A*STAR
Institute for
Infocomm Research

Evidence Gathering for Network Security and Forensics

DFRWS EU 2017

Dinil Mon Divakaran, Fok Kar Wai, Ido Nevat, Vrizlynn L. L. Thing

Talk outline

- Context and problem
- Objective
- Evidence gathering framework (overview)
- Modeling and analysis
  - Regression techniques to detect evidences
- Evidence correlation and decision making
- Performance evaluation

Context

- Detection of anomalies imperative for securing networks
- Anomaly and attack detection -> widely researched topic
  - Applied knowledge from different overlapping spheres: expert system [1], information theory [2], data mining [3], signal processing [4], statistical analysis [5], and pattern recognition [6]
- But often, different solutions developed for different attacks, and classes of anomalies
  - Complicated; and costly for users
  - Anomalies are often detected and analyzed independently
    - e.g., a port scan might not be triggered as anomaly if not statistically relevant; but may be followed by a buffer overflow attack

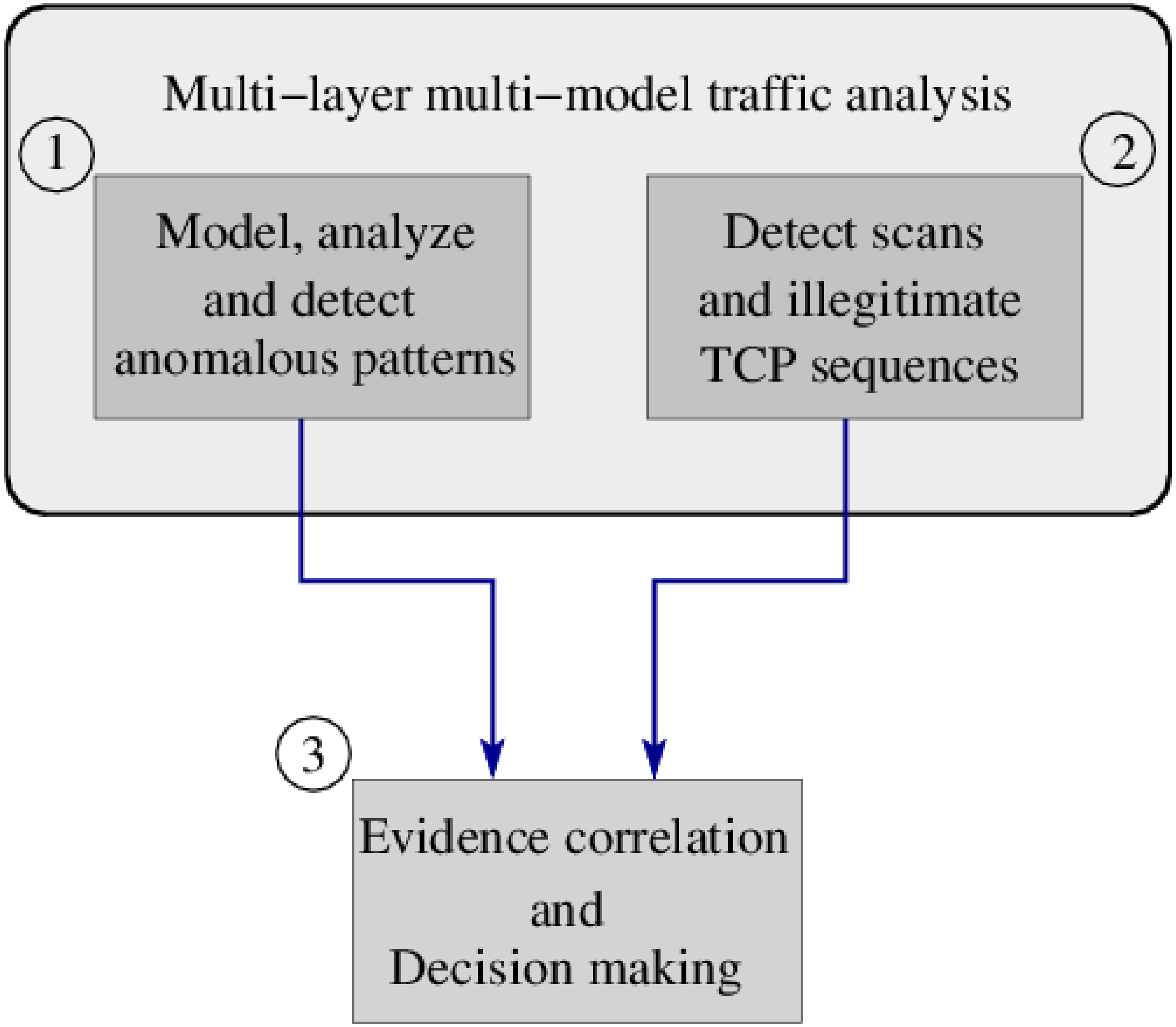
Objective

- <u>Evidences</u>: Fundamental patterns related to suspicious activities (anomalies and attacks)
- Detecting patterns allows detection of anomalies common to multiple attacks

<u>Develop a framework for anomaly detection</u>

- that detect *evidences*
- analyzes and correlates evidences
- to detect an anomalies, without the need to learn from normal traffic

Evidence-gathering framework (overview)



Stage 1: Modeling and analyzing Flows and Sessions

- Flow: A set of packets, localized in time, with the same five tuple of source and destination IP addresses, source and destination ports, and protocol

- Session: A set of flows such that, the inter-arrival time between any two subsequent flows is less than a given value

- Session definition allows coarser aggregation, say, using three tuple (dest. IP addr, dest. port, proto).

Stage 1: Modeling and analyzing Features for traffic representation

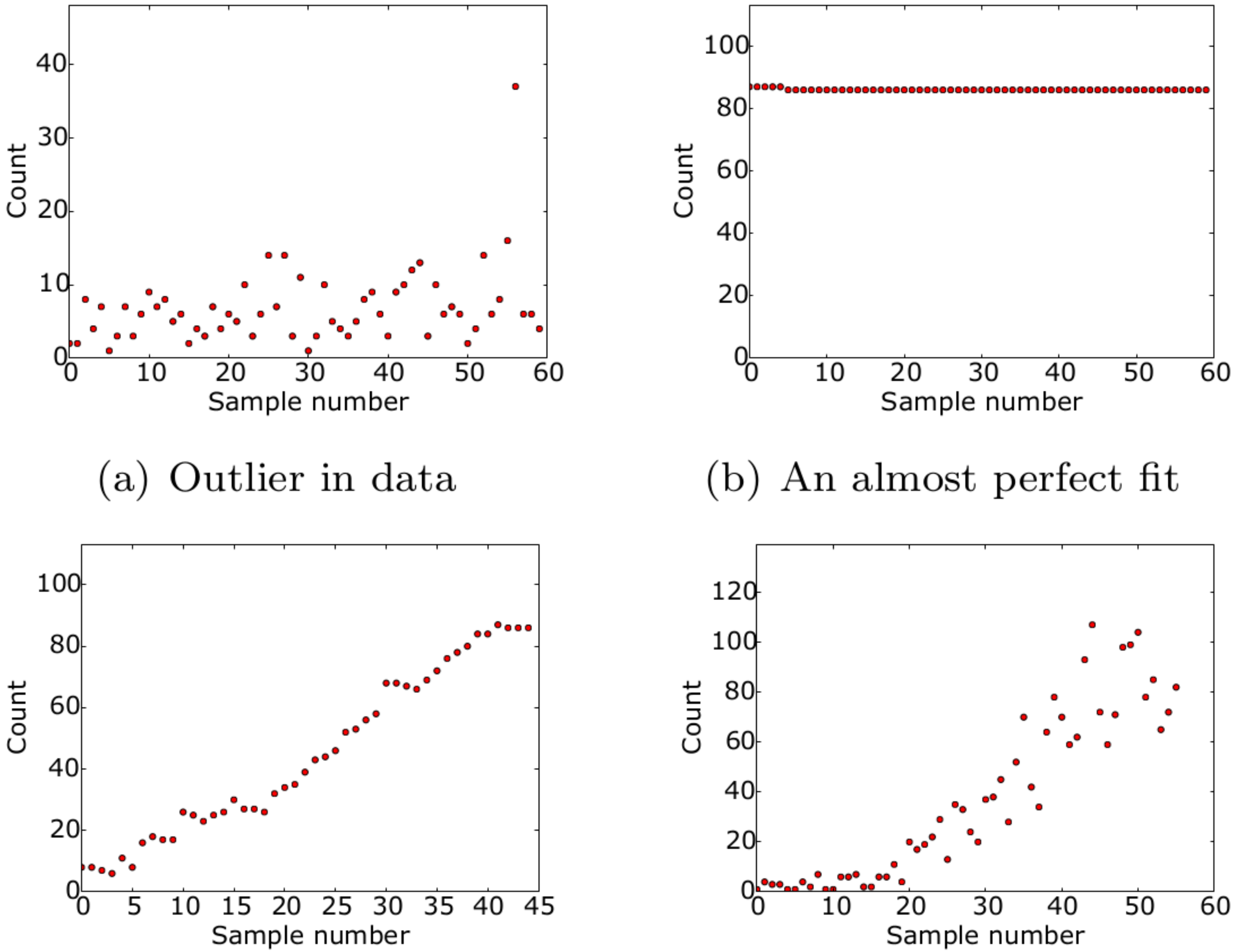
- Inter-arrival times of flows in a session (IAT): define activity measure based on IAT

