
S.Divya¹ and Dr.G.Padmavathi²

¹ Avinashilingam Institute for Home Science and Higher Education for Women University
Department of Computer Science, Coimbatore, India
Email: divya.phd.research@gmail.com

² Avinashilingam Institute for Home Science and Higher Education for Women University
Department of Computer Science, Coimbatore, India
Email: ganapathi.padmavathi@gmail.com

Abstract—Information security is one of the major concerns for military, government, civil and commercial organizations and security risk has been immensely raised on the internet access. Self-duplicating, self-propagating malicious codes known as worms spread themselves without any human interaction and launch the most destructive attacks against networks and cause high security risks. Increasing threats from worms in the network continue to be a challenging task to detect and handle. Various worms exist in network affecting the communication security and different worm handling mechanisms exists to handle the threat created by them. Considering the features and metrics of the existing mechanisms in the detection and containment of worms in the network, this paper surveys the classification of worms and several existing defence mechanism and metrics to detect those worm attacks in the network.

Index Terms—Worm Propagation, Detection Techniques, Containment Algorithms.

I. INTRODUCTION

The worms in the network are computer programs which self-propagate by sending copies themselves from one node to another over a network. Such transmissions occur without any user intervention, thereby allowing them to be spread quickly and easily. Worms use multiple vulnerabilities to spread, such as Remote Procedure Calls (remote execution of a program), Buffer Overflows (data is stored in memory location other than the memory allocated by the programmer), Remote Command Execution (running a shell command remotely on a different host), Cross-site scripting (inject malicious links in the web interface). Worm propagating in the network infects the vulnerable hosts and the affected host acts as the worm injecting host further for spreading. Spread of worms in the network is shown in the Fig. 1 below.

Many real-world worms have caused notable damage on the network and computer. Those worms include “Code-Red” worm in 2001, “Slammer” worm in 2003, “Witty”/ “Sasser” worms in 2004, StuxNet worms in 2010-2012. The recent SQL Slammer worm infected more than 90% of hosts on the Internet within 10 minutes. Within the period of five years, 4,00,000 computers got infected by Blaster worm. Network worms cause immense damage to the network-dependent military, commercial and social service infrastructure of...
Nations throughout the world. Various algorithms are used to detect the unusual traffic patterns indicating the presence of a network worm. Detection and handling of network worm using existing defence mechanisms is still difficult.

Worm performs its process actively and dynamically, its detection is possible only after its spread. Active measures have to be done to prevent the worm initialization in the network through prevention techniques by some signature features, considering the existing techniques. The survey paper is systematized as follows: Section 2 discusses the related works on different worms. Section 3 characteristics of worms are discussed. Section 4 conveys different categories of worms and various worms existing in the network. Different existing worm detection in Section 5 and containment techniques are explained in Section 6. Section 7 describes the proposed method for worm detection. Section 8 briefly presents the performance metrics used for detection. Finally, Section 9 concludes the survey paper.

II. RELATED WORK

Worms classification addressed in Andrika Pratama et al. (Andrika Pratama and Fauzi Adi Rafrastara 2012), where the authors surveyed based on their structure, how attack performed, how defensive measures handled by them and how worms are fought against by system users. Vishrut Sharma (2011), discussed types of worms based on programming and payload, characteristics description on propagation method and payload and programming trends classified into qualitative and quantitative. Yong Tang et al. (Yong Tang, Jiaqing Luo, Bin Xiao and Guiyi Wei 2009), focused on the concept and categorizing based on target identification, characteristics of internet worms by three factors, namely target finding strategy, method of propagation and anti-detection schemes. Various defense mechanisms classified on the finding activities of internet worms, modelling and analysis on different worms and the research trends followed. Pele Li et al. (Pele Li, Mehdi Salour and Xiao Su 2008), presented a survey analysis providing the characteristics of internet worms based on their behaviours, classified on parameter basis, different detection techniques with reference to worm types and exploration of current containment methods by slowing down or stopping its spread.

