IP Traceback for Wireless Ad-hoc Networks

Vrizlynn L. L. Thing, Henry C. J. Lee
Institute for Infocomm Research
Singapore
Email: {vriz, hlee}@i2r.a-star.edu.sg

Abstract - **Denial of Service (DoS) and Distributed DoS attacks constitute a major class of security threats today. As the attackers usually use IP spoofing to conceal their real location, several IP Traceback mechanisms have been proposed to trace the true source of the attackers to institute accountability. In wireless ad-hoc networks, where the nodes are typically devices with limited bandwidth, computational resource and battery power, and unpredictable routing topology, additional constraint is placed on these tracing techniques to locate the attack sources efficiently. In this paper, we conducted qualitative analysis and simulations to investigate the feasibility and evaluate the attack path detection performance of existing IP Traceback techniques (i.e. Source Path Isolation Engine, Probabilistic Packet Marking, and ICMP Traceback) on wireless ad-hoc networks, using proactive (DSDV) or reactive (AODV) routing protocol. Our studies showed that the performance of traceback depended not only on the technique itself but also on the ad-hoc routing protocol chosen and the network size.**

I. INTRODUCTION

Wireless ad-hoc networks are self-organizing systems formed by co-operating nodes. Their topology is dynamic, decentralized, and ever changing with the ability and possibility of nodes moving arbitrarily. Their usage has become increasingly prevalent in emergencies (such as in the cases of disasters and wars) and also in daily life such as university campus and conference settings. The reason is mainly due to easy collaboration and efficient communication on the fly without the need for costly network infrastructure. However, users are concern over the security vulnerabilities at the same time as these wireless networks are becoming an indispensable part of our life.

Denial of Service (DoS) [1] and Distributed DoS attacks constitute one of the major classes of security threats today. Series of DDoS attacks that shut down some high-profile Web sites [2] have demonstrated the severe consequences of these attacks. [3] showed the prevalence of DoS attacks in the Internet, whereby more than 12,000 attacks against more than 5,000 distinct targets were observed in a 3-week long data collection period. As the attackers usually use IP spoofing to conceal their real location, several IP Traceback mechanisms have been proposed to trace the true source of the attackers to institute accountability. However, in wireless ad-hoc networks, where the nodes are typically devices with limited bandwidth, computational resource and battery power, and unpredictable routing behaviors, additional constraint is placed on these traceback techniques to locate the exact attack sources efficiently.

In this paper, we investigated the feasibility of applying the existing IP Traceback techniques on wireless ad-hoc networks. To the best of our knowledge, this is the first work in this area of analyzing specific issues relating to IP Traceback in ad-hoc networks. The IP Traceback techniques considered here were the Source Path Isolation Engine (SPIE) [4], Probabilistic Packet Marking (PPM) [5], and ICMP Traceback (ITrace) [6].

In SPIE, each intermediate router logs digests of the packets it processed. If traffic data for a particular time interval is of interest, these logs are transferred to a central server for longer-term storage and analysis. In this scheme, only one attack packet is required to reconstruct the attack path. In PPM, packets are marked by intermediate routers probabilistically to contain fragments of the path information. When the victim has collected sufficient number of marked packets, it would be able to reconstruct the attack path. In ITrace, each intermediate router generates a new ICMP packet, called the ITrace message at a low probability for each packet it processed. The message contains information about this router and the packet, and is sent to the same destination of the packet. Upon reception of sufficient number of ITrace messages, the victim would be able to reconstruct the attack path.

Ad-Hoc routing protocols are classified into proactive and reactive categories. With proactive routing protocols, each node periodically broadcasts reachability information for each destination. The routing tables are also updated regularly irrespective of whether there is a need for transmitting packets on the route or not. With reactive routing protocols, routes to a destination are only established by a source node when there is a need for transmitting packets to that destination. Route maintenance is carried out till the destination becomes inaccessible or the route is no longer required. In this paper, we selected Destination-Sequenced Distance-Vector (DSDV) [7], as the proactive and Adhoc On-Demand Distance Vector (AODV) [8] as the reactive routing protocol, for carrying out our simulations.

The rest of the paper is organized as follows. Section II provides the background information of the traceback techniques and conduct analysis on their feasibility for wireless ad-hoc networks. Simulations and analysis on the results are presented in Section III. Conclusions follow in Section IV.

