Automatic Camouflaging Worm Detection Over Internet

N. Surya Prakash Raju, M. Anil Kumar
Dept. of CSE, Aditya Engineering College, Aditya Nagar, Surampalem, AP, India

Abstract
Active worms pose major security threats to the Internet. This is due to the ability of active worms to propagate in an automated fashion as they continuously compromise computers on the Internet. Active worms evolve during their propagation and thus pose great challenges to defend against them. In this paper, we investigate a new class of active worms, referred to as Camouflaging Worm (C-Worm in short). The C-Worm is different from traditional worms because of its ability to intelligently manipulate its scan traffic volume over time. Thereby, the C-Worm camouflages its propagation from existing worm detection systems based on analyzing the propagation traffic generated by worms. We analyze characteristics of the C-Worm and conduct a comprehensive comparison between its traffic and non-worm traffic (background traffic). We observe that these two types of traffic are barely distinguishable in the time domain. However, their distinction is clear in the frequency domain, due to the recurring manipulative nature of the C-Worm. Motivated by our observations, we design automatic detection of C-worm and specifying priorities to the different worms over the Internet and the proposed system automatically applied on the worm according to the priority of the worm without disturbing their the real work. Our scheme uses the Power Spectral Density (PSD) distribution of the scan traffic volume and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from background traffic. Using a comprehensive set of detection metrics and real-world traces as background traffic, we conduct extensive performance evaluations on our proposed spectrum-based detection scheme. The performance data clearly demonstrates that our scheme can effectively detect the C-Worm propagation.

Keywords
Worm, Camouflage, Anomaly Detection

I. INTRODUCTION
An active worm refers to a malicious software program that propagates itself on the Internet to infect other computers. The propagation of the worm is based on exploiting vulnerabilities of computers on the Internet. Many real-world worms have caused notable damage on the Internet. These worms include “Code-Red” worm in 2001 [1], “Slammer” worm in 2003 [2], and “Witty”/“Sasser” worms in 2004 [3]. Many active worms are used to infect a large number of computers and recruit them as bots or zombies, which are networked together to form botnets [4]. These botnets can be used to launch massive Distributed Denial-of-Service (DDoS) attacks that disrupt the Internet utilities [5], (b) access confidential information that can be misused [6] through large scale traffic sniffing, key logging, identity theft etc, (c) destroy data that has a high monetary value [7], and (d) distribute large-scale unsolicited advertisement emails (as spam) or software (as malware). There is evidence showing that infected computers are being rented out as “Botnets” for creating an entire black-market industry for renting, trading, and managing “owned” computers, leading to economic incentives for attackers [4], [8], [9]. Researchers also showed possibility of “super-botnets,” networks of independent botnets that can be coordinated for attacks of unprecedented scale [10].

Also, they can be extremely versatile and resistant to countermeasures. Due to the substantial damage caused by worms in the past years, there have been significant efforts on developing detection and defense mechanisms against worms. A network based worm detection system plays a major role by monitoring, collecting, and analyzing the scan traffic (messages to identify vulnerable computers) generated during worm attacks. In this system, the detection is commonly based on the self-propagating behavior of worms that can be described as follows:

In this paper, we conduct a systematic study on a new class of such smart-worms denoted as Camouflaging Worm (C-Worm in short). The C-Worm has a self-propagating behavior similar to traditional worms, i.e., it intends to rapidly infect as many vulnerable computers as possible. However, the C-Worm is quite different from traditional worms in which it camouflages any noticeable trends in the number of infected computers over time. We comprehensively analyze the propagation model of the C-Worm and corresponding scan traffic in both time and frequency domains. We observe that although the C-Worm scan traffic shows no noticeable trends in the time domain, it demonstrates a distinct pattern in the frequency domain. Based on the above observation, we adopt frequency domain analysis techniques and develop a detection scheme against wide-spreading of the C-Worm. In this paper, we are introducing the concept called set priority to different worms using modeling detection methods that the system automatically works on the worms without disturbing the current work on the net. We follow the Power Spectral Density (PSD) distribution of scan traffic volume in the frequency domain and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from nonworm traffic (background traffic).

