Distributed Denial of Service Attacks

Lecture seventeen

Outline of Lecture

- Approaches to defending against DDoS
  - Preventive
  - Survival
  - Responsive

Learning objectives

- By the end of this lecture you should be able to discuss techniques that assist in
  - Minimizing the damage during attacks
  - Stopping DDoS attacks and restore normal operation
  - Finding those who are responsible for the attacks

DDoS Defense Approaches

- Preventive
  - Stop attacks before they occur
    - Passive and active approaches
- Survival
  - Enlarge resource
    - replicating target, not effective, expensive.
  - Make targets more robust
    - Limit number of half-open TCP connections
- Responsive (at early or late stage)
  - Practical, effective but need an accurate detection and a quick, selective response.

Preventive

- Can we stop the DDoS attack before it starts?
- Passive measures
  - Software patched to prevent being used as a Zombie
    - Good security practices
  - Modifications to protocols (long term, IPv6, DNSSec)
- Active measures
  - Identification and suppression of messages that initiate a Zombie attack
  - Infiltration of attackers
    - crime-research.org/library/grcdos.pdf

Preventive

- ‘Moving target’ defense
  - Sometimes IP hopping
- Regularly change the IP address of servers that might be subject to an attack
- Can be done with DHCP and dynamic DNS
  - Lease expires on IP address
  - Request new address
  - Automatically register new address with DNS
- But some limitations to this defense
  - DDoS usually based on domain name rather than IP address

Survival

- Can we survive the DDoS attack?
- Install or modify systems that will enable your systems to survive an attack
  - Perhaps with significant downgraded service levels
- Backup procedures
  - Alternative ISPs
- Firewall configurations
  - Filter packets with internal source addresses (probably spoofed)
  - Filter packets sent to unused ports
  - Filter packets from known attack hosts
  - Dynamic firewalls can protect against some attacks
    - Only allow a DNS response if there has been a DNS request

Responsive defenses

- Can we respond and defend ourselves against the DDoS attack once it has begun?
- Cisco suggest the following phases:
  - Preparation
    - Tools, procedures to respond to attacks
  - Detection
    - Identify that an attack is occurring
  - Classification
    - Determine attack type
  - Traceback
    - Determine where the attack is coming from
  - Mitigation
    - Apply firewall rules (or similar) to prevent attack or have procedures in place for upstream ISPs to impose rules on the traffic they admit to your network
  - Post-mortem

Responsive defenses

- We will examine the following
  - Detection
    - Identify that an attack is occurring
  - Traceback
    - Determine where the attacks are coming from

Passive defense

- Actions are taken after the attack traffic reaches the victim.
- Attacks are usually detected by monitoring inbound traffic and other performance metrics locally.
- Actions can be
  - Filter or block attack traffic at the ingress routers
  - Trace back to find the attackers
    - Block attacks at their sources

Detection methods

- Monitor traffic volumes.
  - Sudden spikes in traffic of particular types might indicate a UDP attack
    - Lots of messages to unused ports
    - Lots of UDP packets to port 80
    - Lots of DNS responses to port 80
- Monitor the ratio between incoming and outgoing flow rates
  - Lots of DNS responses but no DNS requests
    - Can be prevented by a dynamic firewall

Detection methods

- Look (check) for signatures of known attack traffic.
  - Eg large UDP packets to Port 666 a well known attack signature
- Identify anomaly traffic which does not match a pre-built traffic profile (usually statistical based).
  - Traffic to webserver mostly small packets
  - TCP connections mostly have Poisson distributed arrivals
  - Sudden change in traffic characteristics may indicate a DDoS attack

Detection methods

- Monitor source addresses because DDoS attacks usually use spoofed (random) IP addresses
  - In normal operation most IP addresses have been seen before
  - During an attack most IP addresses are new
  - Most reliable and simple to deploy method of detecting DDoS attacks
- Monitor packet content for specific traffic
  - e.g. web traffic, UDP traffic.
  - Similar to method of Identify Anomaly traffic.