A = (Median of IAT of flows x No. of flows) / total duration of session

- Sizes of flows: flow-size in packets (FSP) and flow-size in bytes (FSB)
- Degree of an end-host: no. of distinct IP addresses that an end-host communicates to, within an interval

Stage 1: Modeling and analysis
Regression

Suspicious patterns of interest



(a) Outlier in data (b) An almost perfect fit

(c) Linear relationship (d) Steep rise in count

Regression modeling

- Mainly based on linear regression
- Assume, a first order linear model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon, \qquad i = 1, 2, \ldots, n;$$

slope

Error modeled using
Normal distribution

intercept

Discrete time points / indices

Observed values

Regression modeling (cont.)

- Classical method for line fitting: **Least squares**

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

s. t.

$$r_i = Y_i - \hat{Y}_i$$

– coefficients obtained as solutions by minimizing,

$$\mathrm{SSE} = \sum_{i=1}^{n} r_i^2$$

- Four techniques for detection of patterns

1. Outlier detection

- LS regression sensitive to outliers
  - breakdown point is 1/n for n data points
- Theil-Sen estimator [7], a robust regression
  - breakdown point, b, of 29.3%
  - slope estimated as median of all slopes
- Given j = 1 –b, hypothesis test for detecting outliers:

$$\begin{cases} \mathcal{H}_0 & : r_i^{\mathrm{TS}} < Q(j(1+\kappa)), \ \text{inlier} \\ \mathcal{H}_1 & : r_i^{\mathrm{TS}} \geq Q(j(1+\kappa)), \ \text{outlier} \end{cases}$$

Quantile

control parameter

2. Goodness of fit

- If SSE is zero, there exists a functional relationship between the variables
  - $Y = f(X)$
  - suspicious, as we expect statistical relationship
  - functional relationship likely due to automated communications
- Testing involves checking for zero (or close to zero) slope

3.Inference on slope

- Detect steep linear slope; hypotheses:

threshold

$$\begin{cases} \mathcal{H}_0 : |\beta_1| \leq \theta; & \text{not an anomaly} \\ \mathcal{H}_1 : |\beta_1| > \theta; & \text{anomaly} \end{cases}$$

