Naïve Bayes Classifier to Mitigate the DDoS Attacks Severity in Ad-Hoc Networks

K. Ganesh Reddy¹, and P. Santhi Thilagan²

¹School of Computer Science and Engineering, VIT-AP University, Amaravathi, India
²Department of Computer Science and Engineering, NITK Surathkal, India

Abstract: Ad-Hoc networks are becoming more popular due to their unique characteristics. As there is no centralized control, these networks are more vulnerable to various attacks, out of which Distributed Denial of Service (DDoS) attacks consider as more severe attacks. DDoS attack detection and mitigation is still a challenging issue in Ad-Hoc Networks. The existing solutions find the fixed or dynamic threshold value to detect the DDoS attacks without any trained data. Very few existing solutions use machine learning algorithms to detect these attacks. However, existing solutions are inefficient to handle when DDoS attackers perform this attack through bursty traffic, packet size, and fake packets flooding. We have proposed DDoS attack severity mitigation solution. Our DDoS mitigation solution consists of a new network node authentication module and naïve Bayes classifier module to detect and isolate the DDoS attack traffic patterns. Our simulation results show that naïve Bayes DDoS attack traffic classification outperforms in the hostile environment and secure the legitimate traffic from DDoS attack.

Keywords: DDoS, Ad-Hoc, Naïve Bayes, mitigation, flooding, legitimate traffic.

1. Introduction

Ad-hoc networks also called an infrastructure-less network, here the network nodes have more flexibility, and no node has superior to others. Any time nodes can join/leave the network. In general, Ad-Hoc nodes have cooperative behavior. Also, each node acts as a router. In addition to that, self-organizing, self-configuring, and self-healing features reduce the network establishment and maintenance cost. Based on cost-effective deployment and easy maintenance, many wireless networks like sensor networks, mobile ad-hoc networks, Vehicular networks, and IoT adapt Ad-hoc environments to perform their network functionalities. Typically, wireless network nodes have the bandwidth, communication range, buffer size, battery, computing, and memory resource limitations.

Routing protocols use to build the routing tables, and these tables use to route the packets from source to destination. In ad-hoc network, packet forwarding is supported by node cooperative behavior, in which each node has trust in other network nodes to forward its data. Ad-hoc routing protocol’s main objective is to find the best routes. To find the best route, routing metrics hop_count, throughput, congestion, end-to-end delay, packet delivery ratio, queue size, jitter, reliability, bit error rate, link failures, retransmission rate, processing delay, packet size, source address, destination address, port number, and link availability considered. While processing too many routing metrics to establish a route, lead to control overhead and buffer overflows[8][9]. Existing routing protocols consider the two or three metrics considered as the pivot metrics to establish the best route. In general, nodes in link-state routing protocols share and update the routing information only when routing metric information is updated. On the other hand, nodes in distance vector routing protocols periodically share and update the routing information. In hybrid routing protocols, nodes share and update routing information through both regularly and when the routing metric is updated.

However, we can't guarantee that each node cooperates to forward the packets and share the legitimate routing information; this problem creates more vulnerability in the Ad-Hoc network. These vulnerabilities mainly classified into two: routing vulnerabilities and data forwarding vulnerabilities.

Routing vulnerabilities lead to various network layer attacks such as routing loops, Sybil, blackhole, gray hole, wormhole, colluding, and rushing attacks. Network attacks lead to data plane attacks such as byzantine attack and jellyfish attack [10]-[13]. A malicious node makes unavailability of the network resources by flooding the fake, replay, or stale packets, also called the Denial of Service (DoS) attack[7][8]. The creation of DoS attacks in an infrastructure-less network is much easier than an infrastructure network. Any node sends false information about other nodes, and upon receiving this information, validation is still a challenging task at each node. However, identify these attacks is easier. DoS attack can be a more sophisticated attack if malicious nodes use the network resources on behalf of the legitimate nodes. These attacks take place because of the internal (compromised) attacker or a spoofing technique. As shown in Figure 1, to detect these attacks, we need machine learning algorithms and cryptographic methods.