Nir Nissim et al. (Nir Nissim, Robert Moskovitch, Lior Rotash and Yuval Elович 2012), applied support vector machine for reducing features for classification and active learning is used to improve classifier performance and robustness in misleading occurrences. Also, RBF kernel function produced effective result and achieved improved accuracy in detecting unknown worms. Wei Yu et al. (Wei Yu, Xun Wang, Adam Champion, Dong Xuan and David Lee 2011), designed Distribution Entropy-based Dynamic detection scheme (DED) to utilize the attacks distribution and provide robust statistical aspects to detect Varying Scan
Rate worm (VSR) traffic from normal. Maximal infection ratio and detection time is achieved effectively to defend against new worms threats and vulnerabilities. Wei Yu et al. (Wei Yu, Xun Wang, Prasad Calyam, Dong Xuan and Wei Zhao 2011), investigated Camouflaging worm whose intelligent manipulation scans traffic volume overtime and these worms are distinguished by frequency domain. Power Spectral Density distribution and its consequent Spectral Flatness Measures are used to distinguish the C-worm traffic from normal background traffic and performance evaluated based on infection ratio and detection time. Chao Chen et al. (Chao Chen, Zesheng Chen and Yubin Li 2010), effective countermeasures were used for identifying the weakness and defending mechanisms like removal of infected hosts and honeypots are studied for the divide-conquer scanning worms. The various surveys and detection techniques convey the existing issue by internet worms is still a big challenge.

III. CHARACTERISTICS OF WORMS

Worms are recognized based on their behaviours and they are characterized into four main categories [10]. Once the worms enter into the network for its progress, it moves to different stages. They are:

i. Target Finding
ii. Propagation Format
iii. Transmission Methods
iv. Payload Schemes

Target finding scheme and propagation are two schemes, where the detection algorithms to be developed. Transmission methods and payload are schemes, where containment algorithms are to be developed. The propagation of worm is classified under four phases. Fig. 2 gives the process of worm progress.

![Figure. 2. Overview of Internet Worm progress](image)

A. Target Finding
When the worm enters into the network, its initial step is finding target to spread and exploit. Various target finding schemes are blind target scanning, hit list scanning, topological scanning, passive and web search. Many current worms target servers in the web.

B. Propagation Format
After the target is identified by the initial worm, the worm spreads copies of itself to other victims through various schemes. Some of the propagation formats are self-carried, second channel, embedded and Botnet schemes.
C. Transmission Methods
Transmissions of worms are performed through TCP and UDP worms. TCP worms are connection oriented and latency limited. These worms block the thread progress. UDP worms infect through self-carried. They are connectionless and bandwidth limited. UDP worms block resources in the network.

D. Payload Schemes
Payloads are referred to as the worm code. If worms are encoded, then worms are harder to detect in the network. Monomorphic, Polymorphic and Metamorphic worm schemes are different worm payload formats. Internet worms based on the above factors, discovers targets, drives copies, and chooses transmission media and payload schemes for safe spread. Anti-virus detection techniques are to be updated for better detection based on their spreading factors.

IV. Categories Of Worms
The worms are broadly classified into four different categories [1][12]. The worms are categorized based on their payload, created by respective authors of the worms. The following table I lists the categories of worms.

<table>
<thead>
<tr>
<th>Worms</th>
<th>Characteristics</th>
<th>Existing Defence Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer-to-peer</td>
<td>-Severe threat and live streaming function</td>
<td>-Multicast</td>
</tr>
<tr>
<td></td>
<td>-For the purpose of content distribution</td>
<td>-Scanning</td>
</tr>
<tr>
<td></td>
<td>P2P users use these networks</td>
<td>-Non-Coercive Approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Peer-based Monitoring Approach</td>
</tr>
<tr>
<td>E-Mail</td>
<td>-Mailboxes contain large number of spam</td>
<td>-Machine learning algorithms using</td>
</tr>
<tr>
<td></td>
<td>-Instant Messaging occupied with the advertisement bots.</td>
<td>statistical representations</td>
</tr>
<tr>
<td></td>
<td>-Publicity included in wikis</td>
<td>-Vector Space Model</td>
</tr>
<tr>
<td></td>
<td>-Redundant SMS interactions</td>
<td>-Honeynet</td>
</tr>
<tr>
<td>Instant Messaging(IM)</td>
<td>-Home IM users an organizations that allow IM in workplace suffer with these</td>
<td>-HoneyIM</td>
</tr>
<tr>
<td></td>
<td>security threats.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-Enterprise like Networks lack to safeguard the network from IM malware like zero day attack.</td>
<td></td>
</tr>
<tr>
<td>Internet</td>
<td>-Network security communities major concern</td>
<td>-Signature based schemes</td>
</tr>
<tr>
<td></td>
<td>-To spread in the network, these worms use P2P vulnerabilities</td>
<td>-Anomaly based schemes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-GRIDs</td>
</tr>
</tbody>
</table>

