



Homeland
Security

DDoS Defense Benchmarks

Jelena Mirkovic, University of Delaware

Sonia Fahmy, Purdue University

Peter Reiher, UCLA

Roshan Thomas, SPARTA



Project Goals

- ❖ Develop DDoS defense benchmarks for evaluation of:
 - New defenses in **today's**, **tomorrow's** and **stress** scenarios
 - Network's resilience to/damage from a DDoS attack
 - Comparison of existing defenses
- ❖ Define defense performance measures
- ❖ Define how to use benchmarks and how to interpret results

Dimensions of A Test Scenario

- ❖ Attack traffic
 - Its strength and type define which resource is targeted, how large is attack impact, how well can a defense handle this attack
- ❖ Legitimate traffic
 - Its mix defines how well it competes with the attack and how much it suffers when it loses the battle
 - Defenses also interact with legitimate traffic and may inflict collateral damage
- ❖ Network topology and resources
 - Determine weak points that an attack will target

Two Approaches to Testing

- ❖ Define a set of features for each dimension
- ❖ Approach 1: Typical scenario testing

Example: ICMP flood with random spoofing and constant rate is a frequent attack, we must include it in our tests

- ❖ Approach 2: Comprehensive testing

Example: Attack rate for UDP flood should be 0.5, 1, 2 and 4 times the bottleneck bandwidth, number of attackers doesn't matter. Attack packet features only matter for defenses that model legitimate traffic by training.

Two Approaches to Testing

- ❖ Define a set of features for each dimension
- ❖ Approach 1: Typical scenario testing

Example: ICMP flood with random spoofing and constant rate is a frequent attack, we must include it in our tests

- ❖ Approach 2: Comprehensive testing

Example: Attack rate for UDP flood should be 0.5, 1, 2 and 4 times the bottleneck bandwidth, number of attackers doesn't matter. Attack packet features only matter for defenses that model legitimate traffic by training.

- ❖ Both are important and explored in our project: a typical and a comprehensive suite for each dimension.

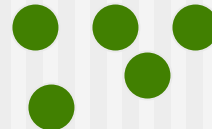
Creating Typical Scenarios

Traffic traces
1.1.0.3:2022 > 5..6.7.8:80 P 1000121:21211441
1.1.0.3:2022 > 5..6.7.8:80 P 1000441:21211461
1.1.0.3:2022 > 5..6.7.8:80 P 1000461:21211481
1.1.0.3:2022 > 5..6.7.8:80 P 1000481:21211500
1.1.0.3:2022 > 5..6.7.8:80 P 1000500:21211600

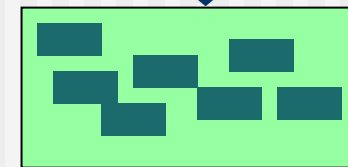
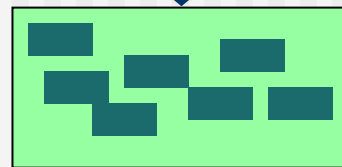
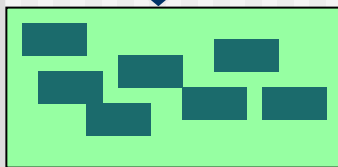
Internet



Traffic samples



Topology samples

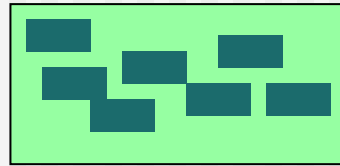


Typical Attacks

Typical Legitimate Traffic

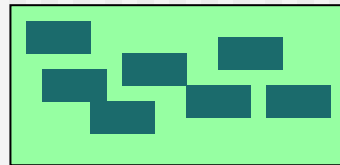
Typical Topologies

Comprehensive Tests



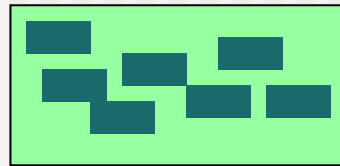
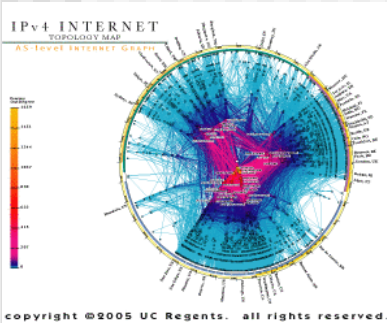
Sophisticated Attacks

What attacks have been proposed in literature?
What attacks would be particularly harmful for certain defenses?



