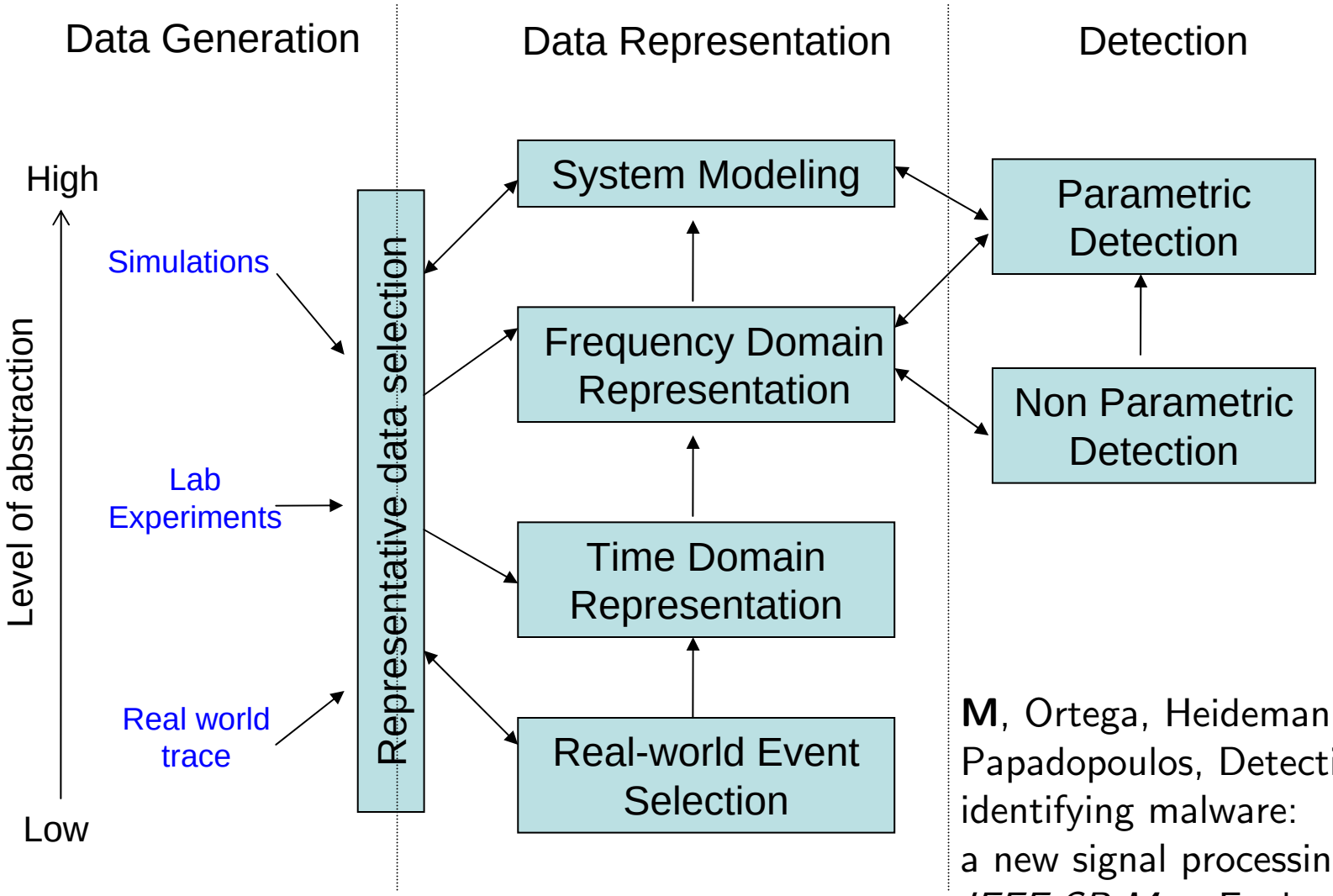
Signal Processing Methods for Denial of Service Attack Detection

Urbashi Mitra
Ming Hsieh Department of Electrical Engineering
Viterbi School of Engineering
University of Southern California
Los Angeles, CA
ubli@usc.edu

with **Xinming He** (Cisco), John Heidemann (ISI), **Gautam Thatte** (USC), Antonio Ortega (USC), and Christos Papadopoulos (CSU)

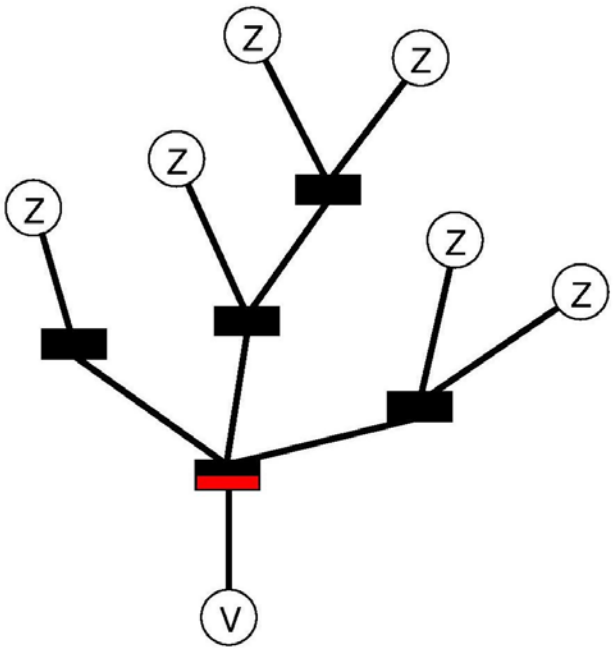
thanks to: NSF CNS-0722073 and Cisco

``Vision''



M, Ortega, Heidemann, & Papadopoulos, Detecting and identifying malware: a new signal processing goal, *IEEE SP Mag.* Exploratory DSP column, 9/06

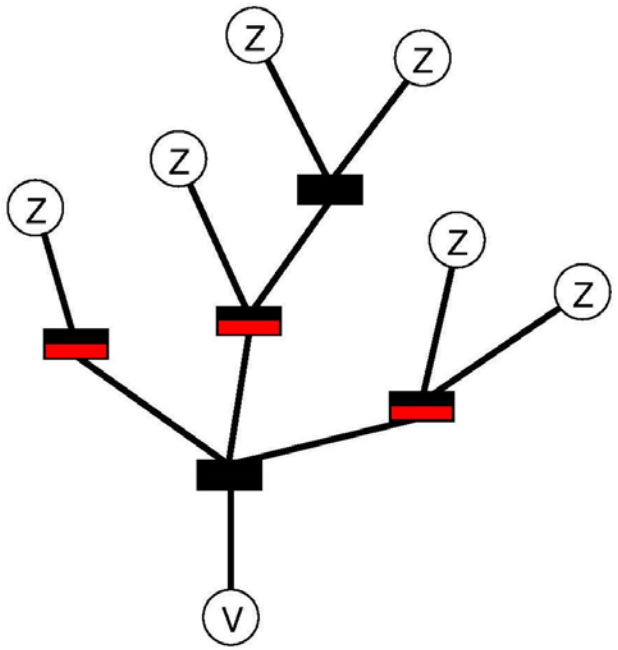
Low-rate Denial-of-Service (DoS) Attacks



Early detection of DDoS attacks, link congestions and traffic anomalies

- Victim overwhelmed with traffic
  → *easy detection*

Low-rate Denial-of-Service (DoS) Attacks



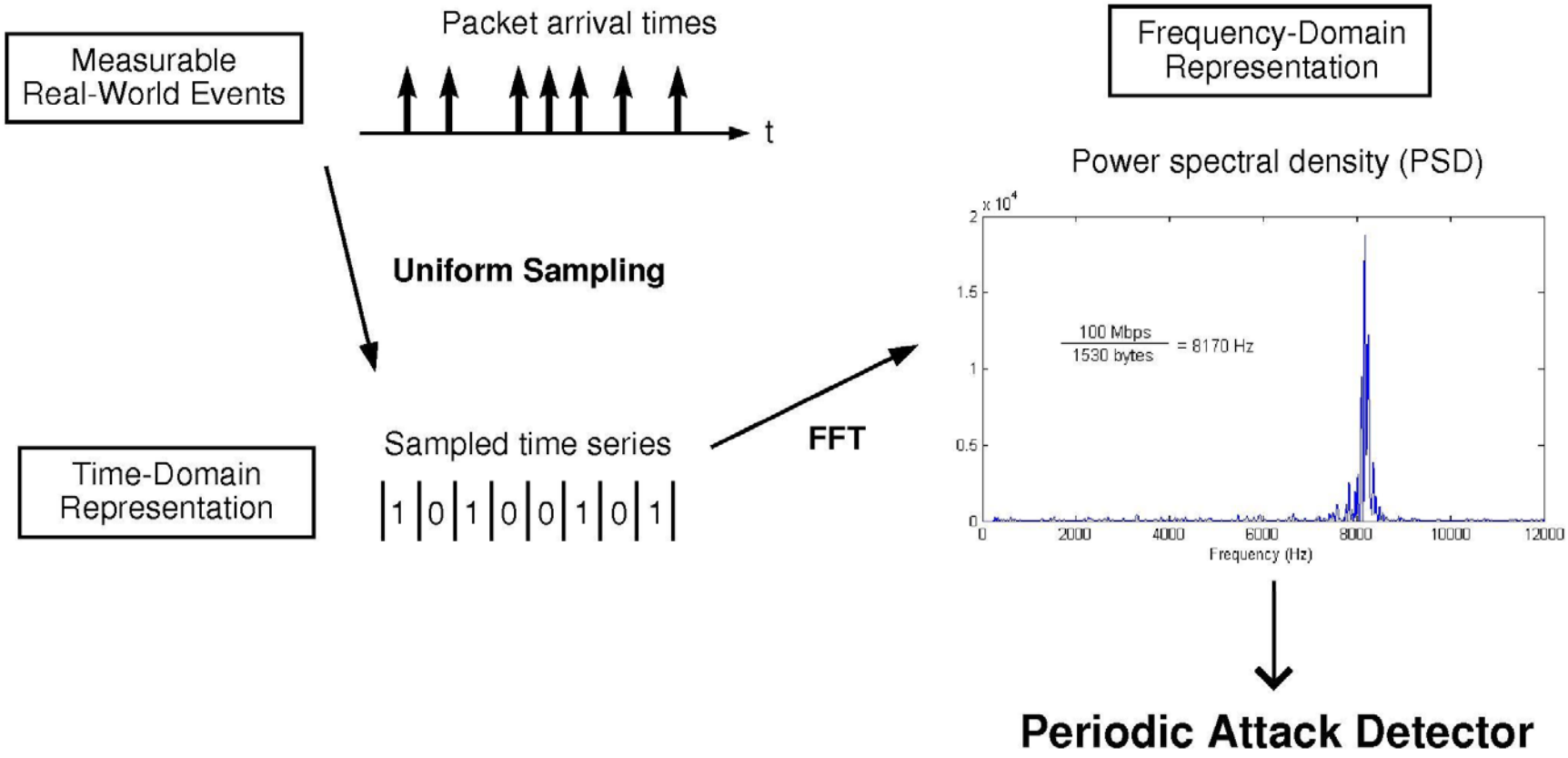
Early detection of DDoS attacks, link congestions and traffic anomalies