II. Related Work

A. Active Worms
Active worms are similar to biological viruses in terms of their infectious and self-propagating nature. They identify vulnerable computers, infect them and the worm-infected computers propagate the infection further to other vulnerable computers. In order to understand worm behavior, we first need to model it. With this understanding, effective detection and defense schemes could be developed to mitigate the impact of the worms. For this reason, tremendous research effort has focused on this area [12, 14, 16, 24-25]. Active worms use various scan mechanisms to propagate themselves efficiently. The basic form of active worms can be categorized as having the Pure Random Scan (PRS) nature. Other worms propagate themselves more effectively than PRS worms using various methods. Different from the above worms, which attempt to accelerate the propagation with new scan schemes, the Camouflaging Worm (C-Worm) studied in this paper aims to elude the detection by the worm defense system during worm propagation. The C-Worm also shares some similarity with stealthy port-scan attacks. Such attacks try to find out available services in a target system, while avoiding detection [35-36]. It is accomplished by decreasing the port scan rate, hiding the origin of attackers, etc. Due to the nature of self-propagation, the C-Worm must use more complex mechanisms to manipulate the scan traffic volume over time in order to avoid detection.
B. Worm Detection

Worm detection has been intensively studied in the past and can be generally classified into two categories: “host-based” detection and “network-based” detection [a]. In order to rapidly and accurately detect Internet-wide large scale propagation of active worms, it is imperative to monitor and analyze the traffic in multiple locations over the Internet to detect suspicious traffic generated by worms. Besides the above detection schemes that are based on the global scan traffic monitor by detecting traffic anomalous behavior, there are other worm detection and defense schemes such as sequential hypothesis testing for detecting worm-infected computers [44], payload-based worm signature detection [34, 45]. Despite the different approaches described above, we believe that detecting widely scanning anomaly behavior continues to be a useful weapon against worms, and that in practice multifaceted defense has advantages.

C. C-Worm: Modeling of the C-Worm

The C-Worm camouflages its propagation by controlling scan traffic volume during its propagation. The simplest way to manipulate scan traffic volume is to randomly change the number of worm instances conducting port-scans. worm propagation over the Internet can be considered a dynamic system. When an attacker launches worm propagation, it is very challenging for the attacker to know the accurate parameters for worm propagation dynamics over the Internet. As more and more computers get infected, they, in turn, take part in scanning other computers. Hence, we consider the C-Worm as a worst case attacking scenario that uses a closedloop control for regulating the propagation speed based on the feedback propagation status. In order to effectively evade detection, the overall scan traffic for the C-Worm should be comparatively slow and variant enough to not show any notable increasing trends over time. On the other hand, a very slow propagation of the C-Worm is also not desirable, since it delays rapid infection damage to the Internet. Hence, the C-Worm needs to adjust its propagation so that it is neither too fast to be easily detected, nor too slow to delay rapid damage on the Internet. A control parameter called attack probability P(t) for each worm-infected computer. P(t) is the probability that a C-Worm instance participates in the worm propagation (i.e. scans and infects other computers) at time t. Our C-Worm model with the control parameter P(t) is generic. P(t) = 1 represents the cases for traditional worms, where all worm instances actively participate in the propagation. For the C-Worm, P(t) needs not be a constant value and can be set as a time varying function. In order to achieve its camouflaging behavior, the C-Worm needs to obtain an appropriate P(t) to manipulate its scan traffic. Specifically, the C-Worm will regulate its overall scan traffic volume such that [a]. There are other approaches to achieve this goal, such as incorporating the Peer-to-Peer techniques to disseminate information through secured IRC channels [49], [50].