- Coefficients need to be estimated
- Rejection criterion for the null hypothesis is

estimate of error

`reject` $\mathcal{H}_0$ `if`

$$\hat{\beta}_1 \notin \left[-t_{n-2,1-\alpha/2} \frac{s}{\sqrt{s_{xx}}} - \theta, t_{n-2,1-\alpha/2} \frac{s}{\sqrt{s_{xx}}} + \theta \right]$$

variance of X

significance level

4. Quadratic regression

- Is simple linear model good enough?
  - Would exponential curve fit better?
- Final test – compare LS fitted model with a higher order polynomial fit (quadratic)

$$Y_i = \alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2 + \varepsilon$$

- Test statistic – coefficient of determination, R^2

estimated

mean of response variable

$$R^2 = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2}{\sum_i (Y_i - \bar{Y})^2}$$

- Hypothesis test

$$\begin{cases} \mathcal{H}_0 & : {R^2}_{\mathrm{QR}} - {R^2}_{\mathrm{LS}} \leq \theta_r, \ \text{normal} \\ \mathcal{H}_1 & : {R^2}_{\mathrm{QR}} - {R^2}_{\mathrm{LS}} > \theta_r, \ \text{anomaly} \end{cases}$$

R^2 for Quadratic

R^2 for Least-squares

Stage 2: Detecting scans and illegitimate TCP state sequences

Stage 2: Detecting scans and illegitimate TCP state-sequences

- Scans common to determine services running
  - For example, to exploit zero-day vulnerability
- TCP state sequences
  - A set of states taken by a TCP flow in its FSM*
  - A legit state sequence conforms to FSM
    - For example, ShA{Da}*FafA is of a TCP data connection (S stands for SYN, F for FIN, etc.)
- Illegitimate TCP state sequence
    - A state-path that do not conform to TCP FSM

* FSM: Finite State Machine

Stage 3: Evidence correlation and Decision making

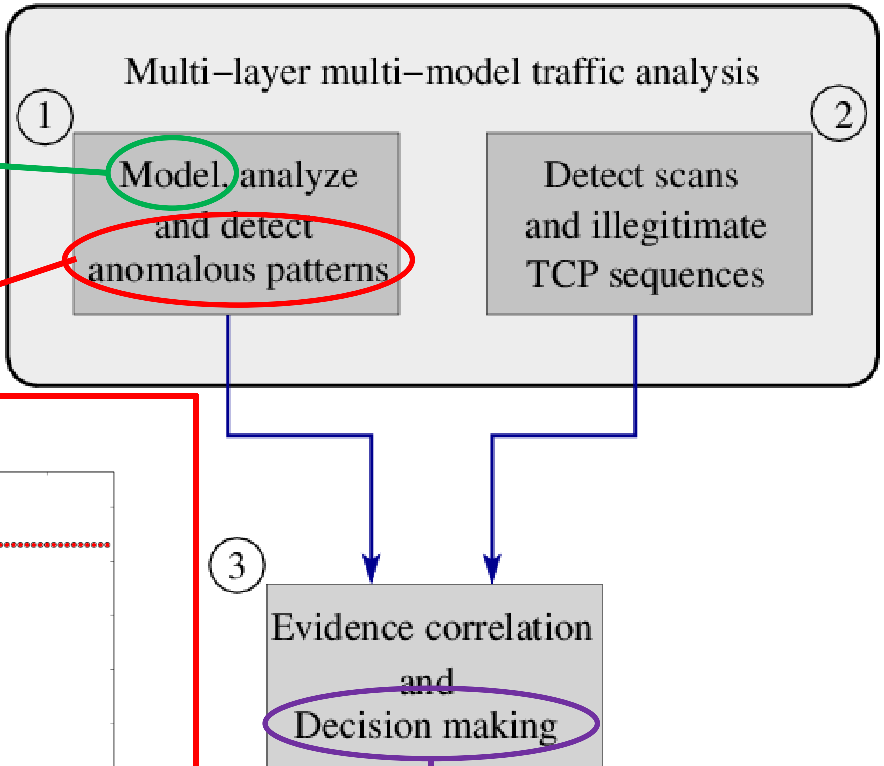
Stage 3: Evidence correlation and Decision making

