

Practical Network Support for IP Traceback

Authors: Stefan Savage, David Wetherall, Anna Karlin and Tom Anderson

Published: SIGCOMM 2000

Presenter: Christian McArthur

Papers of Related Interest

- “Inferring Internet Denial-of-Service Activity,” D. Moore, G. Voelker, and S. Savage, USENIX Security 2001. (Optional paper)
- “Large-Scale IP Traceback in High-Speed Internet,” J. Sung, J. Xu, L. Li, IEEE Symposium on Security and Privacy 2004.
- “Defense Against Spoofed IP Traffic Using Hop-Count Filtering,” H. Wang, C. Jin, K. Shin, IEEE/ACM Transactions on Networking, Feb. 2007.

Denial of Service (DOS) Attacks

- End Goal of Attack: To prevent people from accessing a targeted website.
- Ways to Perform DOS attacks:
 - Overload systems or network equipment (flood target with high volume of traffic)
 - Exploits that crash operating systems, protocol stacks, or equipment (ping of death)
 - Interfere with normal communications (TCP reset)
 - Altering routing information (black holing)
 - Physical attacks against target infrastructure

DoS Attacks Statistics

- From optional paper:
 - 12,805 attacks over the course of three weeks in 2001
 - 5,000 distinct IP addresses
 - 2,000 distinct DNS domains
 - 46% of all attacks send 500 packets per second or more.
 - Cited a 2000 study that suggests commodity network equipment can handle a max of 500 TCP SYN packets per second.
 - 50% of attacks last less than 10 minutes; 2% last longer than 5 hours.
- From <http://atlas.arbor.net/summary/dos>:
 - 126 attacks in 24 hour period Monday – Tuesday
 - 142 attacks Wednesday - Thursday

DOS Attack Concerns

- Discovering who is attacking you.
 - Single source or distributed attack?
 - Using real IP address or spoofed IP address?
 - Derive the attack route for forensic needs
- Mitigating the attack.
 - Depends on how attack is being performed
 - Filtering at border router or firewall
 - Contact source network(s)
 - Planning for attack (fall back IP space, different routes)

Deriving the route of the attack

The “Traceback Problem”

- Knowing the route the flood of packets takes can help in mitigating the attack as well as providing forensic evidence for law enforcement.
- “Input Debugging” / Off-Line Traceback
 - Requires contacting administrators of each router in the attack path to discover the next hop.
 - Inefficient and requires cooperation of many admins.
 - Attack may need to occur for an extended period of time before the source can be discovered.

Deriving the route of the attack

The “Traceback Problem”

- **Controlled Flooding**
 - The victim floods specific links and analyzes the effects on the attacker's traffic.
 - Requires a good knowledge of Internet topology to know what links to flood.
 - Controlled flooding itself is a DOS attack.
- **Ingress Filtering**
 - Routers block packets from propagating if it has invalid IP addresses.
 - Most effective at edge of networks and ISPs.
 - Requires deployment across all networks to be effective.

Deriving the route of the attack

The “Traceback Problem”

- ICMP Traceback
 - A router forwarding traffic copies the contents of a sample of packets into an ICMP packet.
 - ICMP packet also contains information about routers in the area (where packet received, router making ICMP packet, where packet is going).
 - ICMP may have lower QoS priorities, thus may not always be forwarded to victim.
 - Possible for attacker to forge its own ICMP traceback packets.
- Authors' Solution
 - Augment the IP protocol to enable traceback of any stream of packets.

Node Append

- Each node appends its address to the end of the packet.
- Advantage:
 - Every packet contains the complete path traveled.
- Disadvantages:
 - May increase fragmentation due to additional information being stored.
 - Attacker could put its own invalid route information.
 - High overhead for every router to add extra information to every packet it handles.

Node Sampling

- Augment the IP header to add a “node” field.
- Router handling the packet puts its address in the node field with some probability.
- Reconstruct route based upon distribution of addresses in the node field.
- Advantages:
 - More efficient for routers.
 - If probability is high enough ($p > 0.5$) a single attacker cannot effectively hide the path.
- Disadvantages:
 - Requires high number of packets to determine the path (assuming 15 hop route; $p = 0.51$; 294,000 packets)
 - Ineffective against multiple attackers as multiple routers could be the same distance.

