Northeastern University
Network Science Institute

Northeastern University
Khoury College of Computer Sciences

Detecting Self-Propagating Malware in Cyber Networks

Timothy Sakharov sakharov.t@husky.neu.edu
Benjamin A. Miller miller.be@husky.neu.edu
Lisa Friedland l.friedland@northeastern.edu
Talha Ongun ongun.t@husky.neu.edu
Alina Oprea a.oprea@northeastern.edu
Tina Eliassi-Rad tina@eliassi.org

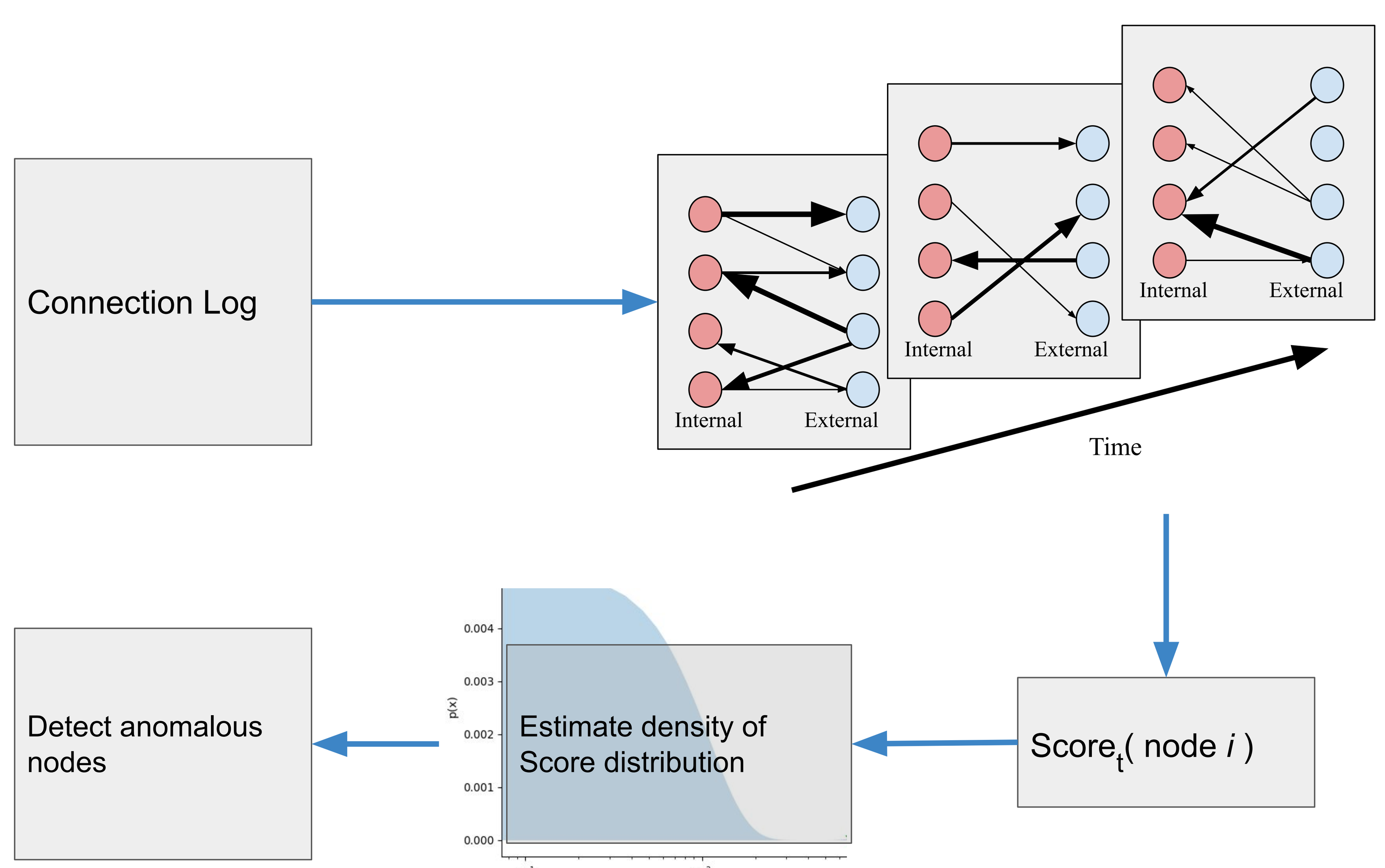
INTRODUCTION

- Self-propagating malware, such as WannaCry, poses a significant threat to enterprise cyber networks
- The malware typically engages in scanning behavior, where it connects to many randomly-generated IP addresses