There are many different types of worms existing and they all fall under the above categories. Worms are classified into Active and Passive worms. Active worms [10] propagate themselves into the network and impose threat to network security without user intervention. Various active worms are code Red II, Slammer, Witty, Nimda, C-Worm and Morris. Passive worms[10] are similar to virus and they require user intervention or any mechanism behavior to start their propagation. Different existing passive worms are NElissa, VBS-Gnetella, W32.Gnuman, Fizzer, worm.Lolol.b and worm.Kitro.
Different types of worms are existing in the network and causing latency and bandwidth limitation. Some of the worms are Hitlist worms, Polymorphic worms, Benign worms, AutoRun worms, Divide-Conquer-Scanning worms, Importance Scanning worms and Self-Disciplinary worms.

A. Hitlist worms

The huge groups of vulnerable hosts are first selected, prior to their spreading progress rather than the general random scanning process. Before the process of releasing worm for spreading, Hitlist worm author groups the vulnerable host through slow scan of ports. Hit list consists of all vulnerable hosts to be affected. The list consists of index of IP addresses, group of DNS names and collection of hash table entries.

B. Polymorphic worms

These worms exploit the buffer overflow vulnerability and they have their structure summarized in network protocol frame. Polymorphic worms [2] at each execution of infection alter their byte sequences. These worms initialize their progress through network Protocol commands and exploit the target code.

C. Benign worms

Benign worms exploit software vulnerabilities. Various types of benign worms are passive, hybrid, active and IDS based on spread strategies. SWORM and RWORM are two different benign worms.

D. AutoRun worms

These worms exploit and affect the removable devices. They are dynamic threats. Controlling the usage of removable device decreases the spread of AutoRun worms and maximize the recovery rate.

E. Divide-Conquer-Scanning worms

Divide-Conquer-Scanning Worms [5] spread through the traditional process of random scanning. Through various infected hosts using different scanning rate, probability and space, these worms infect their targets. These worms are strong epidemic attacks in the Internet.

F. Importance Scanning worms

In internet, irregularly distributed vulnerable hosts are exploited by the importance Scanning worms. Vulnerable host distribution is to be calculated and determined to identify and analyze these worms. Static optimal scanning, dynamic optimal scanning and self-learning worms are different importance scanning worms.

G. Self-Disciplinary worms

Detection probability[13] is decreased by these worms by adapting propagation traffic patterns in infectious computers. To defend against suspicious countermeasures, achieving its exploitation and delay detection, these worms implements their own prorogation patterns. Dynamic and static self-disciplinary worms are of this category.

Worms exploit the vulnerabilities and destructs the hosts and networks. The characteristics of various worms are listed in the table II below

<table>
<thead>
<tr>
<th>Name of worms</th>
<th>Defensive Mechanisms</th>
<th>Speed</th>
<th>Security Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pro active</td>
<td>Re active</td>
<td>Fast</td>
</tr>
<tr>
<td>Hitlist</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>Polymorphic</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>Benign</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>AutoRun</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Divide-Conquer-Scanning</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>Importance Scanning</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Self-Disciplinary</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
</tbody>
</table>
The above listed worms are listed, based on their propagation factors. These worms still exist in the network exploitation and those are prevented using existing techniques. The existing and new worms entering the network should be detected earlier based on the propagation factors.

V. WORM DETECTION

Based on the parameters used for detection, algorithms can be roughly divided into Signature Based (Known Signature) and Anomaly Based Schemes (Unknown signature). Some of the existing worm detection techniques are shown in Fig. 3.

A. Signature Based Detection

Signature-based detection[15] is normally used for detecting known attacks. Knowledge of normal traffic is not required in this case. A signature database is needed for this type of detection system. Signature-based systems take a look at the payload and identify whether or not it contains a worm.

B. Anomaly Based Detection

The anomaly-based detections[15] check the headers of packets to define the type of connection to which the packet belongs to. They observe the network traffic volume, the monitored hosts’ behavior and detect abnormal behaviors. While a packets header information is useful to detect attacks exploiting vulnerabilities of the network stack or probing hosts for vulnerable services, a packets payload information can be used to detect attacks directed at vulnerable applications.

C. Honeypots

Virtual Honeypot was used for worm detection. In an emulator, minimal honeypot uses virtual machines and multihome to cover a large address space called Honey Stat. It is used to gather information about worms as well as capture worms.

**HoneyIM:** HoneyIM is a framework for automating the process of IM malware detection and suppression in an enterprise network. HoneyIM uses decoy accounts to trap IM malware by leveraging malware spreading characteristics.

