*A Report On*

# IP Traceback: Using Packet Marking & Logging Schemes

*Submitted in requirement for the course*

**Advanced Computer Networks (CSN-503)**

*of Bachelor of Technology in Computer Science and Engineering*

Submitted By:

**Akash Gupta** (Enroll no.15114004)

**Gautam Choudhary** (Enroll no.15114027)

**Shivam Jindal** (Enroll no.15118079)

Submitted To:

**Dr. P. Sateesh Kumar**

Assistant Professor, IIT Roorkee

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

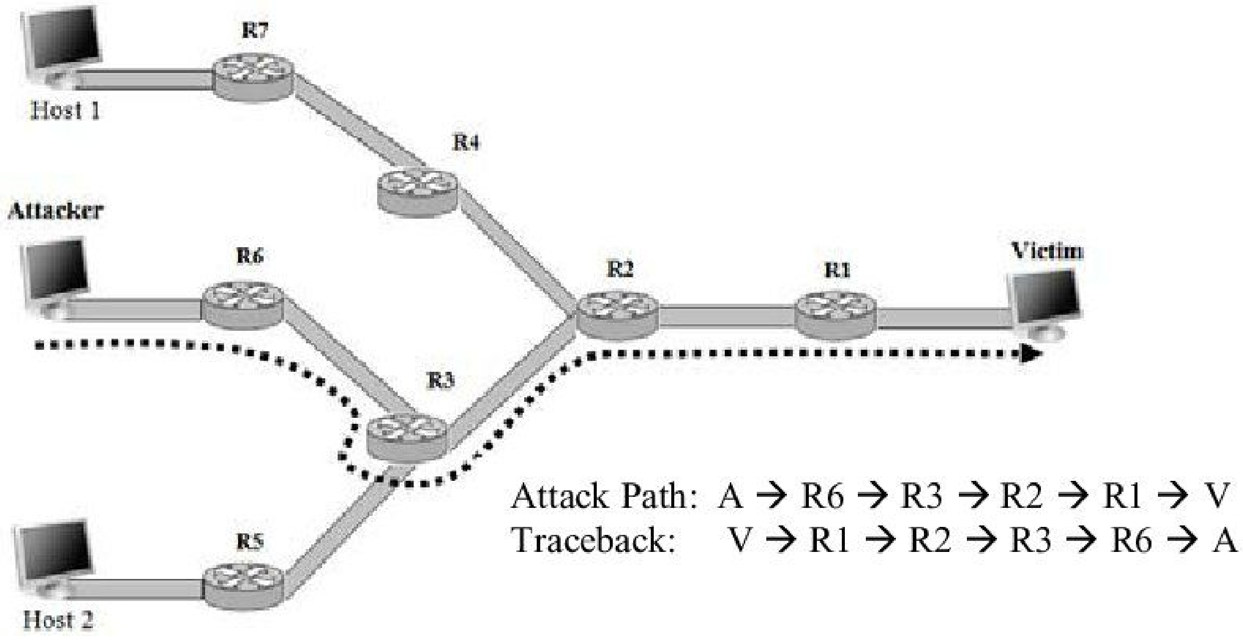
**INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE**

ROORKEE- 247667 (INDIA)

# Table of Contents

# Introduction

The purpose of IP tracing is to trace the route of an IP packet to its origin. The most important use of IP tracking is to deal with certain denial of service (DoS) attacks, where attackers falsify the source IP address. Identifying the sources of the attack packages is an important step for the attackers to be responsible. In addition, discovering the network path that the attack traffic follows can improve the effectiveness of defense measures, such as packet filtering, since they can be applied further away from the victim and closer to the source.


R7
Host 1
R4
Attacker
R6
R2
R1
Victim
R3
R5
Host 2
Attack Path: A → R6 → R3 → R2 → R1 → V
Traceback: V → R1 → R2 → R3 → R6 → A

Two main types of IP tracking techniques have been proposed in two orthogonal dimensions: packet marking and packet registration. In packet marking, the router marks the IP packets forwarded with their identification information. Due to the limited space in the packet header, routers decide to probabilistically mark packets so that each marked packet carries only partial route information. The route of the network can be reconstructed by combining a modest number of packages containing the mark. This approach is known as probabilistic packet marking (PPM). The PPM approach incurs little overhead on the routers. But it can only track traffic consisting of several packets due to its probabilistic nature.