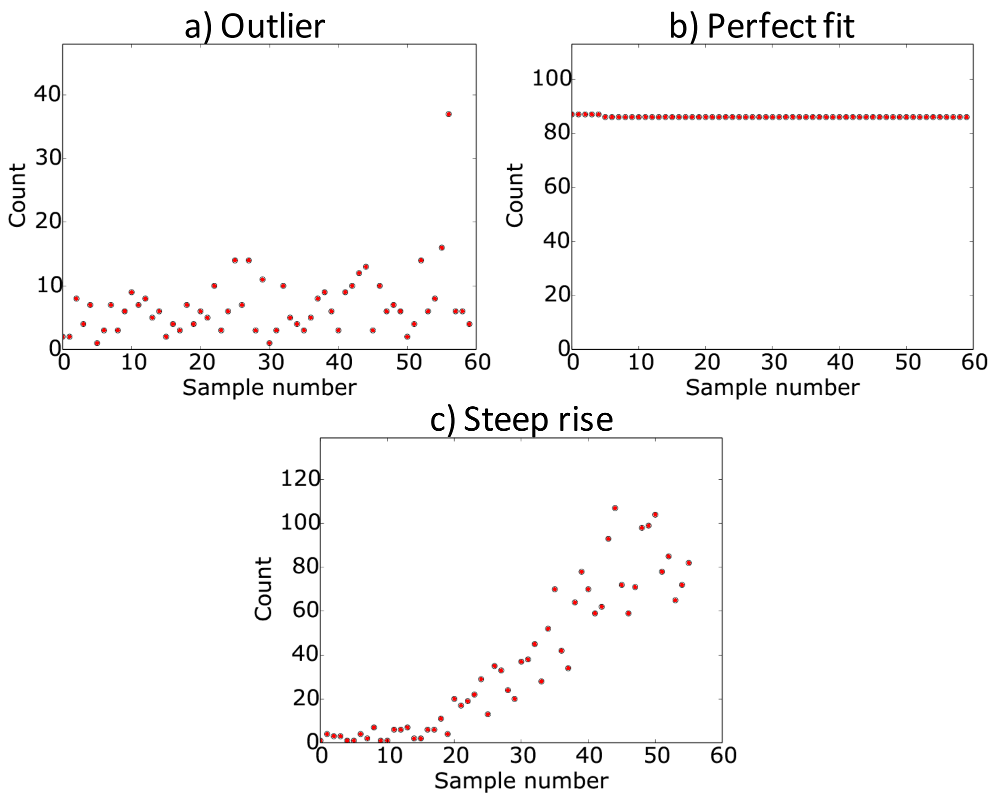
- Correlate evidences detected
  - Based on time or space
- Meta Decision Maker: Decide based on multiple evidences on a set of traffic flows or sessions
  - An anomalous pattern
    - A specific feature, and a specific technique
  - Normalize threshold and output score for each technique to [0-1]
  - Define low, medium, and high score ranges
  - Detection based on number of evidences and scores

Evidence gathering framework (recap)

With features such as:

- Inter-arrival times of flows
- Sizes of flows
- Degree of an end-host





Meta-decision maker to decide whether a traffic session is anomalous

Performance evaluation

Data

- Consists of both benign and malicious traffic
  - 969 benign and 1397 malware traffic sessions
- Benign traffic: ISCX IDS Dataset [8], LBNL Datasets [9], and Internet traffic of two secured Linux machines
- Malicious traffic generated by malware
  - Obtained from Stratosphere IPS Project [10]
  - consisting of traffic from 11 different botnets (Andromeda, Barys, Emotet, Geodo, Htbot, Miuref, Necurse, Sality, Vawtrak, Yakes and Zeus)

