A Study of Non-Gaussian Data Assimilation for Volumetric Network Anomaly Detection

Intern: **Alen E. Golpashin**
University of Illinois at Urbana-Champaign

Mentor: **CDR Chad Bollmann, USN**
Naval Postgraduate School

Summer 2021

Background

What is a volumetric anomaly?

Distributed Denial of Service Attack (DDoS)

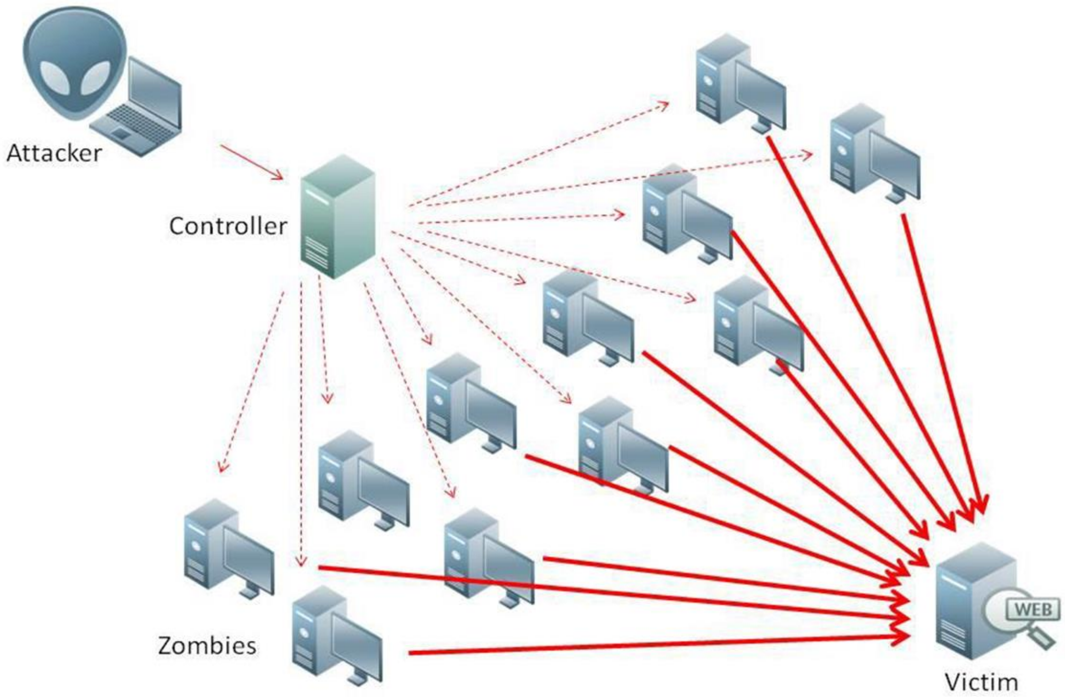
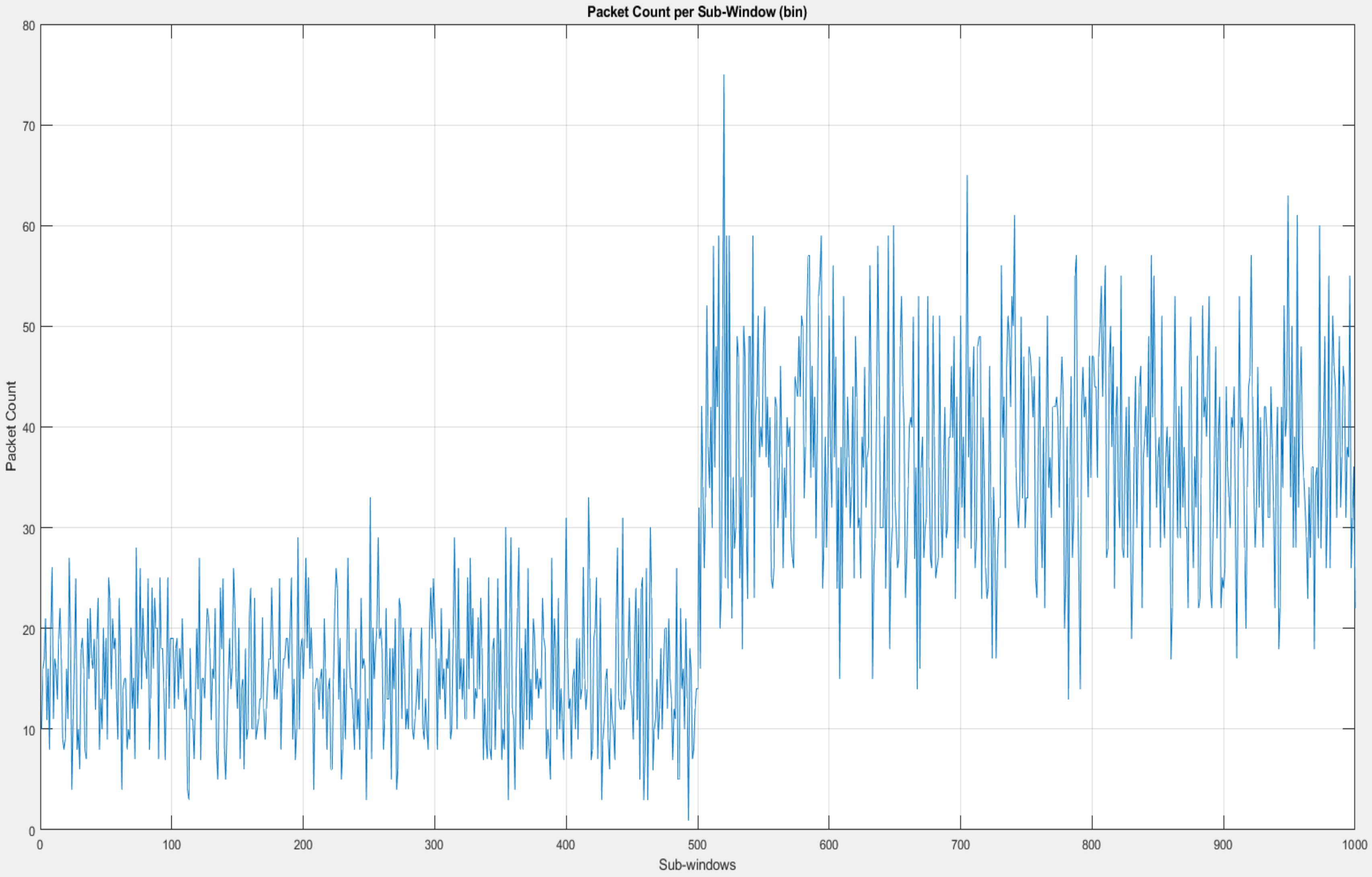