- Victim overwhelmed with traffic
  → *easy detection*
- Detect low-rate attacks *further back* in the network

The Case for a Spectral Approach

- A rich set of periodic patterns in Internet traffic
  - Congestion along bottlenecks
  - DoS attack streams
  - TCP windowing behavior
- Applications for charactering periodicities
  - Better understanding of traffic dynamics
  - Detecting bottlenecks, DoS attacks, useful for traffic engineering, capacity planning, and network security
- Spectral techniques appear promising
  - Mature, and widely used in many other fields
  - Recently been applied to Internet traffic, e.g. detecting DoS and network anomalies

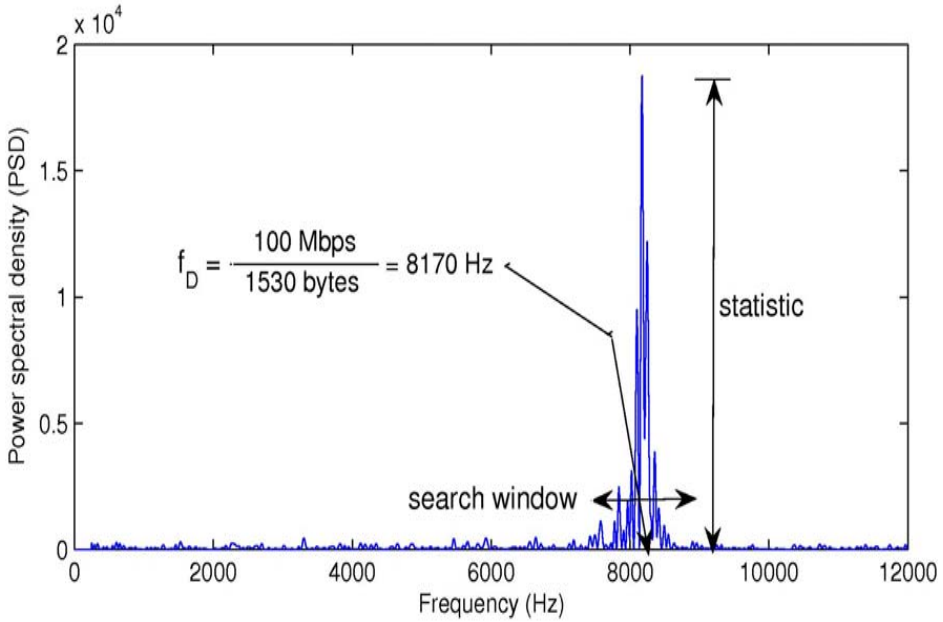
Time Series Representations



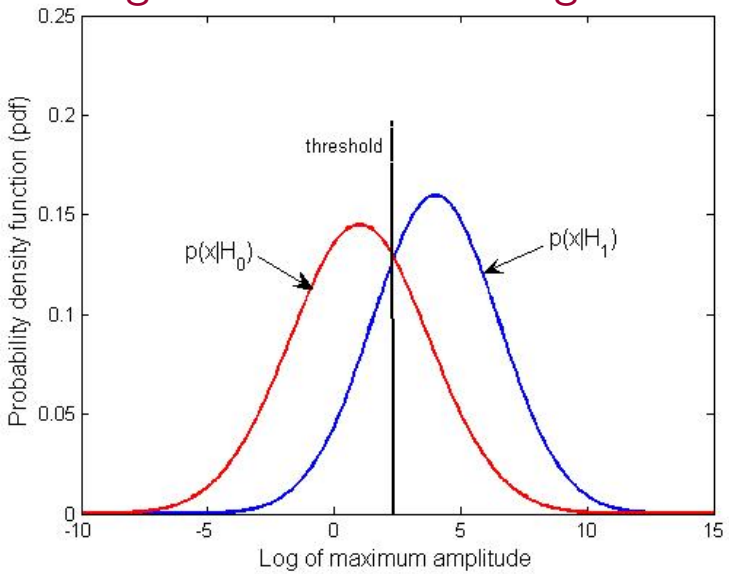
Hypothesis Testing

- Maximum Likelihood Detection
  - determine empirical statistics of key features
  - binary hypotheses test: H_0 (without bottleneck) and H_B (with bottleneck)
- Classify a new trace as either H_0 or H_B via

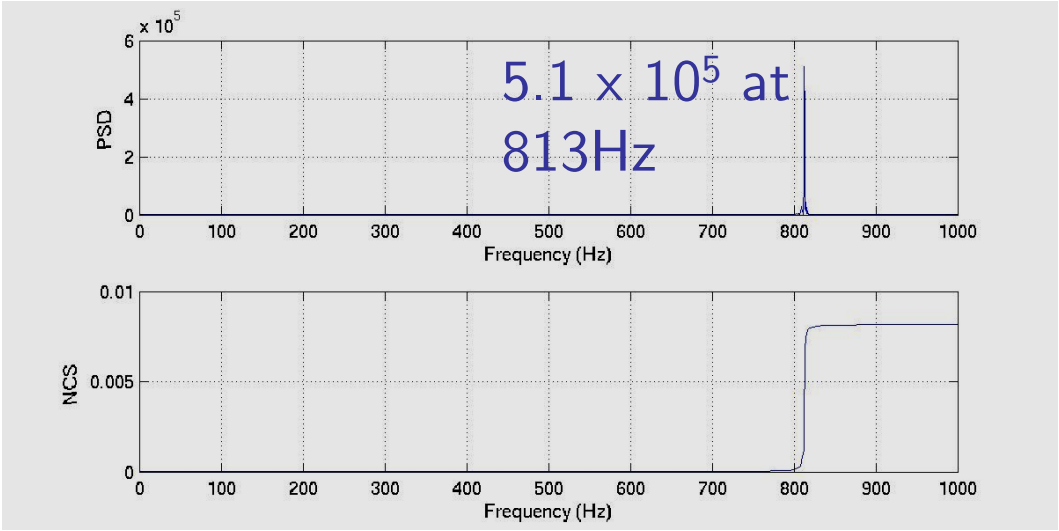
  if $Pr(x/H_0) > Pr(x/H_B)$, *select* H_0 *otherwise, select* H_B
- Our features and statistical models, optimal test is a **threshold** test
  - feature > threshold → bottleneck exists



log-normal for FFT magnitude



Bottleneck Signatures



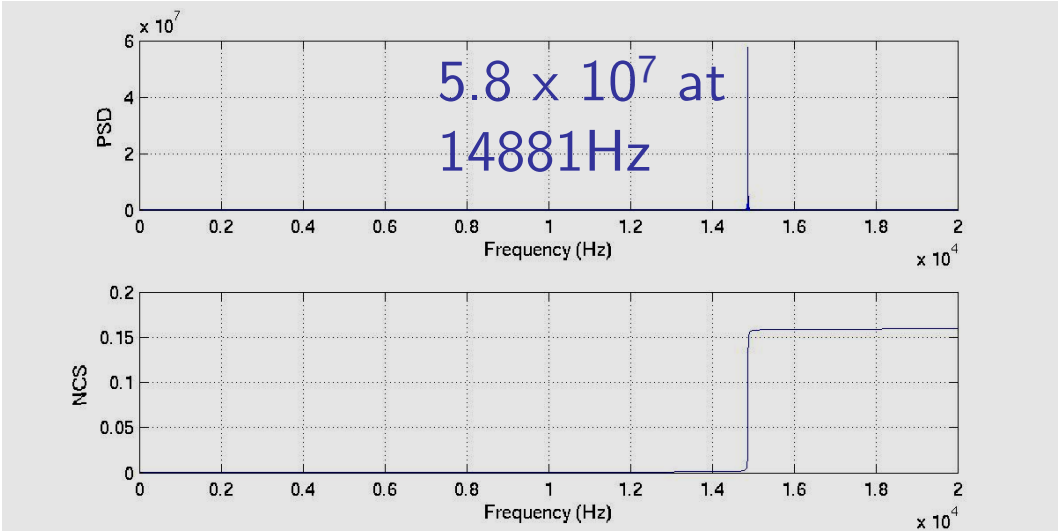
TCP flow through 10Mbps link

- Connected through a 10Mbps switched hub
- A TCP flow with 1500-byte packets
- No cross traffic

