Bots, DDoS and Ground Truth

One Year and 5,000 Operator Classified Attacks

Craig Labovitz

Chief Scientist, Arbor Networks

in collaboration with Jose Nazario

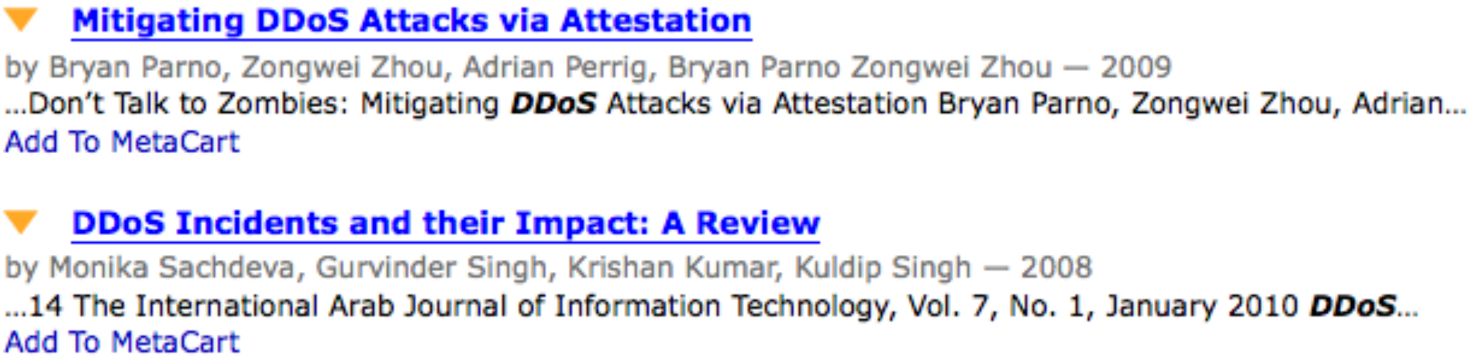
[labovit, jose]@arbor.net

This Talk

- **“Ground-truth” about security is hard…**
  - True in enterprise
  - But especially so in carrier / national infrastructure
- **Most infrastructure attacks go unreported**
  - Less than 5 percent surveyed ISPs reported one
- **Significant anecdotal reports / surveys**
  - including Arbor, Cisco, etc.
- **But no validation**
  - e.g. do providers really know the size of botnets?

This Talk

- **Also no shortage of research**
  - 100+ published papers and counting



▼ Mitigating DDoS Attacks via Attestation

by Bryan Parno, Zongwei Zhou, Adrian Perrig, Bryan Parno Zongwei Zhou — 2009

...Don't Talk to Zombies: Mitigating ***DDoS*** Attacks via Attestation Bryan Parno, Zongwei Zhou, Adrian...

Add To MetaCart

▼ DDoS Incidents and their Impact: A Review

by Monika Sachdeva, Gurvinder Singh, Krishan Kumar, Kuldip Singh — 2008

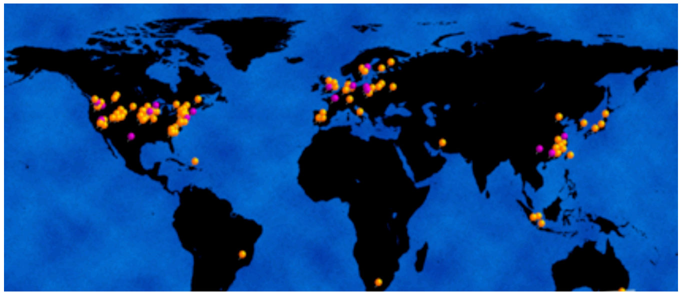
...14 The International Arab Journal of Information Technology, Vol. 7, No. 1, January 2010 ***DDoS***...