Edge Sampling

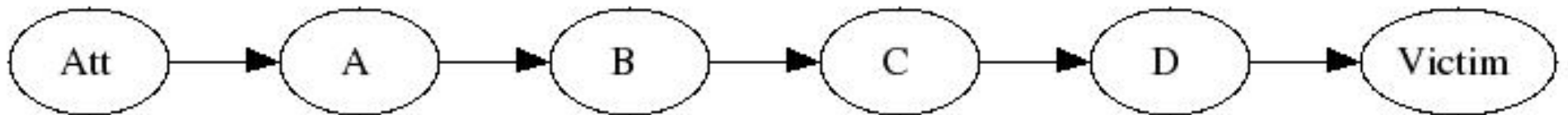
- Packets are marked with two router IPs and the distance from the first router to the victim.
- Assuming no path exceeds 25 hops, path can be determined with 108 packets.

A B 3

D 0

C D 1

B C 2



Edge Sampling

A B 3

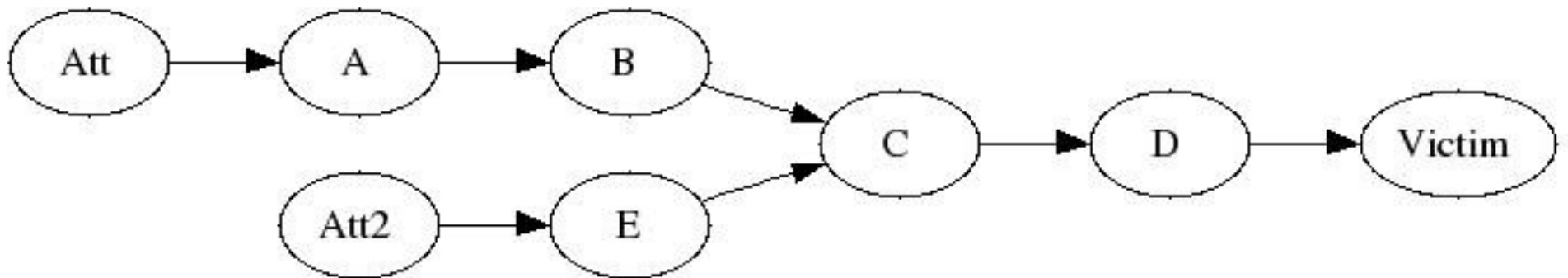
E C 2

D 0

C D 1

B C 2

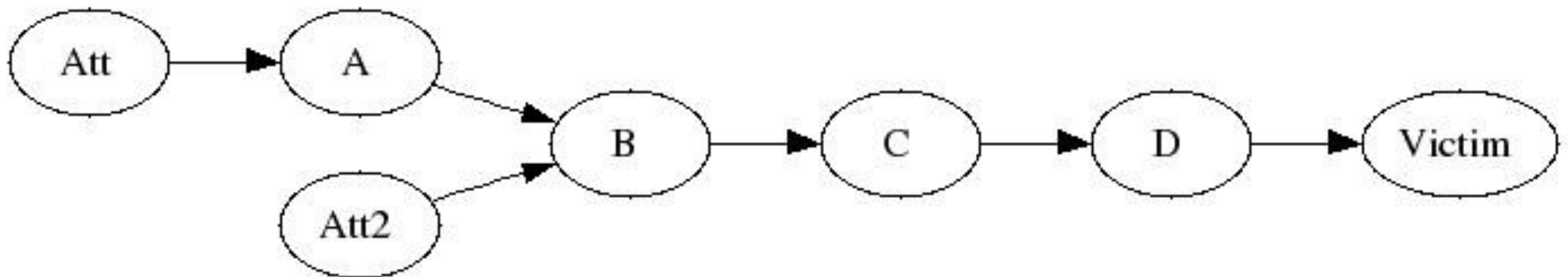
- Capable of determining multiple angles of attack.
- Number of packets required to recreate path is linear based on number of attackers. (216 packets for 2 attackers)