In the packet register, the IP packet is registered in each router through which it passes. Historically, it was thought that packet registration was impractical due to the huge storage space for packet records. The hash-based IP tracking approach records package digest in a space-efficient data structure, flowering filter, to significantly reduce storage overload. The routers are consulted to reconstruct the route of the network. This approach can track a single IP packet. However, the requirements for the storage of summary tables and the access time to record packets according to their arrival are prohibitive for routers with high-speed links.

This document proposes to develop a hybrid IP tracking approach based on both packet marking and packet registration. The motivation is to develop an IP tracking approach that has advantages in both package marking and packet registration. Our goal is to maintain the ability to track a single packet as in the hash-based IP tracking approach, but at the same time reduce the overhead of storage and access time in routers with the help of packet dialing.

## Classification of Traceback Methods

Trackback methods are basically classified into two types; preventive and reactive.

- Preventive methods are used as preventive measures against DoS attacks. A wide range of solutions are as well as have been proposed, however, this problem still remains as an open one.
- Reactive methods solutions try to identify the source of the attacks. This is a very important move because attackers generally spoof their addresses, thus these techniques are necessary to trace back to the source of the attack.

Some of the Reactive Methods are discussed below:

### Link Testing

This test starts from the router closest to the victim and interactively tests their uplinks until they verify that one is used to transport the attacker's traffic. Therefore, this procedure is perennially recursively in the ascending router until the source is reached.

## Logging

Registration is suggested to register packets on the key routers and, therefore, use data extraction techniques to see the path that the packets went through. It has the valuable property that it will track an attack long after the attack is over. This system has drawbacks, and probably the enormous needs of resources and the integration of interprovider information on a large scale are difficult.

## ICMP Traceback

The Internet Control Message Protocol (ICMP) wants to track the full path of attacks. Typically, this scheme is for each router to present an ICMP trace message or scope addressed to the identical destination due to the elite package. The trace message consists of data of previous and subsequent jumps and a timestamp.

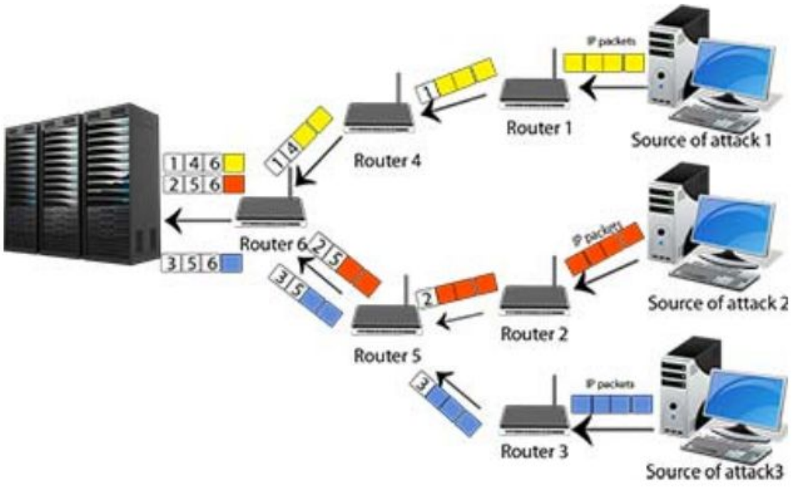
## Packet Marking

During this scheme, each router within the count to forward a packet additionally inserts a mark into the packet. This mark could be a distinctive orthodox symbol to the current specific router. As a result, the victim will verify all shift jumps for each package by looking at the inserted marks. There are two variants of this marking scheme. First, the deterministic packet marking scheme (DPM) in which each router marks all the packets that pass through it with its unique identifier. Second, the packet probability (PPM) and DoS attacks can be avoided if the forged IP address goes back to its origin, which allows the distribution to the wrong part to be penalized or the hosts and compromised domains to be isolated. reminder of the network.

# RELATED WORK

Depending on the vulnerability exploited, DoS attacks can be divided into brute force and semantic attacks. Brute force attacks work by flooding some limited resources with large amounts of traffic, which prevents legitimate users from accessing that resource. Semantic attacks exploit some specific feature or implementation error of operating systems or routers to disable services with one or more packages. An IP tracking approach that can track an individual packet is a must for defense against semantic DoS attacks.

