# Attacker Traceback and Countermeasure with Cross-layer Monitoring in Wireless Multi-hop Networks

**Yongjin Kim, Ahmed Helmy**
Electrical Engineering Dept. – Systems
University of Southern California, California, U.S.A.
{yongjkim, helmy}@usc.edu

***Abstract***
DoS/DDoS attack can cause serious problems in wireless multi-hop networks due to limited network/host resources. Attacker traceback is a promising solution to track down DoS/DDoS attacker and take countermeasure near attack origin. Existing attacker traceback schemes developed for the Internet cannot be directly applied to wireless multi-hop networks due to autonomous nature of wireless multi-hop networks. To efficiently track down DoS/DDoS attacker, we propose cross-layer (MAC, and network layer) monitoring-based traceback scheme. We compare the advantages of using cross-layer information over using only network/MAC layer information. In addition, we propose a novel traceback-assisted countermeasure scheme that is taken at the closest nodes to the attacker. We show that our scheme successfully (98% in DDoS attacker traceback) tracks down attacker under diverse network environment (e.g., high background traffic, DDoS attack, and partial node compromise) with low communication overhead.

## 1. INTRODUCTION

DoS/DDoS attack can cause serious problem in wireless multi-hop networks (e.g., Ad-hoc network, sensor networks, etc) since (1) it is easy to perform using popular tools and (2) wireless multi-hop networks are severely limited in network (e.g., bandwidth) and host resources (e.g., battery, memory, etc). The different types of denial of service attacks can be broadly classified into software exploits and flooding attacks. In software exploits (e.g., Land attack), the attacker sends a few packets or even single packet to exercise specific software bugs within the target's OS or application, disabling or harming the victim. On the other hand, in flooding attacks [2], one or more attackers send incessant streams of packets aimed at overwhelming link bandwidth or computing resources at the victim. We mainly focus on flooding-type DoS/DDoS attack since it cannot be fixed with software debugging and propose a novel protocol for attacker traceback and its countermeasure. In flooding-type DoS/DDoS attack, an attacker transmits a large number of packets towards victim with spoofed source address. For instance, in SYN Flood, at least 200-500 pps (packet per second) of SYN packets are transmitted to a single victim. UDP Echo-Chargen and Smurf also attacks victim using a large amount of packets with spoofed address. In general, we can say that the following are some characteristics of flooding-type DoS/DDoS attacks: (I) Traffic volume is abnormally increased during attack period. (II) Attackers routinely disguise their location using incorrect/spoofed addresses. (III) Such attacks may persist for tens of minutes and in some case for several days. The goal of attacker traceback [1] is to identify the machines that directly generate attack traffic and the network path this traffic subsequently follows. The first efficient attacker traceback scheme geared toward wireless multi-hop networks is proposed in SWAT [4], which pays attention to network layer abnormality (packet count increase). SWAT finds attack path and attacker by tracking intermediate nodes that observe similar abnormality as victim. However, the problem of SWAT is that traceback success rate drastically goes down when there is high background traffic that leads to abnormality mismatching. In addition, under DDoS attack, reduced abnormality is observed near the edges of branch attack route. We show that by using cross-layer information (MAC and network layer) and minimal packet/frame content information, we can drastically increase traceback success rate even under low abnormality (i.e., DDoS attack) and high background traffic. We also propose a novel traceback-assisted countermeasure. That is, we take efficient countermeasure using cross-layer information at the nearest point to the attack origin minimizing harm to the legitimate traffic.

## 2. ATTACKER TRACEBACK WITH CROSS-LAYER MONIOTORING

Once attack is detected by the intrusion detection system of a victim, attack signature is characterized by the victim (refer to 2.1). Then, query with the attack signature is sent to neighbor nodes and contact (refer to 2.3) to find intermediate nodes that observe similar abnormality as attack signature. The searching process is continued recursively towards attack origin. Detailed procedure is explained in the following.

