SWAT: *S*mall *W*orld-based *A*ttacker *T*raceback in Ad-hoc Networks

[1]Yongjin Kim, [2]Ahmed Helmy
[1,2]Electrical Engineering Dept. – Systems
University of Southern California, California, U.S.A.
[1]yongjkim@usc.edu, [2]helmy@usc.edu

Abstract

Mobile Ad hoc NETworks (MANETs) provide a lot of promise for many practical applications. However, MANETs are vulnerable to a number of attacks due to its autonomous nature. DoS/DDoS attacker traceback is especially challenging in MANETs for the lack of infrastructure. In this paper, we propose an efficient on-the-fly search technique, SWAT, to trace back DoS and DDoS attackers in MANETs. Our scheme borrows from small worlds, utilizes the concept of Contacts, and use Traffic Pattern Matching (TPM) and Traffic Volume Matching (TVM) techniques. We also propose multi-directional search, in-network processing and query suppression to reduce communication overhead in energy-constrained MANETs and increase traceback robustness against spoofing and collusion. Simulation results show that SWAT successfully traces back DoS and DDoS attacker under reasonable background traffic. In addition, SWAT incurs low communication overhead (22% compared to flooding-based search).

1. Introduction

Flooding-type/direct DoS [1-4] and DDoS [8] attacks consume the resources of a remote host or network, thereby denying or degrading service to legitimate users. There are several characteristics of such attacks: (I) Traffic volume abnormally increases during the attack period. (II) Attackers routinely disguise their location using incorrect/spoofed addresses. (III) It is reported that such attacks may persist for tens of minutes and in some case for several days [1].

IP traceback in the Internet, which tracks down attacker(s), is a useful technique for forensics and to discourage attackers. There are several IP traceback schemes proposed for the Internet such as packet marking [14], logging [11], ICMP traceback [6,13], and others [5]. Such traceback schemes developed for the fixed networks are not directly applicable to MANETs due to the following reasons.

- In MANETs, there is no fixed infrastructure. Each node acts as an autonomous terminal, acting as both host and router.
- Node mobility frequently changes network topology.
- In general, network bandwidth and battery power are limited.

To perform efficient DoS/DDoS attacker traceback under such a harsh environment in MANETs, we propose an efficient on-the-fly taceback technique. For that, we leverage the small world model. The concept of small worlds was studied in the 60's in the context of social networks [12], during which experiments of mail delivery using acquaintances resulted in an average of 'six degrees of separation', i.e., on average a letter needed six acquaintances to be delivered. Recent research by Watts [16] has shown that in relational graphs adding a few number of random links to regular graphs results in graphs with low average path length and high clustering. Such graphs are called small world graphs. Helmy [9][10] established the applicability of small world graphs to MANETs. Helmy found that path length of wireless networks is drastically reduced by adding a few random links (resembling a small world). Establishing a small world reduces the degrees of separation between victim and attacker and provides a solid basis for efficient traceback mechanism. In this paper, we effectively extend Contacts [10] to build small world in MANETs.

Address of DoS/DDoS attack packets is commonly spoofed to disguise its identity and prevent traceback. Hence, source address of attack packet bears no clues that could be used to determine their originating host. To deal with the address spoofing problem, we use attack traffic signature. A traffic signature is defined by the sequence of number of packets in time slots when traffic is abnormally increased. By finding neighbor nodes of victim that observe *similar* attack traffic signature within a given timeframe and performing the process recursively to attack origin, we can track down attacker(s). To define the similarity of traffic signature, we use Traffic Pattern Matching, TPM [11], and traffic volume matching, TVM. TPM is the process to check the similarity of *traffic pattern* and TVM is the process to check the similarity of *traffic volume*.

To reduce the impact of collusion (i.e., false information reporting), we take majority-matching approach, in which we decide based on the reports from majority nodes.

We also use *directional search, in-network processing and query suppression* to reduce communication overhead in DoS/DDoS attacker

Proceedings of the Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'05)
0-7695-2375-7/05 $20.00 © 2005 IEEE

traceback. Our scheme reduces the traceback search overhead significantly compared with flooding since SWAT has directionality in searching.

Our paper is organized as follows. In section 2, we briefly review existing IP traceback schemes and qualitatively analyze the limitations of the existing IP traceback schemes to be applied to MANETs. In section 3, we mention design goals and requirements. In section 4, we provide SWAT architecture overview and comparison with existing IP traceback schemes. In section 5, we explain TPM and TVM as matching-in-depth techniques. In section 6, we describe how to build small world in MANETs for efficient search. Overall DoS and DDoS attacker traceback mechanisms are provided in section 7. We show simulation results in section 8 and provide discussions in section 9. Finally, in section 10, we conclude our paper.

2. Related works

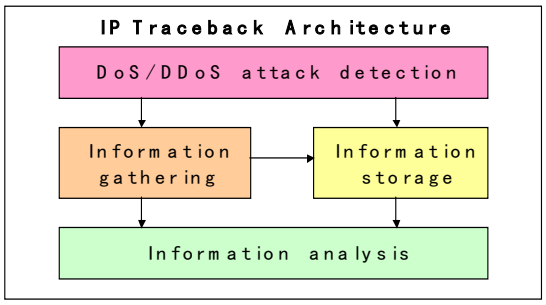
To the best of our knowledge, SWAT is the first work for attacker traceback in MANETs. In this section, we briefly review existing IP traceback schemes developed for the Internet and investigate the applicability of existing IP traceback schemes to MANETs.

Probabilistic Packet Marking (PPM) [17], and ICMP Traceback Message (iTrace) [6] attempt to distribute the burden of storing state and performing computation for IP traceback at the end hosts rather than in the network. For instance, in ICMP-based notification, router generates ICMP message containing information regarding where each packet came from and when it was sent. Then, routers notify the packet destination of their presence on the route. Collection of these messages can be used to trace the attack origin. ICMP traceback message uses ICMP but limits to generating only ICMP message for every 20,000 packets (recommended). In Probabilistic Packet Marking (PPM), routers insert traceback data into each packet probabilistically so the number of packets that are marked at each router is enough for the reconstruction of attack path at the victim.

Logging scheme requires the routers to log meta-data in case an incoming packet proves to be offensive. Audited packet flow is logged at various points throughout the network and then used for appropriate extraction techniques to discover the packet's path through the network. To reduce the size of packet log and provide confidentiality, hash-based logging is proposed [13].

Controlled flooding [7] tests network links between routers to determine the origin of the attacker traffic. Downstream node intentionally sends a burst of network traffic to the upstream network segments. At the same time, it checks incoming attack traffic for any changes. From the changes and frequency of the incoming attack traffic, the victim can determine which upstream router the traffic is coming from. The same process is continued a level higher until finally reaching the attacker. Since this is a reactive method, the trace needs to be completed before the attack is over.