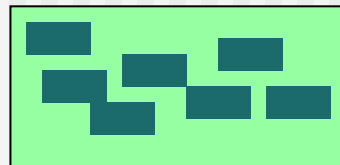
Stress Attacks

What are different application classes and their behavior/needs?



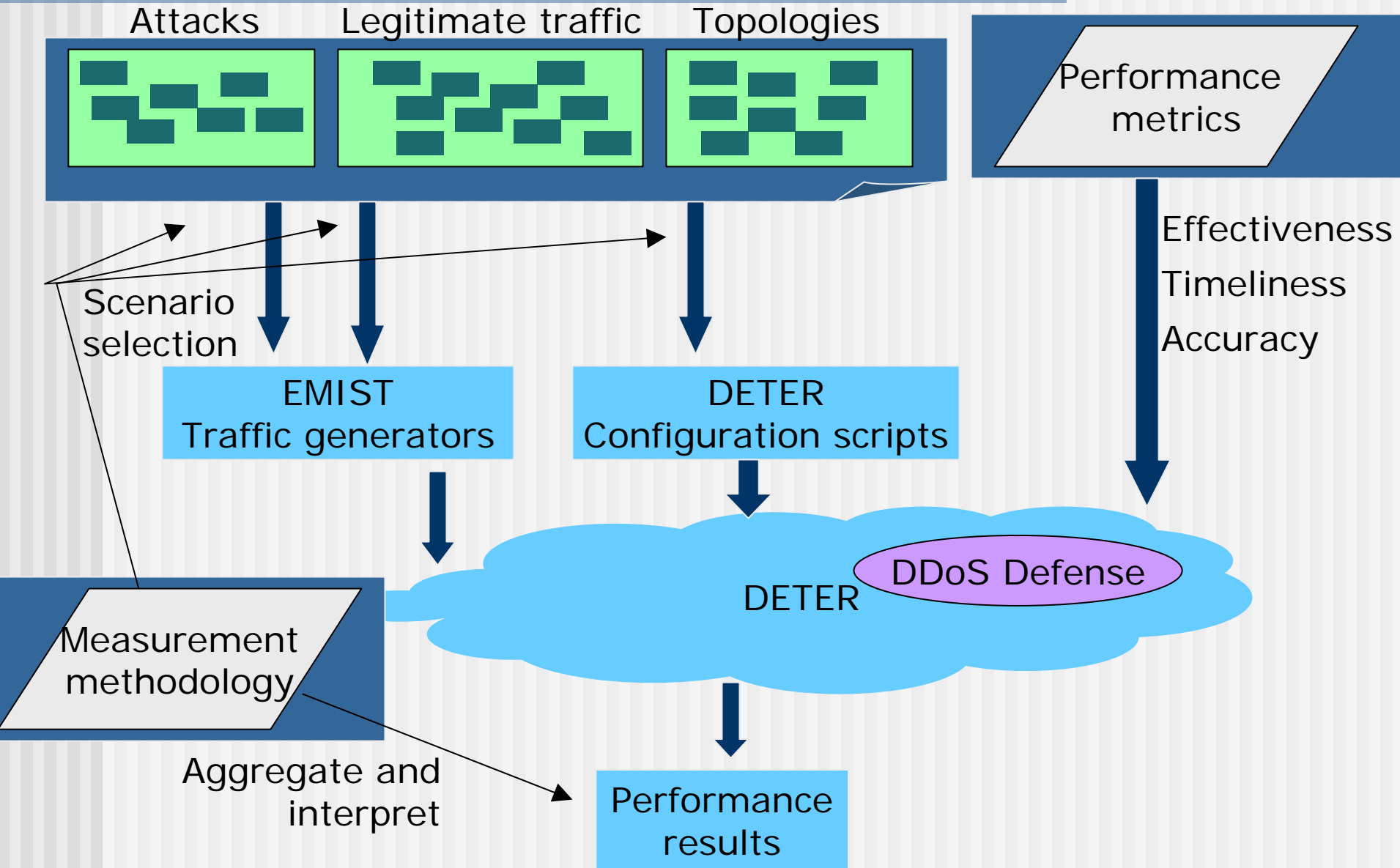
Traffic Classes

What are different network design alternatives?