Figure 1. DoS attack Scenario
In a DDoS attack, internal attackers initially select the target node, and then these attacker nodes flood/replay/stale packets through link/node disjoint paths at the target node as shown in figure 2. Unlike the DoS attack, DDoS attacks happen by multiple attackers. Thus, a single attacker need not have substantial network resources to perform the attack, and one or two link failures will not degrade the severity of the attack. Moreover, individual traffic flows of DDoS attack seem to be normal flows, but the accumulation of traffic flows lead to a DDoS attack at the target node/path. The DDoS attack occurs with bursty traffic within a short time. The individual traffic flow of DDoS attacks is much more similar to regular traffic, and detection of initial DDoS attack flows is complicated and is a challenging problem in the Ad-Hoc network. The DDoS attack mainly targets the limited resources of Ad-Hoc networks such as bandwidth, node energy, buffer, and computing power.

Routing protocols find the best routes to forwards packets from source to destination, in Ad-Hoc networks, Blackhole, Grayhole, wormhole, rushing, and Sybil attacks allow the attackers to join in the active (data forwarding) paths. DDoS attackers will be more prevailing when these attackers are in the active paths. On the other hand, colluding nodes periodically share big routing tables with no valid information to create congestion and buffer overflows.

Intrusion prevention and intrusion detection mechanisms are the solutions for DDoS attacks in MANET[14]-[24]. To isolate unauthorized nodes from the ad-hoc network, the intrusion prevention mechanisms are used, in which cryptographic keys use for node/data authentication and data confidentiality. However, intrusion prevention mechanisms are vulnerable to internal DDoS attackers. These attackers use loopholes in the packet verification process, attackers flood fake/stale authentication packets at the target node, target node computation overhead, and buffer overflow problems become worst to verify received authentication packets. Eventually, the target node drops the legitimate packets instead of the processing of these packets. The intrusion detection mechanism waits until the DDoS attack encounters/suspected based on the traffic flows in the network. The detection mechanism triggers an alarm immediately after it identifies the DDoS attack pattern. In general, we need to fix a few network metrics like packet transmission rate, packet size, number of collisions, etc., to identify the DDoS traffic patterns. Static network metric threshold values are sufficient to handle study traffic flows in the network, but insufficient to handle the dynamic traffic flows. We need machine learning algorithms to handle the dynamic traffic flows in which initially nodes collect the traffic flow data for training. Once the trained data is ready, then the network node fixes the threshold values, and these values further updated based on the availability of network resources.

In this paper, we have implemented the RSA-1024 bit key for network nodes authentication. Further, we have applied the Naive Bayes classifier to classify the normal and DDoS attack traffic patterns in the network nodes by considering the five network metrics. The rest of the paper as follows, section 2 explains the literature survey, and section 3 discusses the proposed work. Section 4 demonstrates the results and analysis, and section 5 ends with conclusions.

2. Literature Survey

S. Ahmed et al. have proposed fuzzy rule-based IDS to detect the DoS and DDoS attacks in Ad-Hoc networks [1]. They have defined six independent rules based on the number of packet transmissions, packet size, packet interval time, number of ack messages, and group packet count/sec and total count. First, four rules use to analyze the DoS attack patterns, and the remaining two rules also applied to analyze the DDoS attack patterns.

In [6], traffic load, number of packet transmissions, delay metrics consider identifying the DDoS attack in MANET. Nodes in the current path the observers these three network metrics by setting up fixed threshold values. If any node receiving the packets more than the threshold values of these three metrics, then the node drops these packets instead of forwarding them to the next router. These approaches have constant threshold values to detect the DoS and DDoS, if we use these approaches for dynamic traffic flows, false positives, and false negatives are very high. Thus, these approaches are inefficient in handling the dynamic traffic flows.