### 2.1 Abnormality Characterization

In our scheme, attack signature is defined as time series data of incoming MAC-layer frame count, $\xi=(n_1,...n_k)$ in [1,k] time slots, which shows abnormal increase. We can use FDM (Fractional Deviation form the Mean) or other statistical technique to capture the abnormality. However, the attack signature can include background traffic, which negatively affects traceback performance. For accurate attack signature (i.e., abnormality) characterization of attack traffic, we need to reduce/remove background traffic (i.e., noise factor) included in the attack signature. We take advantages of both MAC layer and network layer information to achieve the goal. We can reduce forward noise by network layer information (i.e., destination address) and backward noise by using MAC layer information as shown in Fig.1.

By removing forward and backward noise that does not contribute attack traffic, we can drastically increase matching accuracy between abnormality observed at the relay nodes and victim node. Attack signature table is maintained at each node with abnormality $\xi(D\_addr, P\_addr)$, where *D_addr* is destination address and *P_addr* is previous-hop MAC address.

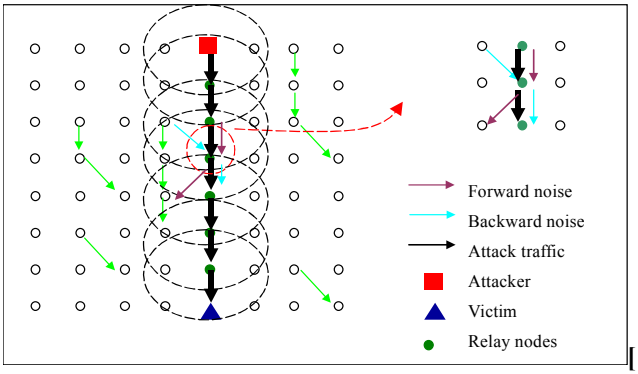



**Figure 1] Illustration of forward/backward noise reduction**

## 2.2 Abnormality Matching using K-S fitness test

We are interested in using the Kolmogorov-Smirnov (KS) statistic $D_n$ to test the hypothesis that the two abnormalities, $F_n(x)$, $F_0(x)$ are matching. $F_0(x)$ corresponds to attack signature characterized by victim, which is included in query message, and $F_n(x)$ is the candidate abnormality observed by intermediate nodes.

$$D_n = \sup_x [\ |F_n(x) - F_0(x)|\ ] \quad \text{(Eq.1)}$$

$$H_0 : F_n(x) = F_0(x)$$

$$H_a : F_n(x) \neq F_0(x)$$

We accept $H_0$ if the distribution function $F_n(x)$ is sufficiently close to $F_0(x)$, that is, if the value of $D_n$ is sufficiently small. The hypothesis $H_0$ is rejected if the observed value of $D_n$ is greater than the selected critical value that depends on the desired significance level and sample size. When the $H_0$ is accepted (sufficiently similar), we can infer that the abnormality is matching, meaning that the attack traffic is traversed the nodes that observe similar abnormality. The process is recursively continued towards attack origin.

## 2.3 Directional/Multi-directional Search

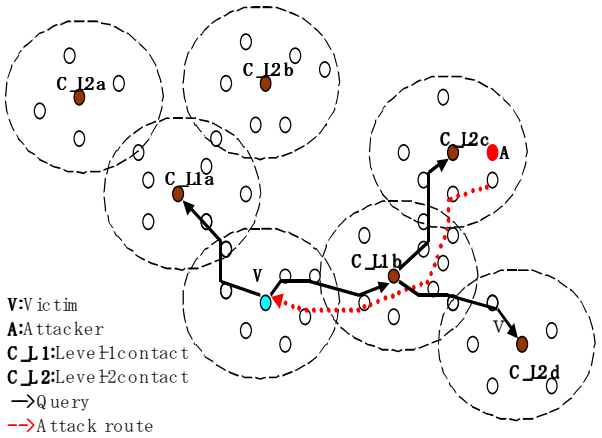
For efficient/robust search, we extend contact architecture [3]. In contact architecture, each node sends query to its vicinity nodes and contact that is outside vicinity to find matching candidate attack signature (Fig.2). Each contact gathers abnormality information from its vicinity node and calculates the following attack signature energy.

