Detecting IP spoofing in blacklist based packet filter in NIDS using Ant algorithm

Selvakumar B\textsuperscript{1}, Umakanth N\textsuperscript{2}, Thirumoorthy K\textsuperscript{3}\textsuperscript{,1}\textsuperscript{umakanthn@mepcoeng.ac.in, \textsuperscript{2}thirumoorthy@mepcoeng.ac.in, \textsuperscript{3}selvakumar.b@mepcoeng.ac.in}
\textsuperscript{1,2,3}Assistant Professor, Mepco Schlenk Engineering College, Sivakasi

Abstract – The current network scenario has many security related problems. Network intrusion detection has been widely deployed to detect the number of intrusive actions. Some of the techniques used context aware filter to detect the blacklisted IP address to reduce the network overload. The blacklist was periodically updated by the NIDS. However, the problem is, if the IP address is spoofed, then it cannot be detected and leads to the denial of service attack. In this paper we proposed the Ant algorithm to detect the IP spoofed packet in network intrusion detection system.

Keywords—Intrusion detection, Ant algorithm, context aware blacklist filter, signature matching, false alarm reduction.

I. INTRODUCTION

In current network communications, network threats are the big challenges. Intrusion Detection Systems (IDS) are used to observing network and system activities to detect the malicious activities. It can be categorized into network based intrusion detection system (NIDS) and Host based intrusion detection system (HIDS). The HIDS is used to detect the anomalies in a computer system. The NIDS is used to detect the anomalies in network environment. Network based intrusion detection system have two categories namely Signature based network intrusion detection system (or) misuse-based NIDS and Anomaly based network intrusion detection system. Signature based NIDS detects an attack by matching the incoming packet payload with signature that is already stored in signature database. Signature is defined as a pattern or rule to detect known attack. But it cannot be detecting the unknown attacks. On the other hand, anomaly based NIDS used to detect the novel attack by defining the normal behavior of the user on the network. A small deviation between the observed event and the normal behavior can be considered as an intrusive. The disadvantages of anomaly based NIDS is because of the diversity of network traffic which is very hard to build the normal behavior.

In real scenario, Signature based NIDS is mostly used than the anomaly based NIDS. The signature based NIDS produces the less false alarms than anomaly based NIDS. False alarm means the normal packets are detected as abnormal or else the abnormal packets are detected as normal. Four type of alarm alert:

- True Positive: detect the attack packet as it is.
- False Positive: detect the normal packet as attack.
- True Negative: no detection of attack packet.
- False Negative: detect the attack packet as normal.

The performance of signature based NIDS is suffered by network packet overload, expensive signature matching and massive false alarm. Due to the IP spoofing packets are overloaded to the destination.

II. LITERATURE SURVEY

The author Weizhi Meng, Wenjuan Li and Lam-For Kwok proposed the Enhanced Filter Mechanism (EFM)\textsuperscript{[1]}. It has the context aware blacklist based packet filter to reduce the packet overload. For improving efficiency signature matching component is implemented within the blacklist based packet filter. They use the single character frequency based exclusive signature matching to identify a mismatch instead of confirming an accurate match. And the KNN based false alarm filter was used to reduce the false alarm rate. It has been proved that EFM improve the performance of NIDS but have the disadvantage of IP Spoofing.

In\textsuperscript{[2]}, Ant based IP trace back is used to detect the spoofed IP packet. The traffic flow information and hop count value are used to find the legitimate source. Among the all possible paths Ant algorithm is used to find the shortest path between destination and source. In this traceback method, flow level information and hop count information are used. Mapping between IP address and their flow level with hop counts, the server can differentiate the spoofed packet from the legitimate ones.

Monowar H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita\textsuperscript{[3]}, categorize existing network anomaly detection methods and systems based various computational techniques used.

Support Vector Machine (SVM) method and the Clustering based on Self-Organized Ant Colony Network (CSOA-CN)\textsuperscript{[4]}, are combined to take the advantage of both SVM and CSOA-CN. It provides the higher average detection rate and faster running time. Also provide the lower rates of both false negative and false positive.

In\textsuperscript{[5]}, describe the all false alarm minimization technique in the signature based Network intrusion detection system.