Detection methods comparison

	Monitor traffic volume / flow rates	Monitor IP addresses	Match traffic profile / attack signature / monitor packet content
Complexity	Simple	Moderate	Complex
Accuracy	Not very accurate	Accurate	Accurate
Speed	Fast	Fast	Slow
Computation	Low	Low	High
Deployment	Easy	Easy	Moderate

Actions after detection

- Assuming an attack can be detected what should be done about it?
  - Filter traffic
  - Trace back and identify where the attack originated
- Filtering attack traffic at
  - Ingress router
  - Upstream ISP networks
  - Further upstream ISP networks
  - Egress router (Source)
- Normal packets could also be dropped in the filtering process.
  - Measure of effectiveness of packet filtering measured by the Normal Packet Survival Ratio
    - Percentage of normal packets that get to the victim during an attack

Ingress filtering

- Filter incoming traffic according to specific rules
  - based on IP addresses
    - internal or specific IP addresses will be filtered
  - can have other rules in combination with detection process (filter on packet basis).
- Ingress filtering is normally integrated with the firewall.
  - Effectiveness depends on the complexity and computational effort.

Upstream ISP filtering

- Customer request upstream ISP to filter attack packets.
- Request is sent through separate communication channel
  - e.g. telephone
- Detector could send intrusion alert message early on with attack signature to ISP.
  - Intrusion alert messages (IAM) have to be protected.
  - IAM itself can be another DoS attack
  - Normal packets might also be dropped in the filtering process.

Further upstream ISP filtering

- Extend filtering beyond the local ISP.
- Victim network is responsible for detecting attacks and notifying upstream filters to filter packets matched with the signatures of the detected attacks.
- Packet filtering can be pushed as far upstream as possible.
- Need cooperation between ISPs.

