On the Effectiveness of Route-Based Packet Filtering for Distributed DoS Attack Prevention in Power-Law Internets

Kihong Park and Heejo Lee
Network Systems Lab, Computer Sciences
Purdue University

In Proc. ACM SIGCOMM 2001

Presented by Brad Burres

Agenda

- Introduction
- Related Work
- Route-Based Packet Filtering
- Performance Evaluation
- Results
- Implementation Issues
- Conclusions

Introduction

- DoS – Denial of Service
  - Attacker demands more resources than are available
  - We’ve talked about this!

- You cannot prevent a DoS/DDoS attack

- Protection takes two forms
  - Proactive – put measures in place to prevent attacks
  - Reactive – put systems in place to react to the attack and minimize its impact

Related Works

- Resource Management (e.g. firewall/detect)
  - Mitigate the impact on the victim
  - Does not eliminate the problem
  - Does not (likely) deter the attacker
- Ingress Filtering
  - Place at all boarder gateways
  - Should limit source IP address spoofing
  - Expensive to implement

IP Traceback (related works)

- Trace back the attacking packets to their source
- Traffic Analysis
  - Use logs at the routers to perform trace
  - High storage and processing costs
- ICMP Traceback messages
  - Variable length marking denotes route path
  - Increased network traffic
  - Now ICMP messages can be spoofed…

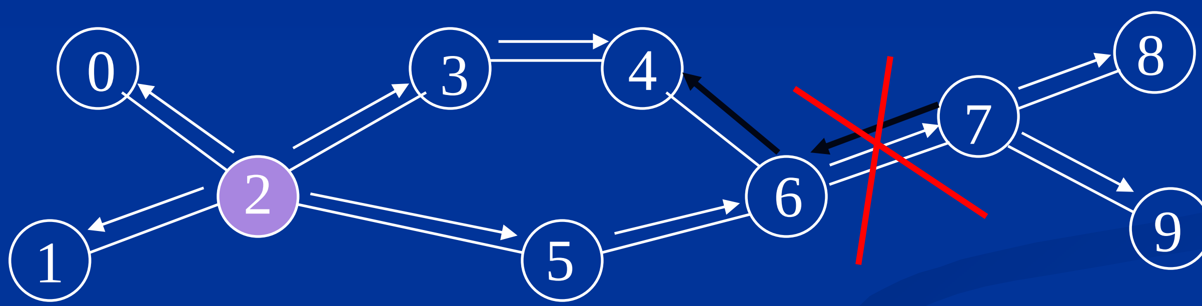
IP Traceback (related works)

- Probabilistic Packet Marking
  - Probabilistically mark a packet by adding route info
  - Constant marking field
  - Efficient to implement
  - Reconstructs the path of the attacker with a high probability
  - Can track attacker to within 5 equally likely sites
  - Reactive Only! Allows initial attack…
  - Doesn’t scale well with lots of attackers

Route-Based Distributed Packet Filtering (DPF)

- Break the name into pieces
- Route-Based Packet Filtering
  - Filter the spoofed packets whenever they are traversing an unexpected routing path
- Distributed Packet Filtering
  - Applying the filtering technique at certain points in the network
- Key Objectives are to 1) Maximize proactive filtering, 2) Minimize the number of possible attackers, 3) achieve 1&2 with smallest number of nodes possible

Illustration of Route-Based Filtering

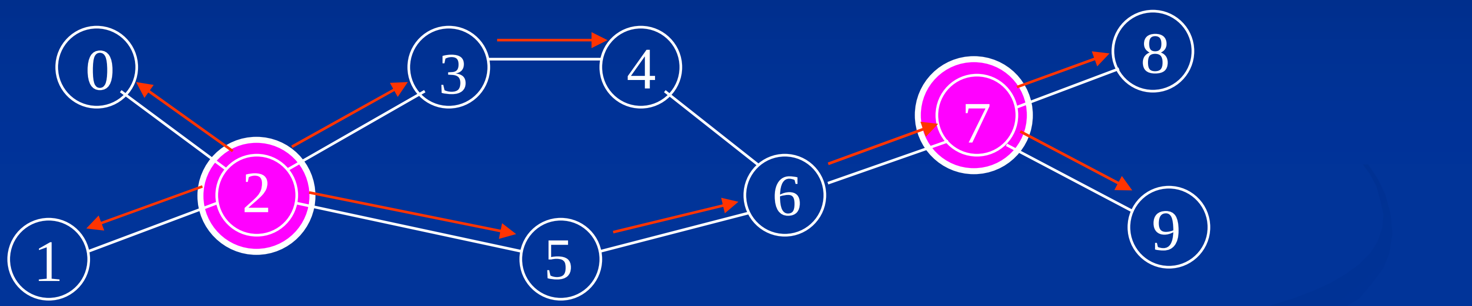


Valid Routing path of node 2

Node 7 Attacks 4 by spoofing Node 2’s address

Node 6 filters the attack

Definition of Terms



- G: network topology
- T: filtering nodes
- R: routing policies
- F: filtering function