Add To MetaCart

- **But almost all research lacks “ground-truth”**
  - Papers compare success only with <u>previous</u> papers

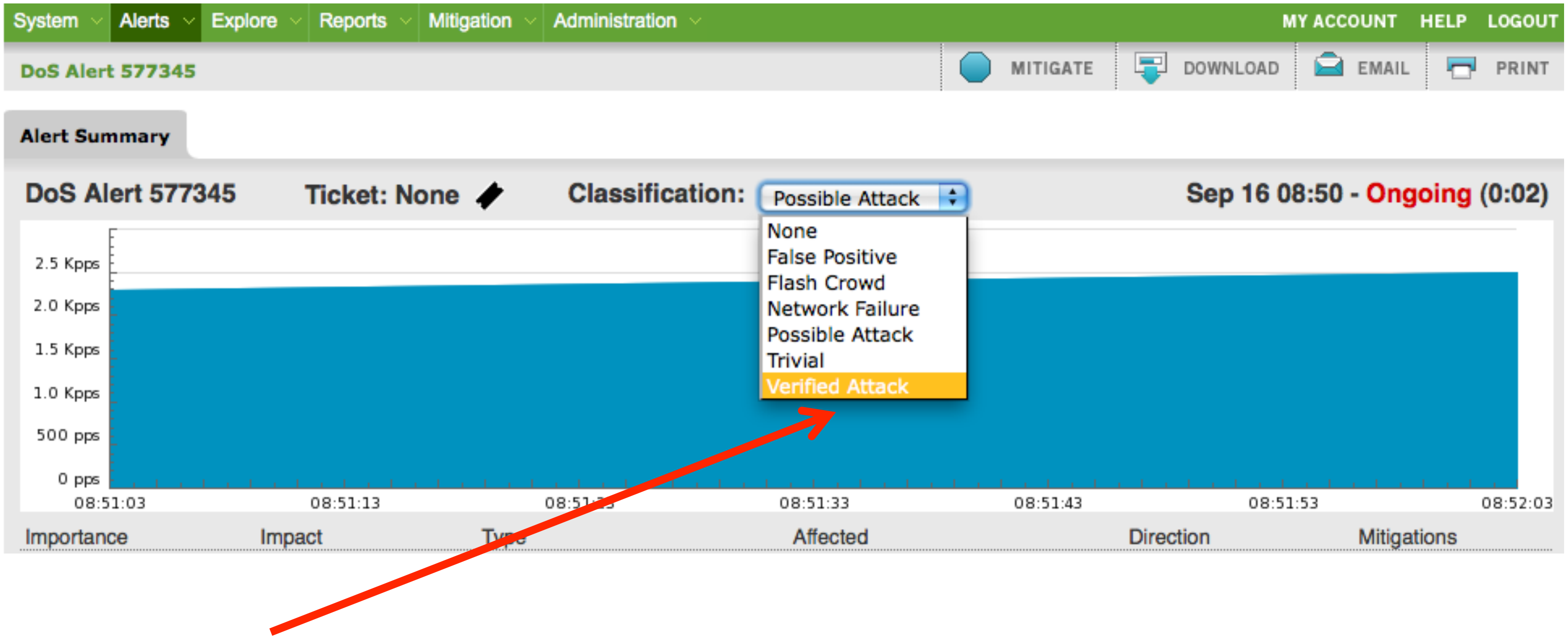
- **Usually impossible to distinguish**
  - “Anomalies” and false positives from true attacks
  - Or compare confirmed security events across providers

This Talk



- **Again leverage ATLAS**
- **But this time focus on security**
  - Added manual classification of events two years ago
- **Gather data from**
  - 37 ISPs over last 12 months
  - Really two overlapping datasets (alerts, mitigations)
    - Not all confirmed attacks are mitigated
    - Not all mitigated attacks have an associated alert
- **And more than 5,000 operator classified events**