The new Modified Wu-Manber (MWM) algorithm\textsuperscript{[6]} used the two-byte unit for pattern matching. Though the number of different signature increases the give the better performance. The comparison time of the MWM algorithm increases when the length of the shortest pattern in a signature group decreases.
Though an attacker can change any field in the IP header but can’t change the number of hops an IP packet takes its destination. Using mapping between IP address and its hops count can easily find the spoofed packet from legitimate request packet [7].

To solve the expensive signature matching process [8] uses the Single character frequency based exclusive signature matching scheme. It computes the single character from stored and matched signatures table. The decision algorithm uses the traditional signature matching algorithm to select the most appropriate character. Table of matched signature has the priorities for highest performance.

For improving detection rate and reduce the false positive rate [9] provide the hybrid intrusion detection system. For misuse detection c4.5 algorithm is used and SVM model is used for anomaly detection.

The denial of service attack with source address spoofing is identified in [10]. Ant based trace back method uses the flow level information to identify the origin of denial of service attack.

III. PROPOSED MODEL

This paper combines the ant algorithm [1] and Enhance Filter Mechanism [2] to overcome the demerits of Context aware filter.

![Proposed system design](image)

**A. IP spoofing detection**

The main aim of ipaddress spoofing is to find out whether the incoming request is from a legitimate user or not. For finding the IP spoofed packet Ant based IP trace back [2] method is used. The hop count value and flow level information are used with the ip address to find the spoofed request.

General Ant algorithm as below,

Input: G K T A

Output: K solution

1. For t=[1,T]
   1. Initialization
      a. For each A to G
   2. Solution construction
      a. For each A randomly remove and add the element until the solution is acceptable.
   3. Iterative improvement
      a. Update all pheromone entries
   2. Return K solution from iteration T

To perform the IP trackback the victim host is the starting point. It is assumed that the legitimate request might reach the victim node in a shortest path. The description of the ant based IP trace back is

**Step 1:** Construct network topology

**Step 2:** Determine the all possible paths between two source node and destination node.

**Step 3:** Find the shortest path using ant algorithm. Ant located at node i and choose the node j by following rule:

\[ j = \begin{cases} \arg \max \{ \tau_{ij}(t)\alpha \eta_{ij}(t)\beta, q \leq q_0 \} & \text{S}, \\ q > q_0 & \text{otherwise} \end{cases} \]

\[ S = p_{ij}(t) = \begin{cases} \tau_{ij}(t)\alpha \eta_{ij}(t)\beta, & j \text{ at time} \\ 0, & \text{otherwise} \end{cases} \]

Where

- \( \tau_{ij}(t) \) = pheromone intensity of trail between router i and j
- \( \eta_{ij}(t) \) = the number of routing packets between router i and j between time (t-1) and time (t).
- \( \alpha \) = weighting factor of pheromone
- \( \beta \) = weighting factor of visibility

Ants update the probability density function of possible attack paths and choose the right one.

Pheromone updating rule:

\[ \tau_{ij}(t+1) = (1-\rho) \tau_{ij}(t) + \rho \Delta \tau_{ij}(t) \]

Where

\[ \Delta \tau_{ij}(t) = \sum_{k=1}^{m} \Delta \tau_{ij}^k(t) \]

\( \rho \) = pheromone decay parameter

\( \Delta \tau_{ij}^k(t) \) = pheromone gain that ant k bring in after this tour

**Step 4:** calculate the hop count value

\[ \text{Hop count} = (\text{initial TTL}) - \text{TTL} \]
Initial TTL value depends on the type of operating system used.

**Step 5:** perform the Mapping between IP address and their flow level with hop counts

If the match is found then accepted as a legitimate request or else partially spoofed packet.

**IV. CONTEXT AWARE BLACKLIST BASED PACKET FILTER**

A Blacklist-based packet filter and a Monitor engine are the components of context-aware blacklist-based packet filter. The blacklist-based packet filter will check the incoming packet IP address as below

- If the IP address is already present in the blacklist then the packet payload will compared with look-up Table. The look-up table have the NIDS signatures.
- If the blacklist does not contain the incoming packet address then the packet is passed to the NIDS.

The monitoring engine will update the blacklist and also collect the alarm information from NIDS.