- Strong energy around the Base Frequency

$$= \frac{\text{Link Bandwidth}}{\text{Packet Size}} = \frac{\text{10Mbps}}{(1500+38) \times 8\text{b}} = 812.74\text{Hz}$$

DoS attack signatures



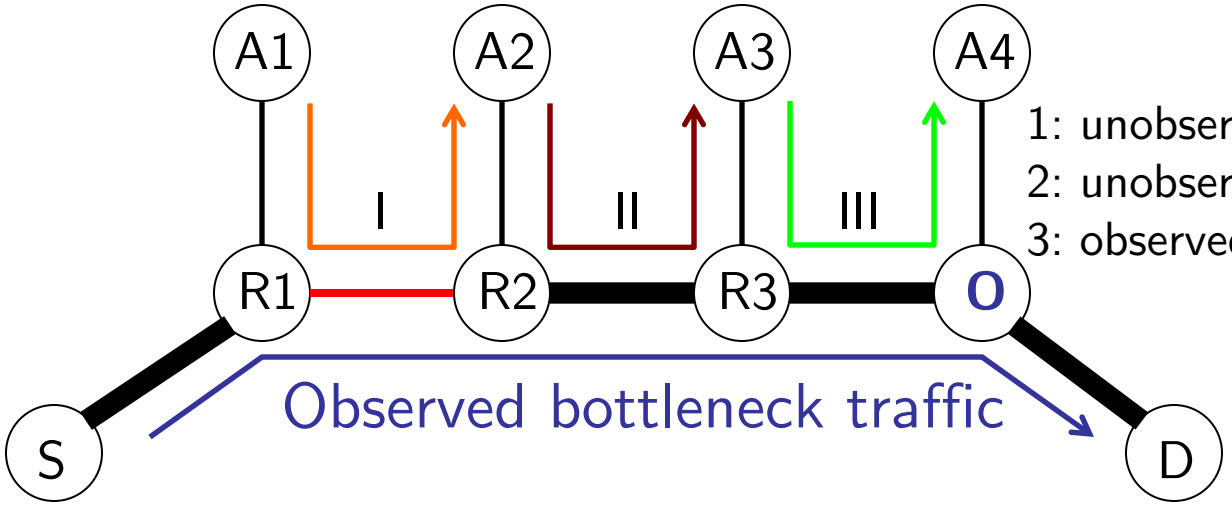
spectrum of an Mstream attack stream

- Connected through a 10Mbps switched hub
- Mstream sends 40-byte TCP packets as fast as possible
- No cross traffic

- Much higher frequency component

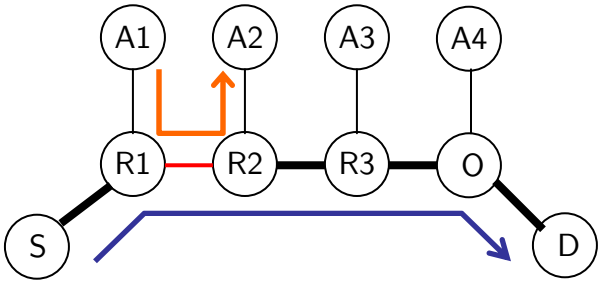
$$= \frac{\text{Link Bandwidth}}{\text{Packet Size}} = \frac{10\text{Mbps}}{(46+38)\text{ x }8\text{b}} = 14880.95\text{Hz}$$

Impact of Cross Traffic



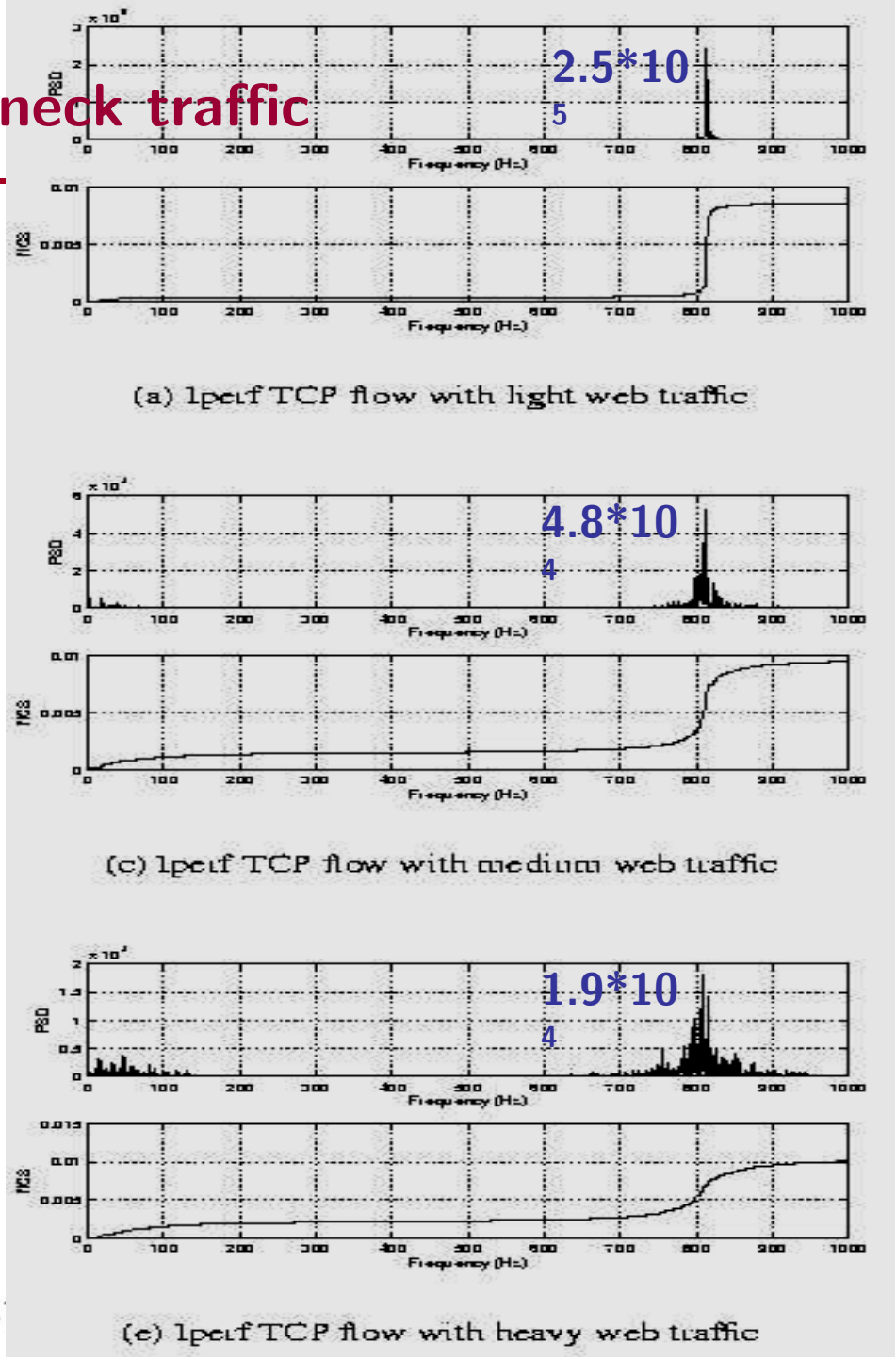
- Unobserved non-bottleneck traffic
  - Least impact on the observed traffic spectrum
- Observed non-bottleneck traffic
  - Introduces its own frequency component to the observed traffic spectrum
- Effects
  - changing spectral peak width
  - shifting peak

Type I: unobserved bottleneck traffic



Web Traffic	Iperf TCP Flow
0.82Mbps, 85pps	8.93Mbps, 745pps
3.56Mbps, 379pps	6.15Mbps, 513pps
5.19Mbps, 606pps	4.45Mps, 371pps

- weaker peak amplitude
- wider spread of energy
- changing location of the peak amplitude