## Marking Method



The basic idea of the IP tracking approach based on packet marking is that the router marks the packets with their identification information as they pass through that router. Since the marking space in the packet header is too small to record the entire route, routers mark the packets with some probability, so that each marked packet carries the information of a node in the route. In addition, depending on the length of the identification of the router and the implementation of the dialing procedure, the router can only write part of its identification information in the marking space. While each marked packet represents only a small part of the route it has traversed, the complete route of the network can be reconstructed by combining a modest number of such packets. This type of approach is known as probabilistic packet marking (PPM). The PPM approach does not generate overhead storage costs on the routers and the dialing procedure (a checksum and write upgrade) can be executed easily and efficiently on the current routers. But due to its probabilistic nature, it can only track traffic consisting of a large volume of packages.

### Logging Method

The main idea of the IP tracking approach based on packet registration is to register the packets on each router through which they pass. To reduce the storage space required, the registration must be done intelligently. Hash-based IP tracking stores packet digest, rather than packets themselves, in a space-efficient data structure, flowering filter. In this way, the storage overload is significantly reduced. The summary table is deleted before saturation, which avoids unacceptable false positive rates. The summary tables are archived for one minute for a possible trace operation. Each summary table is annotated with the time interval that the table covers, and the hash functions used to calculate the packet summaries during that interval. During the trace process, the routers are queried in the manner of reverse route flood (RPF) and the summary tables in the routers consulted are examined to reconstruct the network route. This approach could track a single IP packet and, therefore, is considered to be more powerful compared to the PPM approach. While the hash-based approach requires approximately 0.5% of the total capacity of the link in the storage of summary tables, the storage requirement is prohibitive in routers with high-speed links.

Two approaches have been recently proposed for addressing the deficiencies of hash-based IP traceback:

- T. Lee et al. proposed to digest packet aggregation units (flow or source-destination set) instead of individual packets.
- Li et al. proposed to probabilistically select a small percentage of packets and record the digests of the selected packets. This method reduces both storage and access time overheads at routers.

# SOLUTION

Use a HYBRID approach!

The hybrid IP tracking approach has a similar architecture with the hash-based approach. Trackable routers audit traffic and a tracking server (or multiple servers in hierarchy) that has the attack graph of network topology information constructs when querying routers. The differences are found in the operations of the router in the packets and in the construction procedure of the attack graphic.

## 1. Router Operation

Each router enabled for tracking could commit both packet dialing and packet logging operations. The marking operation in a packet is marking the packet with identification information of the router. The registration operation in a package is to record the summary of the package and the mark (identification of the router) that the package carries. Each router is assigned an ID number of 15 bits in length. In the hybrid IP tracking approach, the router ID number is used to differentiate neighboring routers from a router, rather than all routers within an ISP network. Therefore, the same identification number can be assigned to 2 routers as long as they are more than 2 jumps away. The mark is stamped on the packets by overloading the 16-bit identification field in the IP header.

In the hybrid approach, routers register router IDs transported in packets in addition to packet summaries. The hybrid approach calculates and stores the package summaries using the same method as the hash-based approach. The storage of the identification numbers of the router is implemented in an efficient manner in space. Each router maintains a different summary table for each of its neighboring routers. When a router decides to perform the registration operation on a packet, it examines the ID number of the router carrying the packet to know which neighbor router the packet comes from, then stores the summary of the packet in the summary table corresponding to that neighbor. The summary table is paged before being saturated. Each summary table is annotated with the time interval covered by the table, the hash functions used to calculate the packet summaries during that interval, and the ID number of the neighboring router. Each summary table stores the summaries of the packets that are sent by the same router and carry the same router ID number. The ID number of the router transported by packets is recorded as the annotation of the summary table. In this way, the storage overhead for the router identification numbers is negligible. For each packet that arrives, the current router first examines the router ID number marked in the packet header to verify if it is valid. The identification number of the router that carries the packet p is valid on a router r if it is equal to the identification number of a neighbor router of the router r. That is, the packet p was forwarded from a router neighboring the router r. If the router ID number is valid, depending on the registration indicator bit in the packet, the router may choose to confirm (1) only the dialing operation, or (2) the dialing and registration operations. If the upstream router registered the packet (the log flag is 1), the current router chooses to dial only the packet; If the upstream router did not register the packet (the log flag is 0), the current router chooses to dial and register the packet. If the router ID number is invalid, that means that the packet that arrived came directly from the sender's host or from an attacker who sends packages with a counterfeit brand. In this case, the router chooses to only commit the dialing operation.