Settings

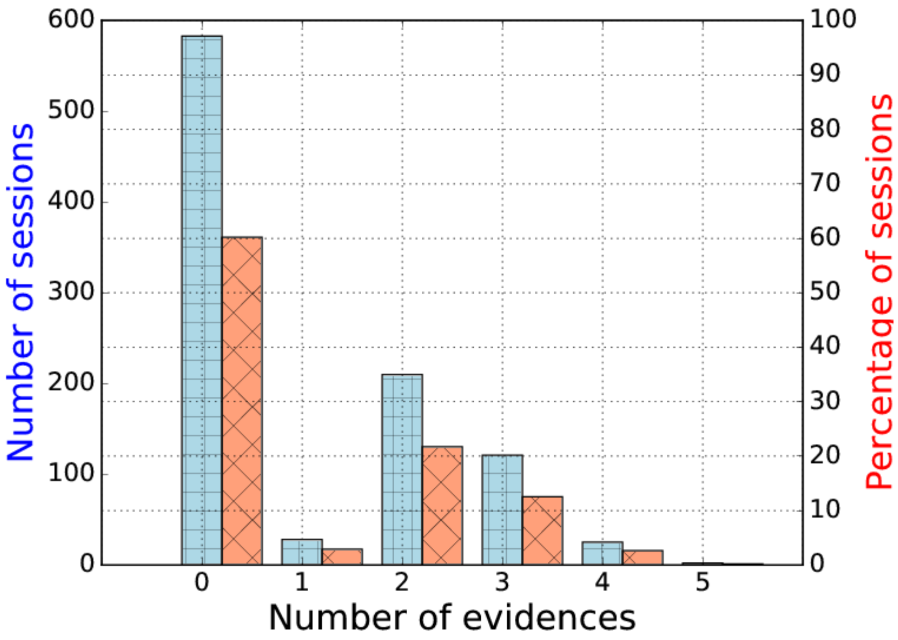
- Conservative values for threshold; and control using test output scores
- Output score *high* if >= 0.7
- Meta decision maker:

> a session classified as anomalous, if at least three anomalous patterns related to this session are detected; moreover, at least two of such patterns should have *high* scores.