The main building blocks of the existing IP traceback schemes can be classified as (I) *information gathering,* (II) *information storage,* and (III) *information analysis.* Information gathering is the process to put or seek clues on the attack packets. For instance in PPM, packet marking is the process for information gathering, in which intermediate routers attach clue on packets and send them to the end host. Information storage is the process to storing the gathered clue in some device for post-analysis. In case of logging, the information is stored inside the network. On the other hand, in PPM or iTrace, information is stored at the end-host. Information analysis is the process to reconstruct the attack path based on clue obtained through information storing process or real-time data provided by information gathering process. Based on this classification and functionality of each building block, we investigate the applicability of existing IP traceback schemes to MANETs.



[Figure 1] Main building blocks of IP traceback. Attack traffic information is first gathered. Then the information is stored or directly used for traceback. Stored attack information is used for post-analysis

(I) Information gathering

PPM, iTrace and logging requires per packet processing for attack information gathering. It is difficult to be directly applied to IP traceback in MANETs due to the following reasons: Intermediate nodes move in/out and may fail due to power outage, frequently changing network topology. Consequently, attack route which is inferred from either packet marking, ICMP message, or logging after receiving enough packets from enough routers are most likely to be obsolete when there exists frequent topology change. On-the-fly technique for traceback is essential in MANETs. Controlled-flooding can provide on-the-fly searching. However, it consumes network bandwidth, which is highly undesirable in resource constrained wireless networks.

(II) Information storage

Clue information, obtained through information gathering process, needs to be stored for post-traceback. Information can be stored at the end-host or inside the network. iTrace or PPM is end-host storage scheme. ICMP or marked packets are stored at the end-host and used for path reconstruction. Obvious drawback of information storage is that large amount of data needs to be stored at either end-host or inside the network. In general, storage capacity of nodes in MANETs is restricted. Hence, it is difficult to store per-packet information for traceback in MANETs. On the other hand, controlled flooding does not require information storage.

(III) Information analysis:

Information analysis is the process to reconstruct the attack path given stored or real-time information. In iTrace and PPM, the path is reconstructed by end-host. Consequently, the burden of end-host is increased. For instance, in iTrace, end host first searches the database, which stores packet information. Then, based on the packet information, end-host should reconstruct the attack path. On the other hand, in case of logging and controlled flooding, analysis burden is put on the network. Information analysis processes of existing schemes require a lot of computation power of either end-host or network to reconstruct the attack path based on accumulated database. In addition, controlled flooding requires lots of bandwidth consumption even if it is short term.

3. Design goals and requirements

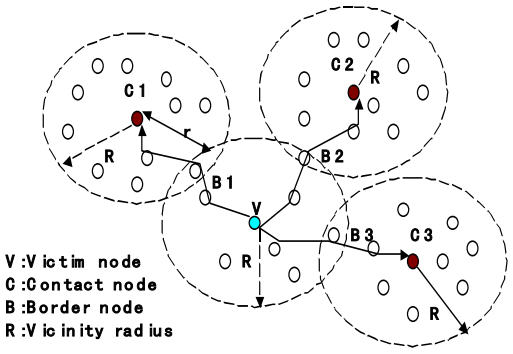
We classify design goals for efficient attacker traceback in MANETs as follows:

I. *Robustness to address spoofing*: It is a common attack technique to spoof attackers' addresses. We should be able to trace attackers in spite of address spoofing.
II. *Robustness against collusion*: In MANETs, intermediate nodes that relay attack traffic can be compromised. Consequently, it is important to have robustness against some node collusion.
III. *Scalability*: Applications of large-scale ad hoc networks involve military and sensor network environments that may include thousands of nodes. Hence traceback mechanism should be scalable in term of communication overhead with increase in network size.
IV. *Efficiency*: Ad hoc networks include portable devices with limited battery power. Traceback mechanism should be power-efficient in terms of communication and computation.
V. *Robustness to topology change*: The mechanism should be robust to handle frequent node mobility and failure due to power outage.
VI. *Decentralized operation*: For the network to be rapidly deployable, it should not require any centralized control.

4. SWAT Architecture overview

In this section, we provide high-level overview of SWAT architecture and compare it with existing IP traceback schemes. DoS/DDoS attack is first identified by intrusion detection system at each node. Once the attack is identified, the victim initiates attacker traceback, which is composed of *attack traffic analysis* and *efficient searching*. Basically, attack traffic is characterized at the victim to be used as attack signature. Then searching process in launched with the attack signature to find relay nodes and final attacker(s).

To characterize attack traffic, we use traffic pattern [11] and traffic volume. Traffic pattern and traffic volume represent abnormal characteristics of attack traffic. Under attack, traffic shows abnormal pattern/volume increase and the abnormality is observed consistently on the route from attacker to victim. Characterizing attack traffic with traffic pattern/volume in SWAT is light-weight in terms of information gathering/storage since it requires only the packet count information in a given time window (refer to section 5 for details).



[Figure 2] Each node has vicinity of radius *R* hops. A victim sends query with attack traffic signature to its vicinity nodes and border nodes B_i. Then, the border nodes choose one of its borders C_i, to be the contact and sends query with attack signature.

Once the attack traffic signature is characterized by traffic pattern/volume, victim node initiates efficient search process. By finding nodes in the neighbor, which observe similar attack traffic signature, we can find nodes that relayed the attack traffic. The process is continued recursively from the neighbor nodes up to the attacker(s). To efficiently search nodes that

[Table 1] Analysis of IP traceback schemes for the applicability to MANETs

	Requirements	PPM	iTrace	Controlled flooding	Logging	***SWAT***
Information gathering	Robustness against route instability	No	No	Yes	No	Yes
	Bandwidth requirement	Low	Low	High	Low	Low
	Battery requirement	Low	Medium	High	Low	Low
Information storage	Storage requirement of intermediate node	Not needed	Not needed	Not needed	High	Low
	Storage requirement of victim node	High	High	Not needed	Not needed	Low
Information analysis	Computation power requirement of intermediate node	Not needed	Not needed	Low	High	Low
	Computation power requirement of victim node	High	High	Low	Not needed	Low

observed similar traffic signature on the attack route, we extend small world-based contact [10]. Contact nodes are a set of nodes outside the vicinity, which are used as short-cut to build small world and provide wide view on entire network to the victim. As shown in figure 2, victim node, *V*, sends query with attack signature to its vicinity nodes (nodes within radius *R*) and contact (*C1, C2*, and *C3*). To send to contacts, the victim node chooses three borders, *B1, B2* and *B3*, to which it sends queries. The borders in turn choose three contacts at *r* hops away to which the borders forward the query. Each contact performs in-network processing to check whether there are vicinity nodes that observed attack traffic. If there is no node that observed (relayed) attack traffic, it suppresses query. Otherwise, it sends next level query to the contact of contact. In doing so, we can perform directional search for DoS attacker traceback and multi-directional search for DDoS attacker traceback, where the search process has directionality towards attacker(s). Directional and multi-directional search significantly reduces communication overhead. We will verify the reduction in the simulation section.