Network Design

Using Benchmarks



Impact

- ❖ Common testing methodology will advance DDoS defense research
 - Provide common base for testing solutions
 - Provide means to compare solutions based on performance, cost and security
 - Provide means to test a network's resilience to various attacks
- ❖ Benchmarks, metrics and testing methodology will become part of DETER/EMIST and will be accessible to a wide research community

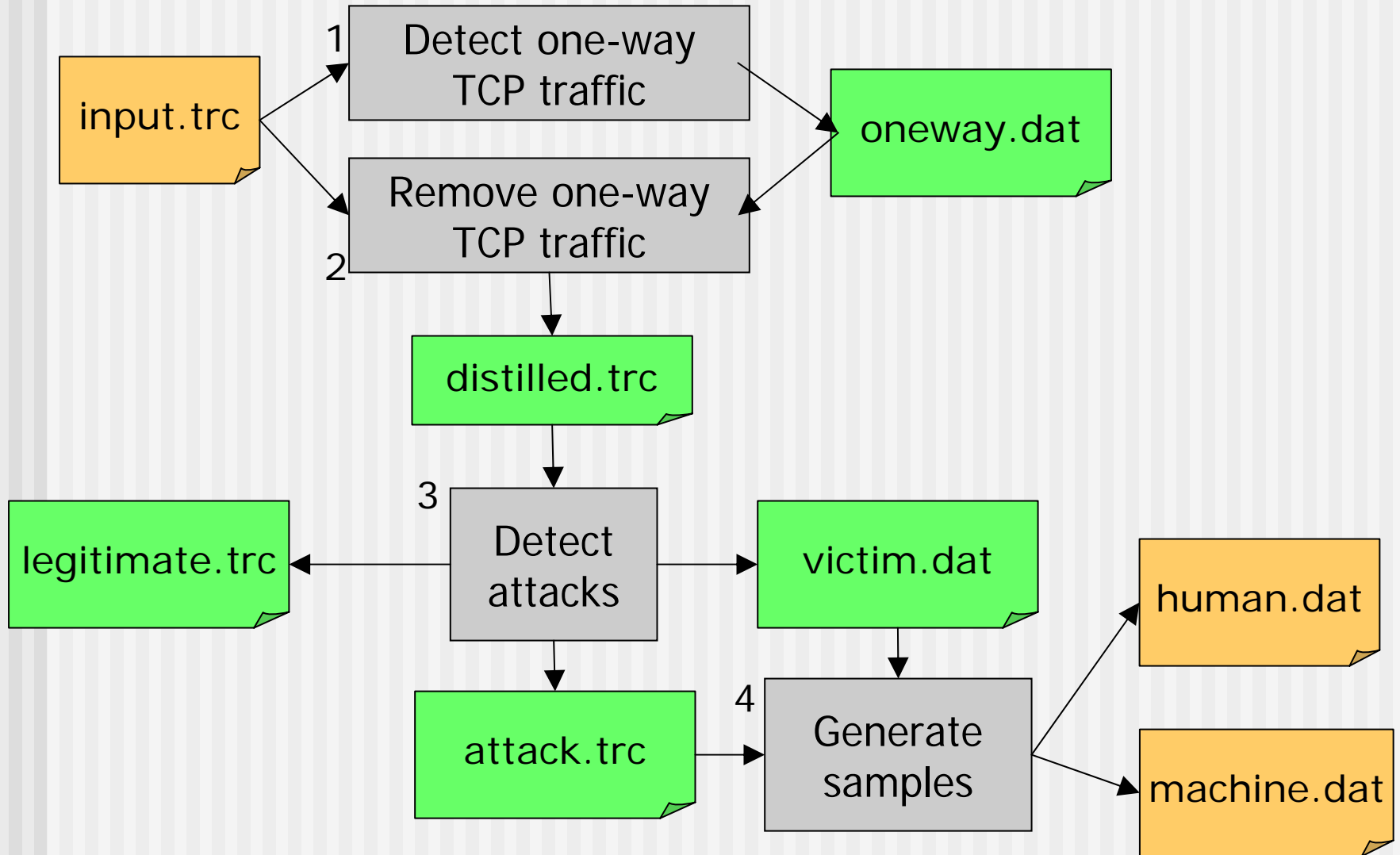
AProf Toolkit

Erinc Arikan, Jelena Mirkovic
University of Delaware

Motivation

- ❖ Learning about typical attacks
- ❖ We harvest attack information from public traffic traces and create attack samples
 - Samples provide information about attack type, victim, duration, number of sources, port distribution, packet and byte rate, etc.
- ❖ Next step: clustering these samples to generate representative attacks