D. Propagation Model of the C-Worm

To analyze the C-Worm, we adopt the epidemic dynamic model for disease propagation, which has been extensively used for worm propagation modeling [2, 12]. Based on existing results [2, 12], this model matches the dynamics of real worm propagation over the Internet quite well. For this reason, similar to other publications, we adopt this model in our paper as well. Since our investigated C-Worm is a novel attack, we modified the original Epidemic dynamic formula to model the propagation of the C-Worm by introducing the P(t) - the attack probability that a worm-infected computer participates in worm propagation at time t. We note that there is a wide scope to notably improve our modified model in the future to reflect several characteristics that are relevant in real-world practice.

E. Effectiveness of the C-Worm

We now demonstrate the effectiveness of the C-Worm in evading worm detection through controlling P(t). Given random selection of Mc, we generate three C-Worm attacks (viz., C-Worm 1, C-Worm 2 and C-Worm 3) that are characterized by different selections of mean and variance magnitudes for MC. Fig. 1 shows the observed number of worm-infected computers over time for the PRS worm and the above three C-Worm attacks. Fig. 2 shows the infection ratio for the PRS worm and the above three C-Worm attacks. These simulations are for a worm detection system discussed [a].

III. Proposed Model

In this paper, we focus on a new class of worms, referred to as the camouflaging worm (C-Worm). The C-Worm adapts their propagation traffic patterns in order to reduce the probability of detection, and to eventually infect more computers. The C-Worm is different from polymorphic worms that deliberately change their payload signatures during propagation [34, 52]. Recent studies also showed that existing commercial anti-worm detection systems fail to detect brand new worms and can also be easily circumvented by worms that use simple mutation techniques to manipulate their payload.
In this paper, we are working on the new concept of setting priorities to the worms over the internet and by running our system over the internet according to the network rate (NR) to model and detect C-Worm. The system first incorporates the Power Spectral Density (PSD) distribution of the scan traffic volume and its corresponding Spectral Flatness Measure (SFM) to distinguish the C-Worm traffic from background traffic. By considering the PSD and SFM, the system automatically detects and assigns values as priorities to distinguish the effect of the worms over the internet. In order to identify the C-Worm propagation in the frequency domain, we use the distribution of Power Spectral Density (PSD) and its corresponding Spectral Flatness Measure (SFM) of the scan traffic. Particularly, PSD describes how the power of a time series is distributed in the frequency domain. Mathematically, it is defined as the Fourier transform of the auto-correlation of a time series. In our case, the time series corresponds to the changes in the number of worm instances that actively conduct scans over time. The SFM of PSD is defined as the ratio of geometric mean to arithmetic mean of the coefficients of PSD. The range of SFM values is [0, 1] and a larger SFM value implies a flatter PSD distribution and vice versa.

To illustrate SFM values of both the C-Worm and normal non-worm scan traffic, we plot the Probability Density Function (PDF) of SFM for both C-Worm and normal non-worm scan traffic as shown in Fig. 5 and Fig. 6, respectively.

The proposed system works on the worms which are prioritized differences the type of attack over the internet, and the system automatically shows the impact over the internet to reduce the runtime of the worm and eliminates the worms from the internet by applying the above methods. The system maintains Worm history in the to specify type, nature, etc.

![Fig. 3: PDF of SFM on C-Worm Traffic](image)

![Fig. 4: PDF of SFM on Normal Non-Worm Traffic](image)

As a matter of fact, other existing detection schemes based on the scan traffic rate [20], variance [21] or trend [19] will also demand a high sampling frequency for ITM systems in order to accurately detect worm attacks. Enabling the ITM system with timely data collection will benefit worm detection in real-time.

### IV. Performance Evaluation

In this section, we report our evaluation results that illustrate the effectiveness of our priority based worm detection and elimination Scheme.