II. BACKGROUND AND FEASIBILITY ANALYSIS

SPIE is an infrastructural hash-based IP logging approach. A Data Generation Agent (DGA) is implemented on each SPIE-enhanced router to produce packet digests of each packet

0-7803-8521-7/04/$20.00 (C) 2004 IEEE

as it departs the router and stores these digests in bit-mapped digest tables. If interest is expressed in the traffic data for a particular time interval, the tables are transferred to a SPIE Collection and Reduction (SCAR) agent for longer-term storage and analysis. Upon an IP Traceback request, each SCAR produces an attack graph for its particular region and these graphs are sent to the SPIE Traceback Manager (STM) for the complete attack graph reconstruction. However, potential large storage requirement and the infrastructural characteristic of the scheme make this approach infeasible for the wireless ad-hoc networking environment.

PPM proposes marking packets with intermediate routers' information probabilistically. In the Edge Sampling approach, the edges of the attack path are encoded in the packet rather than the individual nodes. The number of data packets, *X*, required for the victim to reconstruct an attack path of length, *d*, will then have the following bounded expectation:

$$E(X) < \frac{\ln(d)}{p(1-p)^{d-1}} \quad (1)$$

, where *p* is the probabilistic of marking a packet at an intermediate router.

However, the significant practical limitation of this approach is that it requires additional space in the IP packet header. Therefore, the Compressed Edge Fragment Sampling approach is proposed to store fragments of the edge-id for the edge between two routers in the identification field of the IP packet header. Of the 16-bit identification field, only 8 bits are used for the encoding of the edge-id fragment, while the others are used for encoding the fragment offset (i.e. which of the 8 fragments) and distance. This is necessary to ensure that different fragments from an edge-id can be recombined in the correct order at the victim. If enough packets are sent by the attacker, the victim will eventually receive all fragments from all edge-ids. Due to this deployment consideration, the Compressed Edge Fragment Sampling approach is considered to be suitable for practical usage. The expected number of packets required for path reconstruction is now bounded by:

$$E(X) < \frac{k.\ln(kd)}{p(1-p)^{d-1}} \quad (2)$$

, where *k* is the number of fragments per edge-id.

In ITrace, a new ICMP message type is defined to carry information on routes that an IP packet has traversed. As the IP Marking requires overloading a field in the IP header, which raises backward protocol compatibility problem, ITrace utilizes out-of-band messaging to achieve the packet tracing purpose. As an IP packet passes through a router, an ITrace message is generated with a probability of 1/20000. This ITrace message is then sent to the destination of the IP packet. In the event of a DoS/DDoS attack, the destination node can then use it to traceback the attack path. When a router generates an ITrace message, it may generate one of the followings: back link, forward link, or both. Each link element defines a link along which the packet will or has travelled through. The link element comprises of 3 components: the interface name at the generating router, source and destination IP address of the link, and finally a link-level association string that is used to tie together Traceback messages emitted by adjacent routers. The number of data packets, *X*, required for the victim to reconstruct an attack path of length, *d*, has the following bounded expectation [9]:

$$E(X) < \frac{\ln(d)+\gamma}{p} \quad \text{(for back/forward link)} \quad (3)$$

$$E(X) < \frac{\ln(d)+\gamma}{cp} \quad \text{(for both links)} \quad (4)$$

, where *p* is the probabilistic of generating an ITrace packet at an intermediate router, $\gamma \approx 0.5772$ is the Euler's constant, and *c* ≈ 1.3 for short paths and 1.8 for long paths. *c* can be assumed to be 1.5 for approximation (and is used in this paper).

As the unpredictable routing in wireless ad-hoc networks might result in multiple paths utilized during DoS attacks, fast traceback schemes are required to discover all attack paths before the victim is overwhelmed by the attack traffic. ITrace's both links approach is more suitable in this case as it will result in a faster discovery of intermediate routers since two nodes can be discovered in one ITrace message.

We consider the characteristics of the PPM's Compressed Edge Fragment Sampling approach and ITrace's both links approach (which we shall simply refer to it as PPM and ITrace respectively from here on) to be promising for reconstructing attack paths in wireless ad-hoc networks. They are non-infrastructural based and do not require large storage space. Only a small percentage of attack packets generate/contain traceback information, and this information could be stored longer due to lesser storage space requirement to allow for post mortem tracing. They therefore do not have tight time constraint as in SPIE.