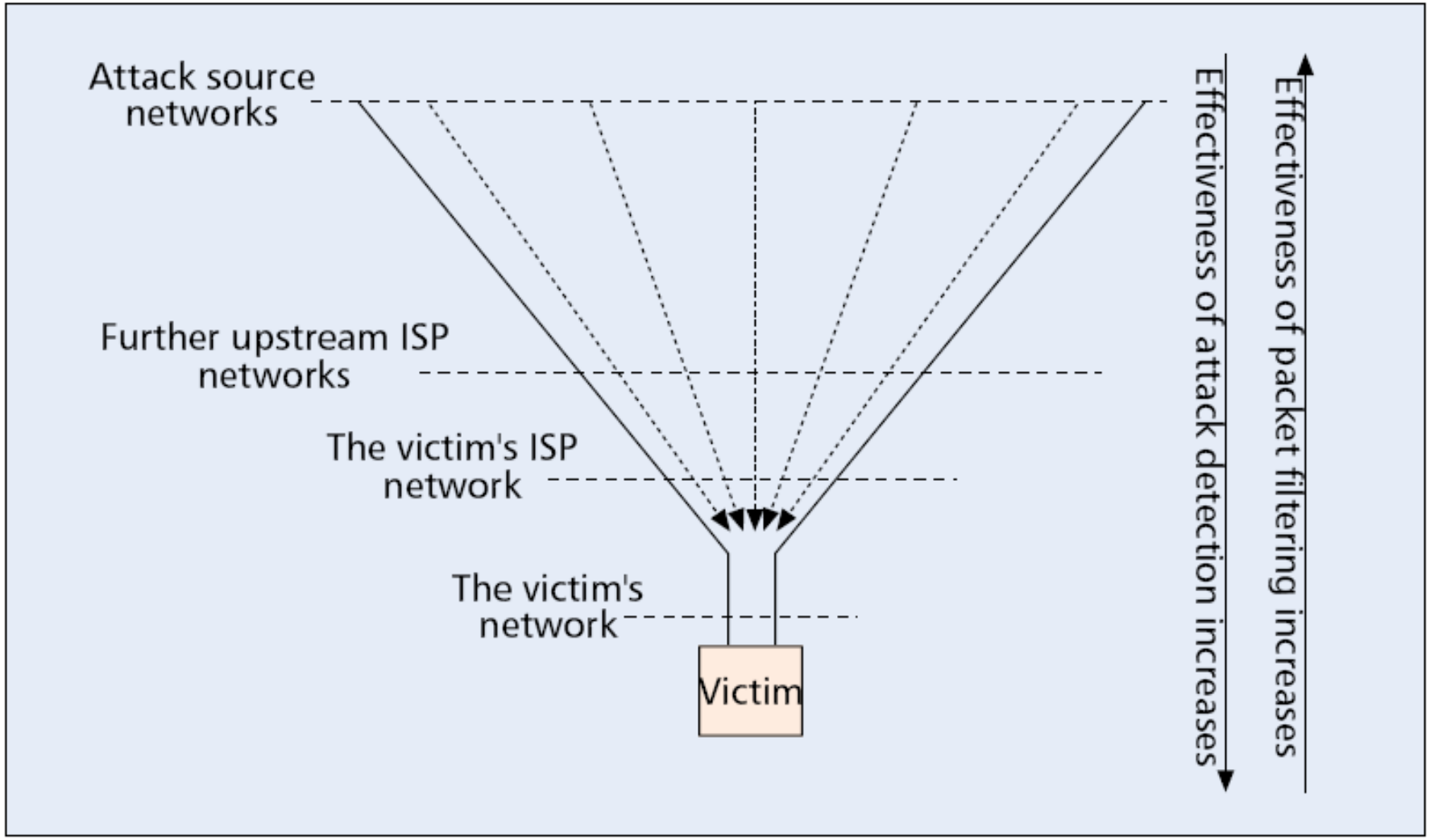
Egress filtering

- Filter as a general policy not because of detecting DDoS attacks.
- Filter attack packets at their source based on
  - illegitimate packets,
  - spoofed IP addresses.
- Can eliminate completely attack packets in a direct attack or from agent to reflector because they usually use spoofed IP addresses
  - if it is not the case then the attack can be easily traced back

Filtering comparison

	Ingress	Upstream ISP	Further Upstream ISP	Egress
Effective	Low	Moderate	High	High
Deployment	Easy	Moderate	Low	Low (large scale)
Damage	High	High	Low	Very Low
Signalling	No	Yes	Yes	No
Cooperation	No	No	Yes	No

Detection and filtering



Detection and filtering

- Accuracy of detection improves the nearer the victim
  - Should be able to detect the DDoS when on the receiving end of it !
  - Much harder to detect the DDoS if you are one of the source networks
- Effectiveness of prevention improves the nearer the source
  - If you are one of the sources of the DDoS then you are very able to prevent it
    - Filter offending packets on your firewall
    - Shutdown offending hosts
  - If you are at the receiving end it is much harder to prevent a DDoS attack

Traceback of attack sources

- Purpose is to identify the source of any packet sent across the Internet without relying on the source information contained within the packet
  - Would like to identify the source of a packet even though the source address has probably been spoofed
    - Tell the source network or host to apply patches to prevent further attacks (if a zombie)
    - Add them to a blacklist within Firewall
- Not easy to identify source because