Manual / Operator Classification of Events

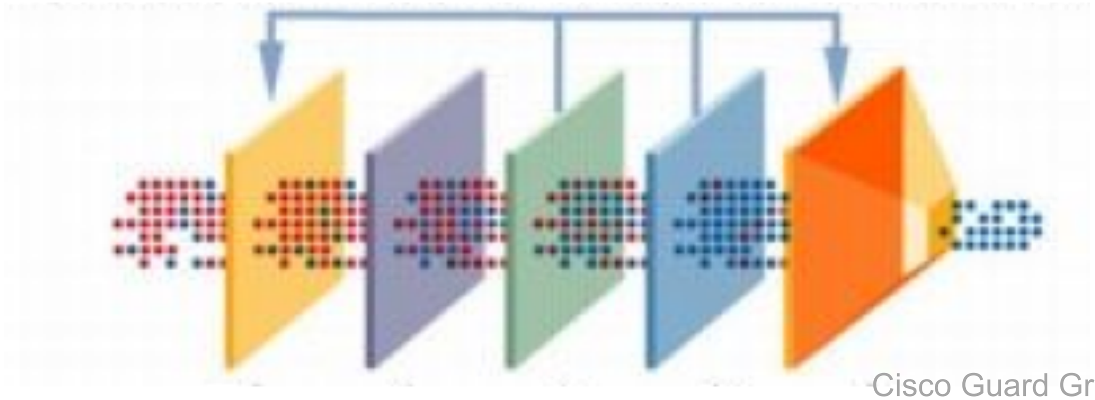


DoS Alert 577345

DoS Alert 577345 Ticket: None Classification: Possible Attack Sep 16 08:50 - Ongoing (0:02)

- Classification of events part of workflow
- Dataset also tracks which events were mitigated
- Goal is research as well as commercial evaluation

Background on Mitigation Data



Cisco Guard Graphic

- **Talks avoids detailed discussion of countermeasures**
- **Most similar technology uses multiple layers**
  - IP validation
  - Botnet IP detection
  - TCP validation
  - Application validation (HTTP, DNS, SIP)
  - Policy (GeoIP, ASN, baselines, filters)
- **Track bandwidth and number IPs caught by each layer**
  - Also connections (different from bps)

Conventional Wisdom

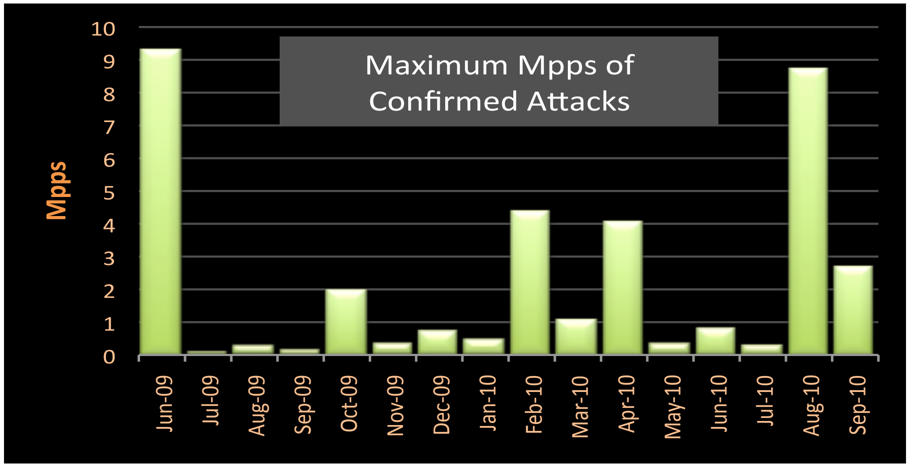
1. **Spoofing DDoS sources no longer common**
   - Botnets so large that no need to spoof

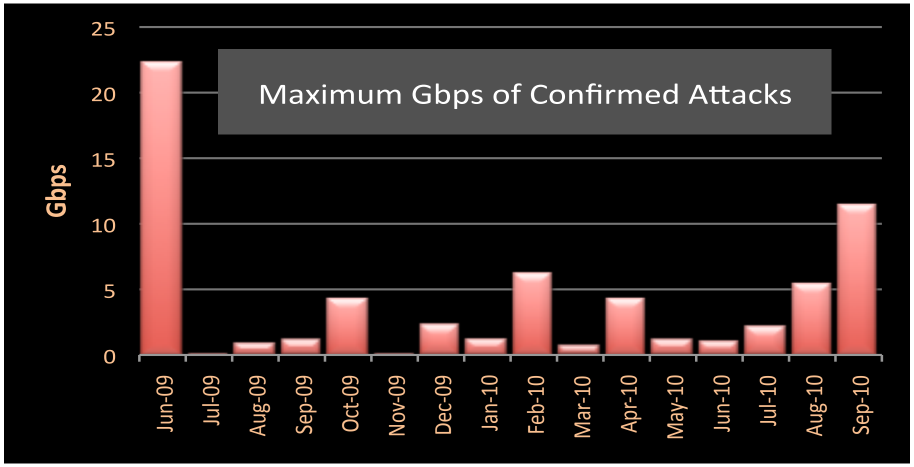
2. **Most botnets used in DDoS are large**
   - e.g. thousand or tens of thousands of hosts

