worm detection interval, $\overline{t_{sd}}$. $m = 4$

(c) Snaphot of $P(B_i|A)$ on the FCE. $m = 4, W = 254, \Phi_f = 0.15$.

(d) Snaphot of $P(B_i|A)$ on the SCE. $m = 4, W = 254$.

Figure 8: Experimental results with simultaneous fast and slow worm attacks.

2hosts/minute (h/m) to 12h/m to emulate fast and slow scanning worms respectively. For each scanning rate, 20 runs of the attack were carried out to minimize statistical errors. The gateway routers, GR-1 and GR-2 ran instances of our proposed fast correlation engine (FCE) and slow correlation engine (SCE).

The experiment was carried out to demonstrate and evaluate detection of simultaneous fast and slow worm attacks using the proposed integrated detection technique.

5.1 Experiment: Multiple simultaneous fast and slow worm attacks

In this experiment, four attacking hosts, two from network-1 and two from network-2 in Fig. 1 were used to launch different attacks (worm-1 and worm-2 respectively) on hosts in the target networks. The scanning rate of a pair of worm-1 and worm-2 attacks were set to 20h/s and 30h/s respectively to emulate fast worms. The scanning rate of the second pair of worm-1 and worm-2 attacks were set to 6h/m and 10h/m respectively to emulate slow worms. The objective of the experiment was to investigate the effectiveness of the proposed detection scheme in detecting simultaneous fast and slow scanning malicious worm attacks on a target network. The target cell sizes (i.e W) were varied by varying the number of vulnerable virtual hosts in the target network to investigate impact of cell size.

Fig. 8(a) shows that average detection interval for the fast worms reduces progressively with increase in worm scanning rate. It also shows that smaller cell sizes result in faster detection. These observations concur with results obtained analytically in Section III. The results show that the proposed detection scheme is capable of detecting malicious intrusion attacks with scanning rate of 20h/s or more within 7 seconds after starting the attack on a network with cells comprising no more than 254 hosts.

Fig. 8(b) shows that average detection interval for slow worms reduced progressively with increase in worm scanning rate. It also shows that a reduction in the W which increases the m/W ratio resulted in faster detection. These observations also concur with results obtained analytically in Section III. Fig. 8(b) shows that the detection scheme is capable of detecting slow scanning worm attacks with scanning rate of over 2h/m within 10 minutes after starting the attack on cells with m/W ratio of $\frac{1}{16}$ (0.0625) or more.

Fig.8(c) and Fig.8(d) show snapshots of results from both the fast correlation algorithm (FCA) and slow correlation algorithm (SCA) respectively. Fig.8(c) shows that though the FCE received both fast and slow worm traffic profiles during the fast worm detection window, only the fast worm traffic profiles exhibited a posteriori probability, $P(B_i|A)$ greater that the chosen threshold, $\Phi_f = 0.15$ and therefore were selected by the FCA. Also, even though both fast and slow worm profiles are captured by the SDA during the slow worm detection window, the adaptive profiler ensures that only the slow worm profiles not selected by the FCA are forwarded to the SCE. On the SCE, Fig.8(d) shows that the slow worm traffic profiles exhibited $P(B_i|A) > 0$ and therefore were selected by the SCA.

6. CONCLUSION AND FUTURE WORK

In this paper, we proposed a unique integrated detection technique capable of detecting and identifying traffic flow(s) responsible for simultaneous fast and slow scanning malicious worm attacks. The detection approach uses a combination of evidence from host-based anomaly detectors, a self-adapting profiler and Bayesian inference from network heuristics for detection. We assumed that fast scanning worm attacks are extreme events which exhibit an extremely high number of connection attempts on vulnerable targets and therefore seemed to be well suited for extreme value theory (EVT). Hence, we used an EVT graphical tool, sample mean excess function to determine appropriate thresholds used for Bayesian inference during fast worm detection.

Worm detection interval was modeled as a stochastic variable and analysis of detection interval for both fast and slow scanning worms using the detection technique was presented. Based on the analysis, a probability model for worm detection interval was developed and validated using Markov's Inequality. The probability model was also used to show that false detection rates can be reduced if certain network parameters are optimized. We experimented with the integrated detection scheme on a live test-bed and the generated results concurred with our analytical model and results.

For future work, we intend to extend our work on integrated detection of worms to peer-to-peer networks. With proliferation of Web 2.0 and peer-to-peer social networks,

a new vulnerability and threat model for large scale infection of unprotected systems and networks by worms seem quite conceivable. We are encouraged by the results of this work and the observation that false detection rates of scanning worms can be improved with optimization of certain network parameters. More work is required to develop adequate detection techniques for both fast and slow worms in peer-to-peer networks.