The effectiveness of IP tracking increases enormously with the widespread deployment of routers enabled for network tracing. However, it is likely that the hybrid IP tracking approach does not require all routers to be enabled for tracking. All routers enabled for tracking form a superimposed network. If the tracking server has knowledge of the topology of that overlay network and each trace-enabled router knows its adjacent trace-enabled routers, the hybrid IP tracking approach still works.

## 2. Attack Graph Construction

If a router performs a registration operation on an attack packet, the examination of the summary tables on that router will not only confirm that the router is on the attack route, but also find its upstream router on the attack route, since each summary table is annotated with an ascending router ID number. Given an attack packet and a victim, the tracking server could infer the last hop router and if the last hop router performed a registration operation based on the registration flag bit that carries the attack packet.

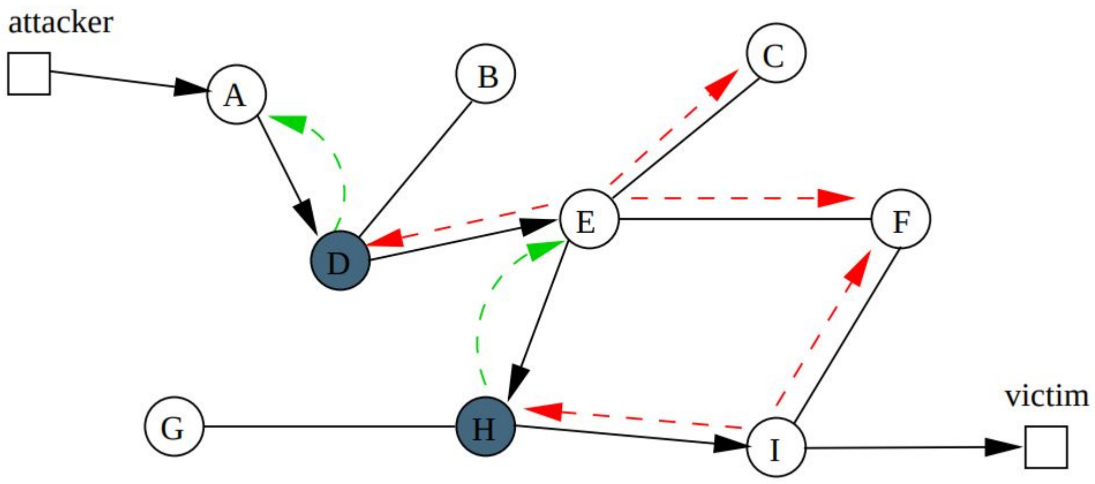
1) If the tracking server infers a registered router the attack packet, the examination of the summary tables on that router would identify its ascending router in the attack route.

2) If the tracking server inferred that a router did not register but dialed the attack packet, query the neighboring routers of that router in an RPF fashion and examine the summary tables on these neighboring routers would identify the upstream router. The attack graph can be constructed using those two methods alternately.

## 3. Transformation & Compatibility

The lethal drawback of any IP tracking based on packet marking is backward compatibility. Because the IP identification field designated for fragmentation is used to overload marks, packet marking collides with fragmented IP traffic. In addition, IP packets can undergo a valid transformation while traversing the network.

With the improvements, the hybrid IP tracking approach is able to track packets that underwent transformation and avoid the problem of backward compatibility. The operations of the router in the packets are improved in the following way. For each package that arrives:



1) If the packet is transformed into the current router, confirm the marking and registration operations in the packet and record the transformation information in the transformation search table. Given a package, consult the transformation query table to know if the package was transformed and the original package can be reconstructed. The implementation of the transformation search table is described in.

2) If the package is a fragmented package, calculate and store the package summary in a particular summary table that is only for fragmented packages and is managed in the same way as the hash-based approach.

3) Otherwise, follow the algorithm described in the Operation section of the router.

The construction of the attack graphic is also improved accordingly. When the crawl server examines the summary tables on a router, it also looks up the transformation lookup table on that router and reconstructs the attack packet to its original form, if possible.

1) If the attack packet provided by the victim is not a fragmented packet, the tracking server uses the procedure to build an attack graph that is similar to that presented in the Construction of the attack graph. The only difference is that it is no longer true that routers in an attack route register an attack pack alternately. It is possible for two adjacent routers, for example, myn, to register the same packet p because the ascending router m performs the registration operation in the packet p, and p undergoes a transformation in the downstream router n. During the trace process, when you move to the ascending router m

from the router n that registered and transformed the packet p, the tracking server can not assume that the router m did not register the packet p. The tracking server needs to examine the summary tables in m to find out if m marked p alone or both marked and registered p, then take appropriate action accordingly.