  - IP source address usually spoofed
  - Routers normally only know the next hop for forwarding packets rather than end-to-end

Traceback approaches

- How do we identify the sources of the attack?
  - Which compromised hosts sent the spoofed packets?
- Three approaches to the problem
  - Router records information of every packet for later traceback request.
  - Router sends additional information to the packet's destination via separate channel (ICMP).
  - Router marks passing packets by inserting (partial) information into the packet header.

Hashed-based IP traceback

- In Hash-based IP traceback a special data structure is used within each router to store partial information about every packet.
  - Which router or host sent it?
- When DDoS is detected, victim can send query to its upstream routers.
- Router check its records, identifies the packet and passes the request to its neighbor routers. Eventually, the packet origin can be located.

iTrace

- In the iTrace approach, routers send ICMP messages (iTRACE messages) to the destinations for a number of packets passing through.
- The iTRACE message consists of the next and previous hop information, and a timestamp.
- In an attack (with large traffic volume) the victim will be able to traceback the source based on the received iTRACE messages
- Not a success
  - Routers and Firewalls commonly block ICMP messages

Packet marking

- The third approach is packet marking including
  - node append,
  - node sampling and
  - probabilistic packet making methods.
- Makes use of identification field in the IP packet header
  - Identification field 15 bits
  - Don’t store full IP address
    - Use some hash function to point to router