Edge Sampling

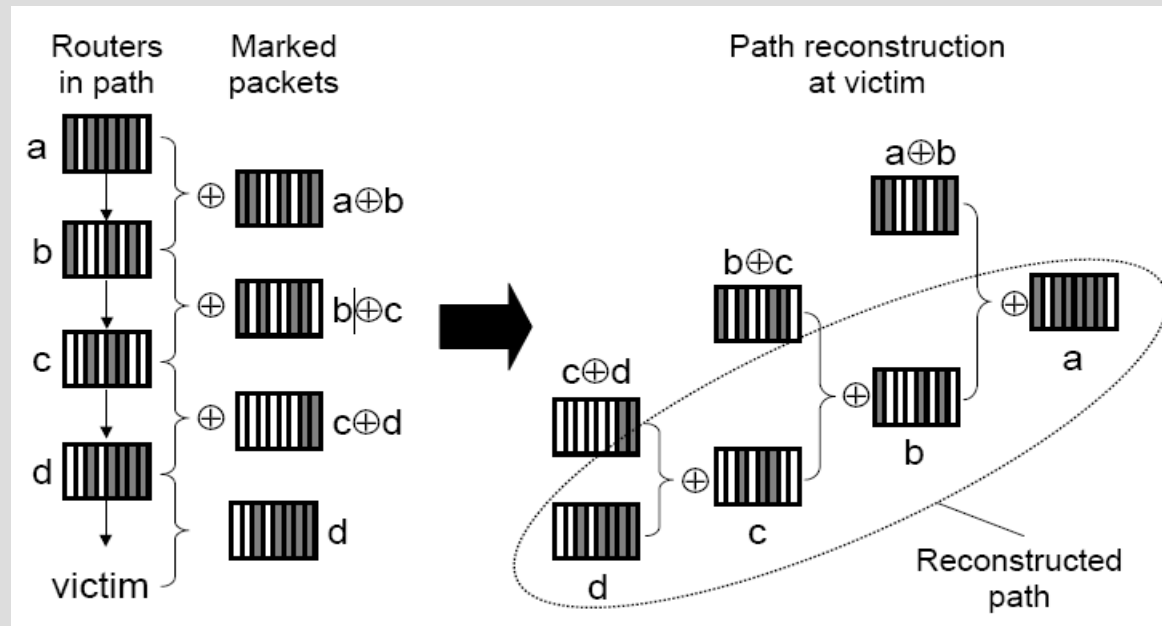
- Disadvantages:
 - Requires 72 bits of space (2 IP addresses @ 32bits, distance 8bits).
 - Less ability to trust longer paths:
 - Can we trust that there is another attacker connected to A?

A B 3
D 0
C D 1
B C 2



Compressed Edge Fragment Sampling

- Halves the space needed for storing IP addresses by XORing the two edges.