More Terms (quickly)

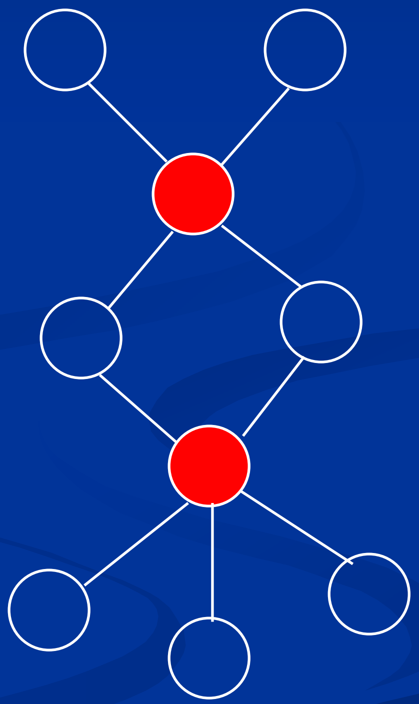
- V – a set of nodes in G (vertices)
- E – a set of links in G (edges)
- U – all non-filtering nodes (so V = U + T)
- S(a,t) – set of nodes an attacker can spoof that won't get filtered (attacker located at a and attacking t)
- R(u,v) – the path from node u to v (in lower case, it's a specific node)
- Routing Policies
  - Tight – there exists a single path between two nodes
  - Loose – any loop free path between two nodes

Maximal and Semi-maximal Filters

- Maximal Filter
  - Use all source and dest routing paths in G
  - If V nodes, then V nodes can be the source, and V-1 nodes can be the dest…
    - $V*(V-1) = V^2 \rightarrow O(n^2)$
    - $F_e(s,t) = \begin{cases} 0, & \text{if } e \in R(s,t); \\ 1, & \text{otherwise.} \end{cases}$
    - If edge e is on the routing path, the filter returns a 0, otherwise return a 1 and filter it.
- Semi-Maximal Filter
  - Use only the source address coming over link e
  - O(n) complexity, storage
  - $F_e'(s,t) = \begin{cases} 0, & \text{if } e \in R(s,v) \text{ for some } v \in V; \\ 1, & \text{otherwise.} \end{cases}$

Final Term: Vertex Cover (VC)

- T=VC
  - Any node in the set U has only nodes in the set T as its neighbors.
- Finding a minimal VC
  - NP-complete problem
  - Two well-known algorithms used for finding a VC



Performance Measures

- Proactive Prevention – limiting (eliminating) the number of nodes from which no spoofed IP packets can be reached

$$\Phi_2(\tau) = \frac{\left|\{a : \forall t \in V, \left|S_{a,t}\right| \le \tau\right|}{n}$$

  - Φ2(1): fraction of AS's from which no spoofed packets coming
- Reactive Traceback – A measure of the percentage of nodes which can – after receiving a spoofed packet (i.e. realizing that it's under attack) – can localize it's true source to within some minimal number