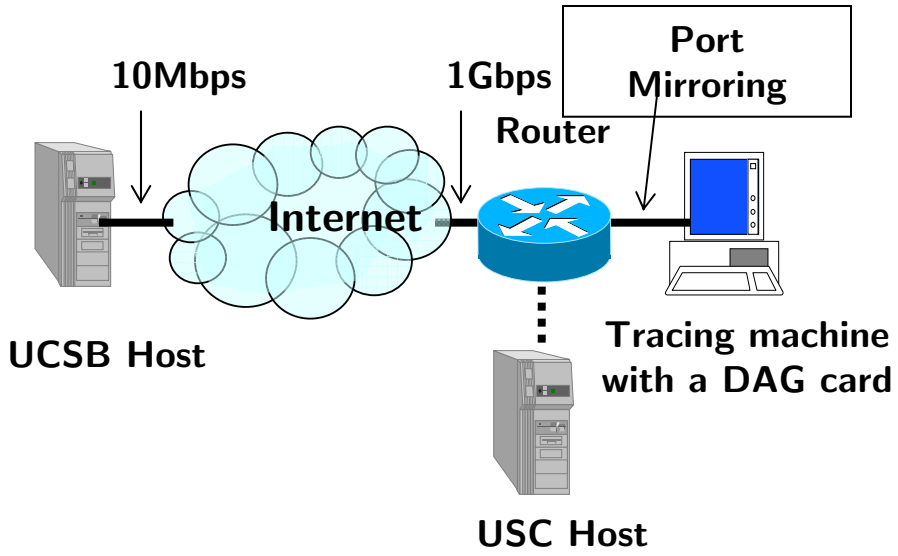


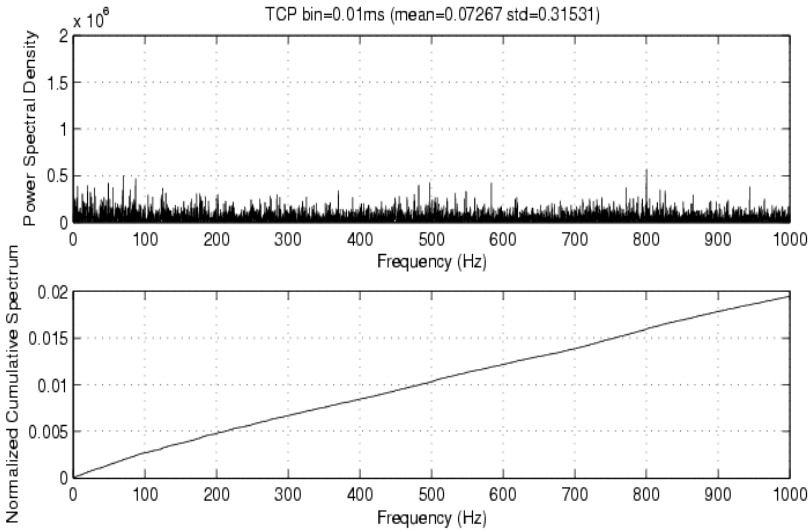
(a) Iperf TCP flow with light web traffic

(c) Iperf TCP flow with medium web traffic

(e) Iperf TCP flow with heavy web traffic

Wide Area Network Experiments

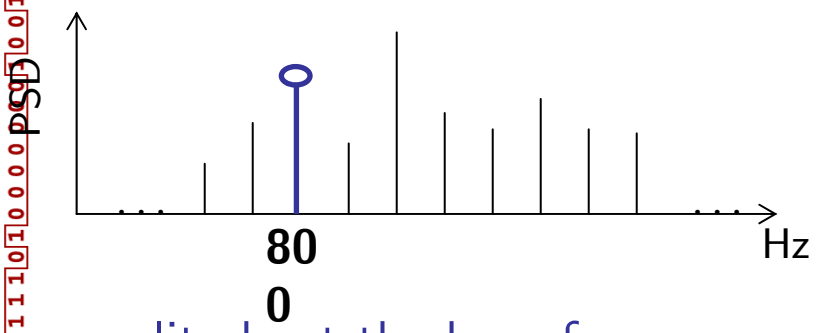




- With bottleneck flow
  - the aggregate has larger amplitudes around predicted base frequency
- Periodic patterns have unique signatures
  - High amplitude around the predicted base frequency
- Cross traffic introduces noise to the spectral signature of periodic patterns
  - The signature is still visually detectable in most cases
  - Need automatic detection

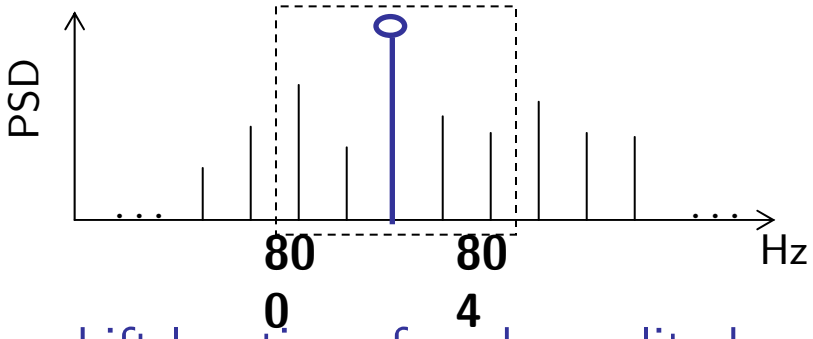
Detection Features

1. Single-Frequency Alg.



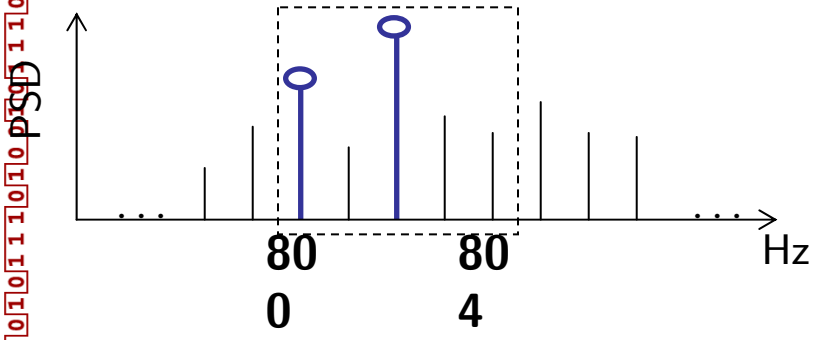
amplitude at the base frequency

2. Top-Frequency Alg.



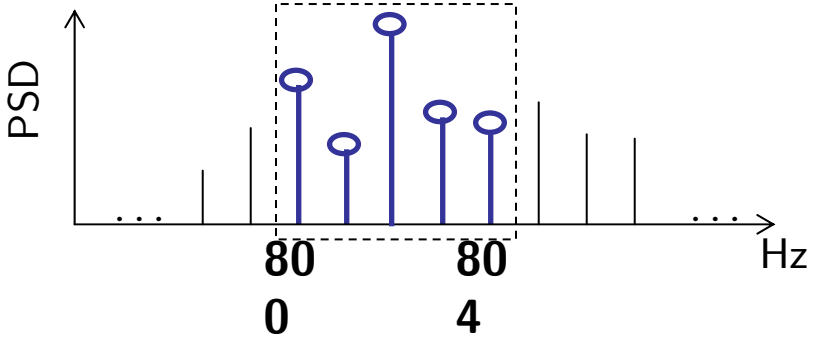
shift location of peak amplitude

3. Top-M Frequencies Alg.



more features → better accuracy

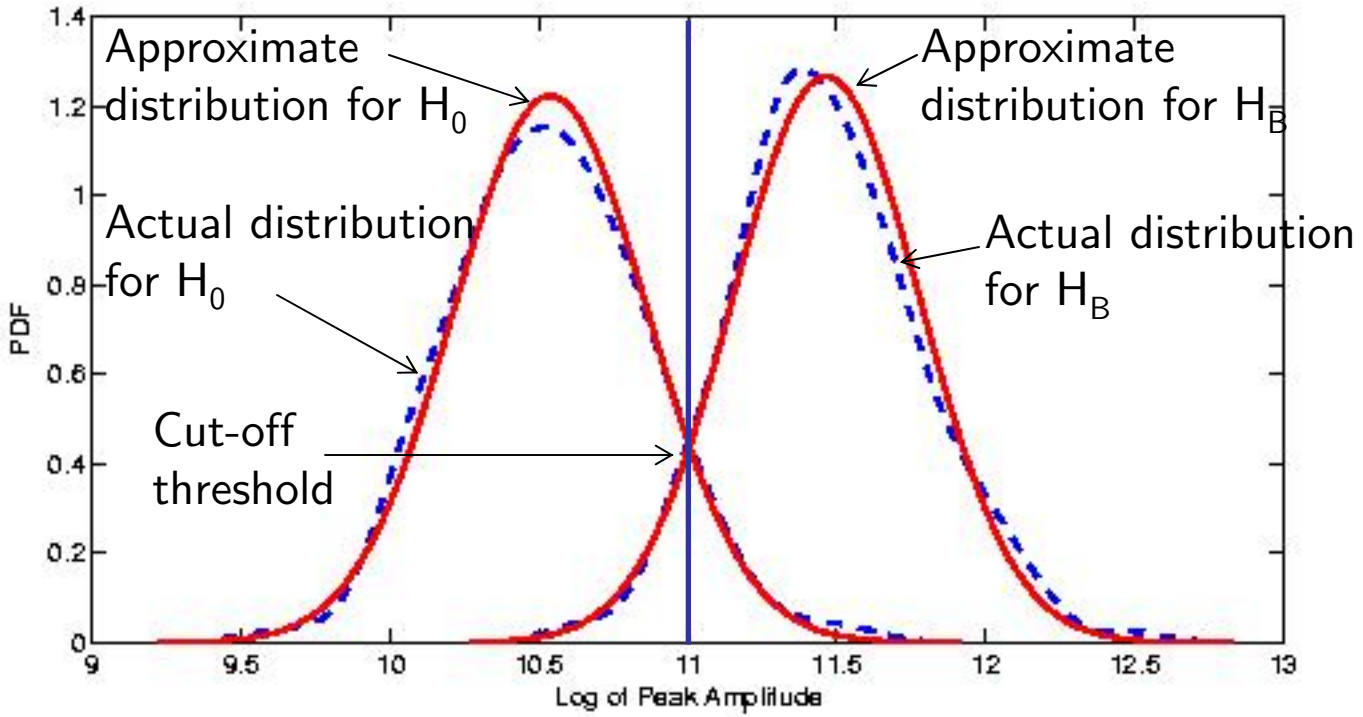
4. All-Frequencies Alg.