We compare existing IP traceback schemes and SWAT for MANETs in table 1. For information gathering, SWAT incurs low battery/bandwidth consumption, since each node counts only the number of packets and it has directionality in searching. In addition, it is robust to topology change since SWAT is performed on-the-fly. For information storage, storage requirement for both intermediate node and victim is low since SWAT only stores packet count information when the number of packet count is abnormally high. Lastly for information analysis, computation complexity is low at both victim and intermediate nodes, since each node performs computationally light traffic pattern matching and volume matching (refer to section 5 for details), which is *O(1)* at each node.

5. Attack signature characterization

In DoS/DDoS attack, a large amount of packets are generated towards the victim. For instance, 200-500 pps of SYN packets are generated in TCP SYN flooding attack [2]. However, in normal case, only one SYN packet is generated per connection. Hence, abnormally large number of SYN packets is considered suspicious.

Each node monitors the number of packets per time slot and if there is abnormal increase of traffic, a node logs the attack traffic signature. Before we define attack traffic signature, we define abnormality of traffic increase. We use simple statistical method, Fractional Deviation from the Mean (FDM), which differentiates abnormal short-term behavior from normal long-term behavior. Let A_S the number of packets in a given time slot and A_R be the average number of packets of the long-term reference model, then the distance of the fractional deviation from the mean statistic is given as follows.

$$Dist = \frac{A_S - A_R}{A_R} \qquad \text{(Eq.1)}$$

The distance, *Dist*, is defined as abnormality level. If the abnormality level is over a threshold (0.5 in SWAT), it is considered suspicious and traffic signature is logged.

Attack traffic signature is defined by the sequence of *number of packets* in *n* time slots, $(a_1,a_2,\ldots,a_n)$, where a_i $(1 \le i \le n)$ is the number of packets at time slot *i*. Sampling window, *D*, is expressed as follows.

$$D = n \bullet d \qquad \text{(Eq.2)}$$

Where d is time slot length.

Traffic Pattern Matching (TPM) defines the correlation coefficient between two traffic signatures at node A and B. TPM captures the variation of traffic volume V_a in Fig. 3. When the traffic observed at node A is given as $(a_1,a_2,...,a_n)$, and the traffic observed at node B is given as $(b_1,b_2,...,b_n)$, the correlation coefficient is obtained as follows.

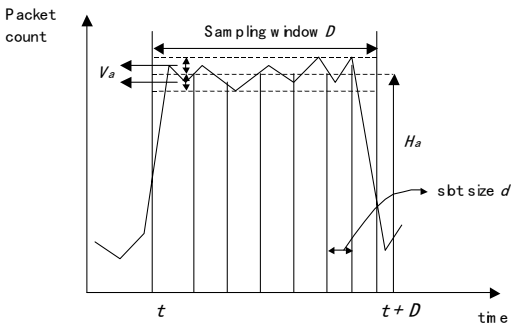
$$r(A,B) = \frac{1}{nS_A S_B} \sum_{i=1}^{n} (a_i - \overline{A})(b_i - \overline{B}) \quad \text{(Eq.3)}$$

Where,

$$S_A = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (a_i - \overline{A})^2} \quad \text{(Eq.4)}$$

$$S_B = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (b_i - \overline{B})^2} \quad \text{(Eq.5)}$$

and $\overline{A}$ and $\overline{B}$ are the averages of $(a_1,a_2,...,a_n)$, and $(b_1,b_2,...,b_n)$. In case correlation coefficient $r(A,B)$ is high (greater than 0.7), the traffic at A is said to match traffic at B.



[Figure 3] Model of attack traffic signature

We also use Traffic Volume Matching (TVM) to reflect H_a factor in Fig. 3 and complement the TPM. We define that traffic volume is matching between two points A and B, when traffic volume at A and B show similar volume size. Mathematically, we use the following equation (least-squares method) to know the matching level.

$$c = \frac{\sum_{i=1}^{N} a_i b_i}{\sum_{i=1}^{N} a_i^2} \quad \text{(Eq.6)}$$

When c is close to 1, the traffic volume at node A and B is matching.

Unlike DoS attack, DDoS attack is performed from multiple nodes. Partial attack traffic is merged at the victim or intermediate node. Consequently, combination of partial traffic from multiple nodes should be compared with the merged traffic to find distributed attack routes. There are two possible scenarios in the merging of partial attack traffic: (i) partial attack traffic shows different traffic pattern from merged traffic. That is $r(P_i,M)$ is low, where P_i is partial attack traffic signature and M is merged attack traffic signature, or (ii) partial attack traffic shows similar traffic pattern with merged attack traffic. That is $r(P_i,M)$ is high. For the second scenario, TVM is especially important to detect multiple attack routes. With TPM only scheme, many false negatives are inevitable since each partial attack traffic may show high TPM with merged attack traffic. We will illustrate this observation in the simulation section. In case correlation coefficient of $r(\sum_{i=1}^{L} P_i, M)$, where L is the total number of partial attack traffic signature, is high, the merged attack traffic, M, is said to match the summation of partial attack traffic of P_i. There can be S number of combinations from K candidate $(L \le K)$ partial attack traffic as follows.

$$S = \sum_{i=1}^{K} {}_K C_i \quad \text{(Eq.7)}$$

In our scheme, the combination that shows the highest TPM/TVM level is selected as the path of distributed DDoS attack traffic.

6. Small world construction

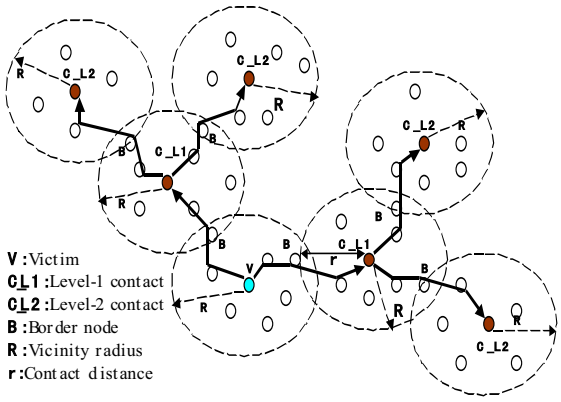
For efficient attacker searching, we use small world model. Helmy [9][10] found that path length in wireless networks is drastically reduced by adding a few random links (resembling a small world). These random links need not be totally random, but in fact may be confined to small fraction of the network diameter, thus reducing the overhead of creating such network. The random links can be established using contacts [10]. We extend the contact architecture to build small world in wireless networks and increase attacker searching efficiency. Contact nodes are a set of nodes outside the vicinity, which are used as short-cut (random links) to build small world. We describe detailed small world construction scheme in the following.