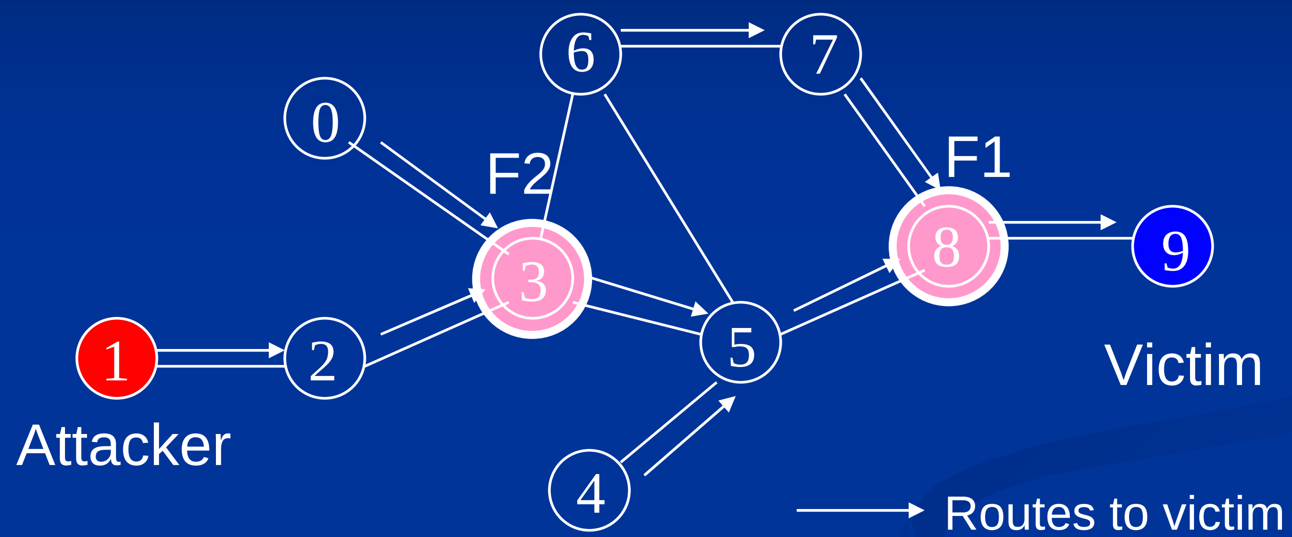
$$\Psi_1(\tau) = \frac{\left|\{t : \forall s \in V, \left|C_{s,t}\right| \le \tau\right|}{n}$$

  - Ψ1(5): fraction of AS's which can resolve the attack location to within 5 possible sites.

Performance Measures (cont)

- Attack Volume reduction
  - Captures the reduction in the volume of an attack, such as when the source IP address is randomly selected
  - $$\Theta = \frac{\left|\{(a,s,t) : s \in S_{a,t}\right|}{n(n-1)^2} = \frac{\left|\{(a,s,t) : a \in C_{s,t}\right|}{n(n-1)^2}$$

Minimizing "Spoofable" Addresses



No filtering: $S_{1,9}=\{0,1,2,3,4,5,6,7,8\}$

Filtering at F1: $S_{1,9}=\{0,1,2,3,4,5\}$

Filtering at F1 and F2: $S_{1,9}=\{1,2\}$

Power-Law Networks

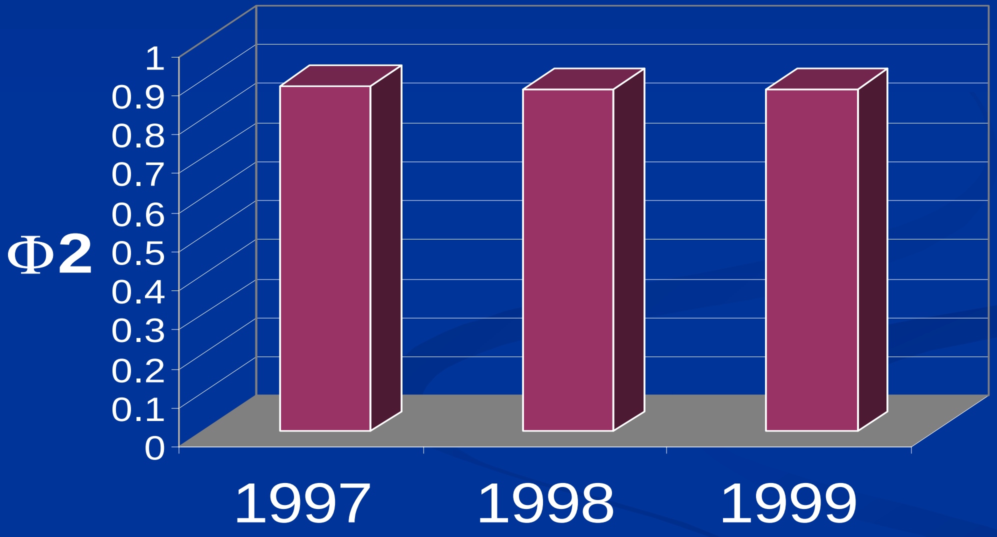
- Mathematically (PDF): $P[X=x] \sim x^{-(k+1)} = x^{-a}$
- Behaviorally. Think of it as "the rich get richer". If a lot of paths go through one node, than as more paths get added to the network, they too will go through that node.
- Like airport hubs – because we made Denver, Chicago, and Atlanta major hubs, now almost all flights of any distance go through one of those hubs.