all features → best accuracy

Top Frequency Algorithm

- Approximate PDFs with log-normal distributions
  - Good fit with simple parameters

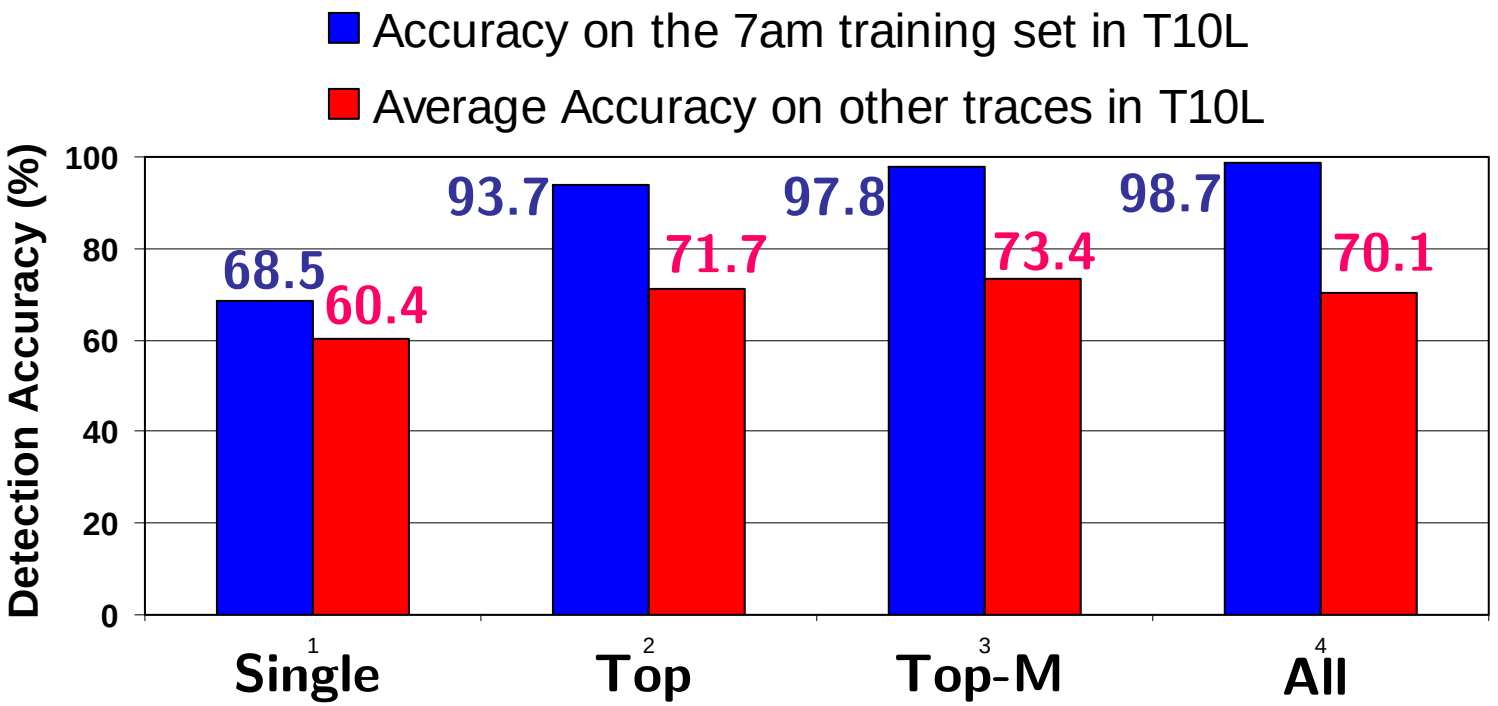


Peak amplitude in [780Hz, 800Hz] for a trace pair in T10L

Experimental Setup

- Variables in experiment setup:
  - Link bandwidth: 10Mbps, 100Mbps
  - Cross traffic volume: low, high
  - Protocol: TCP, UDP
- Four Experiment Scenarios:
  - U10L: 10Mbps UDP flow, low cross traffic
  - T10L: 10Mbps TCP flow, low cross traffic
  - T10H: 10Mbps TCP flow, high cross traffic
  - T100H: 100Mbps TCP flow, high cross traffic
- Pairs of 5-minute long traces
  - One pair every 2 hours for 24 hours
  - One pair for training, and the other for evaluation

Performance Comparisons



- Impact of different parameters
  - Window location, size, Training data, S/N ratio
- Sensitivity to training matching (what time of day)

He, Papadopoulos, Heidemann, **M**, & Riaz , Remote Detection of Bottleneck Links Using Spectral and Statistical Methods, Computer Networks, accepted 9/08

Drawbacks of Alternative Methods

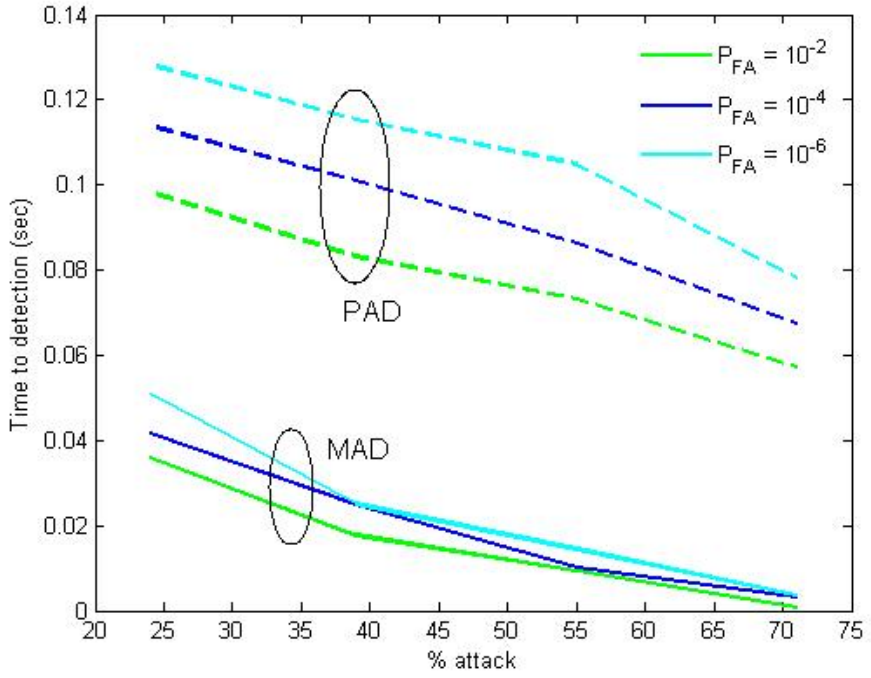
- Packet contents used to detect attack
  - Entropy of source/destination port # and IP address [Feinstein03, Lakhina05, Wagner05]
  - Filter packets using SYN flag in TCP header [Wang02]
  - Detection using TTL field [Jin03, Rodriguez07]

- Aggregate traffic analyzed, but
  - Requires empirical tuning of parameters [Tartakovsky06]
  - Computational overhead [Barford02, He08]
  - sensitivity to training data

Modeled Attack Detector

- Previous spectral method very sensitive to matching between training data and detection data
- An alternative approach
  - apply simple statistical models to traffic
  - estimate key parameters
  - implement a sequential probability ratio test (detect as you)
  - Thatte, **M** & Heidemann, "Detection of Low-Rate Attacks in Computer Networks," IEEE Global Internet Symposium, 4/08

- Poisson/shifted Poisson model (MAD)
  - train for no attack
  - estimate on-line for attack
  - **can't do purely on-line**
  - **issues with false alarms**
- SPRT version of spectral (PAD)