2) If the attack packet is a fragmented packet, starting from the last hop router, the tracking server queries the routers in the RPF fashion and examines the summary tables that record the summaries of fragmented packets to build the attack graph .

## FUTURE WORK

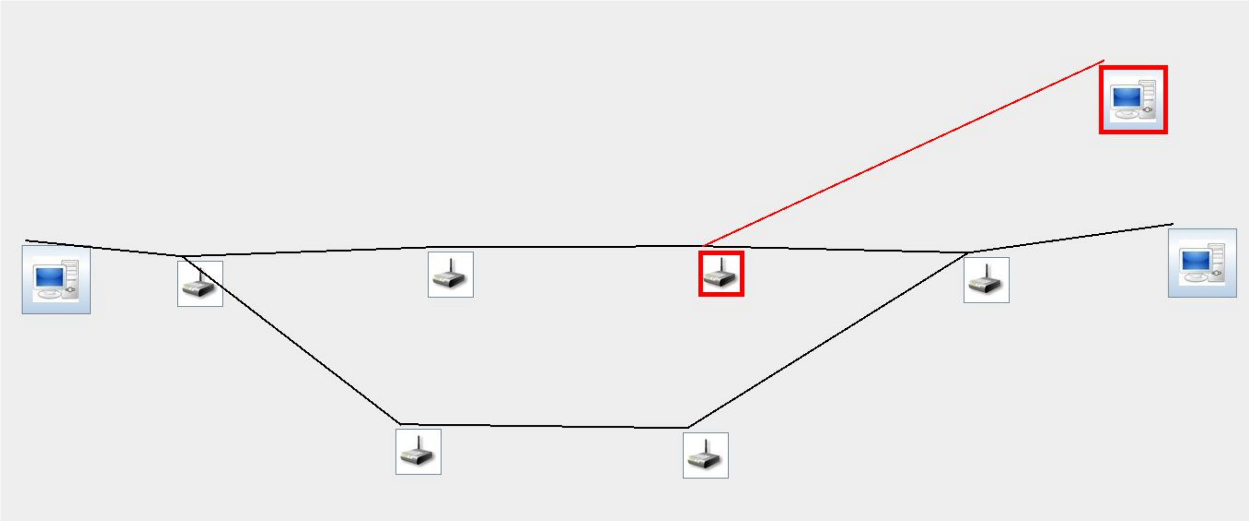
The focus in the tracking scheme has gone from the rapid tracking of the victim to the rapid detection of the attack before the victim is affected, since most of the DDoS attacks take place from the steps (compromised intermediate hosts).

Compared to the hash-based IP tracking approach, this approach reduces storage overload by approximately half and improves access time by a factor of the number of neighboring routers. Some of the proposed schemes can not be used for post-attack analysis, they are resource-intensive, they can overload the network, and they may not be effective against DDoS attacks.

Tracking schemes using the Watermarking technique, information metrics such as entropy, divergence and distance metrics are gaining momentum and a brief study of these techniques will be provided in the near future.

## CONCLUSION

We implemented a new IP tracking approach that is based on both packet marking and packet logging. This approach has the ability to trace a single packet to its origin. Compared to the hash-based IP tracking approach, it reduces storage overload by approximately half and improves access time by a factor of the number of neighboring routers. The implementation shows the trace process by calculating the summary table and the routing table, which is used even more to trace the route back to the sender.



*A snapshot of the screen from implementation*

## REFERENCES

1. Gong, Chao, and Kamil Sarac. "*IP Traceback based on Packet Marking and Logging*."
2. Balyk, Anatolii, et al. "*A survey of modern IP traceback methodologies*." Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2015 IEEE 8th International Conference on. Vol. 1. IEEE, 2015.
3. Kumar, K. Arun, and K. Sai Ashritha. "*Analysis of various IP traceback techniques-A Survey.*" International Journal of Computer Applications 77.13 (2013).
4. Murugesan, Vijayalakshmi, Mercy Shalinie, and Nithya Neethimani. "*A Brief Survey of IP Traceback Methodologies*." Acta Polytechnica Hungarica 11.9 (2014): 197-216.