Performance Results

- Found using a lot of evaluation tools (dpf, inet, brite)
- Proactive Filtering Effect
  - Not viable as a "perfect" filter
  - Does a very good job as DDoS attack prevention technique (limiting which nodes can attack and spoof from where)
  - $\Phi 2(1) = .88$ on real Internet topologies from 97-99

Proactive Filtering on DDoS

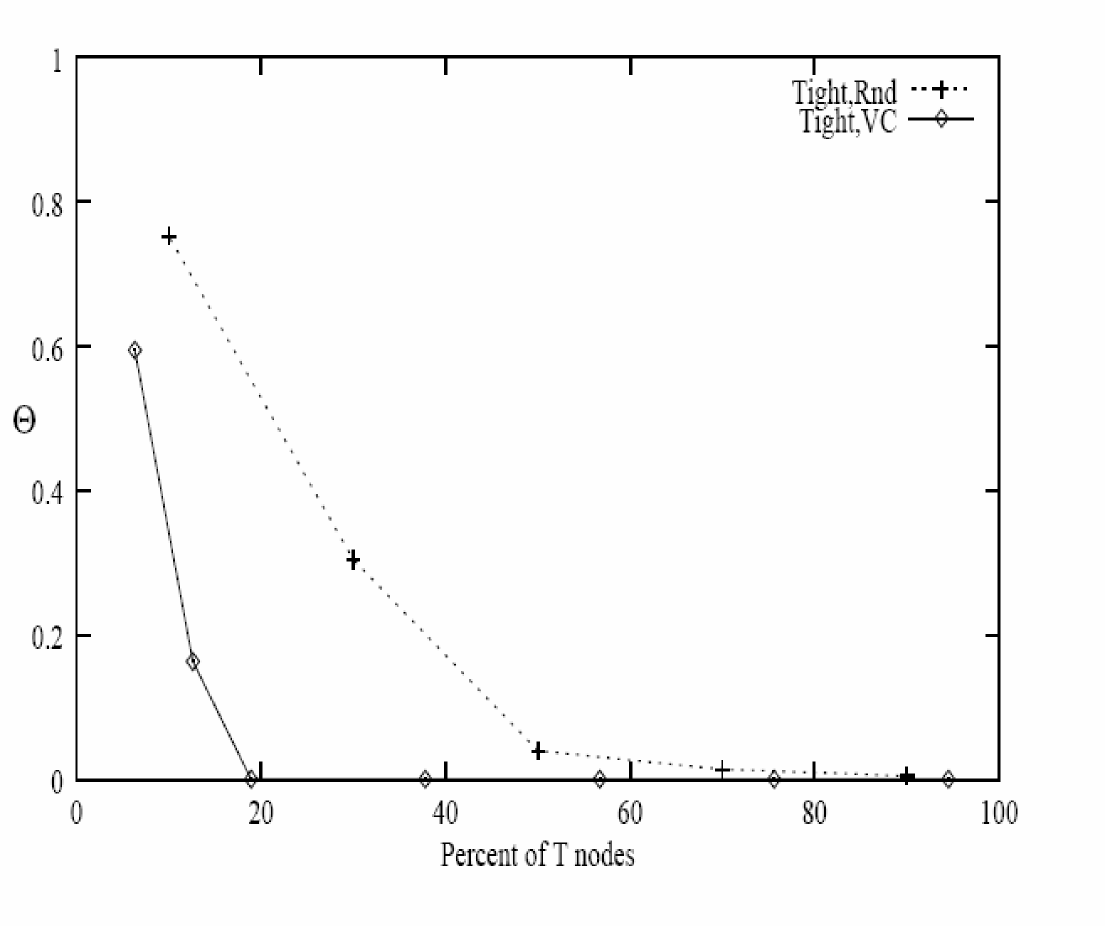
- G: 1997~1999 Internet connectivity
- T: VC
- R: Tight
- F: Semi-maximal



On real Internet topologies from 97-99, DPF makes 88% of internet sites “unspoofable”. This obviously hurts an attackers chances and makes them work much harder to even find valid attack nodes.