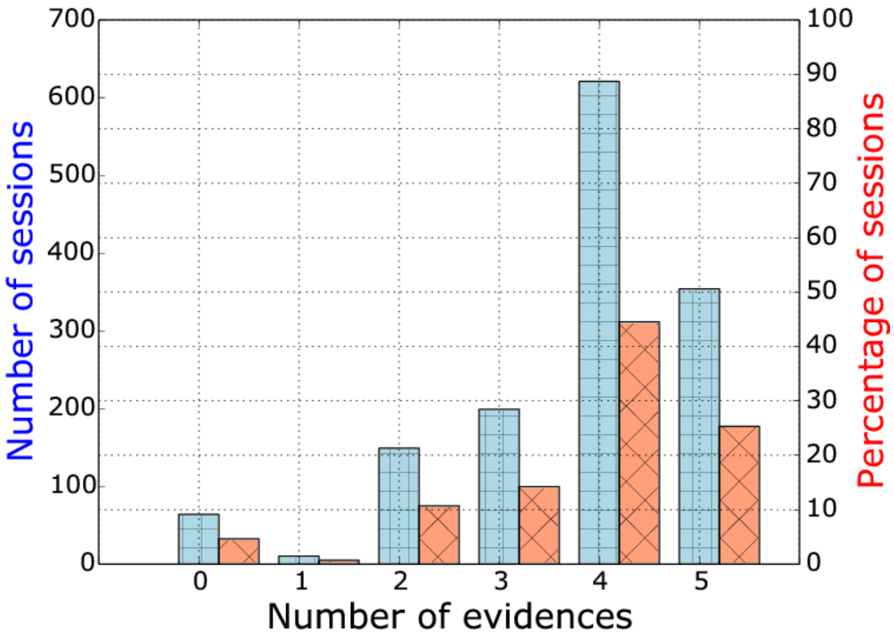
Results

<u>Histogram of evidences</u>

normal traffic sessions



malware traffic sessions

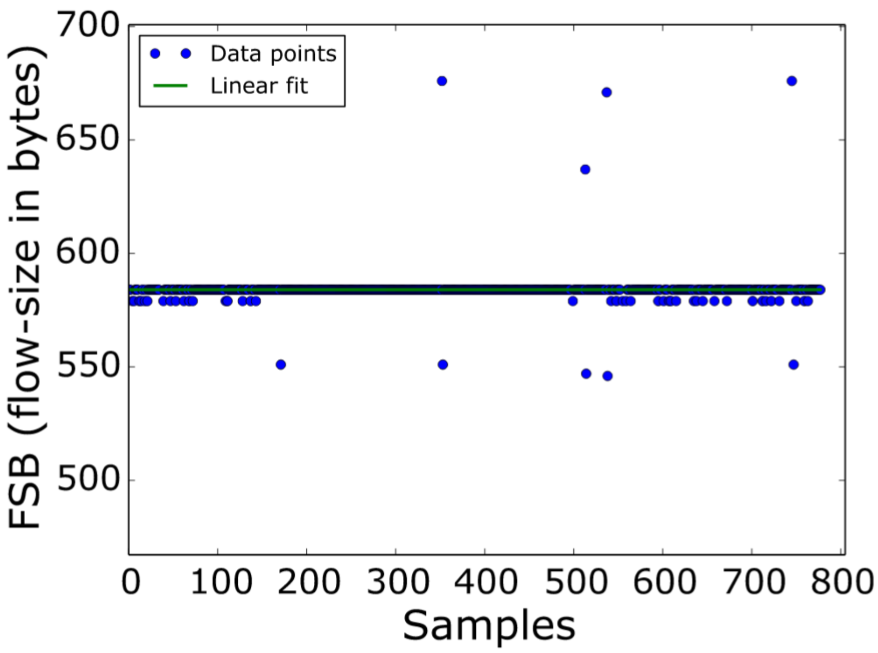


- Observations
  - Normal traffic: 60% sessions have no evidences
  - Malware traffic: 84% sessions have three or more evidences

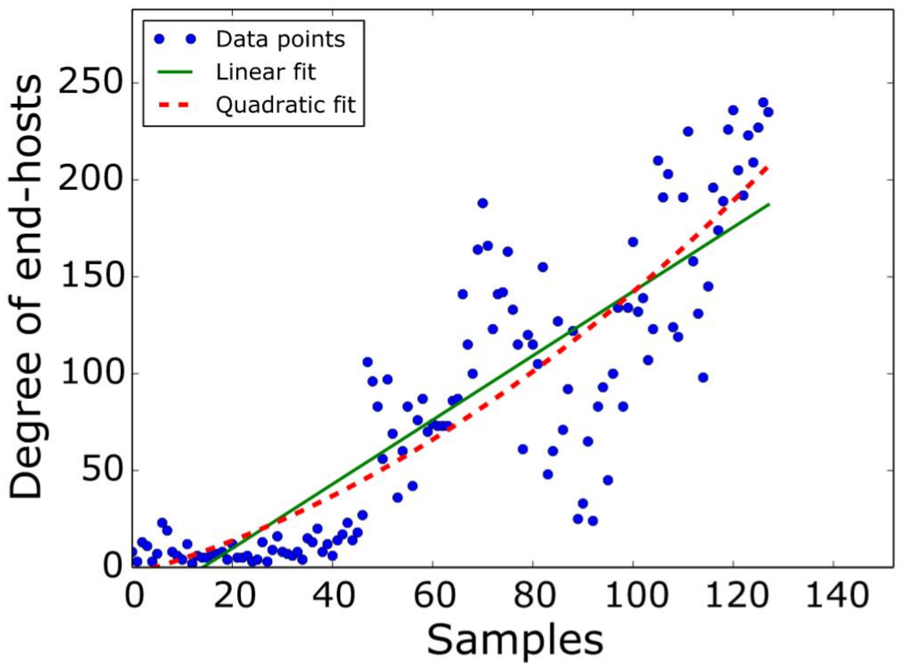
Results (cont.)