In [2], T.Luong proposed a FAPRP solution for detecting DDoS attacks in the Ad-Hoc network. In this solution, route request (RREQ) packets frequency considers as the primary metric. The kNN algorithm uses for classifying the normal and abnormal traffic patterns in route discovery. They have implemented this solution in AODV protocol to train the dataset. The trained dataset results used to define the route discovery frequency vectors in the normal and abnormal scenarios. Authors also propose flooding prevention algorithm in which Euclidian distances calculate for normal vector class, and malicious vector class based on number request messages. If any node no.of requests packets distance equals to malicious vector class, network nodes immediately drop the request packets instead of forwarding them to the next routers. This approach detects the RREQ packet DDoS attacks only when nodes flood the RREQ packets in the network, which is inefficient in detecting sophisticated DDoS attacks.

In [4], fuzzy logic is used for the trust estimation of network nodes in MANETs. Each node initially sets the trust value of its neighboring nodes. A node updates the trust value [-1,1] plus or minus 0.05 of its preceding node and forward node based on their Packet Delivery Ratio (PDR). The normal node trust value range is defined between greater than -1 and less than or equal 1. If a node trust value is -1, then this node is treated as a malicious node. This approach consumes more
time to identify the attack; in some scenarios, it is also not possible to detect the attacks.

In [3], the intrusion detection system is developed and deployed in all network nodes to prevent flooding attacks. In the network layer, routers monitor and validate the number of requests/fake packets flooded by its neighboring nodes. Based on that validation, the neighboring node’s reputation values dynamically increases. If the node/nodes identified as their malicious behaviors, then these nodes/nodes blocked during a specific interval time from the network. This approach fails when a group of nodes plays a DDoS attack through node disjoint paths.

In [5], a random forest algorithm is used to train and classify the network traffic in Vehicular Ad hoc Network (VANET). In this approach, initially collect the large network traffic and suspected traffic, and this data is stored, then this data is processed in HDFS to detect the DDoS attacks. In this approach, a number of false positives and false negatives are low, since it processes the large data to make the decision. However, we need the special monitoring nodes do not have the CPU power, energy, storage, and bandwidth resource constrains.

3. Proposed Work

We have developed two modules, such as a new node authentication module and an automation module to mitigate the severity of DDoS attack in Ad-Hoc Networks. In the node authentication module, public and private keys are issued to network nodes to isolate unauthorized nodes (external attackers). In the automation module, we implement the Naive Baye’s classifier to predict the DDoS attack patterns and mitigate the DDoS attack severity by ignoring these attack patterns.

3.1 New Node Authentication

In Ad-hoc network, nodes can join/leave the network at any time that causes difficulty in identifying the network nodes. In the node authentication module, Certification Authority (CA) issues a public and private key pair to a new node, then new node use this key pair to join/leave the network.

Ad-Hoc network nodes have resource constraints. Thus the maximum number of Authentication requests sent by a new node is limited to 5 per minute. If any new node sends more than five authentication packets per minute, all excessive authentication requests are dropped by the network nodes. This process mitigates the severity of node authentication request flooding attack, shown in Figure 3.

3.2 Automation

In our proposed work, we have implemented the Naive Bayes classifier on the collected network traffic to classify these traffic patterns into two types: Non-attack traffic pattern, and DDoS attack pattern. To do this classification, we mainly consider the following traffic pattern parameters $A = \{a_1, a_2, a_3, a_4, a_5\}$ are considered to predict whether the given traffic patterns belong to DDoS attack or non-DDoS attack traffic patterns.