In our evaluation, we considered both experiments with real-world "non-worm" traffic and simulated c-worm traffic. To make our experiments reflect real-world practice, some key parameters that we used to generate C-worm traffic in our simulation were based on previous results from a real-worm incidence - "Code-Red" worm in 2001 [1]. Specifically, we set the total number of vulnerable computers on the Internet as 360,000, which is the maximum number of computers which could be infected by "Code-Red" worm. Additionally, we set the scan rate S (number of scans per minute) to be variable within a range, this allows us to emulate the infected computers in different network environments. In our evaluation, the scan rates are predetermined and follow a Gaussian distribution. In our evaluation, we merged the simulated C-worm attack traffic into replayed "non-worm" traffic traces and carried out evaluation study.

### Table 1: Evaluation Metrics

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection Rate (IR)</td>
<td>Ratio of worm infection over time without the presence of detection scheme</td>
</tr>
<tr>
<td>Maximum Infection Rate (MIR)</td>
<td>Ratio of worm infection at the moment that worms being detected</td>
</tr>
<tr>
<td>Detection Time (UT)</td>
<td>Time taken to successfully detect a worm spreading worm based on both</td>
</tr>
</tbody>
</table>

Table 1: Evaluation Metrics

A. Performance of Detection Schemes

We evaluate our proposed priority based detection schema by comparing its performance with three existing representative traffic volume-based detection schemes. The first scheme is the volume mean-based (MEAN) detection scheme which uses mean of scan traffic to detect worm propagation [20]; the second scheme is the trend-based (TREND) detection scheme which uses the increasing trend of scan traffic to detect worm propagation [19]; and the third scheme is the victim number variance based (VAR) detection scheme which uses the variance of the scan traffic to detect worm propagation [21]. We define our Priority based detection schema as PDS. Table 2 shows the detection results of different detection schemes against the C-Worm. The results have been averaged over 500 C-Worm attacks. From this table, we can observe that existing detection schemes are not able to effectively detect the C-Worm and their detection rate (PD) values are significantly lower in comparison with our priority based detection scheme PDS. For example, PDS achieves the detection rate of 99%, which is at least 3-4 times more accurate than detection schemes such as VAR and MEAN that achieve detection rate values of only 48% and 14%, respectively.
Table 2: Detection Results for the C-Worm

<table>
<thead>
<tr>
<th>Detection Rate (DR)</th>
<th>VAR</th>
<th>TREND</th>
<th>MEAN</th>
<th>PSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>18%</td>
<td>0%</td>
<td>14%</td>
<td>99.5%</td>
<td></td>
</tr>
<tr>
<td>Maximal Infection Ratio (MIR)</td>
<td>14.4%</td>
<td>100%</td>
<td>7.5%</td>
<td>2%</td>
</tr>
<tr>
<td>Detection Time (DT) in minutes</td>
<td>2367</td>
<td>∞</td>
<td>183</td>
<td>1234</td>
</tr>
</tbody>
</table>

V. Conclusions and Future Work

In this paper, we studied a new class of worm called C-Worm, which has the capability to camouflage its propagation and further avoid the detection. Our investigation showed that, although the C-Worm successfully camouflages its propagation in the time domain, its camouflaging nature inevitably manifests as a distinct pattern in the frequency domain. Based on observation, we developed a novel spectrum-based detection scheme to detect the C-Worm. Our evaluation data showed that our scheme achieved superior detection performance against the C-Worm in comparison with existing representative detection schemes. This paper lays the foundation for ongoing studies of “smart” worms that intelligently adapt their propagation patterns to reduce the effectiveness of countermeasures.

References


Mr. N. Surya Prakash Raju, well known & excellent teacher Received M.Tech (CSE) from JNTU, Kakinada is working as Sr.Asst. Professor, Department of CSE at Aditya Engineering College. He has 10 years of teaching experience in various Engineering Colleges. To his credit couple of publications both national and international conferences/journals. His area of Interest includes Computer Networks, Information Security, Soft Computing Techniques and other advances in computer Applications. And guided many Projects.

Mr. M. Anil Kumar is a student of Aditya Engineering College, Surampalem. Presently he is pursuing his M.Tech from this college and He received his graduation from JNTUK UNIVERSITY in the year 2010.