Each node in the ad hoc network keeps track of a number of nodes in its vicinity within R hops away. This defines the vicinity of a node. The vicinity information is obtained through underlying routing protocol. Each node chooses its vicinity independently, and hence no major re-configuration is needed when a node moves or fails. There is no notion of cluster head, and no elections that require consensus among nodes.

On-demand, a victim node selects a set of contacts outside its vicinity. The main purpose of contact nodes is to act as a short cut. Hence, it is important for contacts to have vicinity that does not overlap significantly with that of the victim node, V, or the other contacts of V. The first kind of overlap (vicinity overlap) occurs between the contact's vicinity and the victim's vicinity. To reduce this overlap, victim node attempts to push the request as far out from the victim's vicinity as possible. Let the borders of victim V be B

(Fig.4). *V* sends a query to the number of (*NoC*) its *B*. *B* constructs a topology view up to *R* hops away using its own vicinity information, and chooses a border in its vicinity that has maximum distance to *V*.

The second type of overlap, route overlap, occurs between vicinities of contacts. To reduce this overlap, *V* selects *NoC* borders with maximum separation. This is done using vicinity information.



[Figure 4] Small world construction with muti-level contacts. Victim, *v*, selects level-1 contacts. Level-1 contacts select its level-2 contacts.

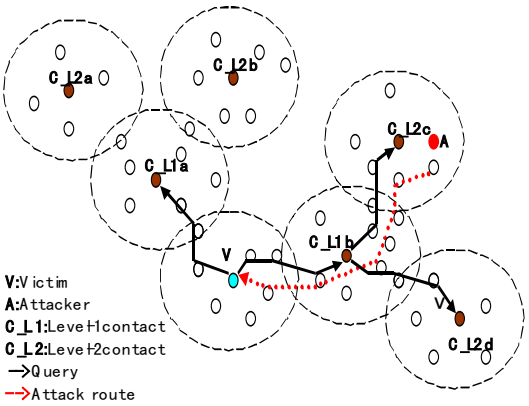
The above contact selection scheme provides a mechanism to select *NoC* contacts that have distances up to *R*+*r* hops away from *V*. We call these contacts level-*1* contacts. To select farther contacts (contact of contact), this process is further repeated as needed at the level-*1* contacts, level-*2* contacts and so on, up to a number of levels called maxDepth, *D*.

Our contact selection and search policy have the following important distinctions from [10]: (1) Contacts are randomly selected every time it launches search to prevent divulgence of contact information to attackers. That is, if contact nodes for a victim are fixed, attacker will try to compromise the fixed contact nodes to prevent traceback. To reduce this risk, we select contacts randomly. (2) Contacts in SWAT perform in-network processing (TPM and TVM test) to check whether attack traffic is traversed through vicinity nodes or not. (3) SWAT performs (multi-) directional search where the search is directed towards the attacker(s) to reduce communication overhead. (Multi-) Directional search becomes possible through query suppression where contacts that do not have attack route in their vicinity suppress further queries. (4) Our contact selection is performed upon underlying ad-hoc routing protocols and independent of any specific ad-hoc routing protocol.

7. Overall traceback mechanism description

In this section, we describe detailed attacker traceback mechanism of SWAT, based on the efficient searching and TPM/TVM techniques described in the previous sections. Basically, victim first initiates traceback procedure by sending characterized attack signature to its level-*1* contacts. Then, level-*1* contacts look for vicinity nodes that observe similar traffic pattern/volume with the attack signature. The contacts that find vicinity nodes with high TPM/TVM level perfrom the search process recursively until it find attacker(s).

7.1 DoS Attacker Traceback



[Figure 5] Victim (*V*) sends queries with attack traffic signature to the first level contacts, (*CL_1a, CL_1b*). Only *CL_1b* that observed matching traffic signature within vicinity sends next level queries to level-2 contacts (*CL_2c, CL_2d*). *CL_1a* suppresses further query. *CL_2c* sends final attack route to the victim.

We describe the DoS attack traceback scheme as follows: (1) when a victim node, *V*, detects attack such as SYN flooding, it first extracts attack traffic signature described by the traffic pattern and volume. It then sends a query to the nodes within its vicinity and level-*1* contacts specifying the depth of search (*D*) large enough to detect an attacker. The query contains sequence number (*SN*) and attack traffic signature. (2) As the query is forwarded, each node traversed records the *SN*, and *V*. If a node receives a request with the same *SN* and *V*, it drops the query. This provides for loop prevention and avoidance of re-visits to the covered parts of the network. (3) In case high TPM and TVM reports are observed by vicinity nodes and contacts, the first step of trace is completed. For instance, victim (*V*) sends query to the vicinity nodes and 2 level-*1* Contacts (*CL_1a* and *CL_1b*) around the victim in Fig. 5 (transmission arrows to vicinity nodes by each contact are omitted in the figures). Then, one level-*1* (*CL_1b*) contact reports to the victim that some of its vicinity nodes observed high TPM/TVM level. To reduce the risk of false matching report from vicinity nodes, contact requests traffic signature observed at the vicinity nodes at given time slots instead of distributing attack traffic signature to all

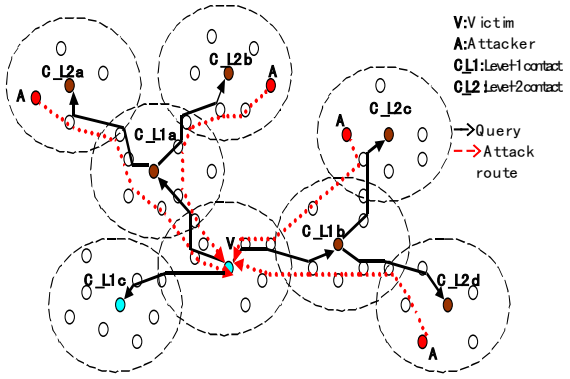
vicinity nodes and waiting for TPM and TVM response. TPM and TVM tests are done at each contact. Although it cannot completely remove the risk of false matching report, it can reduce such risk. (4) Next, only the contact, *CL_1b*, that observes traffic signature matching in its vicinity sends next level query to level-*2* contacts (*CL_2c*, and *CL_2d*) with the partial attack path appended to the query. It also reduces *D* by 1. This processing by contact is called *in-network processing*. Other contacts that do not have relay nodes of attack traffic in their vicinities, suppress forwarding the query *(query suppression)*. This results in *directional search* towards the attacker. (5) When there is no more contact report or no other nodes outside the vicinity, the last contact (*CL_2c*) reports the complete attack route to the victim.