A New Model – back to Hypothesis Testing

- Choose between

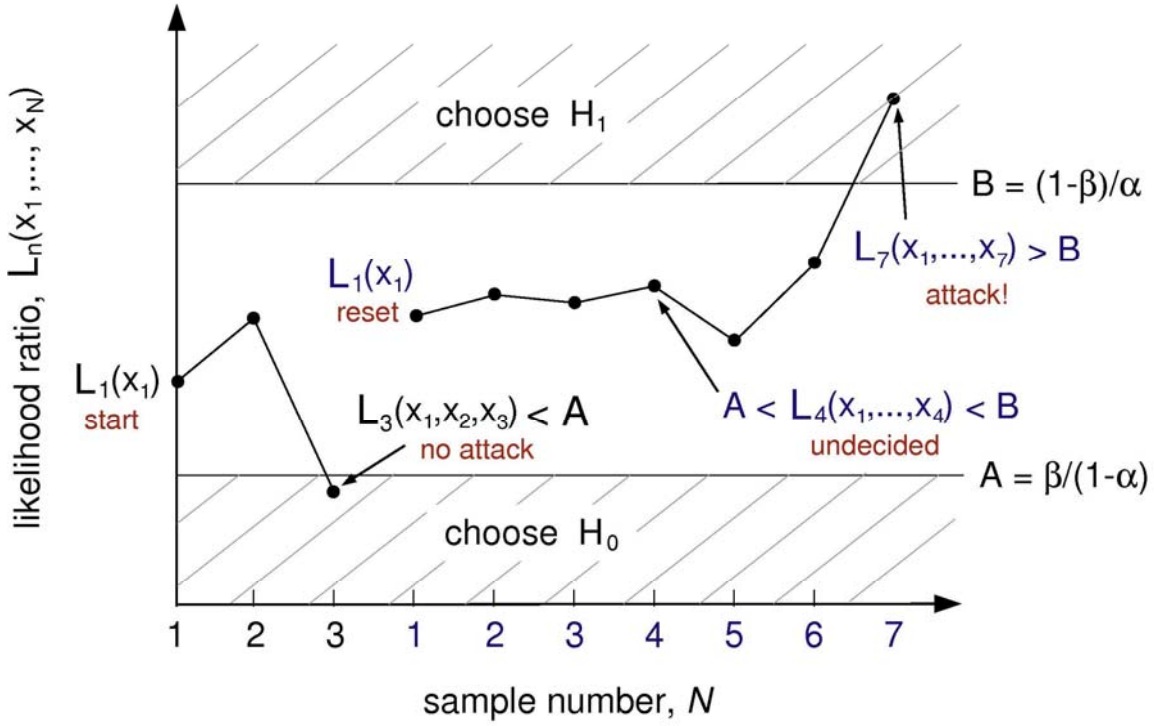
$$H_1: \quad \text{Presence of an attack in traffic}$$
$$H_0: \quad \text{No attack}$$

- Focus on two types of error
  - False positive, $P_{FA} = \alpha$
  - False negative, $P_M = \beta$

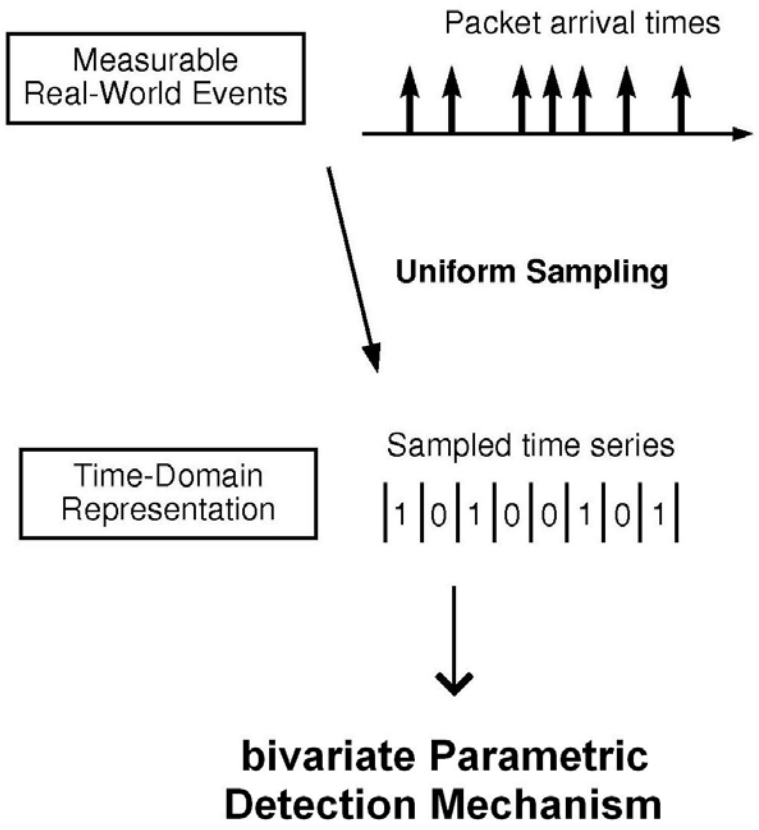
- Performance criteria defined using these two quantities
  - exact design challenging for sequential detector

Sequential Probability Ratio Test (SPRT)

- Given N iid observations $\{x_1, x_2, \cdots, x_N\}$
- Likelihood ratio $L_N(\mathbf{x}) = \frac{p(x_N|H_1)}{p(x_N|H_0)} L_{N-1}(\mathbf{x})$

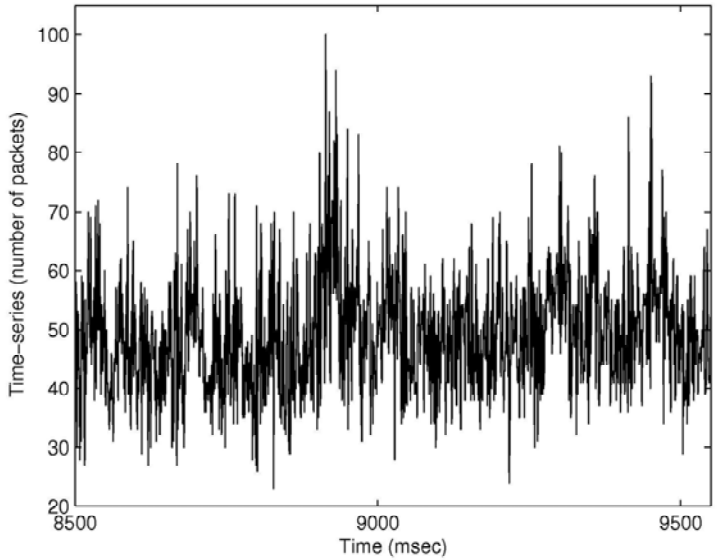


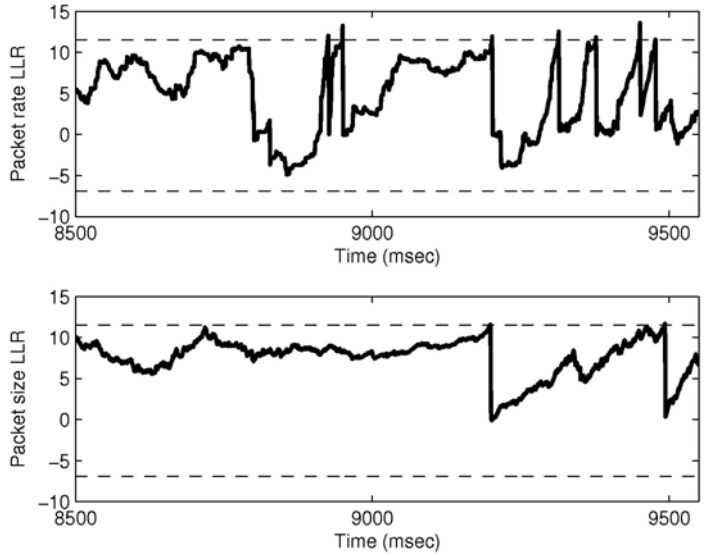
bivariate Parametric Detection Mechanism (bPDM)



- Operates on aggregate traffic with no flow-separation
- No dedicated training-phase
- Develop parametric models for packet rate and sample entropy of packet size distribution

bPDM detects low-rate attacks





- 6% synthetic TCP SYN attack starts at 9.0 seconds; not visually distinguishable
- Both packet rate and packet size SPRTs *simultaneously* cross threshold to detect attack

The packet rate model

- Background packet rate modeled using generalized Poisson distribution [Consul89]

$$p(x|H_0) = \theta(\theta + \lambda x)^{x-1} e^{-\theta-\lambda x}/x!$$

$x = 0,1,\ldots$ is number of packet arrivals

- Attack is modeled as a constant rate attack

$$x|\mathrm{H}_1 = r + x|\mathrm{H}_0$$

modeled using the *shifted GPD*

$$p(x|H_1) = \theta(\theta+\lambda(x-r))^{x-r-1} e^{-\theta-\lambda(x-r)}/(x-r)!$$

$x = r,r+1,\ldots$ is number of packet arrivals

Modeling Internet Traffic?