Image source: https://www.networkworld.com/

A Simulated Signal Mimicking a DDoS Attack

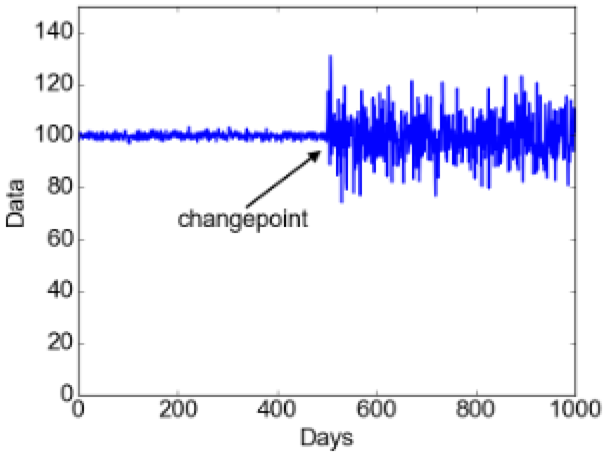


The Problem

Can the attacks be detected before they fully materialize?

Can we detect the volume change-point early?

Challenges?



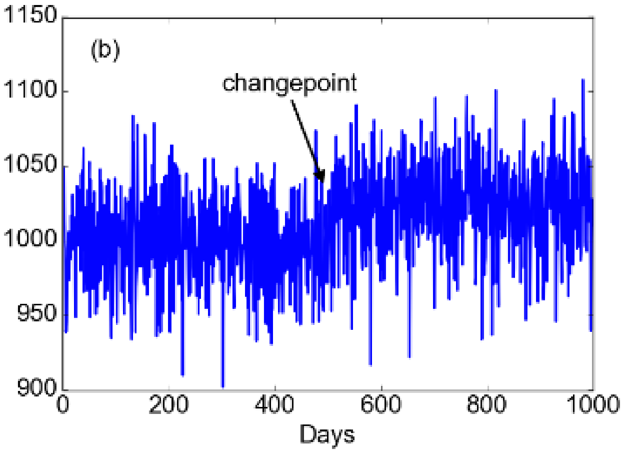
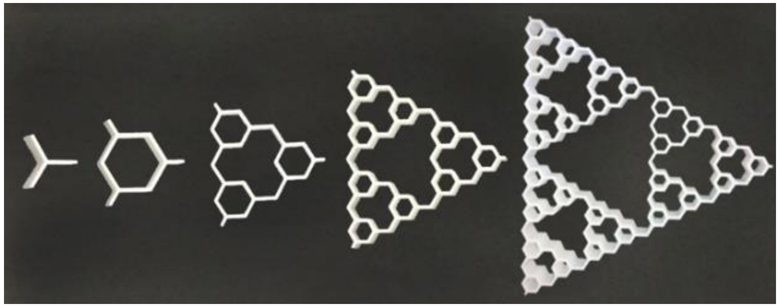


Image source: http://www.claudiobellei.com/

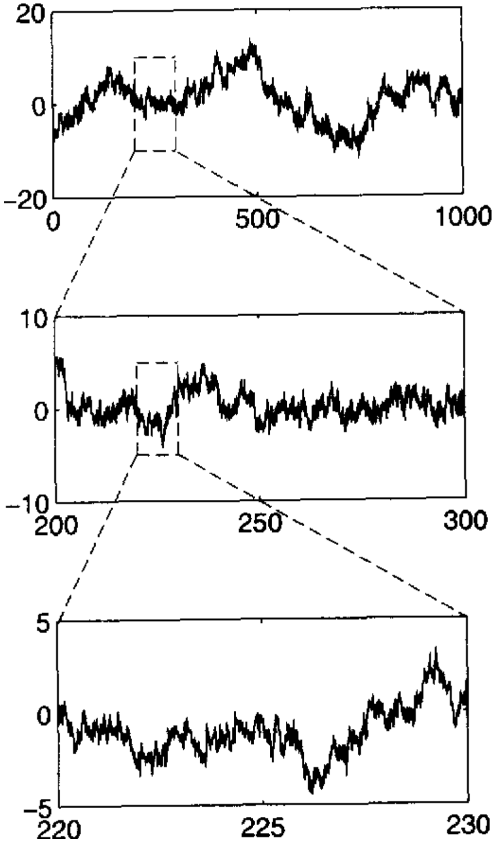
Preliminary Considerations

Properties of Network traffic

-Fractal-like (Self Similar) [1][2]



"Fractal-like"



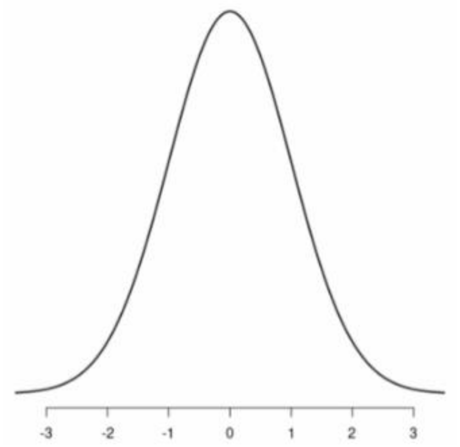
"Self Similar"