PROBLEM DEFINITION

- Given a set of timestamped IP communications, **detect anomalous IP addresses in a subnet of interest**
- The data only shows internal-external traffic (the perimeter)
- At each time interval, want an **anomaly score** for each internal IP

FRAMEWORK



Supported by the DARPA CHASE Program

METHODS

1. **VOLUME** (baseline): Node's score is its degree at time *t*
   - We might expect the malware to communicate a lot

$$\text{VOLUME(i)} = \sum_j a_{ij}$$

2. **KURTOSIS**: Node's score is kurtosis of its tf-idf distribution
   - We might expect the malware to communicate with some very unusual IP addresses

$$\text{tf-idf(i, j)} = \#\text{ communications between } i \text{ and } j \cdot \log\left(\frac{N}{1 + \#\text{ neighbors of j}}\right)$$

$$\text{KURTOSIS(i)} = E\left[\left(\frac{\text{tf-idf}(i,\forall j) - \mu}{\sigma}\right)^4\right]$$

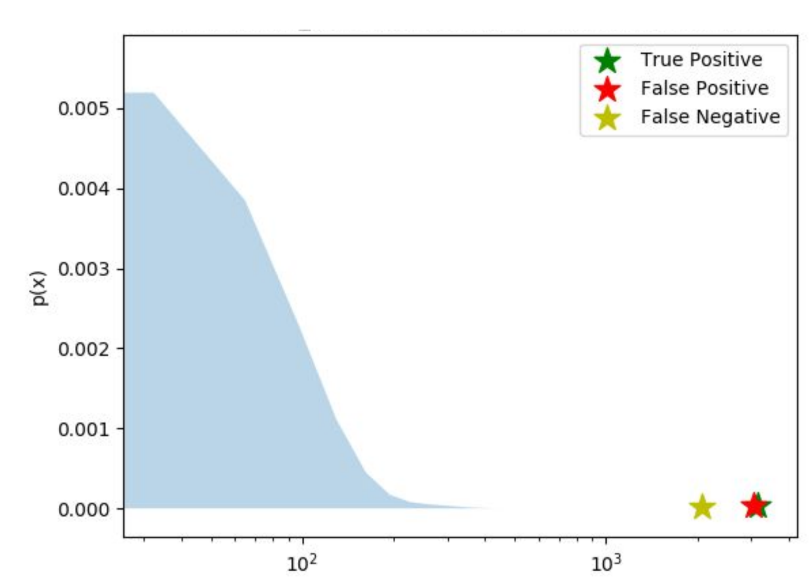
3. **LIKELIHOOD**: Node's score is related to likelihood according to an exponential null model of the bipartite graph
   - We might expect the malware to have unlikely links

$$p_{ij} = \frac{e^{\lambda_i + \sigma_j}}{1 + e^{\lambda_i + \sigma_j}}$$

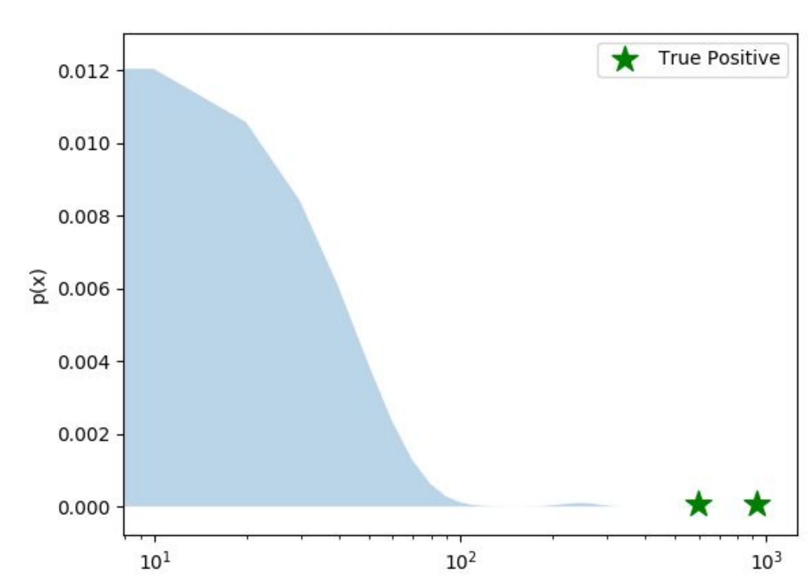
$$\text{LIKELIHOOD(i)} = \prod_{j=1} p_{ij}^{I(a_{ij}=1)} (1 - p_{ij})^{I(a_{ij}=0)}$$

IP address is anomalous if: its score is in the top *k* AND the probability of that score is low