- ***We do not believe that these are accurate models for internet traffic***

- BUT – they do enable high performance attack detection with relatively low rate attacks
  - also enable on-line attack detection

The packet size distribution model

- Sample entropy of packet sizes

$$y_i = -\sum_{j \in S_i} q_j \log q_j$$

q_j denotes proportion of packets of specific size

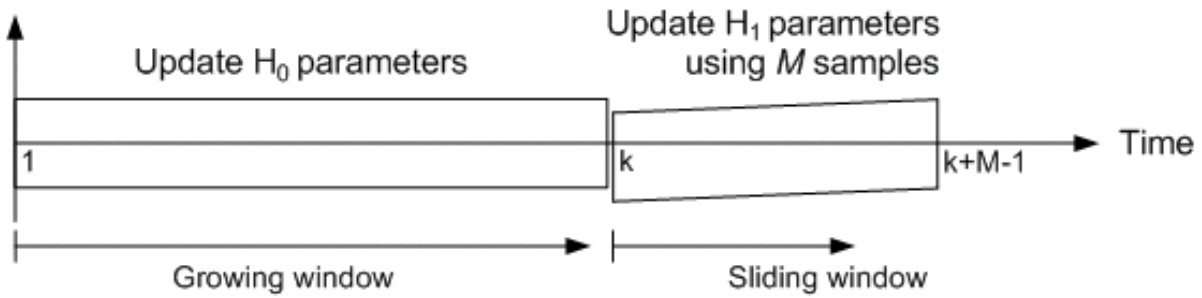
- y_i is modeled as Gaussian

$$p(y|H_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_i^2}(y - \mu_i)^2\right]$$

for both the background ($i=0$) and attack ($i=1$) hypotheses

Eliminating the Training Phase 1/3

- Model parameters estimated in real-time
- Non-overlapping growing and sliding windows used to estimate parameters



- bPDM must be *initially deployed* in the absence of an attack

Eliminating the Training Phase 2/3

- Parameter estimates for the GPD [Consul89]

$$\hat{\theta}_0 = \sqrt{\frac{\overline{x}^3}{s^2}}, \quad \hat{\lambda}_0 = 1 - \sqrt{\frac{\overline{x}}{s^2}}$$

where unbiased estimators are employed

- Parameter estimates for sGPD

$$\hat{r} = \max\{\lfloor -\hat{\theta}_0/(1 - \hat{\lambda}_0 + \overline{x}) \rfloor, \min\{x_1, \cdots, x_M\}\}$$

$$\hat{\theta}_1 = \sqrt{\frac{(\overline{x} - \hat{r})^3}{s^2}}, \quad \hat{\lambda}_1 = 1 - \sqrt{\frac{\overline{x} - \hat{r}}{s^2}}$$

Eliminating the Training Phase 3/3

- Parameter estimates for packet size SPRT

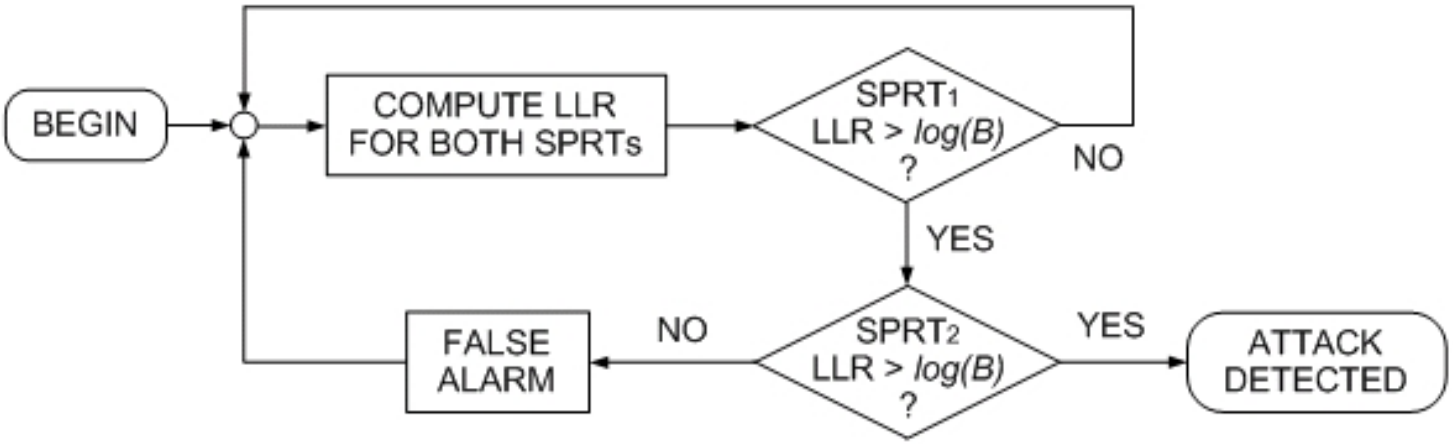
$$\widehat{\mu}_i = \frac{1}{k}\sum_{i=1}^{k} x_i$$

$$\widehat{\sigma}_i^2 = \frac{1}{k-1}\sum_{i=1}^{k}(x_i - \widehat{\mu}_i)^2$$

  are unbiased estimators

- Parameter estimates under null hypothesis are recursively computed

bPDM Operation



- Both SPRTs must *simultaneously* cross threshold for bPDM to declare attack
- Currently incorporating ASN function to develop more robust detection mechanism

Model Validation

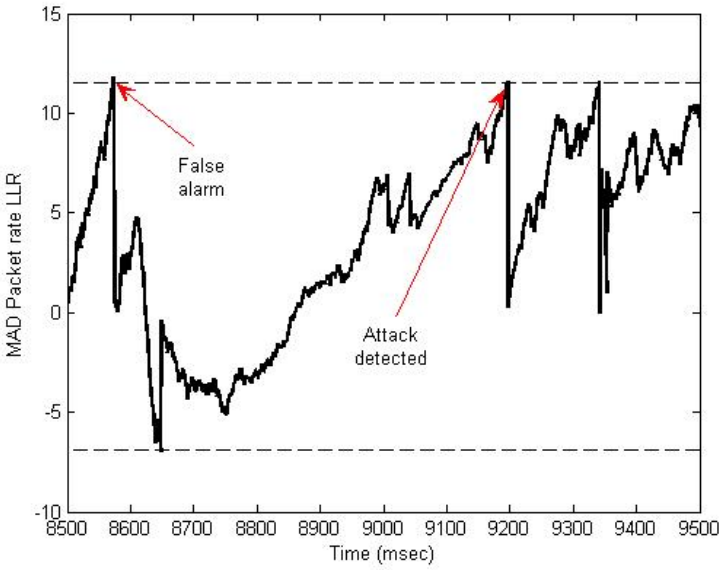
- Use synthetic traces to methodically evaluate sensitivity to low-rate attacks
- bPDM also tested on real network attacks
- Define attack rate as:

$$\% \text{ attack} = \frac{\text{Number of attack packets}}{\text{Total number of packets}}$$

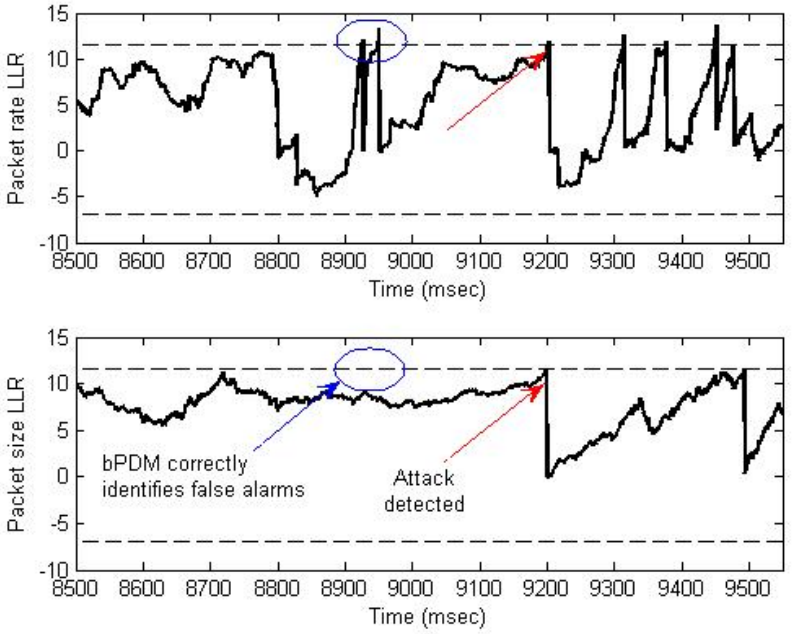
Synthetic Traces

- Synthetic traces created using stream merger application [Kamath02]
- Simulate real attacks (constant attack packet size) with varying attack strength
  - Synthetic TCP SYN attacks use minimum packet size of 68 bytes
  - Synthetic DNS reflector attacks use maximum packet size of 1518 bytes
- Smart adversary considered
  - Attack packet size distribution is bimodal [Sinha07]

Does Packet Size help?



packet rate only



packet rate and size

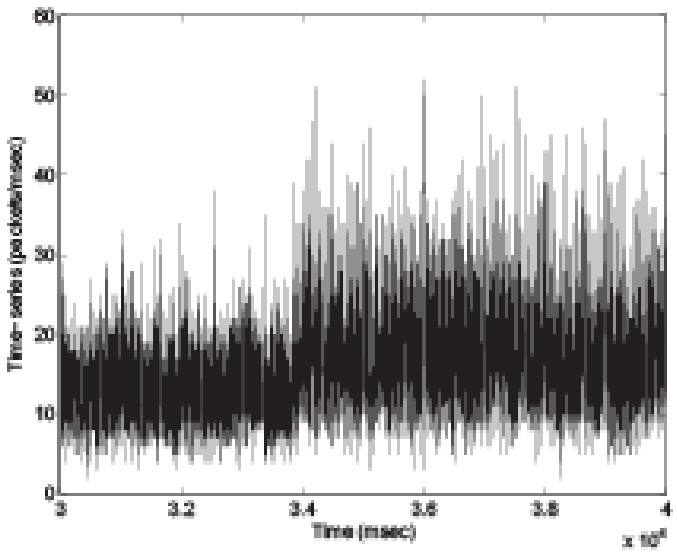
more false alarms with packet rate only

6% synthetic TCP SYN attack starts at 9.0 seconds

Performance Comparisons

Scheme	paper	# FA	TD (msec)	drawback
GPD/sGPD		0	220	first deploy w/ no attack
Poi/sPoi	[1]	2	150	training phases
Change-point	[2]	1	190	empirical parameter tuning
IP Entropy	[3]	0	400	flow-separation
Spectral	[4]	1	210	FFT (↑ complexity) training phases