Image source: Wikipedia

Preliminary Considerations

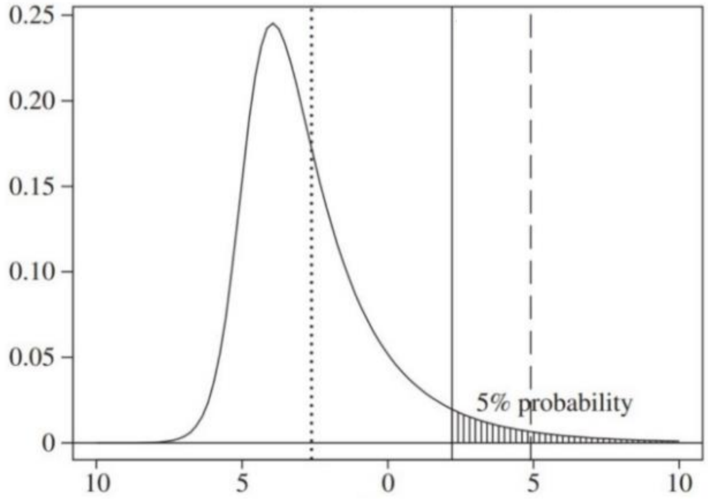
Properties of Network traffic

-Heavy Tail Distribution [3]



"Gaussian"

VS

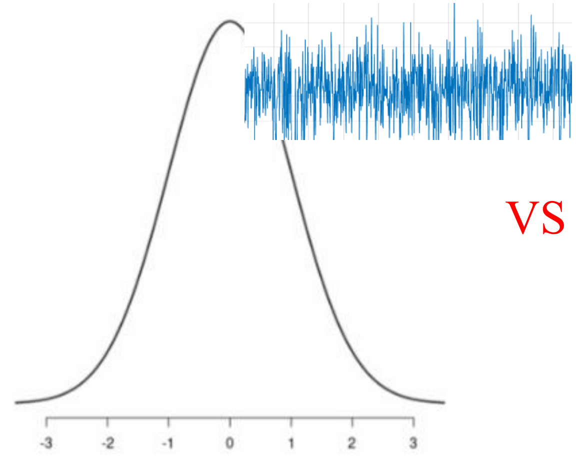


"Heavy-Tail"

Preliminary Considerations

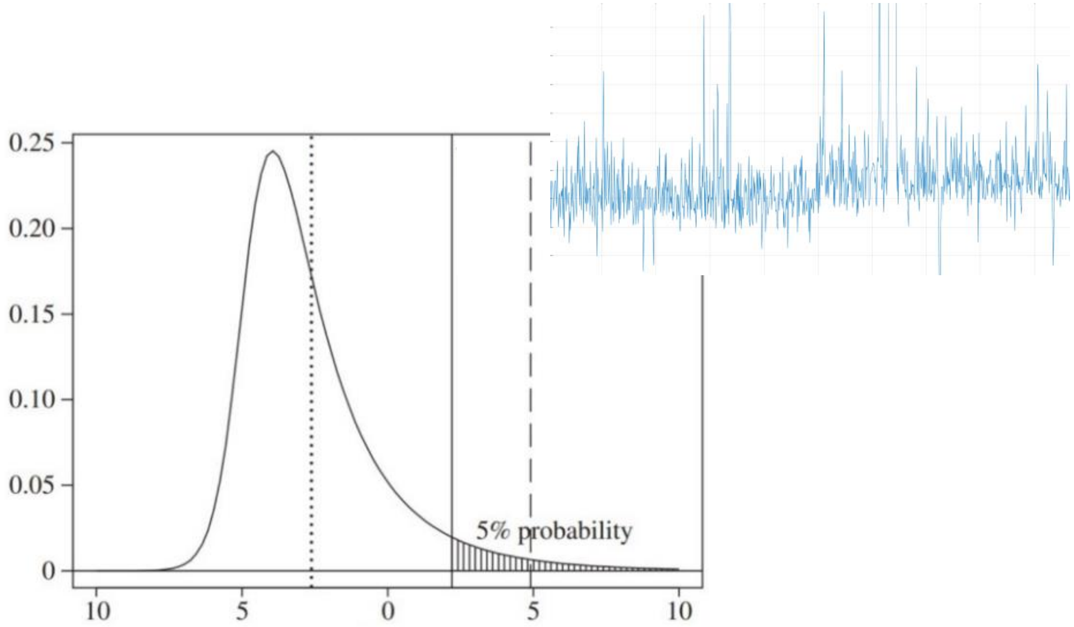
Properties of Network traffic

-Heavy Tail Distribution [3]