III. Simulations and Results

Simulations were conducted using ns-2 [10]. We investigated the performance of PPM (with marking probability of 0.04) and ITrace (with ITrace message generation probability of 1/20000 = 0.00005) in the scenario of a single attacker in static wireless ad-hoc networks with 14, 34 and 54 nodes (inclusive of the attacker and victim). The routing protocols selected for simulations were DSDV and AODV, for the proactive and reactive categories respectively. For an intermediate router to be discovered in PPM, it will require all fragments for the router to be received at the victim. An intermediate router's discovery in Itrace, on the other hand, required that the router, its back-link router or forward-link router generates an ITrace message.

The number of attack packets generated for the simulations in the PPM and ITrace path reconstruction was set to be bounded by the above-mentioned formulas for the average number of intermediate routers along the attack paths. To find the average length of the attack paths in each network of

different sizes running different routing protocols, we send 20000 packets to the victim to obtain total number of paths traversed by the packets, length of each path and calculated the average. The average lengths are presented in Table 1. Using AODV, the packets traversed through only 1 path, whereas for DSDV, multiple paths were selected as the traffic load increased. As the average length of the traversed path/s for the AODV and DSDV is the same for networks with similar number of nodes, we needed to calculate only a set of the expected number of attack packets required to be received by the victim during an attack to reconstruct a single attack path in a wired network (using the above-mentioned formulas), and presented the values in Table 2.

	14-node	34-node	54-node
AODV	3/1 = 3	8/1 = 8	13/1 = 13
DSDV	16/5 ≈ 3	629/77 ≈ 8	10511/824 ≈ 13

Table 1: Average path length for different routing protocols and network sizes

	14-node	34-node	54-node
PPM	690	1107	1516
ITrace	22344	35422	41895

Table 2: Expected attack packets for path reconstruction for different traceback mechanisms and network sizes

At the start of the simulations, attack packets were sent from the attacker to the victim at a rate of 200 packets per second. The interface queue of all the nodes was set to 200 packets. However, at the rate of the attack, packets were dropped due to congestion at the routers during the simulation. Therefore, instead of limiting the packets generation at the attacker, the attack packets were monitored at the victim. When the expected number of attack packets was received by the victim, the rest of the attack packets and their corresponding traceback packets were dropped and not taken into account. The victim kept a record of the paths traversed by the attack packets (bounded by the expected number), and all the intermediate routers of the paths. The marked packets or ITrace messages were kept track of, to monitor which routers were discovered.

Using AODV, only 1 attack path was chosen to be used for routing the attack packets. This was done when the first packet was transmitted out by the attacker. Therefore, the scenario was similar to that of the wired network, and each of the paths was fully reconstructed. Table 3 shows the number of marked or ITrace packets recorded in each simulation scenario.

	14-node	34-node	54-node
PPM	72	318	606
ITrace	2	12	25

Table 3: Traceback packets recorded for AODV simulations

Using DSDV, multiple paths to the victim started appearing as the number of attack packets increasingly swarmed the network. The reasons were most probably route changes due to updates and the occurrence of network traffic congestion. When simulating PPM traceback in the 14 and 34-node networks, the expected attack packets were set to 690 and 1107 respectively. For these 2 scenarios, only 1 attack path was used for the routing and it was fully reconstructed. This was due to the small number of attack packets generated and size of the networks. Therefore, the attacks did not overload the networks to the point of resulting in multiple paths. The number of marked packets received by the victim for the 14 and 34-node networks was 53 and 249 respectively. However, for the 54-node network, 10 paths were traversed by the 1516 attack packets. Figure 1 shows the distribution of the attack packets through each of the 10 paths. 580 marked packets were received by the victim. Based on the marked packets, only attack path no. 5, 7, 8, and 9 were fully reconstructed. These 4 paths were used to route about 61.7% of the attack traffic to the victim and therefore, discovery of these paths would be able to support mitigation of part of the attack. The other 6 paths had either 12 or 13 intermediate routers each with 1 or 2 routers not fully discovered.