Also, we intend to investigate the use of extreme value theory for detection of DDoS attacks.

7. REFERENCES

[1] J. Agosta, C. Diuk-Wasser, J. Chandrashekar, and C. Livadas. An adaptive anomaly detector for worm detection. In *SYSML'07: Proceedings of the 2nd USENIX workshop on Tackling computer systems problems with machine learning techniques*, pages 1–6, Berkeley, CA, USA, 2007. USENIX Association.

[2] F. Akujobi, I. Lambadaris, and E. Kranakis. Endpoint-driven intrusion detection and containment of fast spreading worms in enterprise networks. In *IEEE Military Communications Conference (MILCOM) 2007*, 2007.

[3] A. Balkema and L. de Haan. Residual life time at great age. *The Annals of Probability*, 2(5):792–804, 1974.

[4] M. Burgess. Probabilistic anomaly detection in distributed computer networks. *Science of Computer Programming*, 60(1):1–26, March 2006.

[5] C. A. CA-2001-26. Nimda worm. http://www.cert.org/advisories/CA-2001-26.html, 2001.

[6] M. Costa, J. Crowcroft, M. Castro, A. Rowstron, L. Zhou, L. Zhang, and P. Barham. Vigilante: End-to-end containment of Internet worms. In *Proceedings of the Symposium on Systems and Operating Systems Principles (SOSP*, pages 133–147, 2005.

[7] C. Gates and C. Taylor. Challenging the anomaly detection paradigm: A provocative discussion. In *NSPW '06: Proceedings of the 2006 workshop on New security paradigms*, pages 21–29, New York, NY, USA, 2006. ACM.

[8] J. Jung, V. Paxson, A. Berger, and H. Balakrishnan. Fast portscan detection using sequential hypothesis testing. In *In Proceedings of the IEEE Symposium on Security and Privacy, May 9– 12, 2004.*, 2004.

[9] E. KÃńllezi and M. Gilli. Extreme value theory for tail-related risk measures. FAME Research Paper Series rp18, International Center for Financial Asset Management and Engineering, Oct. 2000.

[10] D. Moore, C. Shannon, and K. Claffy. Code Red: A case study on the spread and victims of an Internet worm. In *ACM SIGCOMM Internet Measurement Workshop*, pages 273–284, 2002.

[11] D. Mutz, F. Valeur, C. Kruegel, and G. Vigna. Anomalous system call detection. *ACM Transactions on Information and System Security*, 9:61–93, 2006.

[12] D. Nicol. The impact of stochastic variance on worm propagation and detection. In *WORM '06: Proceedings of the 4th ACM workshop on Recurring malcode*, pages 57–64, New York, NY, USA, 2006. ACM.

[13] J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131, 1975.

[14] S. Schechter, J. Jung, and A. Berger. Fast Detection of Scanning Worm Infections. In *7th International Symposium on Recent Advances in Intrusion Detection (RAID)*, French Riviera, France, September 2004.

[15] D. Schirmacher, E. Schirmacher, and N. Thandi. Stochastic excess-of-loss pricing within a financial framework. http://www.casact.org/pubs/forum/05spforum/05spf297.pdf, 2005.

[16] V. Sekar, Y. Xie, M. Reiter, and H. Zhang. A multi-resolution approach for worm detection and containment. In *DSN '06: Proceedings of the International Conference on Dependable Systems and Networks*, pages 189–198, Washington, DC, USA, 2006. IEEE Computer Society.

[17] C. Shannon and D. Moore. The spread of the witty worm. In *IEEE Security Privacy, vol. 2, no. 4*, 2004.

[18] S. Singh, C. Estan, G. Varghese, and S. Savage. Automated worm fingerprinting. In *In OSDI*, pages 45–60, 2004.

[19] C. Sullivan. *Cisco Security Agent*. Cisco Press, 2005.

[20] Swsoft. Openvz homepage. http://openvz.org/, 2008.

[21] C. Systems Inc. Cisco Catalyst 6500 Supervisor Engine 32 PISA. http://www.cisco.com/en/US/products/ps7209/index.html, 2008.

[22] N. Weaver, S. Staniford, and V. Paxson. Very fast containment of scanning worms. In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium*, pages 3–3, Berkeley, CA, USA, 2004. USENIX Association.

[23] C. C. Zou, L. Gao, W. Gong, and D. Towsley. Monitoring and early warning for Internet worms. In *CCS '03: Proceedings of the 10th ACM conference on Computer and communications security*, pages 190–199, New York, NY, USA, 2003. ACM.