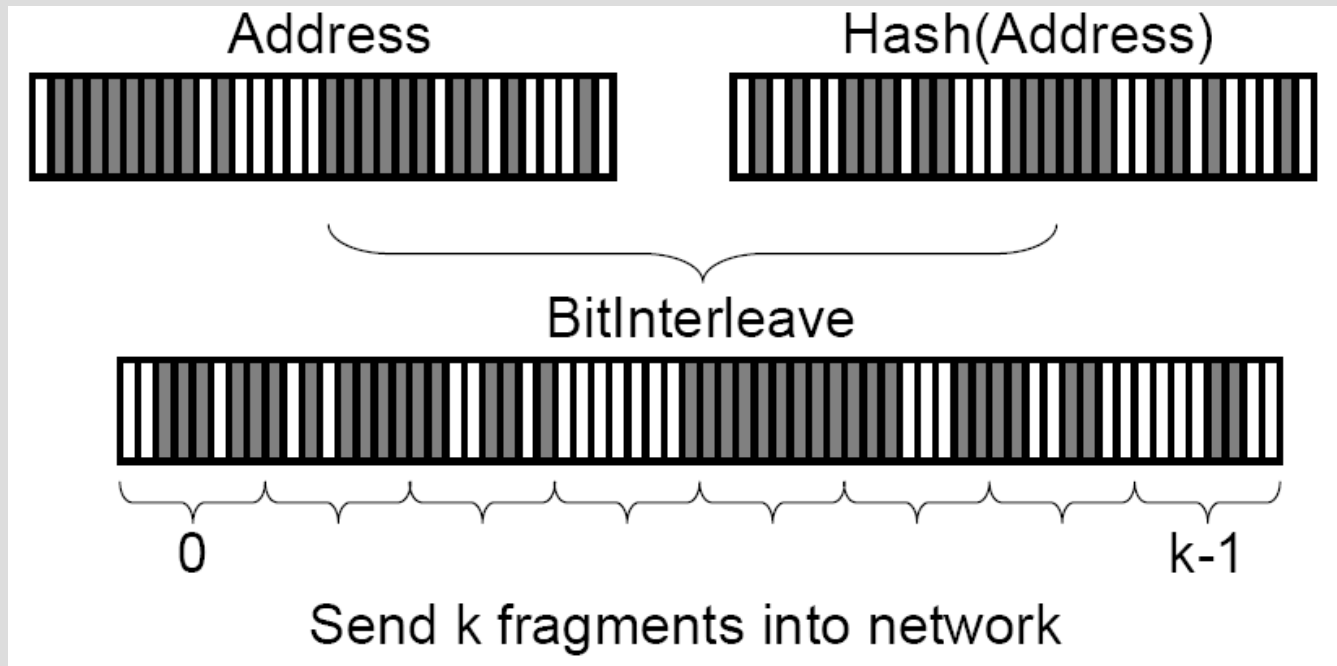


- The victim reverses the process to discover the route of the packet.

Compressed Edge Fragment Sampling

- Further reduction in storage space by transmitting a fragment of the address information.
- Additional space needed to store the offset of the fragment. Overall size still smaller.
- Possibility of matching fragments from different edges.
 - Interleave the bits of the address with bits of a hash of the address.
 - Choose fragments of the interleaved bits to send.

Interleaving Bit Fragments

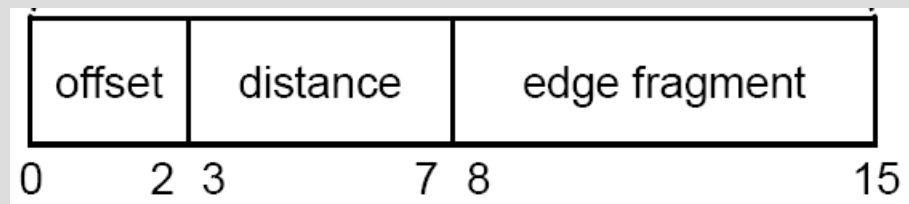


Reconstructing Path From Interleaved Edge Fragments

- Introducing fragments requires additional packets to determine all edges and addresses in the path.
- Edges divided into 8 fragments
- Probability of router marking a packet: $1/25$
- Assume attacker 10 hops away
- Reconstruct the path with 95% certainty requires no more than 2150 packets.