Our scheme is based on majority node reporting. That is, even if some nodes move out from the attack route or are compromised by attackers, we can still find an attack route using available information from good nodes in the vicinity.

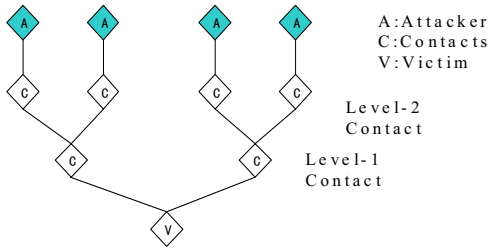
7.2 DDoS Attacker Traceback

In this section, we describe the DDoS attacker traceback scheme. DDoS attacks involve a sufficient number of compromised nodes to send useless packets toward a victim around the same time. The magnitude of the combined traffic is significant enough to jam, or even crash, the victim or connection links.

Similar to DoS case, a victim node sends traffic pattern/volume matching query to its vicinity and level-*1* contacts with its characterized attack traffic signature. In DDoS attacker traceback, multiple candidate attack signatures are observed and returned from multiple contacts. For instance, in Fig. 6, three responses are returned from level-*1* contacts (*CL_1a*, *CL_1b*, and *CL_1c*) and the victim calculates TPM and TVM level from all possible combinations (Eq.7). In this example, TPM and TVM tests show highest value between the summation of two traffic signatures (from *CL_1a*, and *CL_1b*) and attack traffic at the victim. As a result, a victim concludes that attack traffic comes from *CL_1a* and *CL_1b* vicinity nodes. Note that TPM between partial attack traffic (either from *CL_1a* or *CL_1b*) and merged attack traffic at the victim may show high TPM level. We need TVM test to decide whether it is distributed attack traffic or single attack traffic. In case TPM level is high and TVM level is low we conclude that it is partial attack traffic and seek for other partial traffic which forms DDoS attack traffic. Contacts that are determined as attack route by the victim node perform next level query in a recursive manner. Each level-*1* contact finds two other branches of attack route in two level-*2* (*CL_2a, CL_2b, CL_2c*, and *CL_2d*) contacts. Final attack route is reported to the victim by the last contact nodes. Figure 7 illustrates logical view of DDoS attacker traceback tree, from the victim (root) to distributed attackers (leaves). Intermediate contacts have child contacts from which partial attack traffic is coming.



[Figure 6] Victim (*V*) sends queries with attack traffic signature to level -*1* contacts, (*CL_1a,CL_1b, CL_1c*). Two level-*1* contacts (*CL_1a, CL_1b)* that observe matching traffic signature within vicinity sends next level queries to level-*2* contacts (*CL_2a, CL_2b, CL_2c, CL_2d*). Final level-*2* contacts send final distributed attack route to the victim.



[Figure 7] Logical view of DDoS attacker traceback tree

8 Simulation results

We have performed extensive simulations to evaluate the effectiveness of the proposed traceback scheme with varying parameter space (table 2). Transmission range of each node is 140m. We repeated each simulation 100 times in grid topology and calculated the average value. The evaluation metrics that we measured in the simulation are TPM/TVM level, traceback success rate, and communication overhead. We set *NoC (*Number of Contacts*)* = 6, *R* (vicinity radius) = 3, *r* (contact distance) = 3, *d* (search depth) =5 for contact selection. Attack traffic of u[200,240] pps is generated for 10 minutes. Low or no mobility is considered and DSDV is used as underlying routing protocol. We used simulation scenario 1 in table 2 to evaluate the effect of background traffic and success rate and scenario 1,2, and 3 to evaluate the effect of node density. In addition, we used scenario 1,4,5, and 6 to evaluate the effect of network size.

[Table 2] Simulation environment (*Node degree is the number of nodes within transmission range)

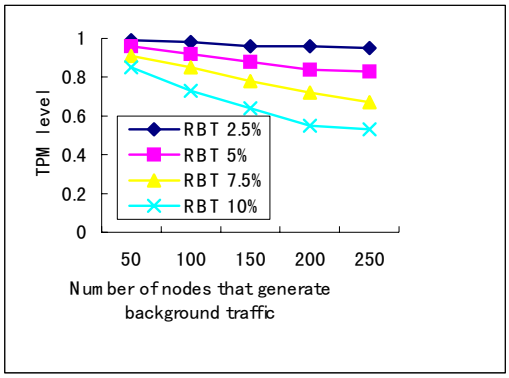
Scenario	Nodes	Area	Node degree*	Distance
1	484	1680m x 1680m	9	80m
2	625	1680m x 1680m	13	70m
3	841	1680m x 1680m	21	60m
4	1089	2560m x 2560m	9	80m
5	1936	3440m x 3440m	9	80m
6	3025	4320m x 4320m	9	80m

8.1 DoS Attacker Traceback

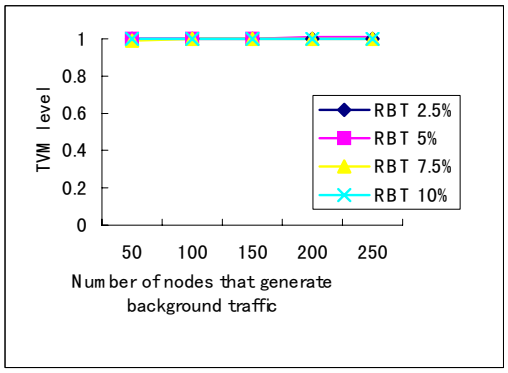
In DoS attack simulation, attacker is randomly located 17 hops away from victim.

■ *Effect of background traffic*

One of the most important factors that may affect the performance of the proposed scheme is the volume of background traffic. Increased background traffic negatively affects TPM/TVM test (i.e., lowering the TPM/TVM level), which prevents successful traceback. To investigate the impact of background traffic on TPM and TVM level, we varied two parameters in our simulation. The first parameter is the number of source nodes at random location that generates background traffic towards randomly selected destinations. The second parameter is Relative Background Traffic (RBT) that each random source node generates. RBT represents the percentage of background traffic relative to attack traffic. For instance, when attack traffic volume is 200pps, RBT of 10% represents 20pps of background traffic. We measured TPM level (Fig.8) and TVM level (Fig.9), which represent the matching levels between attack traffic signature observed at the victim and attack traffic signature observed at the intermediate (relay) nodes.