3. **Most DDoS use brute force flooding**

Two of the three claims are not true…

Overview of Data

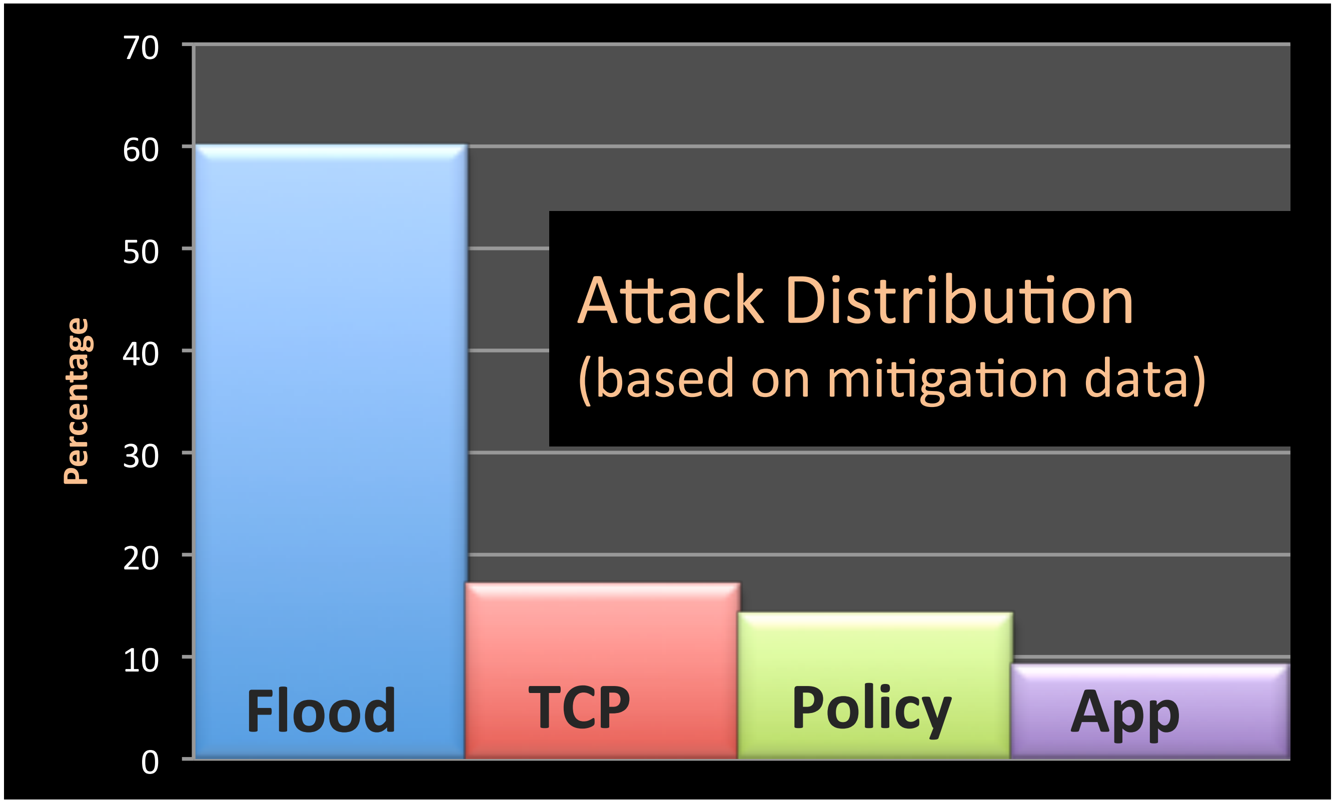




- **Preliminary study**
  - Too small to draw industry-wide conclusions
  - Wide variation of data sources (e.g. from commercial service to protect global online infrastructure)

- **But provides useful initial insights**
  - First large scale study of operator confirmed data
  - Includes 22 Gbps / 9 Mpps (smaller than 45+ Gbps)

Overview of Data



- **Majority of attacks flooding (60%)**
- **Followed by TCP (17%), Policy (12%) and Application (8%)**
  – Policy includes GeoIP, ASN, regular expression, etc. rules

Overview of Attack Statistics

	Mbps	Pps	Hours
Median	34	33,519	0.8
Average	349	258,347	3.1
95th	1,055	703,438	10.7
max	22,000	12,354,606	114.6