- $a_1$: Packet size
- $a_2$: Port number
- $a_3$: Source address
- $a_4$: Destination address
- $a_5$: Jitter (delay inconsistency between each packet)

![Figure 3. New node Authentication process](image)

In Ad-Hoc network, each node calculates the probability of DDoS attack occurrence by considering all incoming links/paths traffic of a node. This process is formulated in equation 1.

$$P\left(\frac{D^i}{t^i}\right) = P(D^i) \prod_{i=1}^{m} P\left(\frac{a_i}{t^i}\right)$$

(1)

Where $D^i$ is DDoS-attack traffic, $T$ is the total number of active links/paths $t_1, t_2, \ldots, t_n$ at the monitor node.

$P(D^i) = \frac{T_p}{T_m}$ where $T_m$-Total number of packets received/sec from all active paths, $T_p$ - number of unprocessed packets such as dropped packets due to a buffer overflow, stale packets, and error packets received/sec from all active paths. Probability of DDoS attack traffic pattern occurrence for the given traffic pattern of $i^{th}$ path

$$P\left(\frac{a_i}{D^i}\right) = \max\left\{P(D^i) * P\left(\frac{a_i}{D^i}\right)\right\}$$

for all $a_i \in A$  (1.1)

$P(D^i) = \frac{T_p}{T_m \_i}$ $t_{add}$ $t_{port}$ $t_{delay}$ $t_{add}$ $t_{port}$ $t_{add}$ $t_{port}$ $t_{add}$ $t_{port}$ $t_{add}$}

- $t_{add}$ - number of packets are generated towards the same port number
- $t_{port}$ - Number of packets are generated by a source node
- $t_{delay}$ - number of packets are transmitted towards a destination node
- $t_{add}$ - number of packets are generated towards the same port number
- $t_{port}$ - Number of packets are generated by a source node
- $t_{delay}$ - number of packets are transmitted towards a destination node
number of arrived packets jitter values are less than \( t_{\text{min} \_\text{delay}} \). In Ad-Hoc network, nodes classify the network traffic by calculating the probability of DDoS attack occurrence value \( P \left( \frac{D}{T} \right) \). Based on these values will take the countermeasures as follows

\[
\text{if } P \left( \frac{D}{T} \right) = 0 \\
\text{Network traffic considered as normal traffic} \\
\text{else} \\
\text{for } i=1 \text{ to } n \\
\text{where } n \text{ is all incoming links/paths at the node} \\
\text{if } P \left( \frac{D_i}{T_i} \right) > 0 \\
\text{ } \text{ } i^{\text{th}} \text{ path } \text{Network traffic considered as DDoS attack traffic} \\
\text{Drop } P \left( \frac{D}{T} \right) \% \text{ packets are dropped to mitigate the DDoS attack} \\
\text{else} \\
\text{ } \text{ } i^{\text{th}} \text{ path traffic considered as normal traffic} \\
\text{end-if} \\
\text{end-for} \\
\text{end-if}
\]

In the above process, the network monitor node calculates the \( P \left( \frac{D}{T} \right) \) value periodically; if this value is zero, then the monitor node considers the received traffic is normal traffic during this particular time interval. If the \( P \left( \frac{D}{T} \right) \) value is more than zero, then the received traffic detected as DDoS attack, in which the packets exceed the limits of the network parameters' (a1,a2,a3,a4,a5) are identified with respect to each incoming link/path. Monitor node punishes the DDoS attack link/path by dropping the\( P \left( \frac{D}{T} \right) \% \) packets to mitigate the severity of the DDoS attack traffic.

4. Results and Analysis

We have implemented our proposed solution in Network Simulator (NS2), and created a scenario with twenty-seven network nodes, out of which twenty-one nodes have legitimate behavior, and six nodes have DDoS attack behavior. We have selected one target node from twenty-one legitimate nodes. All the network nodes deployed in the 2500m X 2500m network coverage area, and each node transmission range is 150m, shown in Figure 4. Network nodes communicate through UDP communication protocol, generate the CBR traffic with packet size of 1500 bytes is normal, and 10000byte packet is abnormal and used to create DDoS attack traffic. In addition to that, the legitimate node generates 1000 packets/sec, and a malicious node generates 10000 packets/sec. Network nodes share the RSA-1024 public keys for authentication. DDoS attacks take place in a very short duration, and to detect and mitigate this attack, we run the total simulation for 10 seconds, shown in Table 1.