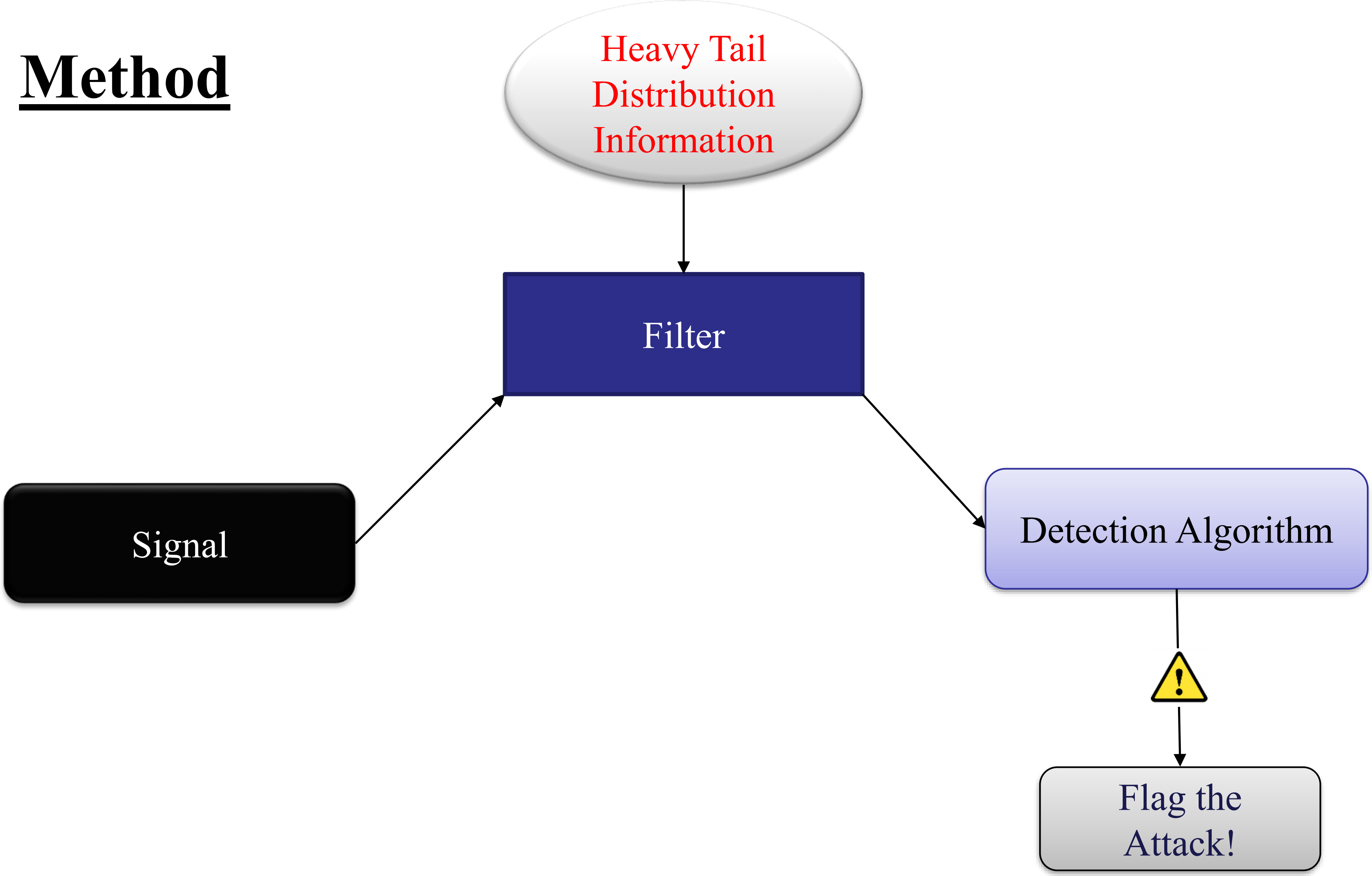
VS

“Gaussian”



“Heavy-Tail”

Method


Heavy Tail Distribution Information
Filter
Signal
Detection Algorithm
⚠
Flag the Attack!

