NREIP Naval Research Enterprise Internship Program

SEAP Science and Engineering Apprenticeship Program

Aggregated Impulses: An Explanatory Model for Self-Similar and Alpha Stable Network Traffic

Jorge L Gonzalez*, Joshua Clymer•, Dr. Chad Bollmann⋆

*Florida Atlantic University, •Monte Vista Christian School, ⋆ Naval Postgraduate School

*jorgegonzale2013@fau.edu, •joshuaclymer@students.mvcs.org ⋆cabollma@nps.edu

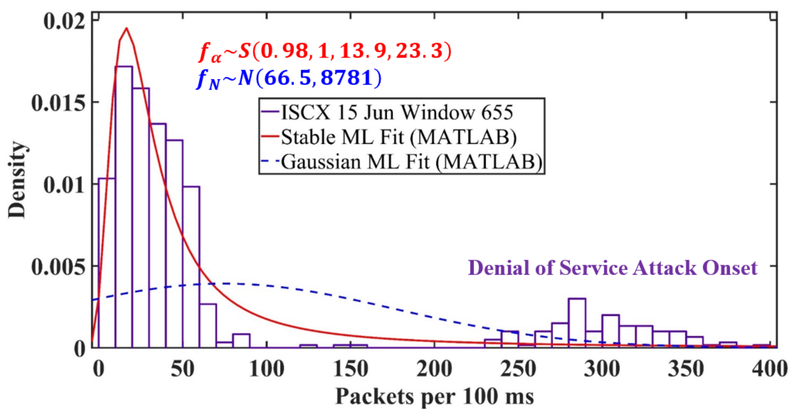
Center for Cyber Warfare

Objective

- To understand how a network's size affects its properties through the aggregation of individual sources.
- To unify documented mathematical properties of networks under the perspective of impulses and renewal processes.

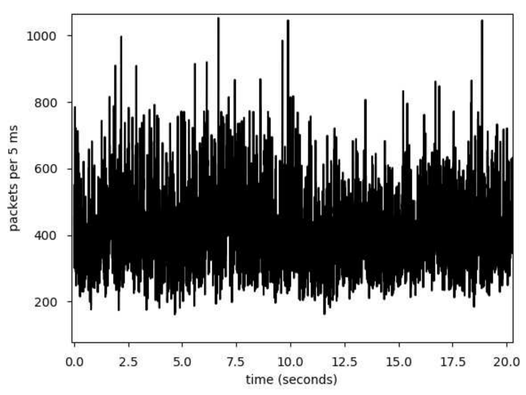
Background

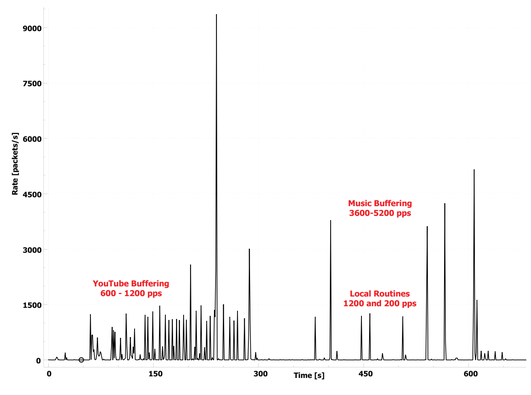
- Previous research has established the accuracy of Alpha-stable distributions for modeling computer network traffic. Gaussian distributions are often sufficient to describe small networks. Our models explain why and when these distributions are effective and are consistent with Long-Range Dependence and Self-Similarity.



Candidate distribution models and their application to cyber-attack detection.

- We propose coherent, explanatory, complementary models for aggregated network traffic, starting at the device level.
- Individual processes can be isolated and divided into impulses with independent, Pareto-distributed volumes.





- The aggregation and perturbation of these inputs gives rise to the self-similarity in network traces.