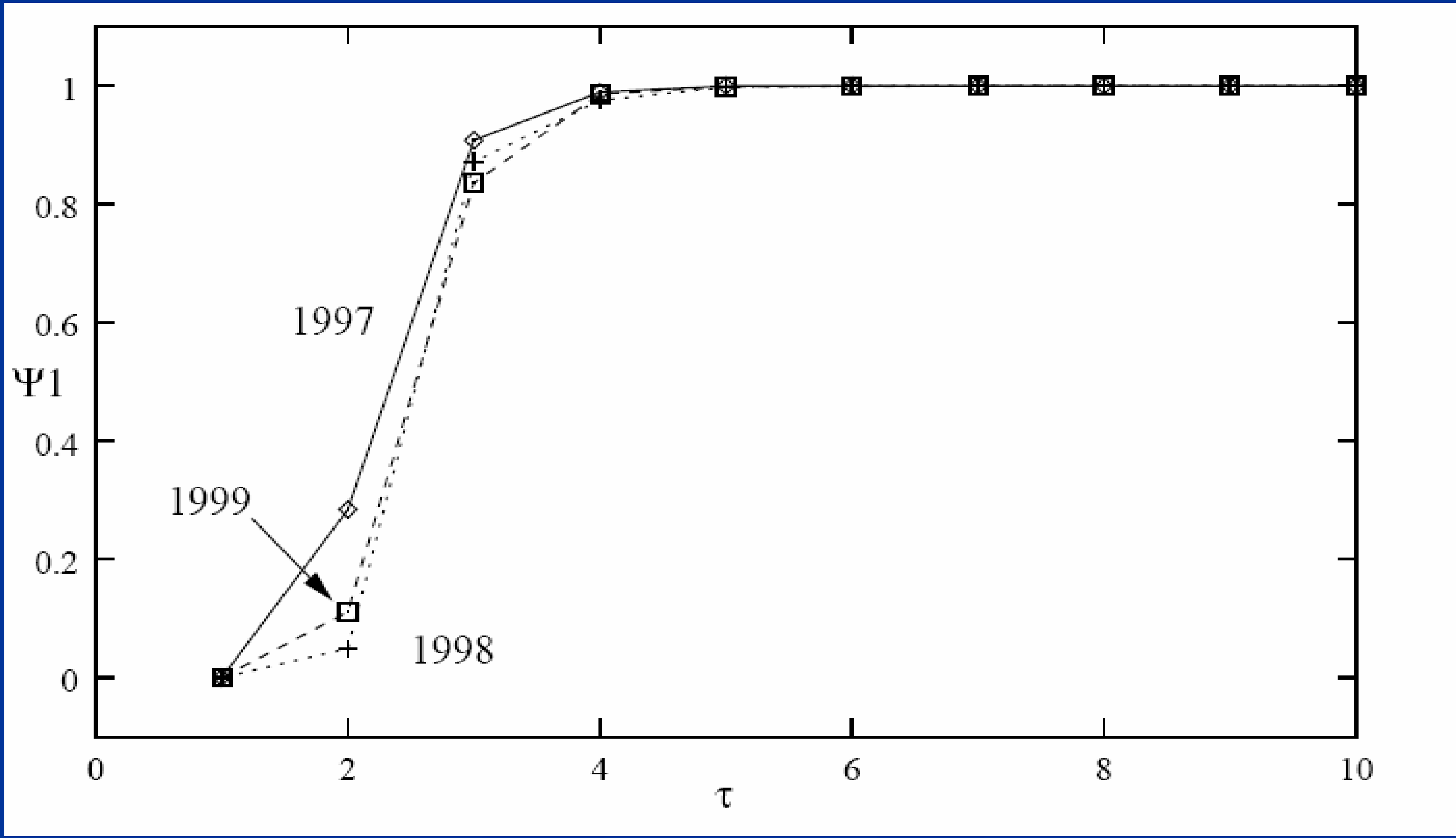
Attack Volume Reduction

- Randomly generated spoofed addresses are filtered 99.96% of the time!!
  - When T=VC, $\Theta = 0.0004$