Hypothesis

- Consideration of heavy tailed-ness of the signal while filtering will improve the results:

Expected benefits:

-Lower false-alarm rate

-Faster volume change-point detection

The Filter Setup

Currently the filter is assuming an autoregressive AR(p) model of the traffic signal with heavy tail residual

Filter's Task:

Detect changes in the coefficients of the autoregressive model

$$Y_k = H_k X_k + V_k$$

$$H = \begin{bmatrix} Y_{k-1} & Y_{k-2} & \dots & Y_{k-m} \end{bmatrix}, \; X = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

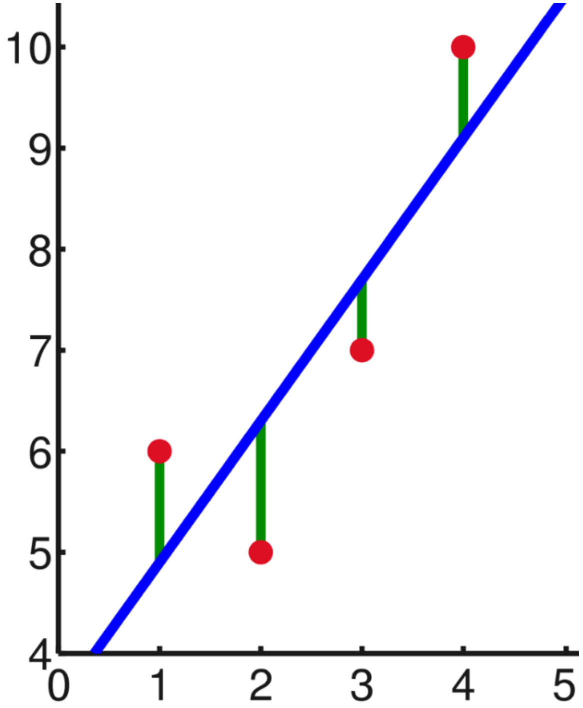
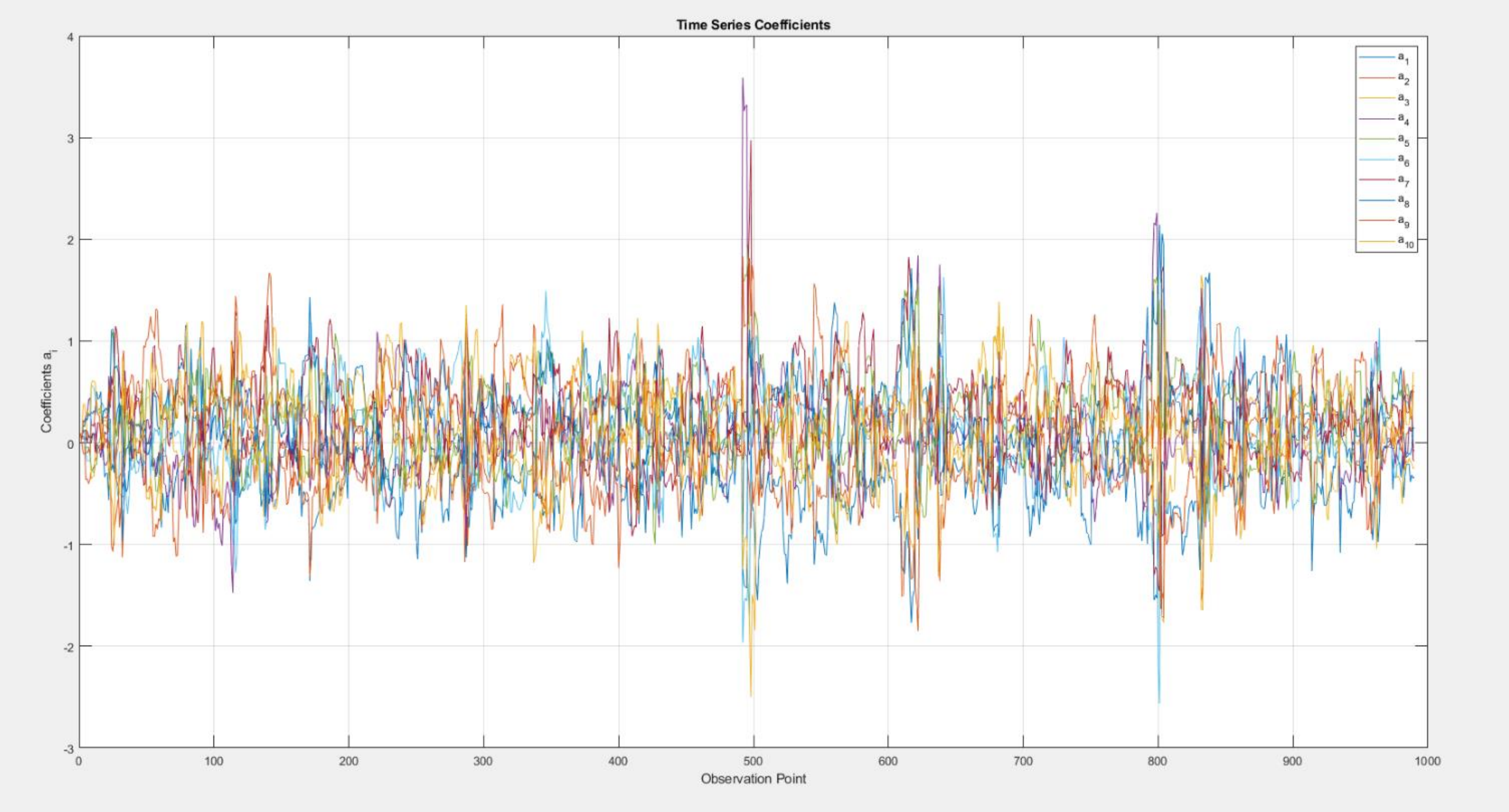
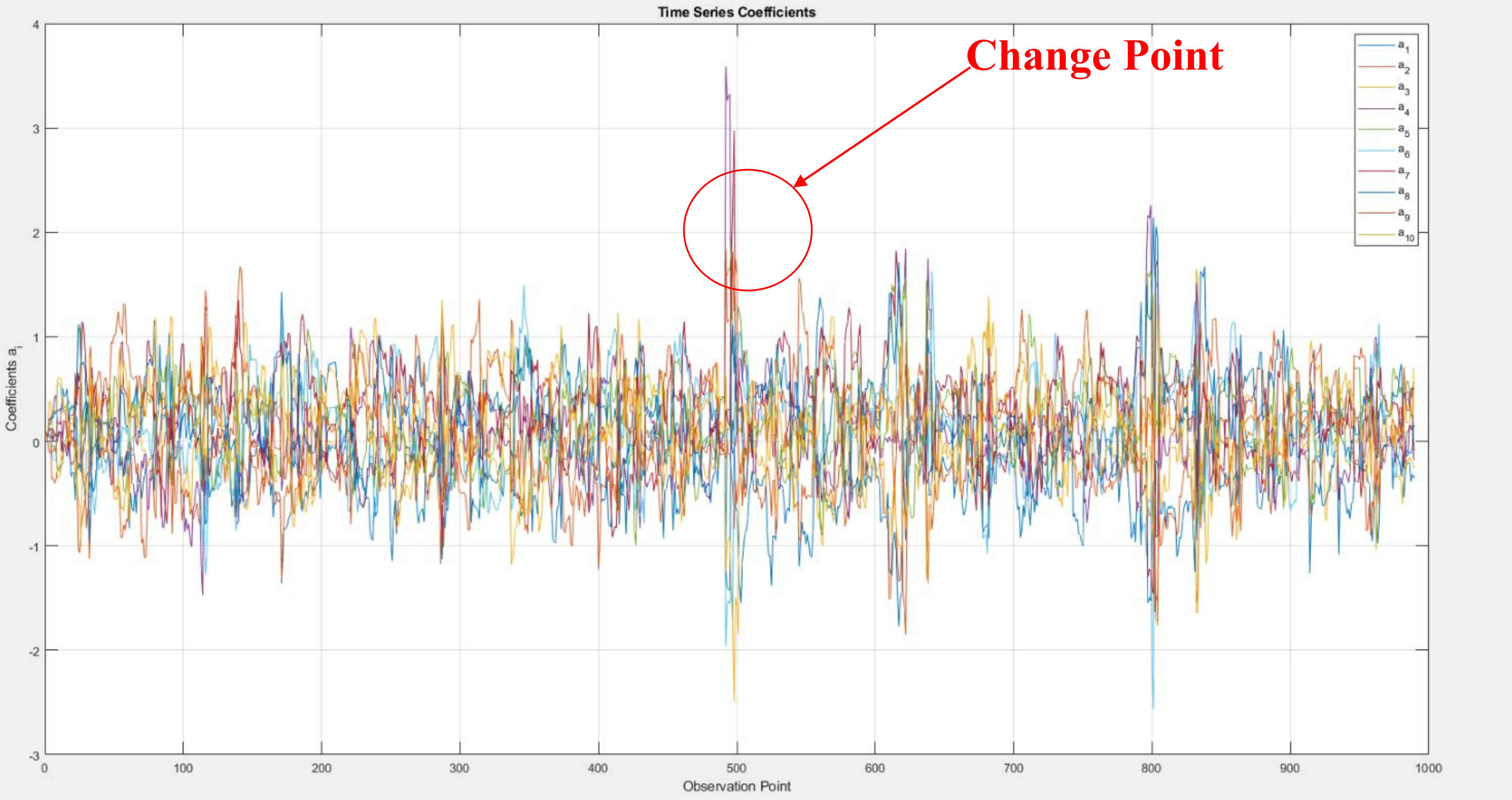