VOLUME Kernel Density Estimate



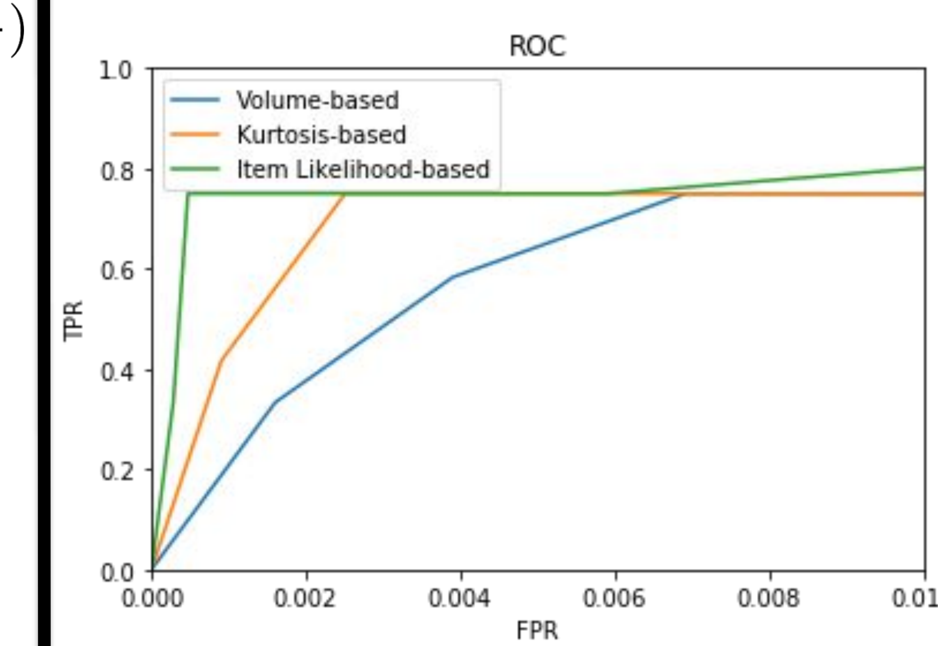
KURTOSIS Kernel Density Estimate



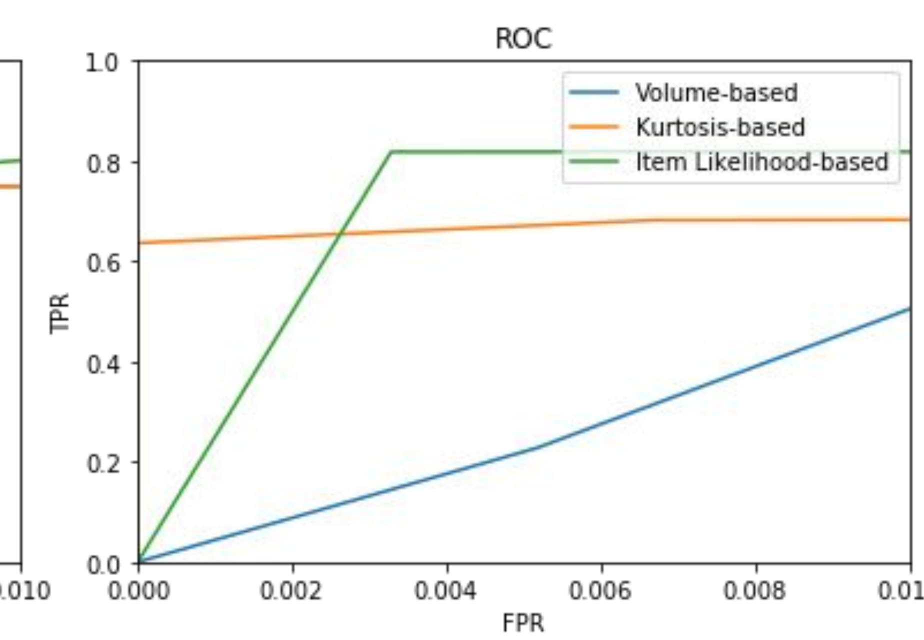
RESULTS

- Fix time window, vary thresholds for score and probability
- Evaluate performance by plotting partial ROC curve, averaged over time
- At very low False Positive Rates, KURTOSIS and LIKELIHOOD outperform the baseline

Varying number of top *k*



Varying probability threshold



With a top *k* set to 2 and a probability threshold of 0.0001, we obtain:

	TPR	FPR
VOLUME	.589	.003
KURTOSIS	.667	**0**
LIKELIHOOD	.75	.0005

CONCLUSIONS

- Self-propagating malware can be detected based entirely on its communication patterns
- When a high rate of false alarms is unacceptable, VOLUME is insufficient for detecting self-propagating malware
- LIKELIHOOD and KURTOSIS both maintain high true positive rates