24% reflector attack



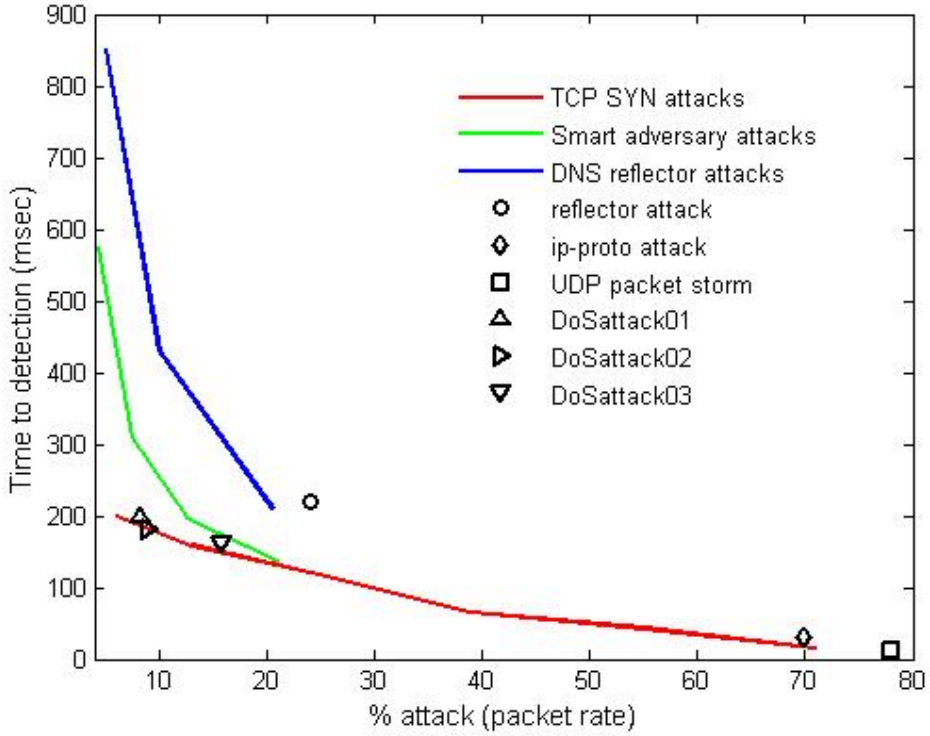
[1] Thatte, **M**, Heidemann, IEEE GI 4/08

[2] Tartakovsky, Rozovskii, Blazek, & Kim IEEE Trans. Sig. Proc., 9/06

[3] Feinstein, Schnackenberg, Balupari & Kindred, Proc. DARPA Information Survivability Conf. and Exposition, '03

[4] He, Papadopoulos, Heidemann, **M**, & Riaz, Computer Networks, accepted 9/08

Detection Results

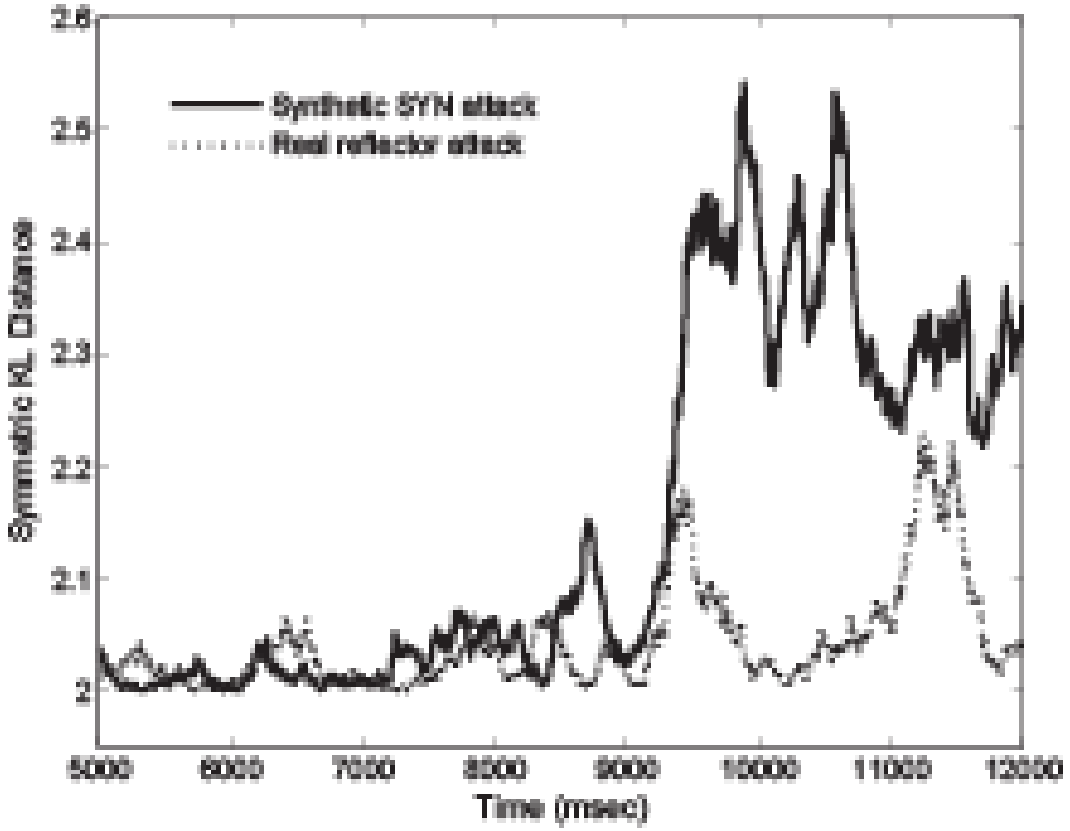


- For the bPDM, synthetic traces are good proxy for real attacks for model characterization
- attack sources (via PREDICT)
  - [a] DoS traces 20020629
  - [b] attack-servpath-udp22-20061106
  - [c] DoS 80 timeseries-20020629
- **reflector attack**: packet size distributions very similar in attack vs. no attack cases

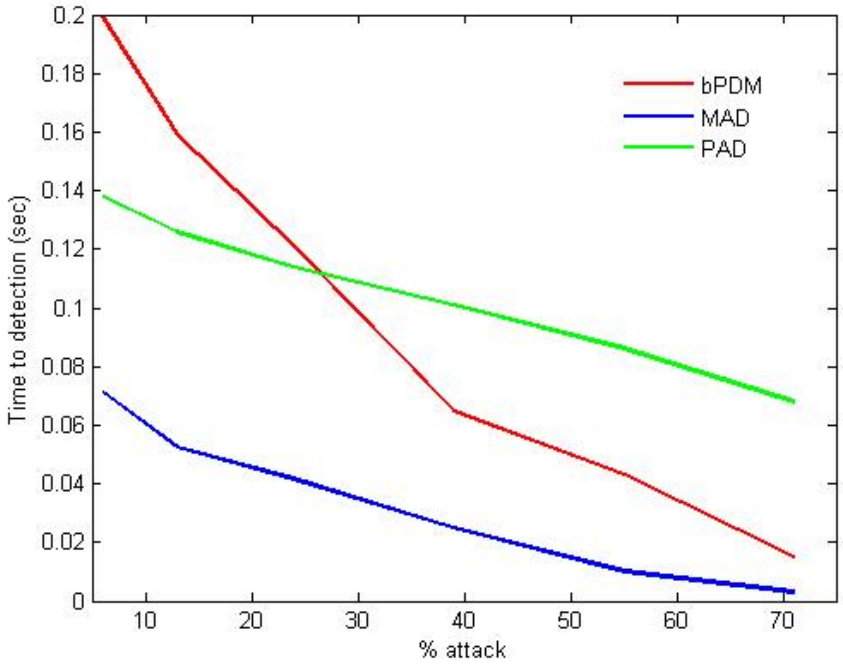
symbol	trace	SNR	TD (msec)	type
◇	[a]	70%	31	ip-proto 255
○	[a]	24%	742	reflector
□	[b]	78%	12	UDP servepath
△	[c]	8%	200	merged real traces
▷	[c]	9%	179	merged real traces
▽	[c]	16 %	163	merged real traces

24% Reflector Attack

- packet size distribution does not change much
- Kullback-Leibler distance between two distributions



Comparisons with Previous Methods



synthetic TCP SYN attack

- bPDM (new/(s) generalized Poisson) appears worse than MAD ((s)Poisson models)
- sometimes better than PAD (SPRT spectral)
- not a fair comparison
  - PAD has no on-line analog has to train for both hypotheses
  - MAD can be on-line for attack parameters if you know the no-attack statistics, still needs a training phase

Future Work

- Incorporating ASN function of SPRT to develop more robust bPDM
- Real-time deployment of bPDM
- Characterizing performance of shifted GPD parameter estimates

Summary

- Developed several attack detection methods
  - spectrally based
  - modeled: Poisson/shifted Poisson; generalized Poisson/shifted generalized Poisson
  - additionally consider packet size distributions
- The bivariate Parametric Detection Mechanism (bPDM)
  - detects low-rate attacks further back in the network
  - achieves fast detection with lower computational overhead
  - does not require flow-separation
  - estimates model parameters in real time – dedicated training phase not required

`http://www.isi.edu/ant/madcat/`