Node Append

- Append address of every nodes along the path to the end of the packet.
- Victim (destination) will have a complete ordered list of the routers it traversed.
- Advantages
  - Robust,
  - Easy to identify sources of attack
  - Quick to converge since only need one packet to find out the source.
- Disadvantages
  - Difficult to append data to packet on the fly
    - Very big processing overhead
  - Difficult to find space in the packet to store the address list
    - 1500 byte packet lengths
    - Can't use identification field

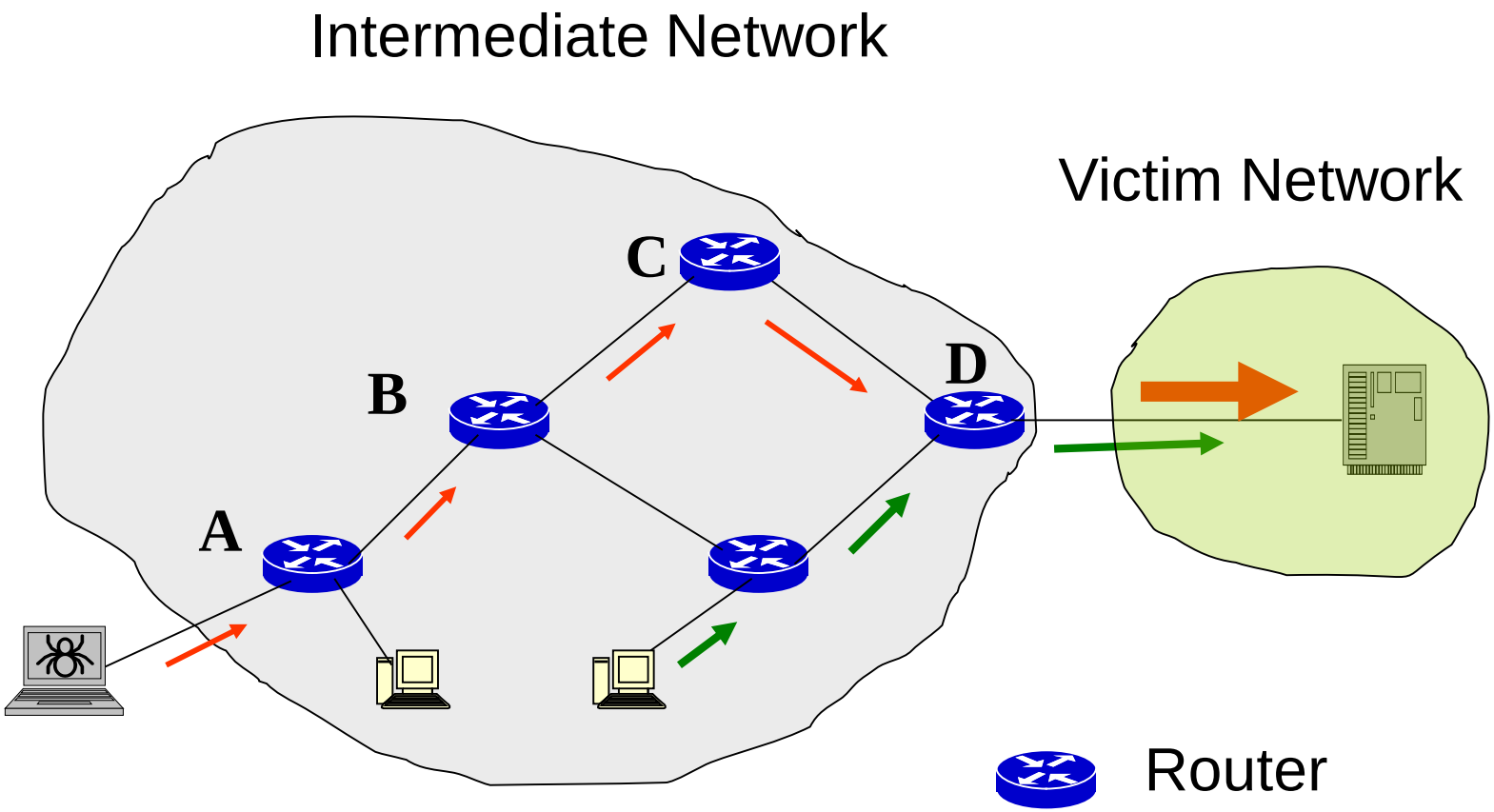
Node Sampling

- Similar to node append, but only sample the path one node at a time (with some probability *p*) instead of recording the entire path.
  - Each node may put its address in the packet
- For large traffic volume, victim will have a complete list of routers
- Disadvantage is that information as to path information will not be ordered
  - If enough packets received will receive packets with all the hops back to the source (or sources) of the attack
  - The order can be determined based on the number of packets received at the destination with the same node's address. It is because the probability that a packet is marked by a router and left untouched by all downstream routers is a strictly decreasing function of the number of hops to the victim.