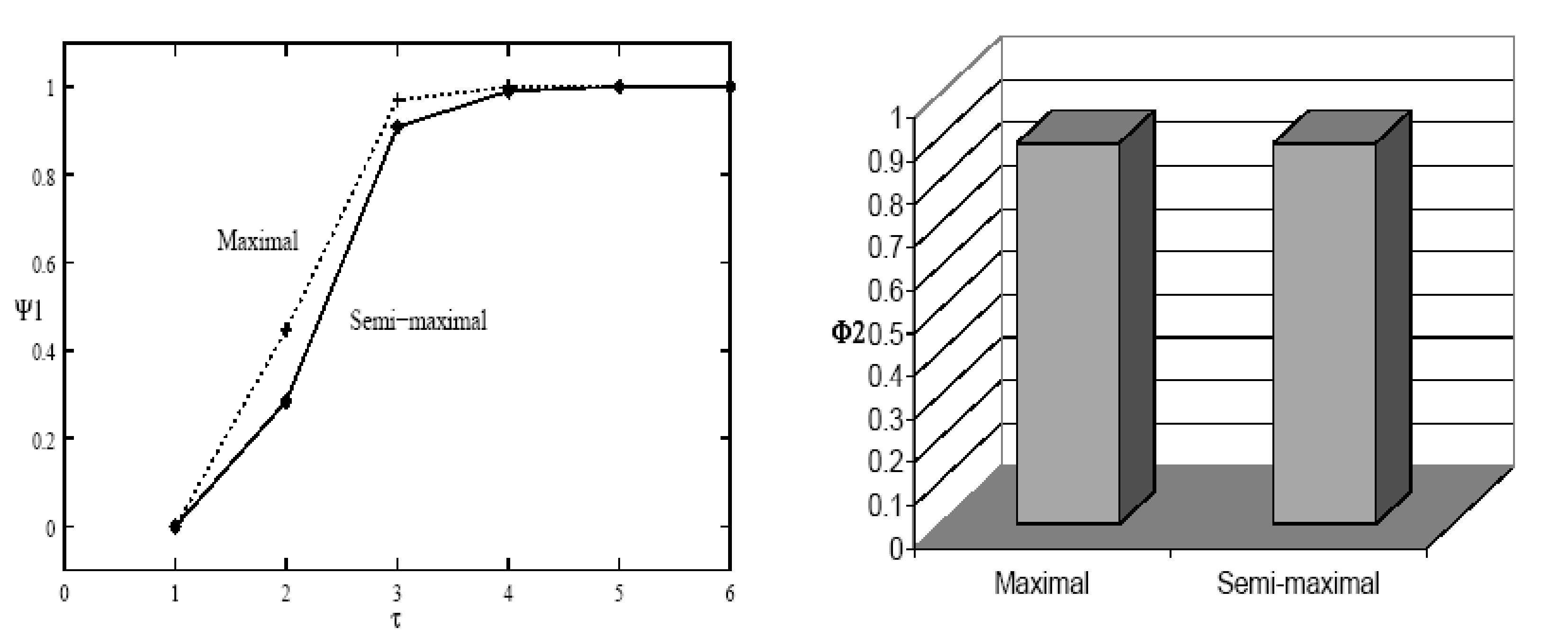
Reactive Performance for Traceback

- $\Psi 1(5) = 1$ for all three real Internet Topologies
  - Means that an attack can be localized to no more than five nodes



Maximal vs. Semi-maximal Filters

- Semi-Maximal filters are almost as good at a fraction of the cost!!
  - Maximal filters require V^2 storage and searching for insignificant gain



Impact of Network Topology

- The authors spent a lot of time here – I will not.
- Random topology (Not Power-Law Network)
  - Really bad performance. Takes lots of filter nodes and still doesn't filter a high percentage of spoofed addresses.
  - VC = 55% of total nodes!
- Inet topology
  - Has power-law characteristics
  - VC = 32% of nodes (real Internet was 18%)
  - Performance close to that reported for 97-99 Internet
- Brite topology
  - Basically, couldn't make it do what we want (or at least give us the results that we want)
  - Why put this in the paper?

Other Miscellaneous Results

- All simulations were done with the “T” nodes doing Ingress Filtering
  - $\Psi 1(5) != 1$ when this is not true
  - $\Psi 1(20) = 1$, and 20 nodes is still managable
- Multipath Routing degrades this solution.
  - For $R=3$, $\Psi 1(10) = 1$

Conclusion

- Distributed Route-Based Packet Filtering is effective
  - Preventative – minimizes the choices available to attackers
  - Reactive – minimizes the nodes which can originate a given attack
- Is it Practical?
  - Can be deployed incrementally
  - Needs protocol support to get source routing information (i.e. BGP needs a face lift)

References

- Info on ICMP traceback: http://www.nwfusion.com/news/2000/0724itrace.html
- Graphs: http://www.cs.cornell.edu/People/egs/syslunch-spring02/syslunchsp02/park-lee.pdf
- Concepts and images: cosmos.kaist.ac.kr/cs540/seminar/hjlee020911.ppt
- Power Law Networks: http://tisu.it.jyu.fi/cheesefactory/documents/PowerLawNetworks.ppt
  http://rio.ecs.umass.edu/~gao/ece697_0.../lect-03.01-properties.ppt