**V. EXCLUSIVE SIGNATURE MATCHING**

Signature matching has the two major components: statistical tables and decision component. There are four statistical tables: a table of stored NIDS signatures (SNS), a table of matched NIDS signatures (MNS) and remaining two tables for single character frequency (SCQ1 and SCQ2). The SNS contains all active signatures and the MNS have ever been matched signature during the detection. The table of SCQ1 calculates the single character frequency according to the table of SNS and SCQ2 calculates the single character frequency based on the table of MNS.

The decision component is used to output the most appropriate single character for comparing with incoming packet payloads.

The detailed algorithm is described below

**Step 1:** input signature s and payload l; rounds r=1

**Step 2:** for each aᵢ in s find the least frequent single character aᵢ (from single characters which have not been output) in the table of SCQ2.
- If aᵢ does not exist, then go to step 4
- If aᵢ exists, then go to step 3

**Step 3:** if aᵢ is not contained in l then no match
- If aᵢ is contained in l then r=r+1, return to step 2
- Else if aᵢ is in l and r=r then go to step 5

**Step 4:** find the aᵢ in the table of SCQ1.

**Step 5:** all single character in s aᵢ contained l then match is found.

**VI. KNN-BASED FALSE ALARM FILTER**

The main aim of K-Nearest Neighbour algorithm is to filter out false alarms. To classify the incoming alarm, KNN algorithm uses the Euclidean distance as the distance metrics.

\[
\text{Distance}\ (p_1, p_2) = \sum_{i=1}^{n} (p_{1_i} - p_{2_i})^2
\]

Where p1 and p2 are the i\textsuperscript{th} attribute of point’s p1 and p2

Data Standardization, Alarm Storage and Alarm Filtration are the three components of KNN-based false alarm filter. Data Standardization is divided into Feature Selection and Format Conversion components. Feature Selection will extracts the pre-defined features from an input alarm. Based on the predefined feature set Format Conversion will convert the NIDS alarm to a standard alarm.

Alarm Storage will store the all incoming alarm and false alarm reduction is performed by Alarm Filtration component. The features set are periodically trained using KNN.

**VII. RESULT AND DISCUSSION**

To simulate the Ant algorithm, Network topology with 5 nodes are created. S be the source node and D act as a destination node. All the possible paths between S and D are identified and shortest path is selected based on Pheromone Intensity using Ant algorithm.
According to the ant system optimization by a colony of cooperating agent, ants follow a path between the source to destination with all possible paths with equal probability. This process continues until all of the ants will eventually choose the shortest path.

**A. Evaluation of spoofed packet accuracy**

It is assumed that the attacker generates packets by randomly selecting source IP addresses among legitimate clients. It is further assumed that the attacker knows the general hop-count distribution for each web server and uses it to randomly generate a hop-count for each spoofed packet. To evaluate the filtering accuracy, it is assumed that the legitimate request has the shortest path between source and destination. To measure the detection accuracy shortest path and hop value are considered. False Positives are the legitimate request that is incorrectly identified as spoofed. False Negative is the spoofed requests that are incorrectly identified as legitimate request.

**Table 1: Experimental values**

<table>
<thead>
<tr>
<th>S.no</th>
<th>Possible path</th>
<th>Hop count</th>
<th>Pheromone Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>S-&gt;A-&gt;D</td>
<td>1</td>
<td>1.8342</td>
</tr>
<tr>
<td>3.</td>
<td>S-&gt;A-&gt;C-&gt;A</td>
<td>2</td>
<td>0.5897</td>
</tr>
<tr>
<td>4.</td>
<td>S-&gt;B-&gt;C-&gt;D</td>
<td>2</td>
<td>0.5438</td>
</tr>
<tr>
<td>5.</td>
<td>S-&gt;B-&gt;C-&gt;A-&gt;D</td>
<td>3</td>
<td>0.3591</td>
</tr>
</tbody>
</table>

**Fig3: Experimental topology with 5 nodes**

**VIII. CONCLUSION**

In existing blacklist based packet filter, only filter out the black listed IP address. But it could not detect the spoofed IP packets. So in this work we proposed an Ant based IP trace back to find the spoofed packet in context aware blacklist based packet filter using Ant algorithm and it produced the better accuracy. The network packet overload is reduced by the Context aware blacklist based packet filter. The exclusive signature matching component improved the traditional process of signature matching. The KNN-based false alarm filter reduced the analysis burden on NIDS alarms by filtering out a number of false alarms generated by both the EFM and the NIDS.

**IX. REFERENCES**