[Figure 8] TPM level comparison with varying background traffic

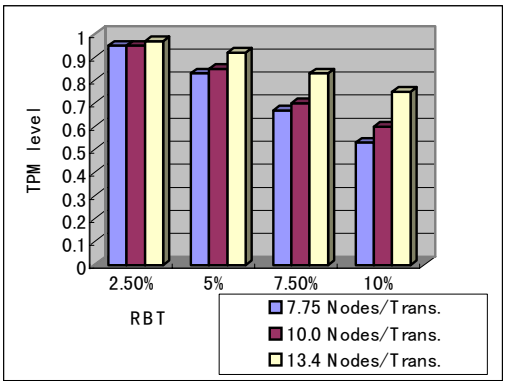


[Figure 9] TVM level comparison with varying background traffic

As shown in Fig.8, when RBT is less than 5%, we can constantly observe high TPM across varying number of nodes that generates background traffic. In case RBT is 10% and the number of nodes that generates background traffic is over 150, TPM level is less than 0.7 (average of 0.57). It is because high background traffic affects attack traffic pattern. However, from practical point of view, it is reasonable to assume that normal traffic (background traffic) is less than 7.5% of attack traffic, in which case TPM level is consistently high. For instance, 200pps-500 pps of attack traffic is observed [2] in SYN attack case. In normal case, assuming less than 15-38 pps, which is less than 7.5% of attack traffic, of SYN packets generated by good nodes is reasonable range. We can observe high TVM level across varying background traffic size in Fig. 9. That is, regardless of background traffic, traffic volume of victim and relay nodes is closely matching.

■ *Effect of node density*

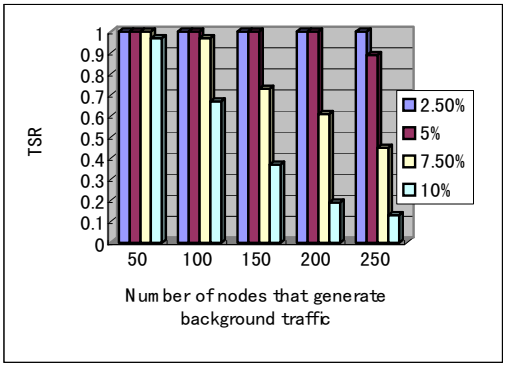
Fig. 10 shows the effect of node density on TPM level. In the simulation, we fixed the number of nodes that generate background traffic as 250. When RBT is greater than 7.5% and node degree is 7.75, the TPM level goes below 0.7. However, when node degree is 13.4, TPM level goes beyond 0.7. It is because as node density increases, routing diversity is increased. That is, background traffic is relayed by diverse different nodes and each node relays reduced amount of traffic. Consequently, the negative impact of background traffic on attacker traceback decreases as node density increases.



[Figure 10] TPM level comparison with varying node density

■ *Traceback success rate*

Fig. 11 and 12 show the overall traceback success rate. We define traceback success as two categories: *Traceback Success Rate (TSR),* and *Perfect Traceback Success Rate (PTSR).* TSR is the rate at which we can trace back attacker and *partial* intermediate nodes that relay attack traffic. In TSR, we can trace back an attacker if there is at least one TPM and TVM matching report from vicinity nodes of each intermediate contact. PTSR is the rate that we can find attacker and *all* the relay nodes on the attack route. As shown in Fig. 11, SWAT shows perfect TSR when RBT is less than 2.5% and RBT is 5% and when number of nodes that generate background traffic is less than 200 (around 40% among entire nodes). In case maximum background traffic is more than 7.5%, the success rate is shapely decreased as background traffic increases.

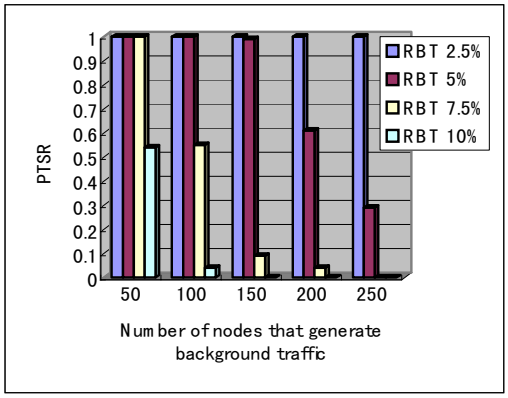


[Figure 11] Traceback success rate with varying background traffic

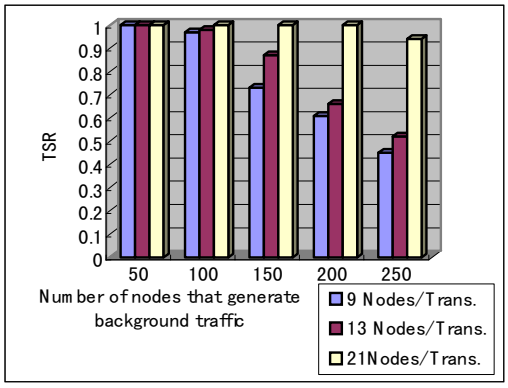
Fig. 12 shows perfect traceback success rate, PTSR. PTSR is decreased drastically as background traffic increases except in case RBT is less than 2.5%. It is not always necessary to get high PTSR for the purpose of attacker traceback. That is, we do not need to detect all the intermediate nodes that relay attack traffic. We can trace back attacker successfully if we can achieve high TSR.

False positive in attacker traceback was not observed since other nodes (not relaying attack traffic) show low TPM, and TVM level.

In Fig. 13, we performed simulation to evaluate the effect of node density on TSR with 7.5% of RBT. TSR is largely increased (avg. 97% for 21 Nodes/Trans.) as density increases. It is due to increased routing diversity and consequent high TPM level as explained in Fig. 10.



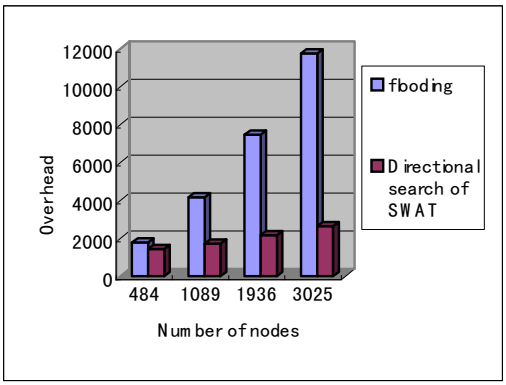
[Figure 12] Perfect traceback success rate with varying background traffic



[Figure 13] Traceback success rate with varying node density

■ *Communication overhead analysis*

We compared communication overhead (the number of transmitted/received packets) to trace back an attacker in Fig.14. A victim is located at the center of network and an attacker is located at random position (17 hops away) on the edge of network. In flooding, query message with attack signature is flooded to the entire network. Consequently, communication overhead shows exponential growth as network size increases. Our scheme shows very low communication overhead (22% in case network size is 3025 nodes) since it deploys directional search and query suppression to reduce communication overhead. Note that the energy saving becomes significant especially when network size increases.