Image source: Wikipedia

Some Results




Time Series Coefficients
Change Point
Coefficients a_i
Observation Point
4
3
2
1
0
-1
-2
-3
0
100
200
300
400
500
600
700
800
900
1000
a_1
a_2
a_3
a_4
a_5
a_6
a_7
a_8
a_9
a_{10}

Future Work

- Develop a more sophisticated detection algorithm to accompany the filter

- Carry out extensive Monte Carlo testing of the presented filtering method

- Compare the (filtering + detection algorithm) to the state-of-the-art or other broadly used detection methods

Image source: https://www.horizoninternettechnologies.com/

References

[1] Gonzalez, J., & Bollmann, C. A. (2019, December). Aggregated impulses: Towards explanatory models for self-similar alpha stable network traffic. In *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)* (pp. 1-10). IEEE.

[2] Willinger, W., Govindan, R., Jamin, S., Paxson, V., & Shenker, S. (2002). Scaling phenomena in the Internet: Critically examining criticality. *Proceedings of the National Academy of Sciences*, *99*(suppl 1), 2573-2580.

[3] Willinger, W., Paxson, V., & Taqqu, M. S. (1998). Self-similarity and heavy tails: Structural modeling of network traffic. *A practical guide to heavy tails: statistical techniques and applications*, *23*, 27-53.

Acknowledgments:

- This research was supported in part by an appointment to the Department of Defense (DOD) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the DOD. ORISE is managed by ORAU under DOE contract number DE-SC0014664.

- All opinions expressed in this presentation are the author's and do not necessarily reflect the policies and views of DOD, DOE, or ORAU/ORISE.

- Alen Golpashin is also thankful to his mentor for his guidance, daily availability, and investment in this project. Alen Golpashin would also like to thank Dr. Hoong Yeong for the technical discussions in data assimilation methods and stochastics that were useful to this research.