Data Flow



Attack Detection

- ❖ We collect statistics for each connection and each destination, save them in hash tables
- ❖ Connection is identified as pair of IP addresses and ports
- ❖ Destination is identified by the IP address
- ❖ We apply multiple heuristics to detect attacks, looking for known anomalies and for aggressive traffic
 - We use packet data and context provided by destination and connection tables

Detection Heuristics

- TCP attacks:
 - Sequence number mismatch (junk traffic) connection
 - Packets-to-replies ratio (aggressive traffic) destination
 - Syn-to-ack ratio (SYN flood) destination
 - Absence of flags packet
- UDP attacks:
 - Increase in UDP rate, decrease in TCP rate destination
 - If there is a TCP connection between source and destination, recognize UDP traffic as legitimate (media) connection
 - If there are sufficient UDP replies, recognize as legitimate (DNS) destination
- ICMP attacks
 - Packets-to-replies ratio (aggressive traffic) destination
- Malformed headers packet
- Usage of non-existing protocol numbers packet
- High fragmentation rate destination

Results

- ❖ All synthetic attacks detected unless too small to create DoS effect
- ❖ Detected all labeled attacks in Los Netos traces and several unlabelled ones
- ❖ Detected about 5 attacks per hour in AUCK-VIII traces
 - Mostly SYN attacks, some short-lived but some last for several hours

AUCK VIII Sample Attack

1070348402.360687 attack on 10.0.0.206 type SYN flood
duration=3597.611205 pps=2.673163 Bps=134.212391
packets 9617 sources 2206 spoofing NO_SPOOFING
source ports 5635 dst ports 1

```
03:06:18.355216 IP 10.3.128.17.2581 > 10.0.0.206.25: S 1514458360:1514458360(0) win 8192 <[|tcp]>
03:06:18.355248 IP 10.0.0.206.25 > 10.3.128.17.2581: R 0:0(0) ack 2570469168 win 0
03:06:18.697523 IP 10.0.28.139.53046 > 10.0.0.206.25: S 2710130753:2710130753(0) win 5840 <[|tcp]>
03:06:18.697557 IP 10.0.0.206.25 > 10.0.28.139.53046: R 0:0(0) ack 2710130754 win 0
03:06:18.754734 IP 10.0.33.17.45454 > 10.0.0.206.25: S 646511622:646511622(0) win 5840 <[|tcp]>
03:06:18.754771 IP 10.0.0.206.25 > 10.0.33.17.45454: R 0:0(0) ack 646511623 win 0
03:06:18.874758 IP 10.5.159.222.59297 > 10.0.0.206.25: S 2385167281:2385167281(0) win 64240 <[|tcp]>
03:06:18.874791 IP 10.0.0.206.25 > 10.5.159.222.59297: R 0:0(0) ack 2385167282 win 0
03:06:19.059898 IP 10.3.128.17.2581 > 10.0.0.206.25: S 1576598835:1576598835(0) win 8192 <[|tcp]>
03:06:19.059931 IP 10.0.0.206.25 > 10.3.128.17.2581: R 0:0(0) ack 2632609643 win 0
03:06:19.598937 IP 10.5.159.222.59297 > 10.0.0.206.25: S 3881444780:3881444780(0) win 64240 <[|tcp]>
03:06:19.598970 IP 10.0.0.206.25 > 10.5.159.222.59297: R 0:0(0) ack 1496277500 win 0
03:06:19.605728 IP 10.0.212.103.50680 > 10.0.0.206.25: S 1976657513:1976657513(0) win 65535 <[|tcp]>
03:06:19.605765 IP 10.0.0.206.25 > 10.0.212.103.50680: R 0:0(0) ack 1976657514 win 0
03:06:19.794482 IP 10.3.121.203.57424 > 10.0.0.206.25: S 3722075728:3722075728(0) win 65535 <[|tcp]>
03:06:19.794522 IP 10.0.0.206.25 > 10.3.121.203.57424: R 0:0(0) ack 3722075729 win 0
03:06:20.240384 IP 10.5.159.222.59297 > 10.0.0.206.25: S 2572975221:2572975221(0) win 64240 <[|tcp]>
```