**Honeynet:** It is a high-interaction honeypot designed to capture extensive information on threats. A honeynet also called honeywall, is a gateway device that separates honeypots from the rest. Any traffic from the honeypots must go through the honeywall and its purpose is to minimize risk.

D. Spectrum Based Detection

A spectrum based smart worm detection scheme[13] is based on the idea of detection of worm in the frequency domain. Spectrum detection scheme uses the power spectral density of the scan traffic volume and its corresponding flatness measure to distinguish the smart worm traffic from background traffic.
E. Support Vector Machines
Support vector machines (SVM)[9] is a technique used in an individual computer host to detect unknown worms based on their behavior (measurements) extracted from the operating system. A feature-selection method enables to identify the most important computer features in order to detect unknown worm activity.

F. Destination-Source Correlation
The Destination-Source Correlation(DSC) algorithm [6] is a two-phase local worm detection algorithm, based on the correlation between incoming and outgoing traffic. DSC keeps track of SYN packets and UDP traffic of the source and destination. It works for both TCP and UDP worms.
There are also some existing techniques [9][13][7][14][11][4][3] to detect the different worms surviving in the network. Those methods prove their detection performance and illustrated their evaluation achievements. Some of those techniques are listed below in the table III.

| Table III. Different Existing Techniques And Their Evaluation Measures |
|---------------------------|-------------------|-----------------|-----------------|
| Technique                  | Worm Type          | Detection Rate  | Para-meters     |
|                           |                   |                 |                 |
| Signature-Based Method     | Internet Worms    | -High Detection Rate | -Low False Rate |
|                           |                   | -True Positives | -False Positives |
| Distributed Honeypot       | Divide-Conquer-Scanning Worms | -High Scanning Rate | -No of Infected hosts -Time |
| Distributed Entropy based Dynamic detection Scheme(DED) | Varying Scan Rate Worms | -Maximal Infection Ratio | -Detection Rate -False Positive Rate -Dynamic Threshold Parameter |
| Frequency Detection based Filtering(FDF) | Scanning Worms | -High Detection Rate | Time Slot Length -Duration Threshold |
| Conjunction of Combinational Motifs | Polymorphic Worms | -False Detection Time | -False Positive Rate -False Negative Rate |
| Game Theory                | Self-Disciplinary Worms | -Maximal Infection Ratio | -False Positive Rate -Growth Rate |

From the table above, various techniques updated for the detection progress. Yet there exists limitations that the anomaly is to be detected faster and should be updated as new signature to block its spread further. With the existing parameters, threshold constant added will result in fast detection. Based on the behaviour of unknown worms and the traffic made by them can be detected earlier using Time to Live measure.

VI. Containment Techniques
Containment refers to blocking the traffic created by worm on suspected port used in propagation of detected worms [15]. Various containment techniques are used to slow down and block the spread of worms. There are various containment techniques existing as in Fig.4.

A. Per-Process Based Containment
A Per-Process- based containment[16] framework consists of two building blocks: behavior analysis and containment. The behavior analysis component includes several system monitors and a suspicion-level
generator. For containment, the mapping function optimizer generates the most appropriate function of transforming the suspicious level to a containment threshold. Both the suspicious-levels and the mapping function are taken as the input to the containment model which then outputs a customized threshold for each process.

B. Vigilante Containment

Vigilante[8] uses an end-to-end approach to contain worms automatically. Vigilante leverages information that includes worms, which break out network level detection and to eliminate false positives. Filters generate low overhead, have no false positives, and block all worms that follow the same execution path to gain control.

C. Proactive Worm Containment (PWC)

PWC can stop, instead of slowing down an infected host from releasing worm scans. PWC developed the following two novel detection techniques for containment: (a) PWC exploits a unique vulnerability window lemma to avoid false initial containment; (b) PWC uses a relaxation analysis to uncontain (or unblock) those mistakenly contained(or blocked) hosts within a few seconds.

This section conveys that there exists various detection and containment techniques to detect and containment of spread of worms in the network. To improve its fast detection and accuracy of its performance, epidemic spreading modelling to be framed for speed detection and containment of worms.

VII. PROPOSED SYSTEM

The host in the network once get infected with worm will affect the complete network by spreading. Epidemic spreading model has various states to detect the hosts status and helps to perform the recovery steps. Modelling an effective detection system helps the network and computer from the internet worms propagation. Below figure 5 proposes the novel framework for effective internet worm detection.