- **Average attack 300 Mbps and 200 Kpps**
  - Mean skewed by high-end attacks
  - Median is 30Kpps (relatively effective small server)
- **Largest attack**
  - 22 Gbps / 9 Mpps IP fragment
  - 4 days and targeting one /32

Flooding Attacks

	Zombie IPs	Avg Mbps per IP	Avg Pps per IP
Median	33	11	6,021
Average	80	162	48,084
95th Percentile	311	731	124,946
Max	1,286	4,327	891,737

- **Most attacks involve relatively few unique src IPs**
  - Median is 33 IPs generating aggregate 200kpps
- **Sources (if real) are well-connected**
  - Average is 162 Mbps and 48 Kpps per source IP
- **Unrealistic per IP traffic**
  - 4Gbps per same IP!
  - Suggests IP spoofing (dumb tools) or mega-proxy
  - About 10% of attacks fall into this category

TCP / Spoofing Attacks

	Connections per Second	Validated Hosts
Median	77,418	1
Average	223,431	27
95th	880,326	191
Max	1,710,676	268

- **Significant rate of bogus TCP connections**
  - Orders of magnitude gap between connection attempts and validation hosts
- **Suggests**
  - Significant levels of spoofing
  - Or incomplete client attack stacks

Application Attacks

	Mbps	Src IPs	Hours
Average	4	231	52
95th	10	414	119
Max	90	6,983	1488

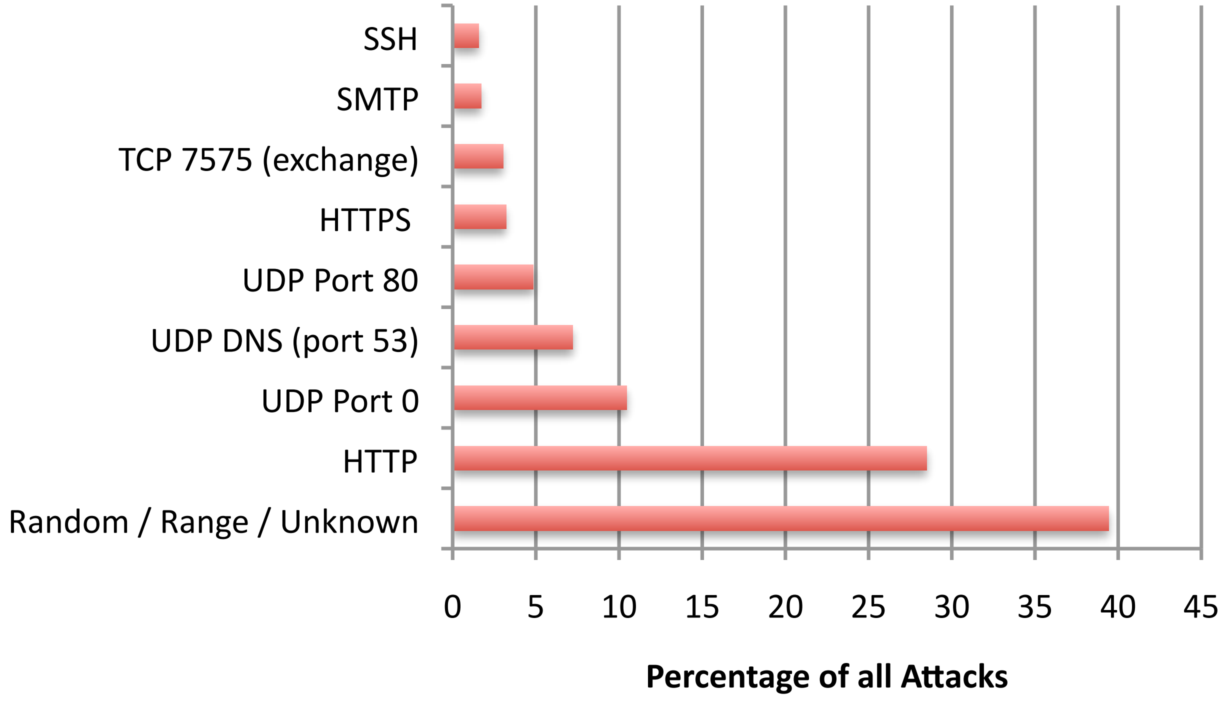
HTTP Attacks

	Mbps	Pps	Hours
Average	40	11,913	3.4
95th	262	77,516	7.1
Max	385	114,216	9.1

SIP

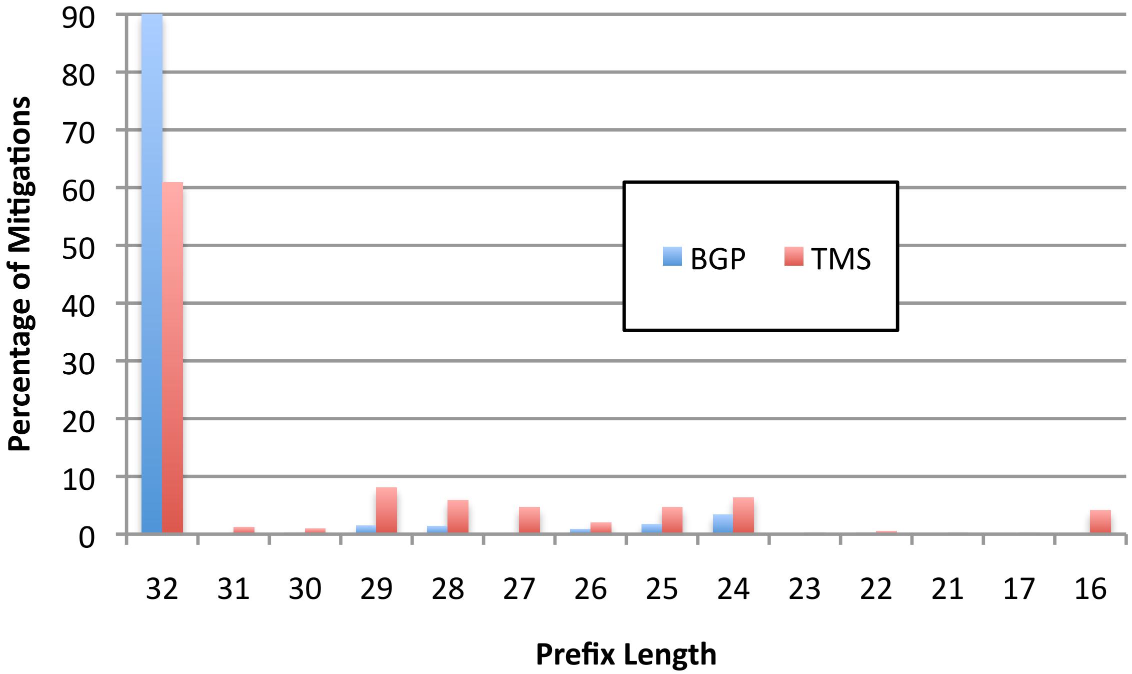
- **Approximately 8% of attacks application**
  - HTTP URLs, DNS, SIP
- **HTTP attacks relatively low bps / pps**
  - 95th is 1.2Kpps from 414 hosts
  - Generally focused on expensive back-end computation
  - But many constant attacks over days, weeks
- **SIP tends to more resemble flooding attacks**