Node Sampling

- Reduces router overhead (data appending) and space requirement per-packet.
- Attacker can still
  - insert “false” router into the path by contributing more samples than that are marked by the down stream routers,
  - reorder valid routers by contributing more samples than the difference that are marked by any two down stream routers.
- Set *p* large enough (> 0.5) will eliminate these problems.

Node Sampling



Node Sampling

- Referring to previous diagram
  - Each node marks with a probability of 0.5
  - 50% of packets leaving node A are marked with ‘A’ or unmarked
  - 50% of packets leaving node B are marked with ‘B’, 25% are marked with ‘A’ and 25% are unmarked
  - 50% of packets leaving node C are marked with ‘C’, 25% are marked with ‘B’, 12.5% are marked with ‘A’ and 12.5% are unmarked
  - 50% of packets from node C leaving node D are marked with ‘D’, 25% are marked with ‘C’, 12.5% are marked with ‘B’, 6.25% are marked with ‘A’ and 6.25% are unmarked

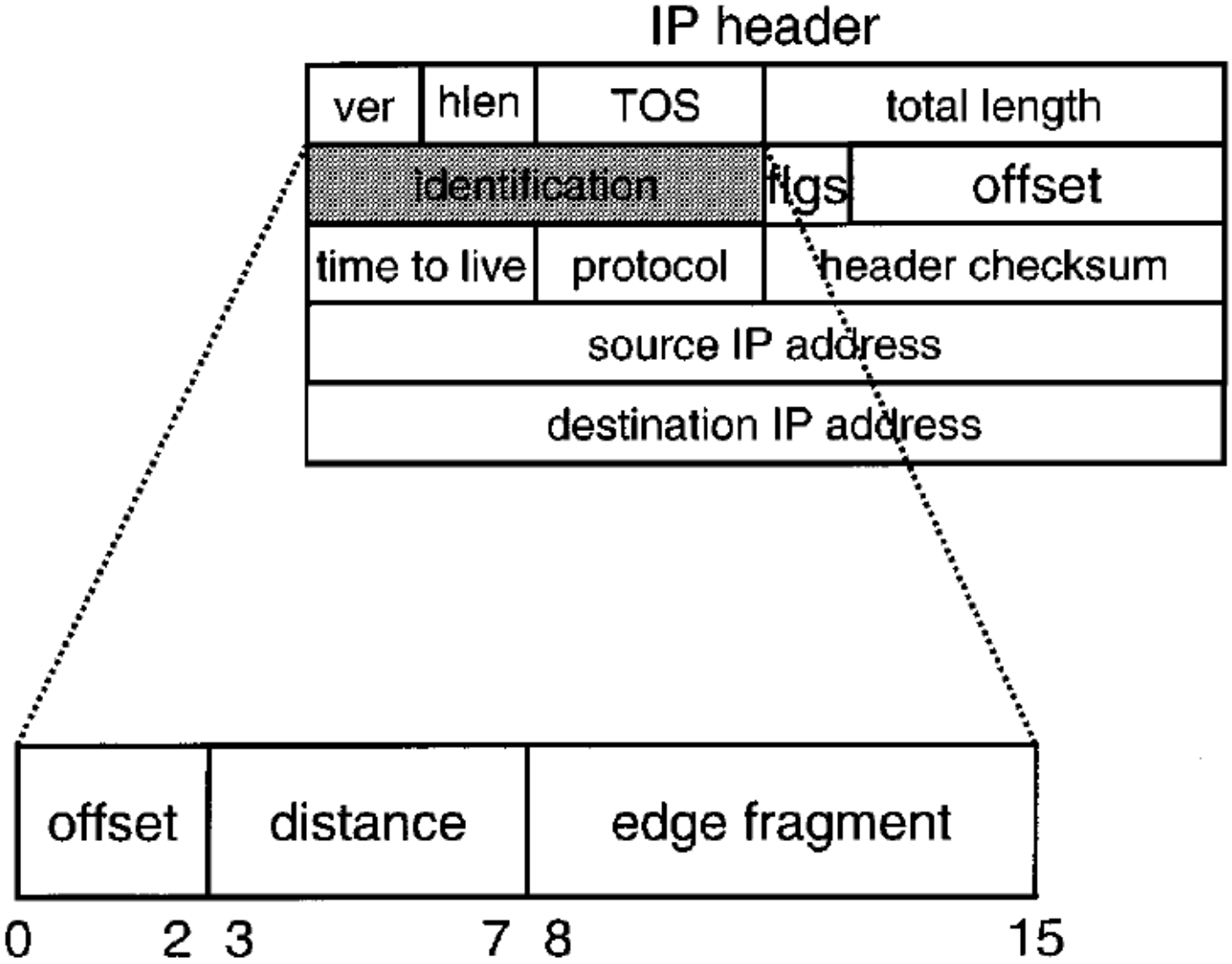
Node Sampling

- Process of determining the order of routers is slow (need a certain number of packets from every router).
- Not robust when there are multiple attackers and hence multiple paths
  - Nature of DDoS is that there will be multiple attackers and multiple paths
  - Consequently more than one router can be at the same distance from the victim and are marked with the same probability
  - Makes it difficult to identify paths of attackers
- Edge sampling (also known as probabilistic packet marking – PPM) avoids these problems.

Probabilistic Packet Marking

- PPM adapts the idea of node sampling but records information about edges
- An edge is an identifier of a path between two routers
- By recording information about edges we can say an attack packet traversed this path
  - With Node sampling where there are multiple attack paths we cannot easily obtain this information
- The edge information is written into the ID field of the packet header (with a low probability of *p*)
- If the edge information is not written into the ID field the distance information is incremented
- With enough packets the victim can construct all the attack paths

IP Header Marking



Probabilistic Packet Marking

- With probability *p* a router put part of its own IP address into the edge field and zero into the distance field of the marked packet.
- If the distance field is already zero, which means this packet has already been marked by the previous router, it processes the packet as follows:
  - Put a combined value of its IP address and the existing value in the edge field,
  - Increases the distance value by 1.
  - This produces an **edge id**
- If router does not mark the packet, it always increments the distance field.
- The victim will receive a number of edge information packets that enables reconstruction of the attack paths