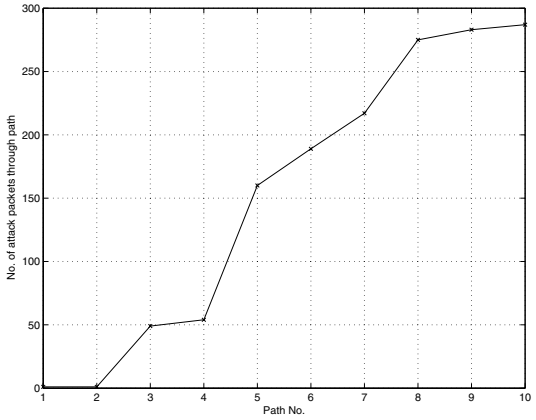

Figure 1: Attack packets distribution (using DSDV routing protocol and PPM traceback in 54-node network)

When simulating ITrace in the 14-node network, 22344 attack paths received by the victim traversed through 5 different paths. These 5 paths had between 2 to 4 intermediate routers each. Figure 2 shows the distribution of the attack packets through the 5 paths. Only 1 ITrace message was received by the victim and that resulted in 1 attack path (Path No. 5) been fully reconstructed. Path no. 5, with 2 intermediate routers, carried 20787 attack packets, which was about 93% of the attack traffic. The other 4 paths had between 1 to 3 intermediate routers not been discovered.

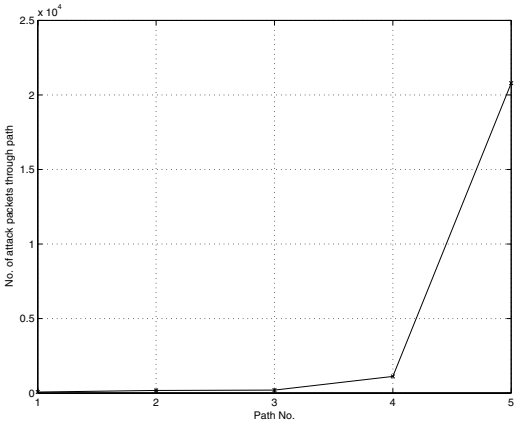

Figure 2: Attack packets distribution (using DSDV routing protocol and ITrace traceback in 14-node network)

When simulating ITrace in the 34-node network, 35422 attack paths received by the victim traversed through 272 different paths. Figure 3 shows the distribution of the attack packets through the different paths. 11 ITrace messages were received by the victim and 15 attack paths (i.e. Path no. 13, 41, 58, 67, 70, 71, 161, 165, 173, 186, 199, 254, 256, 259, and 264) were fully reconstructed. A total of 2020 attack packets (≈5.7%) travelled through these 15 paths. Path no. 272, which routed the highest percentage of the attack packets (i.e. 10010 packets or ≈28.3%), had 6 intermediate routers with 5 of them been discovered (i.e. ≈83.3% of the path).

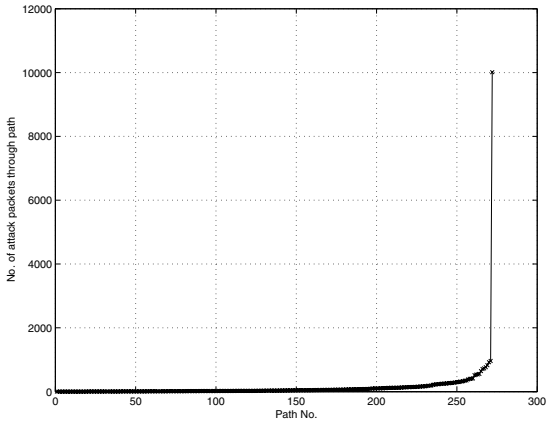

Figure 3: Attack packets distribution (using DSDV routing protocol and ITrace traceback in 34-node network)

When simulating ITrace in the 54-node network, 41895 attack paths received by the victim traversed through 4753 different paths. Figure 4 shows the distribution of the attack packets through the different paths. 22 ITrace messages were received by the victim and 176 attack paths were fully reconstructed. A total of 3719 attack packets (≈8.9%) travelled through these 176 paths. Path no. 4753, which routed the highest percentage of the attack packets (i.e. 1348 packets or ≈3.2%), had 8 intermediate routers with only 4 of them been discovered (i.e. 50% of the path).