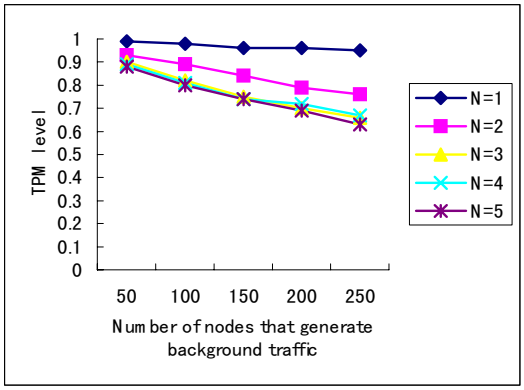
[Figure 14] Communication overhead comparison between directional search of SWAT and flooding

8.2 DDoS Attacker Traceback

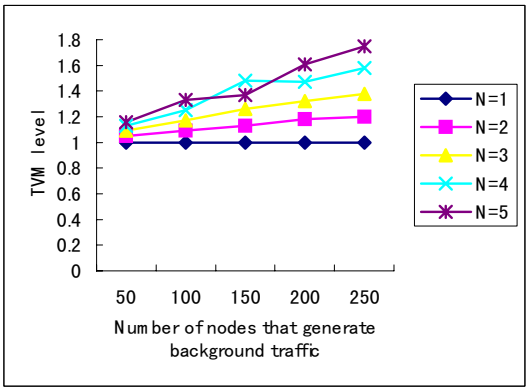
In the DDoS attack simulation, attackers are located at random positions with the average distance of 10 hops between victim and each distributed attacker. Total attack traffic observed at the victim is same as DoS case (u[200,240] pps). However, the attack traffic comes from distributed attackers and merged at the victim.

■ *Effect of merging traffic*

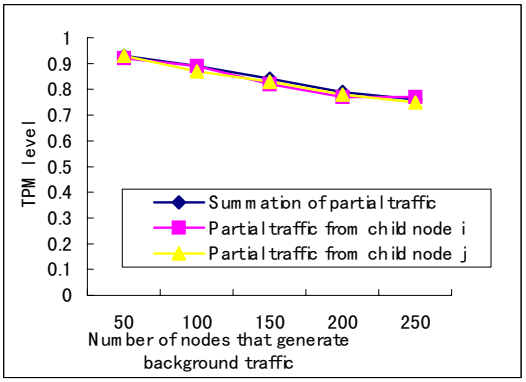
In DDoS attacker traceback, attack traffic is generated by distributed nodes. Partial attack traffic is merged at different part of network depending on the location of distributed attackers. As the partial attack traffic gets closer to the victim, partial traffic is more actively merged with other partial attack traffic. We first evaluate how the merging traffic affects TPM and TVM level test. As shown in the Fig. 15, TPM level (between the summation of partial traffic and merged traffic) decreases largely as N (the number of partial attack traffic) increases up to N=3. However, the TPM level does not show drastic decrease after N>3. TVM level shows, in Fig. 16, linear increase as background traffic and N are increased. It is due to the increased volume of background traffic included in each partial attack traffic. Fig. 17 shows TPM level between (1) merged traffic and summation of partial traffic, and (2) merged traffic and each partial traffic at child node i and j. Child node represents each node from which partial traffic comes. As shown, partial attack traffic also shows high TPM level. However, TVM level is separated in Fig. 18 clearly. TVM level between summation of partial attack traffic (traffic from child node i+traffic from child node j) and merged traffic is around 1. On the other hand, TVM level between partial traffic (either from child node i or j) and merge traffic is around 0.5, which implies that the partial attack traffic is merged with other partial attack traffic. It allows us to effectively separate the two clusters (TVM ≈1, and TVM≈0.5) and track down distributed attackers.



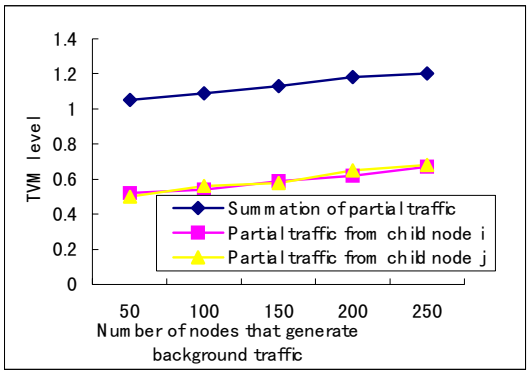
[Figure 15] TPM level comparison with varying number of partial attack traffic



[Figure 16] TVM level comparison with varying number of partial attack traffic



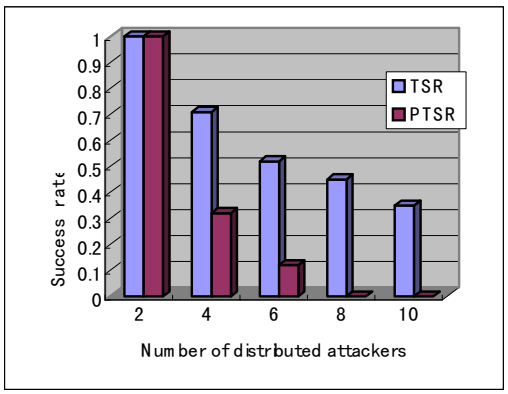
[Figure 17] TPM level comparison with partial attack traffic



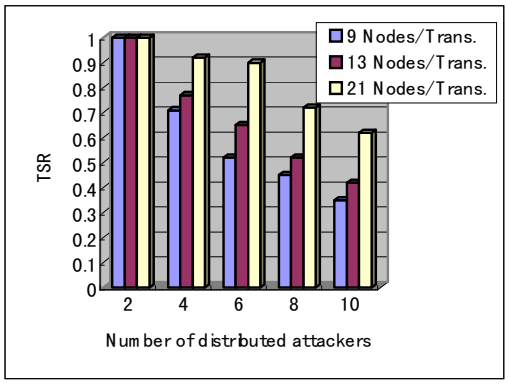
[Figure18] TVM level comparison with partial attack traffic

■ *Traceback success rate*

Fig. 19 shows traceback success rate (TSR and PTSR) with varying number of attackers. We set RBT as 2.5% and the number of node that generate background traffic as 100 nodes. As the number of attackers increases, the success rate (both TSR and PTSR) is decreased. It is because as the number of attackers increases, the abnormality of attack traffic (increased packet count) decreases, meaning that it is hard to differentiate attack traffic from background traffic. Fig. 20 shows traceback success rate with varying node density. Similar to DoS case, we can see higher traceback success rate with high node density across varying number of nodes that generate background traffic due to increased routing diversity.