Distribution of Targets



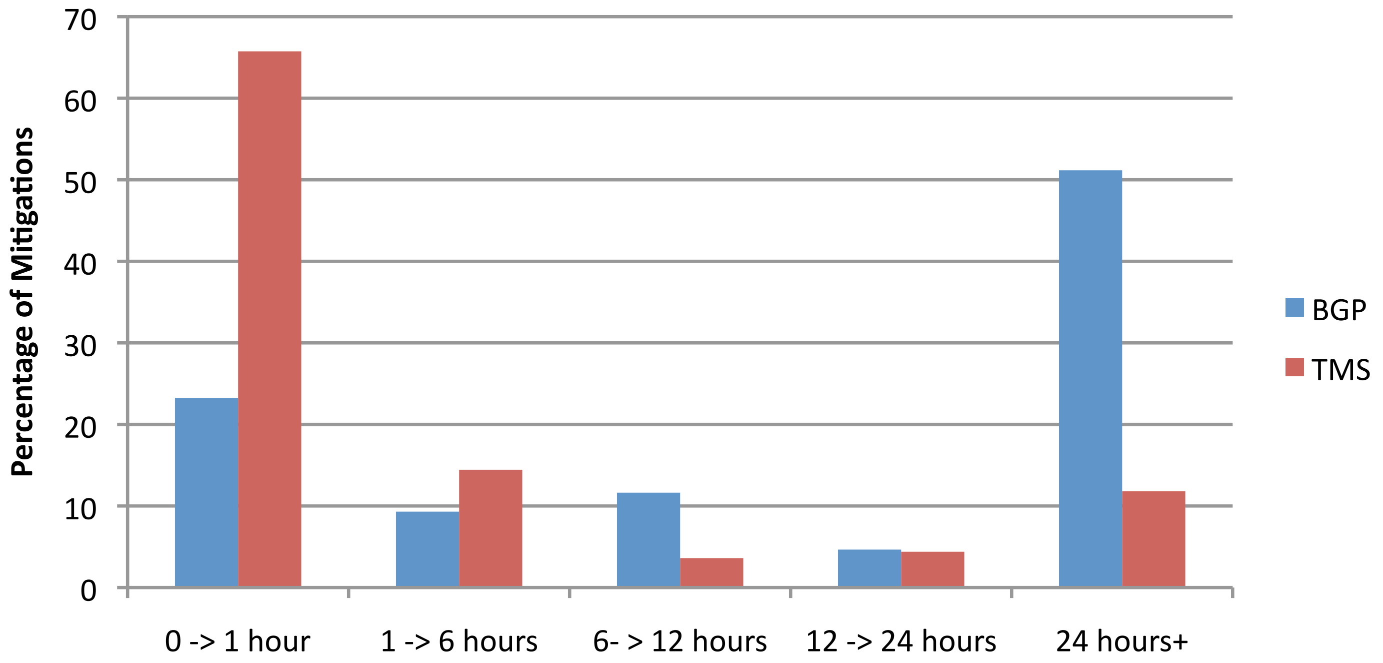
- **Growing number of services use port ranges**
  - Also represents randomized / flooding attacks
- **Historic victims of DDoS remain unchanged**
  - Web, DNS, Mail

Distribution of Target Size



- **Most attacks (and resultant mitigations) for /32**
  - Commonly representing LB / NAT infrastructure
- **Especially so for BGP blackhole**
- **TMS appears to be used for infrastructure protection and covers wider CIDR range**

Distribution of Durations



- **Approximately half mitigations short-lived**
  - less than an hour
- **Heavy tail with some attacks lasting multiple days**
  - And some constant (mitigations running months)
- **Once BGP blackhole begins, likely to remain for days**
  - Some providers have hundreds ongoing

Observations

- **Preliminary analysis**
  - One of first validated attack / mitigation data sets
- **Suggests**
  - Spoofing is still prevalent in DDoS
  - Most attacks involve hundred or fewer hosts
  - Hosts are well-connected (or bad tools)
  - Significant incidence of application / service attacks
- **Real goal of this talk is to encourage participation**
  - Less than 1/4 have enabled anonymous statistics
  - Data is useful for community / research
  - Please participate