Sophisticated and Stress Attacks

Sonia Fahmy
Purdue University

Sophisticated Attacks

- ❖ From research literature
 - Pulsing attacks on TCP
 - Attacks on control media connections (SIP)
 - Attacks that increase rate slowly
 - Flash-crowd type attacks

Stress Attacks

- ❖ Attacker is familiar with the target network and with the defense and tailors the attack
 - Attacks that attempt to blend in
 - Attacks on the defense
 - Attacks on routing and DNS services

Internet Host Behavior Models

Songjie Wei, Ezra Kissel, Jelena Mirkovic
University of Delaware

Objectives

- ❖ Creating typical models of legitimate traffic from normal Internet hosts
 - For faithful simulation of Internet events
 - For realistic tests of cyber defense in DETER
 - For instant detection of host-behavior anomaly

Our Approach

- ❖ Profiling host behaviors based on Internet traffic traces
 - Public traces, packet header information only
 - Only profile frequent and active senders
- ❖ Characterizing hosts by their behaviors
 - Host features either directly obtained or derived from packet header
- ❖ Clustering hosts based on their behaviors, to reveal common behaviors
 - Unsupervised learning
 - Hierarchical clustering

Host Profiling

```
ip address  
daily_dst_num  
daily_byte_num  
average_ttl_value  
<tcp_service>  
    port1  
    ...  
</tcp_service>  
<udp_service>  
    port1  
    ...  
</udp_service>  
  
communication_similarity
```

```
<tcp_communication>  
    destination_address  
    daily_byte_num  
    daily_connection_num  
  
average_duration_time  
    <port>  
    ...  
</port>  
</tcp_communication>  
<udp_communication>  
    destination_address  
    daily_byte_num  
    daily_packet_num  
    <port>  
    ...  
</port>  
</udp_communication>  
  
</host>
```

Clustering Strategies

- Representing each cluster with a centroid
Distance of two clusters = distance between centroids
- Agglomerative algorithm

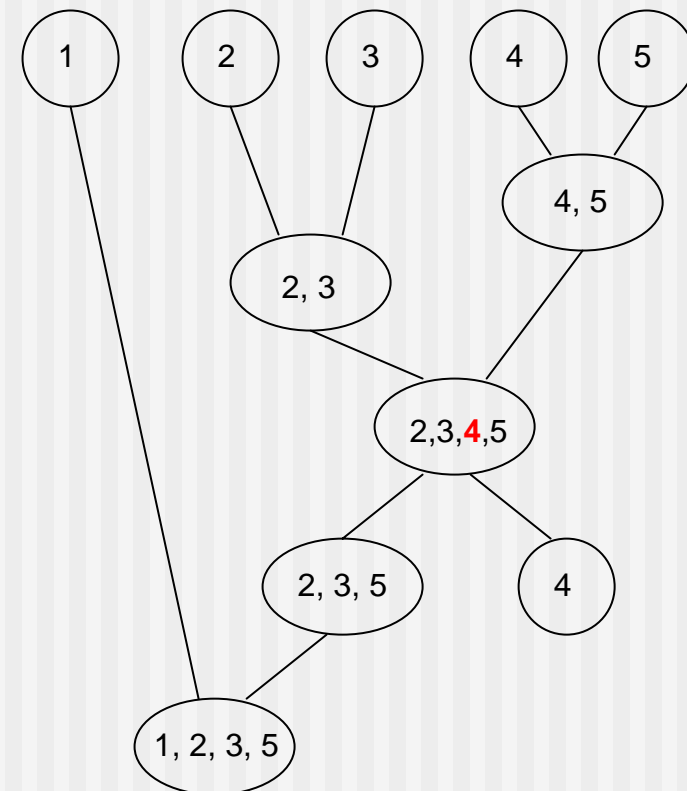
Step1: measuring distance between any two clusters