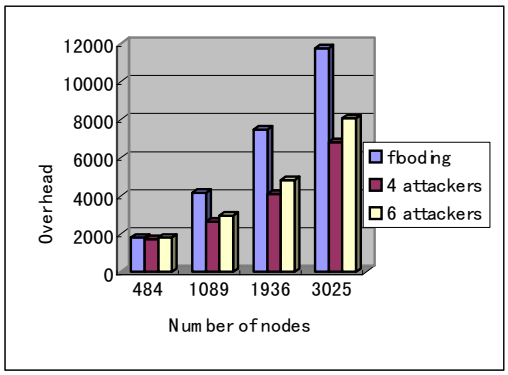
[Figure 19] TSR and PTSR with varying number of attackers



[Figure 20] Traceback success rate with varying node density

■ *Communication overhead*

Similar to DoS case, SWAT incurs low communication overhead in DDoS traceback. In the simulation, a victim is located at the center of network and attackers are located at random positions on the edge of network (average 10 hops away). As the number of attackers increase, communication overhead to search distributed attackers is also increased. However, compared with flooding type query, our scheme incurs very low communication overhead as shown in Fig. 21. The improvement (44% reduction in 4 attacker case) becomes significant as the network size increases.



[Figure 21] Communication overhead comparison

9 Discussion

In this section, we discuss some of issues related to our proposal.

■ Identification of real attacker

Using SWAT, we can detect the next hop node to the origin of attack traffic. However, SWAT is not able to detect the real identity of attackers if attackers disguise both its IP address and MAC address. In addition, the node(s) that generates attack packets might be compromised innocent nodes. However, it is still important to trace back the closest point to the

attack origin to take appropriate actions (e.g., filtering, isolation, etc).

- Attacker mobility

SWAT has tolerance against the presence of partial intermediate node mobility since our vicinity region of contact is larger than single hop and contact gathers attack signature from multiple nodes within vicinity, which observes attack traffic. So, even if some relay nodes move out or fail due to battery outage, we can leverage information from other good nodes that remain in the vicinity region. However, in case attacker is moving very fast to disguise its location, it is hard to know even next hop neighbor of attacker. Fast moving DoS/DDoS attacker that disguise its identity is considered one of the hardest problems in traceback in MANETs.

- DDoS without attack traffic abnormality

As we confirmed in the simulation, if attackers are orchestrating attack so that attack traffic from distributed attackers does not show abnormal traffic increase, our scheme fails in identifying attackers. However, as partial attack traffic is actively merged as it comes closer to the victim, traffic abnormality is observed. Consequently, SWAT can trace back up to upstream points where traffic abnormality is observed. It is useful to know the closest point to the attackers because the efficacy of measures such as packet filtering improve as they are applied further from the victim and closer to the source.

- Trust on contact and message integrity

In our scheme, a contact is selected independently by each node. Each node selects its contacts randomly to prevent divulgence of contact information and consequent compromise. Message between contacts - victim, contacts - contacts, and contact - vicinity nodes are relayed through intermediate nodes. Bad intermediate node can tamper the integrity of message. To provide message integrity, we can leverage secure protocol such as [12].

10 Conclusions

We proposed efficient DoS/DDoS attacker traceback scheme based on small world model. Small world model reduces the degree of separation between attacker and victim, and provides a useful way for attacker traceback. We used TPM, TVM and majority-based traceback to tolerate address spoofing, intermediate node mobility and collusion. To reduce communication overhead, we proposed (Multi-) directional search /in network processing/query suppression. Through simulation, we confirmed that our scheme successfully (97% success rate in DoS attacker traceback) traces back attacker under reasonable background traffic. Communication overhead reduction (78% in DoS and 44% in DDoS compared with flooding) is significant in SWAT especially when network size is large.

11 References

[1] CERT Advisory CA-97.28, IP Denial-of-Service Attacks, May 26, 1996.

[2] CERT Advisory CA-96.21, TCP SYN Flooding and IP Spoofing Attacks, Sept. 24, 1996

[3] CERT Advisory CA-98.01, Smurf IP Denial-of-Service Attacks, Jan. 5, 1998.

[4] CERT Advisory CA-96.01, UDP Port Denial-of-Service Attack, Feb. 8, 1996.

[5] A. Belenky and Nirwan Ansari, "On IP Traceback", IEEE Communication Magazine, July 2003

[6]S.M.Bellovin, "ICMP Traceback Messages," IETF draft 2000; http://www.research.att.com/smb/papers/draft-bellovin-itrace-00.txt.

[7] H. Burch, et al, "Tracing Anonymous Packets to Their Approximate Source", Proc. 2000 USENIX LISA Conf., pp.319-327, Dec. 2000

[8] R.K.C.Chang, "Defending against Flooding-Based Distributed Denial-of-Service Attacks: A Tutorial," IEEE Communication Magazine, Oct. 2002

[9] A.Helmy, "Small Worlds in Wireless Networks", IEEE Communication letters, Vol.7, No.10, Oct. 2003

[10] A.Helmy, "Contact-extended Zone-based Routing for Transactions in Ad Hoc Networks", IEEE Transactions on Vehicular Technology, July 2003

[11] G.Mansfield, et al., "Towards trapping wily intruders in the large", Computer Networks, Vol.34, pp.650-670, 2000

[12] S.Milgram, "The small world problem", Psychology Today 1, 61 (1967)

[13] A. Perrig, et al., "SPINS: Security Protocols for Sensor Networks", ACM MOBICOM, 2001

[14] Alex C. Snoeren, et al, "Single-Packet IP Traceback", IEEE/ACM Trans. Net., Dec. 2002

[15] Stefan Savage, et al., "Network Support for IP Traceback", IEEE/ACM Trans. On Nets. June 2001

[16] D.J.Watts. In Small Worlds, The dynamics of networks between order and randomness. Princeton University Press, 1999

[17] A.D. Wu et al., "On Design and Evaluation of Intention-driven ICMP Traceback," Proc. 10th Int'l. Conf. Comp. Commun. And Nets., 2001