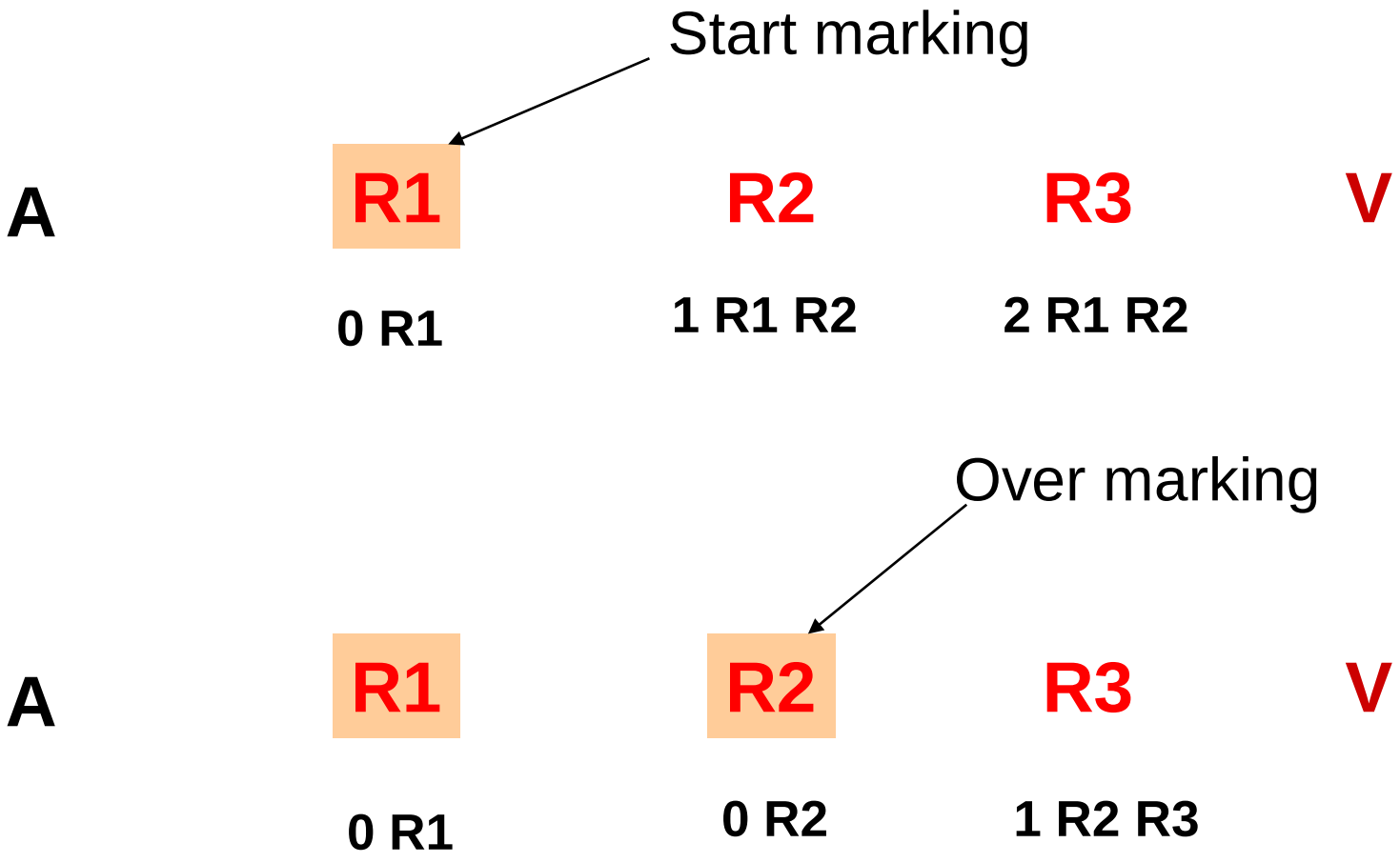
PPM Marking Procedure

Marking procedure at router R:

```
for each packet w
    let x be a random number from [0..1)
    if x < p then
        write R into w.start and 0 into w.distance
    else
        if w.distance = 0 then
            write R into w.end
        increment w.distance
```

[Source: *S. Savage et al, Network Support for IP Traceback, IEEE Trans. on Networking, 2001*]

PPM Illustration



PPM Advantages

- Any packets written by the attacker will have a distance field greater than or equal to the length of the real attack path.
- Therefore single attacker is unable to forge any edges between themselves and the victim
- Probability *p* can be arbitrary since it is not based on the sampling rank approach as in node sampling.
  - Typically 1 in 20,000 packets is marked ($p = 0.00005$)
- PPM is robust, but may converge slowly (needs certain number of samples to converge).

PPM Reconstruction Procedure

Path reconstruction procedure at victim v:

let G be a tree with root v
let edges in G be tuples (start,end,distance)
for each packet w from attacker
    if w.distance = 0 then
        insert edge (w.start,v,0) into G
    else
        insert edge (w.start,w.end,w.distance) into G
remove any edge (x,y,d) with $d \neq$ distance from x to v in G
extract path (R_i..R_j) by enumerating acyclic paths in G

PPM Improvement

- To speed up the convergence and minimize the time needed to reconstruct the attack path one can use non-uniform marking probability.
- It is because packets marked by upstream router can be over-marked again by a down stream router closed to the victim.
- To reconstruct the path we need a certain number of packets from each router along the attack path.
  - Marking upstream routers with higher probability will speed up convergence of the process.
- Still an active area of research

Conclusion

- Approaches to defending against DDoS
  - Preventive
  - Survival
  - Responsive
- Responsive ways of dealing with DDoS attacks
  - iTrace
  - Node append
  - Node sampling
  - Probabilistic Packet Marking