In Figure 5, six malicious nodes have created the DDoS attack with a number of fake packets, size and packets interval time at a target node, during this attack legitimate nodes also send 1000 request packets/sec.

<table>
<thead>
<tr>
<th>Table 1. Simulation metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Parameters</strong></td>
</tr>
<tr>
<td>Total number of nodes</td>
</tr>
<tr>
<td>Number of nodes</td>
</tr>
<tr>
<td>Number of Malicious nodes</td>
</tr>
<tr>
<td>Target node</td>
</tr>
<tr>
<td>Node Authentication</td>
</tr>
<tr>
<td>Number of ping packets</td>
</tr>
<tr>
<td>Packet Size</td>
</tr>
<tr>
<td>Network Traffic</td>
</tr>
<tr>
<td>Attack Scenario</td>
</tr>
<tr>
<td>Communication Media</td>
</tr>
<tr>
<td>Routing Protocol</td>
</tr>
<tr>
<td>Network Coverage Area</td>
</tr>
<tr>
<td>Node transmission range</td>
</tr>
<tr>
<td>Simulation time</td>
</tr>
</tbody>
</table>

Figure 4. DDoS attack scenario

Figure 5. Legitimate network traffic on Malicious and Non-Malicious Scenarios

Target node not able to reply to none of the legitimate requests due to the excessive fake packets arrived at the node—another hand, non-DDoS attack target node process all legitimate packets.
We have created the flash events (i.e., legitimate traffic which heavy legitimate traffic generated during a peak interval) along with the DDoS attack, and all network nodes applied our proposed approach to detect the DDoS attack patterns and drop these networks traffic. In this hostile environment, out of 190000 requests packets, the target node is able to process the 120000 request packets in the worst case, which is shown in Figure 6.

DDoS attackers transmit the very large packet size to depletion of the network resources, and legitimate nodes transmit the packets which have less than or equals to 1500bytes. In this hostile environment, all the network nodes applied our proposed approach to detect the size of the packet, which has more than the 1500bytes and drop these packets instead of forwarding them.

**Figure 6.** Performance analysis of Flash Events

The target node is able to process the 80% of requests in our proposed approach, shown in Figure 7. Other hand, the target node is not able to process any request packets without our proposed work.

**Figure 7.** Performance Analysis on DDoS attack using Packet Size

In Figure 8, the x-axis shows the time and y-axis show the number of packets processed. We have observed that our proposed approach network performance in a hostile environment. DDoS attackers use very large packets, excessive ping packets, and fix a target node to create a hostile environment. However, our proposed approach detects the DDoS attack traffic patterns and mitigates the DDoS attack severity. Thus 82% of legitimate network traffic is processed by other network nodes. Based on the Figure 5, 6, 7 & 8, we have observed that our proposed IDS outperforms in a hostile environment.

**Figure 8.** Performance Analysis on proposed solution Vs. different types of DDoS attack vs. non-attack scenario

5. Conclusions

In this paper, we proposed the naive Bayes classifier to classify the DDoS attack traffic patterns by considering the five highly influencing DDoS attack network parameters. Our proposed DDoS attack classifier implemented on all monitor nodes to process the legitimate traffic and drop the DDoS attack traffic of a link/path based on the probability of DDoS attack $P \left( \frac{D}{T} \right)$ value. We have implemented the DDoS attack, non-attack, and DDoS attack with proposed approach simulations scenarios in NS2 to test the performance network performance. Based on our simulation results, our proposed approach mitigates the severity of DDoS attacks, and network nodes process 80% of the legitimate traffic. On the other hand, without our proposed approach, network nodes process 0% legitimate traffic in the hostile environment.

References