$$LE(t) = \frac{\sum_{i=1}^{P} E_i^u(t)}{P} \bullet \frac{1}{\mu_{1/2}} \quad \text{(Eq.2)}$$

Where, $E_i^u(t)$ is the inverse of $D_n$ of node $i$ and contact $u$. at time $t$. $P$ is the total number of node that observes abnormality at contact region $u$. $\mu_{1/2}$ is the median value of distance (hop count) between contact and the nodes that observes abnormality. The reason that we take median value instead of average is to prevent false distance report from malicious or compromised node. By finding the largest *LE(t),* we can infer the region that attack traffic traverses. In addition, *LE(t)* should satisfy the following condition.

$$\alpha = \frac{n}{N} > \delta \quad \text{(Eq.3)}$$

α is majority-voting factor (*N*: total number of vicinity nodes of the contact, *n*: number of nodes that observe abnormality). When, α is extremely low, we can infer that there is high chance of false reporting and consequently attack signature energy becomes small.




**[Figure 2] Victim (*V*) sends queries with attack 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.**

Spatial region around attacker shows high attack signature energy value. In addition, spatial region around attack path of attack traffic also shows high energy value. The energy is affected by percentage of nodes observing signature energy, median distance from the target, and average individual signature energy in a spatial region. Intuitively, we can infer that attacker is residing or attack traffic is traversing the region where high attack signature energy is observed.

For illustration with Fig.2, 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 signature described by network/MAC layer abnormality information. It then sends a query to the nodes within its vicinity and level-*1* contacts. Contacts are the nodes that relay query to its vicinity node as shown in Fig.2. Note that vicinity between contacts is minimized. The query contains sequence number (*SN*) and attack 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 KS test is passed and high *LE(t)* is observed, meaning that there exist vicinity nodes of contacts that observe similar attack signature, 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. 2 (transmission arrows to vicinity nodes by each contact are omitted in the figure). (4) Next, only the contact, *CL_1b*, that observes high *LE(t)* 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. We can use multi-directional search for DDoS attacker traceback.

## 3. TRACEBACK-ASSISTED COUNTERMEASURE

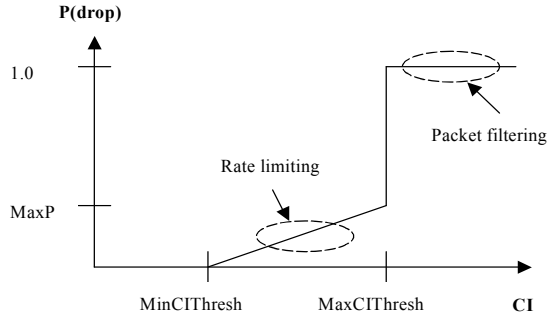
Existing countermeasure (i.e., packet filtering, rate limiting) against DoS/DDoS attack has the following drawbacks: (1) it is taken at the nodes where attack is detected. For instance, it is taken at the ingress point of victim. However, it is inefficient since attack traffic exhausts valuable network/host resources of intermediate nodes. (2) Packet filtering is challenging since it is hard to distinguish bad and good traffic. (3) It is hard to know how much rate should be limited to reduce negative traffic against legitimate traffic and increase rate-limiting efficiency against attack traffic. By using our cross-layer information (destination address, previous MAC address), we can detect attack traffic with high accuracy. In addition, we propose hybrid scheme between packet filtering and rate limiting. That is, when abnormality matching is high, we apply packet filtering. On the other hand, when abnormality matching is medium level, we apply rate limiting. To determine optimal rate limiting level under medium matching level, we use Confidence Index (CI). CI is normalized value of inverse of matching level as follows.