Some of the advantages of the proposed method are as follows.
- Improves detection accuracy
- Reduces false alarm rate
- Provides increased detection rate
- Overcome the misclassified negative instances
- Improves sensitivity and specificity of detection

VIII. PERFORMANCE METRICS

The performance of worm detection is measured using various parameters[9]. The parameters should achieve its target of performance for effective detection. The accuracy of detection rate is measured using True
Positive Rate or Recall value, False Positive Rate, Precision Value, Detection Rate and Error Rate. Accuracy of detection is the main metric to achieve best detection and classification models.

A. True Positive Value or Recall Value
Recall value rate is calculated to find and identify that the actual worms are actually detected. Ratio classifies the worm and the normal data. Number of positive instances is classified correctly to measure true positive rate. An alarm is generated to represent that attack has occurred and IDS triggered.

\[
Recall = \frac{True \ Positive}{(True \ Positive + False \ Negative)}
\]

B. False Positive Rate
False positive occurs, when a false alarm is generated that an intrusion has occurred. Negative instances that are misclassified are generated as false positive. Represents the alerts produced incorrectly. Here the worms are classified converse from normal data.

\[
False \ Positive \ Rate = \frac{False \ Positive}{(False \ Positive + True \ Negative)}
\]

C. Precision Value
Ratio of the calculated positive cases considered right is referred to as precision value.

\[
Precision \ Value = \frac{True \ Positive}{(True \ Positive + False \ Positive)}
\]

D. Detection Rate
True positive detection for the intrusions occurred has been detected correctly are referred to as detection rate. Detection time should be minimum for the increased performance. Summation of total positive and total negatives are referred to as detection rate.

\[
Detection \ Rate = \frac{Total \ worm \ data \times True \ Positive + Total \ normal \ data \times False \ Negative}{Total \ worm \ data + Total \ normal \ data}
\]

E. Total Accuracy
Accuracy is classified based on the intrusion propagation predicted correctly. The total Accuracy calculates the total of complete correctly classified requests as whether positive or negative and is divided by total number of requests.

\[
Accuracy = \frac{(True \ Positive + False \ Negative)}{(Positive + Negative)}
\]

Error rate is defined as 1-Accuracy.
Confusion Matrix is measured for every class for classification to construct the detection better. True Positive rate, False positive rate, precision rate providing its best results reveals the best evaluation measures for detection of worms in the network. Based on True Positive, False Positive and Accuracy, ROC curves performed for better result evaluation.

IX. CONCLUSION
Worm is a standalone malware computer program that replicates itself in order to spread to other computers. Active worm refers to a malicious software program that propagates itself on the Internet to infect new computers. The propagation of the worm is based on exploiting vulnerabilities of computers on the Internet. Many active worms are used to infect a large number of computers and recruit them as bots or zombies. This survey categorized the categories of worms and existing handling mechanisms for worm detection and containment.

In Support Vector Machine(SVM), operating system measurements are monitored and mean detection accuracy is achieved exceeding 90% and accuracy with 94%. Constant quarantine model and pulse quarantine model recovered network at 90% by alleviating infection-free equilibrium points. Using Digital Signal Processing(DSP) Scheme using Power Spectral Density(PSD) achieved 96.4% detection rate. Neetwork Address Space Randomization(NASR) is effectual of 90% above in bounding hitlist worms and achieved maximum effectiveness of 80%. SWROM and RWORM secures 34.4% with lesser bandwidth resources consumption. Existing methods fail due to lack of intelligent approaches. Bio-inspired approaches handle most of the challenging tasks effectively. Epidemic Spreading system model under computational Intelligence model can be devised for successful detection and immunity improvement of the network.
REFERENCES


S.Divya received her MCA from PARK College of Engineering and Technology, Coimbatore and M.Phil degree in Computer Science from Narasam Arts and Science College for Women, Erode in 2008 and 2010 respectively. She is pursuing her PhD at Avinashilingam Deemed University for women, Coimbatore. She has 2 years of teaching experience. Her areas of interest include Network and Communication Security.

Dr.G.Padmavathi is the Professor and Head of computer science of Avinashilingam Deemed University for women, Coimbatore. She has 24 years of teaching experience and one year of industrial experience. Her areas of interest include Real Time Communication, Network Security and Cryptography. She has 198 publications in her research area. In presently she is guiding M.phil researcher and PhD’s Scholar. She has been profiled in various Organizations her academic contributions. She is currently the principal investigator of four projects funded by UGC and DRDO. She is life member of many preferred organizations of CSI, ISTE, WSEAS, AACE, and ACRS.