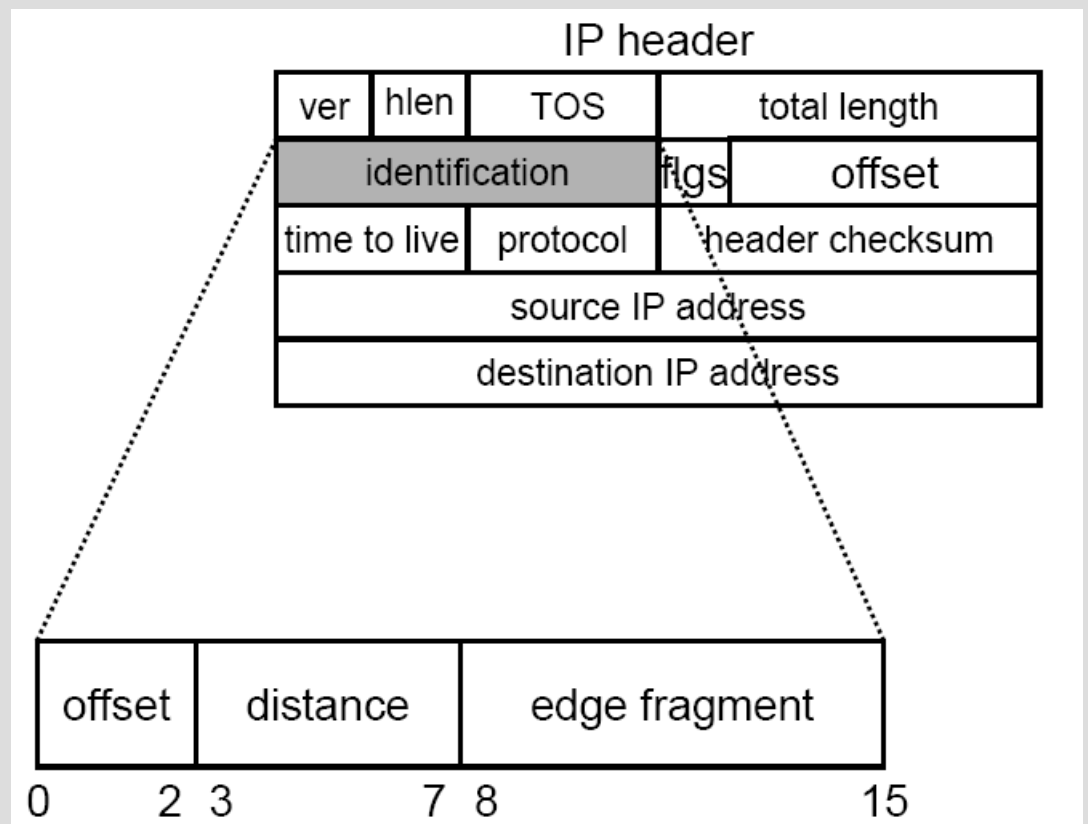
IP Header Encoding

- Overload the “identification” field of IP header.
 - 16 bits
 - Intended to differentiate fragments of a single packet.
- Field divided into three sets of information:
 - 3 bits offset of 8 possible fragments
 - 5 bits for distance from marking router to victim
 - 8 bits for the fragment itself



IP Header Encoding

- In many cases, less computation needed by routers for marked packets.
 - distance field incremented
 - TTL decremented
 - Results in the checksum staying the same.



Backwards Compatibility

- Identification field used for fragment identification
 - Studies show 0.25% of packets are fragmented.
- Set “Don't Fragment” bit for marked packets.
- Marking fragmented packets not previously marked.
 - Routers will mark packet fragments with a smaller probability than non-fragments.
 - Prepend an ICMP “echo reply” header
 - Contains full edge data (2 IPs + distance)
 - Truncate the tail of the packet
 - “The packet is consequently 'lost' from the standpoint of the receiver” ... How is this useful?

Limitations

- Fragmented packets
- IPsec protects identification field from modification
- IPv6 implementation
 - Lacks an identification field of IPv4
 - Proposes overloading the 24bit “flow label” field.
- Distributed Attacks
 - As number of routers at the same distance increases so does the possibility of grouping edge fragments together.

Limitations

- Attacker could attempt to spoof additional edges not really in use.
 - Use knowledge of Internet topologies and routing structure to verify whether links are valid.
 - Adding additional 'secret' information in header for verification off-line.
- Tracebacks may reveal the source router for the attack, not the attacker itself.

BIGGEST Limitation

- Implementation!
 - Traceback method effective only if most/all routers implement the solution.
 - Otherwise, holes will be found in the reconstructed path.
 - Uncertainty of origin:
 - Are routers at beginning of path not marking packets?
 - Is the attacker marking packets suggesting a longer path?