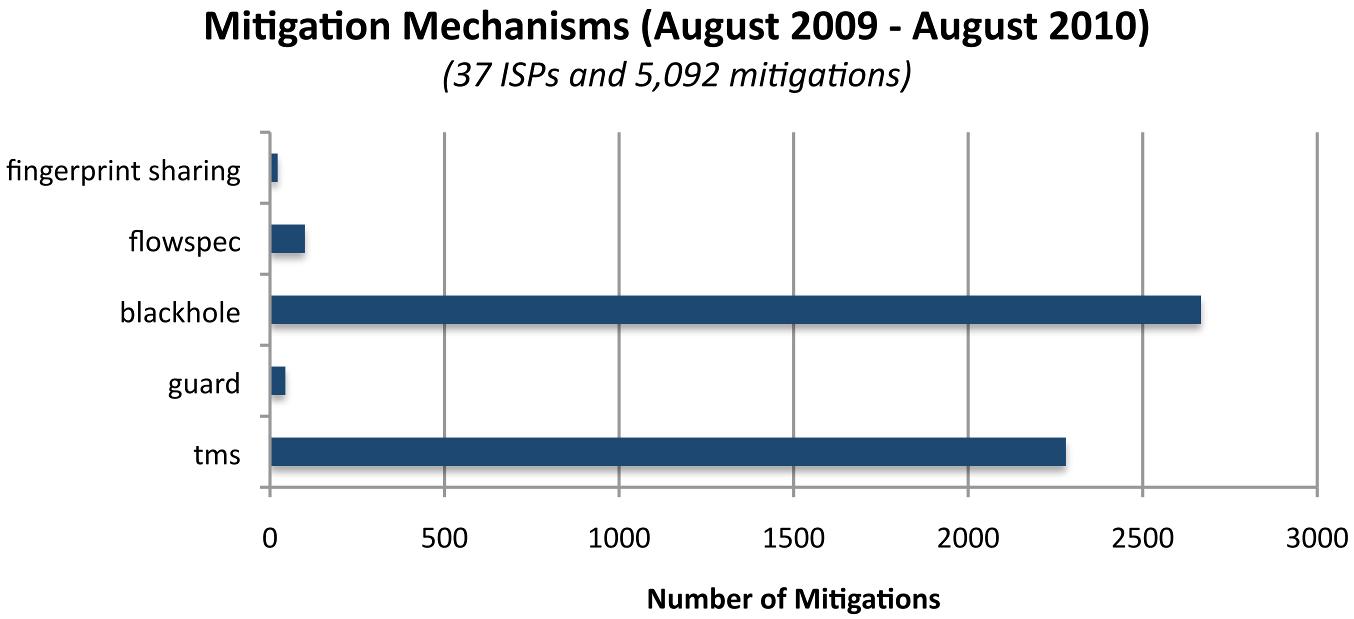
Questions

labovit@arbor.net

http://www.monkey.org/~labovit

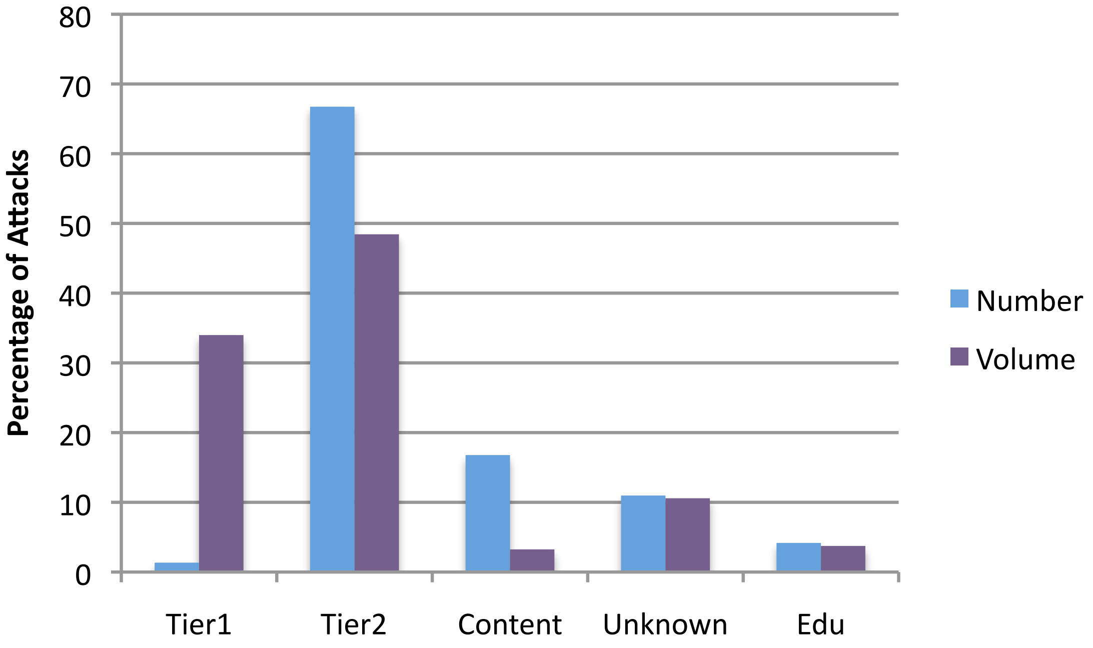
EXTRAS

Mitigation Mechanisms



- **Biased dataset**
  - Undercount Guard and Blackhole (not visible via ATLAS)
- **Even with bias, blackhole dominates as preferred mitigation**
  - No visibility into whether src or dst blackhole
- **At least four ISPs using flowspec**
  - (I had only been aware of two)

Where are the Attacks?



- **Tier2 dominate both by volume and number of attacks**
- **When Tier1 is attacked, attack is large**
- **Again this is a preliminary / small dataset**