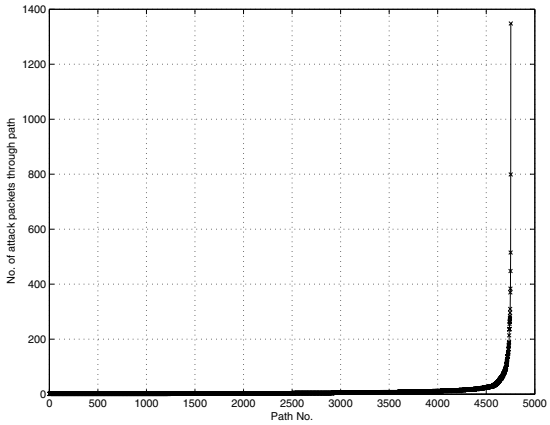

Figure 4: Attack packets distribution (using DSDV routing protocol and ITrace traceback in 54-node network)

As the expected number of packets required for path reconstruction in ITrace is much more than PPM (due to lower probability of traceback message generation), significant route changes leading to a larger number of traversed paths were expected. This was evident in the simulation results obtained.

Therefore, we experimented with increasing the probability of ITrace message generation in the 54-node network by setting the expected number of attack packets to be the same as that in PPM (i.e. 1516 packets). The probability was calculated to be 0.00138 (i.e. ≈27.6 times the original probability of 1/20000). 9 paths were traversed by the packets. Figure 5 shows the distribution of the attack packets through the different paths. 23 ITrace messages were received by the victim and 4 attack paths (i.e. Path no. 1, 2, 8, and 9) were fully reconstructed. A total of 916 attack packets (≈60.42%) travelled through these 4 paths. Path no. 9, one of the fully reconstructed path, was the route with the highest percentage of the attack packets (i.e. 497 packets or ≈32.8%).

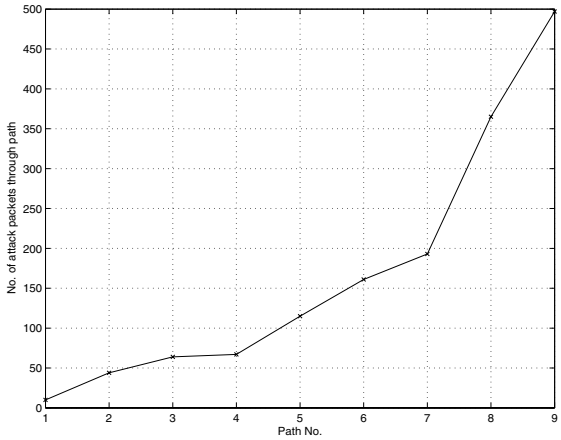

Figure 5: Attack packets distribution (using DSDV routing protocol and ITrace traceback in 54-node network, with same expected number (i.e. 1516) of attack packets as PPM, and corresponding lower probability at 0.00138)

It was shown that in wireless ad-hoc networks, the performance of the traceback mechanisms depended on the routing protocols used and the speed of the attack paths reconstruction, even in a static network and a single attacker. If routing was maintained throughout the length of the attack (e.g. in AODV), resulting in only a single attack path, the reconstruction of the path would be possible and the speed of the traceback depended only on the technique used and parameters (e.g. probability) applied. When using dynamic routing protocol (e.g. in DSDV), an efficient traceback mechanism with appropriate parameters setting will be required to make attack paths reconstruction (even partial in a large network) possible. In the event of multiple attackers (e.g. DDoS attack), multiple paths will be traversed even when routing protocols such as AODV was used and traceback would be difficult without an efficient scheme.

For the dynamic case, whereby the attacker (or even the other nodes) will be moving at a constant rate, the different routes traversed by the attack packets increased and the routes also became unstable. Therefore, in such a situation the detection of the routes would be even more difficult then in a DDoS attack, even when using routing protocols such as

AODV, as the routing would not be able to be maintained in such a dynamic scenario.