Step2: combining two closest clusters

Step3: characterizing the new cluster and expelling (randomly) conflicting hosts

Stop criteria:

- Clusters are far from each other
(min inter-cluster distance > threshold)
- Predefined number of clusters



Iterations

Sample Result

- ❖ Auckland-VIII data set from NLANR PMA
 - Two-week **anonymized** IP header trace
 - Captured in December 2003
 - Between Auckland University and the rest of the Internet
 - With 62,187 **active** host profiles
 - 0.15 minimum distance as clustering stop criterion
- ❖ In the following slides we show host models
 - Models include data for **both** the inside and outside hosts, but because of anonymization we can not tell them apart
 - Need to be distinguished for legitimate traffic models

Host Categories

Category	Percent.	Description
Servers	58%	Open services on well-known ports 11% SMTP service (top 8 clusters/88% hosts) 7.5% Web service (top 6 clusters/74% hosts) 76% DNS service (top 6 clusters/96% hosts)
NATs	9%	TTL varies a lot for the same IP address (top 7 clusters/67% hosts)
Scanners	5%	Identical outgoing traffic to different dsts (top 5 clusters/99% hosts)
Clients	28%	No open services, initiate communication (top 6 clusters/90% hosts)

Comprehensive Traffic Scenarios

Jelena Mirkovic

University of Delaware

Traffic Features

- ❖ TCP features that determine congestion response:
 - Connection window size, life, RTT, TCP type
- ❖ Application behavior
 - Sending dynamics, retransmissions or not, timeout
- ❖ Coupling of media traffic with control traffic
- ❖ “Special” packets vs “ordinary” packets

Typical Topologies

Sonia Fahmy
Purdue University

From Internet to DETER

- ❖ **NetTopology** tool (similar to Rocketfuel) collects AS topology samples
 - User can select a subset of these topologies
- ❖ **Rocketfuel-to-ns** tool converts them to DETER topology scripts
 - Link bandwidth and delay are also determined
- ❖ **RouterConfig** tool takes a topology input and produces BGP and OSPF configuration scripts
 - AS relationships are determined using [Gao01]
 - Works also with GT-ITM topologies
- ❖ Much more about this in tomorrow's session on tools and methodologies

Comprehensive Topologies

Roshan Thomas
SPARTA

Enterprise Topologies

- ❖ CISCO's three-layer model
 - Core layer provides Internet access
 - Distribution layer connects the core to the access layer, provides policy-based connectivity
 - Access layer connects individual departments and buildings
- ❖ Six major aspects of topology
 - Multi-homed or single-homed
 - Do inside machines have public Ips
 - Design of subnet and VLANs
 - Degree of redundancy at distribution layer
 - Load sharing across links and servers
 - Placement of VPNs and firewalls

Performance Metrics

Jelena Mirkovic, University of Delaware

Sonia Fahmy, Purdue University

Peter Reiher, UCLA

Roshan Thomas, SPARTA

Alefiya Hussain, SPARTA

Steven Schwab, SPARTA

Calvin Ko, SPARTA

Defense's Performance

- ❖ Main question: How well it removes DoS effect?
 - To answer this we must be able to measure DoS effect
 - More about this in the next session
- ❖ How comprehensive is the protection?
- ❖ How quickly it detects/responds?
- ❖ What is its deployment and operational cost?
- ❖ What is its security model?

Measurement Methodology

Jelena Mirkovic, University of Delaware

Sonia Fahmy, Purdue University

Peter Reiher, UCLA

Roshan Thomas, SPARTA

Alefiya Hussain, SPARTA

Steven Schwab, SPARTA

Calvin Ko, SPARTA

Measurement Methodology

- ❖ How do we scale down DDoS experiments?
- ❖ How do we aggregate results from various tests?
 - It would be useful to produce a single number that expresses the defense performance
 - Misleading, since we are comparing apples and oranges
 - Is a defense that protects against 90% of attacks completely better than a defense that protects against 100% attacks but inflicts collateral damage to HTTP traffic?

Conclusions

- ❖ Many factors influence an attack's impact and a defense's performance
- ❖ We have just scratched the surface, much work remains to be done
- ❖ Since DDoS attacks are adversarial, comprehensive evaluation is a necessity, so we must carefully design all test aspects

Questions? Comments?
sunshine@cis.udel.edu