$$CI = \frac{1}{D_n} \quad \text{(Eq.4)}$$

Rate limiting level (P) is determined with the following equation. (refer to Fig.3)

$$P = MaxP \bullet \frac{CI - MinCIThresh}{MaxCIThresh - MinCIThresh} \quad \text{(Eq.5)}$$

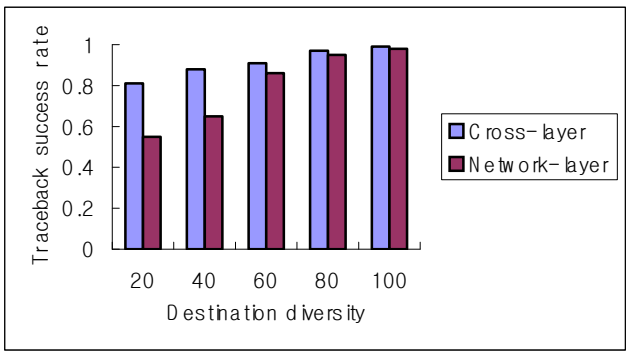
As shown in the Fig.3, when CI is very high it reduces to packet filtering since it implies that there is no background traffic. On the other hand, when CI is medium, it becomes rate limiting based on CI level to reduce negative impact on legitimate traffic.



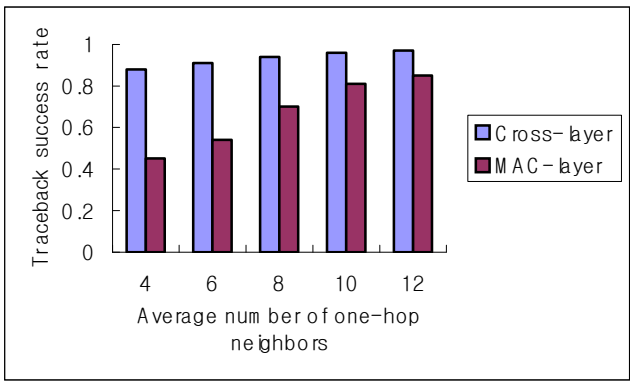

**[Figure 3] Hybrid Countermeasure**

## 4. SIMULATION RESULTS

We performed simulation with *ns-2* on DDoS attacker traceback. Fig.4 shows traceback success rate with various destination diversity (i.e., number of destination). We set number of one-hop neighbor as 6 and percentage of nodes that generate background traffic as 50%. When destination diversity is low (<20), traceback success rate is low when we only use network layer information since much traffic goes to the same destination and abnormality matching level is decreased. However, our scheme shows high success rate (>80%) across different diversity level. It is because MAC layer information complement Network layer information, which reduces backward noise traffic. Fig. 4 shows success rate with various number of one-hop neighbor. Our scheme shows greater improvement compared with the scheme using only MAC-layer information. It is because Network layer information can reduce more noise traffic (i.e., forward noise).




**[Figure 4] Traceback success rate comparison between cross layer-based scheme and network layer-based scheme**




**[Figure 5] Traceback success rate comparison between cross layer-based scheme and MAC layer-based scheme**

## 5. CONCLUSION FUTURE WORKS

We proposed attacker traceback scheme with cross-layer (MAC and network layer) monitoring. Noise traffic can be drastically reduced by using cross-layer information, which leads to high traceback success rate (Avg. 98% success rate under diverse environment). We also proposed a novel countermeasure assisted by traceback procedure. We show that abnormality matching level can be effectively used to reduce negative impact on legitimate traffic and increase attack limiting/filtering efficiency. We will perform extensive analysis with a rich set of network environment to evaluate the efficiency of the proposed scheme and to find optimal parameters (e.g., optimal value of MaxP for efficient countermeasure, etc.)

**[REFERENCES]**

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

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

[3] A.Helmy, et al, "A Contact-based Architecture for Resource Discovery in Ad Hoc Networks", ACM Baltzer MONET Journal, 2004

[4] Yongjin Kim, A.Helmy, "SWAT: Small World-based Attacker Traceback in Ad-hoc Networks", IEEE/ACM Mobiquitous, July 2005