Examples of sessions detected:

Due to Goodness of Fit test



Quadratic model being a better fit



Results (cont.)

Overall detection rate of malware generated traffic sessions: 82.6%
False positive rate: 7.9%.

Related botnet	Total number of sessions	Detected sessions	Detection rate
Andromeda	148	132	89.2%
Barys	16	16	100.0%
Emotet	95	95	100.0%
Geodo	63	44	69.8%
Htbot	287	171	59.6%
Miuref	82	44	53.7%
Necurse	19	19	100.0%
Sality	440	435	98.9%
Vawtrak	40	40	100.0%
Yakes	39	25	64.1%
Zeus	168	133	79.2%

Results (cont.)

Effectiveness of features

Detected sessions		FSP	FSB	IAT	Degree	Illegitimate TCP flows
#	1154	1090	969	1129	978	609
%		**94.5%**	**84.0%**	**97.8%**	**84.7%**	**52.8%**

Effectiveness of techniques

Detected sessions		Outliers	Goodness of Fit	Linear or Quadratic
#	1154	1008	1154	94
%		**87.3%**	**100.0%**	**8.1%**

Results (cont.)

- Changing the decision criteria to detect (more)
  - *any session with two or more evidences, with at least one of them having high scores*
  - Detection accuracy of 93.9%, but false positive rate of 26.8%
- Computational time
  - Configuration: Intel Xeon W3690 CPU @ 3.47GHz and 12 GB RAM
  - close to 3,000 flows processed per second

Conclusions

- Developed a framework for gathering evidences to detect malicious network activities
- No learning of characteristics of normal traffic
- Regression modeling and analysis to detect fundamental patterns related to malicious activities
- Experiments using diverse dataset demonstrated the effectiveness of using evidences for detection of malware sessions
- Next steps:
  – Enhance the solution to work on live real-time traffic
  – Experiment with other relevant features

References

[1] Koral Ilgun, R. A. Kemmerer, and Phillip A. Porras. State Transition Analysis: A Rule-Based Intrusion Detection Approach. IEEE Trans. Softw. Eng., 21(3):181–199, March 1995.

[2] George Nychis, Vyas Sekar, David G. Andersen, Hyong Kim, and Hui Zhang. An Empirical Evaluation of Entropy-based Traffic Anomaly Detection. In Proc. 8th ACM SIGCOMM Conf. on Internet Measurement, IMC, pages 151–156, 2008

[3] Anukool Lakhina, Mark Crovella, and Christophe Diot. Mining anomalies using traffic feature distributions. In Proc. ACM SIGCOMM, pages 217–228, 2005.

[4] M. Thottan and Chuanyi Ji. Anomaly detection in IP networks. IEEE Trans. on Signal Processing, 51(8):2191–2204, Aug 2003.

[5] Gautam Thatte, Urbashi Mitra, and John Heidemann. Parametric Methods for Anomaly Detection in Aggregate Traffic. IEEE/ACM Trans. Netw., 19(2):512–525, April 2011.

[6] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita. Network Anomaly Detection: Methods, Systems and Tools. IEEE Communications Surveys Tutorials, 16(1):303–336, 2014.

[7] Pranab Kumar Sen. Estimates of the Regression Coefficient Based on Kendalls Tau. Journal of the American Statistical Association, 63(324):1379–1389, 1968.

[8] UNB ISCX Intrusion Detection Evaluation DataSet, 2016. http://www.unb.ca/research/iscx/dataset/iscx-IDS-dataset.html

[9] LBNL Enterprise Trace Repository, 2005. http://www.icir.org/enterprise-tracing

[10] Stratosphere IPS Project, 2016. https://stratosphereips.org/category/dataset.html

Thank you!