IV. CONCLUSIONS

In this paper, we investigated the feasibility of applying the existing IP Traceback techniques on wireless ad-hoc networks. The IP Traceback techniques considered were SPIE, PPM, and ITrace. Potential large storage requirement and the infrastructural characteristic of the SPIE scheme made it infeasible for the wireless ad-hoc networking environment. Characteristics of the PPM's Compressed Edge Fragment Sampling approach and ITrace's both links approach were considered to be promising for reconstructing attack paths in wireless ad-hoc networks. They are non-infrastructural based and do not require large storage space. Probabilistic nature of these schemes allowed traceback information to be stored longer due to lesser storage space requirement (as compared to SPIE) to cater for post mortem tracing too. Therefore, these 2 schemes were chosen for further quantitative analysis. The routing protocols selected for simulations were DSDV and AODV, for the proactive and reactive categories respectively.

Simulation studies were conducted using network sizes of 14, 34, and 54 nodes. Using AODV, only 1 attack path was chosen to be used for routing the attack packets from the attacker to the victim. Therefore, the scenario was similar to that of the wired network, and each of the paths was fully reconstructed using both PPM and ITrace in the 14, 34, and 54-node networks. However, multiple paths were used to route the attack packets to the victim using DSDV when the number of packets and network size increased. The reasons were most probably route changes due to updates and the occurrence of network traffic congestion. When the network size and the expected number of attack packets required for path reconstruction was small (as in 14 and 34-node networks using PPM), a single path was used for routing and it was fully reconstructed. However, as the network size increased to 54-node and the expected number of attack packets to 1516, multiple paths began to appear and reconstruction of all the paths was not possible. The successfully reconstructed attack paths were responsible for routing ≈61.7% of the attack traffic.

As the expected number of packets required for path reconstruction in ITrace is much more than PPM (due to lower probability of traceback message generation), significant route changes leading to a larger number of traversed paths were expected. This was evident in the simulation results obtained. Only a small percentage of the attack graph (i.e. consisting all the attack paths) was reconstructed when ITrace was used as the traceback mechanism. 20%, 5.5% and 3.7% of all the attack paths were fully reconstructed in the 14, 34 and 54-node network, respectively. These discovered attack paths were responsible for routing 93%, 28.3% and 8.9% of the attack traffic in each of the respective networks. We also experimented with setting the expected number of packets for the 54-node network to be the same as PPM (i.e. 1516 attack packets). By changing the probability of ITrace message generation to 0.00138, the simulations showed that 44.44% of all the attack paths were fully reconstructed, and they were responsible for routing 60.42% of the attack traffic. This included the path with the highest percentage (≈32.8%) of attack packets traversing through it.

It was shown in this paper that the performance of traceback in wireless ad-hoc networks depended on the network size, and the routing protocols and traceback mechanism used (even in a static network with a single attacker). If routing was maintained throughout the length of the attack (e.g. in AODV), resulting in only a single attack path, the reconstruction of the path would be possible and the speed of the traceback depended only on the technique used and parameters (e.g. probability) applied. When using dynamic routing protocol (e.g. in DSDV), the traceback mechanisms would have to be applied with higher probability settings to make attack paths reconstruction (even partial in a large network) possible.

REFERENCES

[1] K. J. Houle, G. M. Weaver, "Trends in Denial of Service Attack Technology", CERT Coordination Center, Oct. 2001

[2] L. Garber, "Denial-of-Service attacks rip the Internet", IEEE Computer, Vol. 33, No. 4, pp. 12-17, Apr. 2000

[3] David Moore, Geoffrey M. Voelker, Stefan Savage, "Inferring Internet Denial-of-Service Activity", Usenix Security Symposium, Aug. 2001

[4] Alex C. Snoeren et al, "Hash-Based IP Traceback", ACM Sigcomm 2001, Aug. 2001

[5] Stefan Savage et al, "Practical network support for IP traceback", ACM Sigcomm 2000

[6] Steve Bellovin et al, "ICMP Traceback Messages", IETF Internet Draft, Version 4, Feb. 2003 (Work in progress)

[7] Charles E. Perkins, Pravin Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers", ACM Sigcomm 1994

[8] C. Perkins, E. Belding-Royer, S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing", IETF RFC3561, July 2003

[9] Vadim Kuznetsov, Helena Sandstrom, Andrei Simkin, "An Evaluation of Different IP Traceback Approaches", ICICS 2002, Springer LNCS Vol. 2513, pp. 37-48, 2002

[10] The Network Simulator (ns-2), http://www.isi.edu/nsnam/ns