Relevant Theorems and Definitions

- For a random variable X. We define the *domain of attraction* of X, denoted by $\mathscr{D}(X)$ to be the set of i.i.d random variables Y such that there exists $d_n > 0$, $a_n \in \mathbb{R}$ and

$$\frac{Y_1 + Y_2 + ... + Y_n}{d_n} + a_n \xrightarrow{\text{d}} X \tag{1}$$

- A random variable X is a *stable* if $\mathscr{D}(X) \neq \varnothing$. Specifically, when the tail of Y_1 obeys a power-law, $Y_1 \in \mathscr{D}(X)$.
- A Lévy Process is stable if and only if it is self-similar.
- Second-order self-similar processes are long-range dependent.

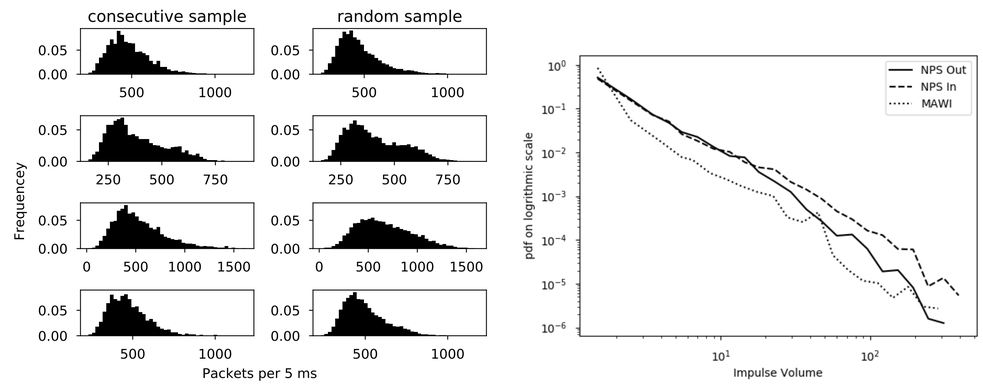
Results

Impulses

- Traffic in a sub-window can be represented by the volume aggregation of impulses.

$$S = Y_1 + Y_2 + ... + Y_e \tag{2}$$

Where e increases with the network's "size"(bottom left figure), and Y_i's are i.i.d., distributed according to a power law (top right figure). The Ergodic property implies that the distribution of a big enough window approximates the *sample* distribution of the aggregation S (top left figure).

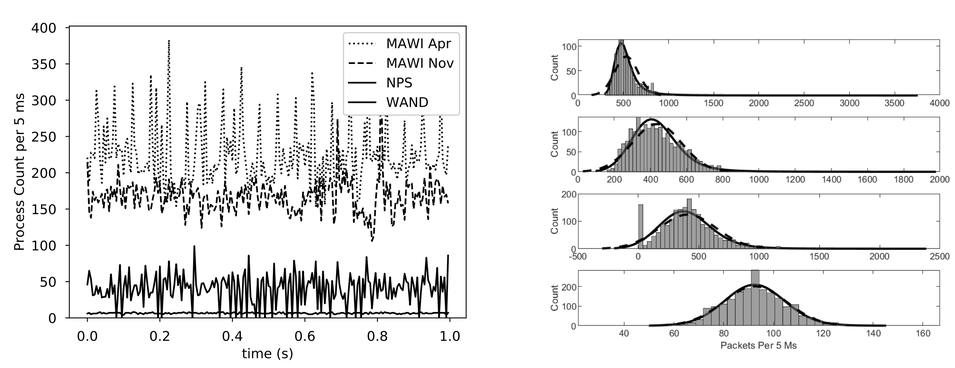


This mechanism of aggregation justifies the emergence of Alpha-Stable and Gaussian distributions from the Central Limit Theorem and its generalization.

Renewal Processes

$$\mathrm{X}^*(T, M) = \sum_{t=1}^{T} \sum_{m=1}^{M} X_{m,t} \tag{3}$$

- We apply that when $T << M$ (typical in small networks), fractional Brownian motion emerges and when $T >> M$ (large networks) the process exhibits stable (non-Gaussian) characteristics.



Conclusion

- Established conditions for the alpha-stable aggregation of network traffic from individual device processes in larger networks.
- Traffic can be modeled as the aggregation of impulses with large variations in amplitude.
- Renewal process characterization is associated with physical properties of the network.

Future work

- Quantify convergence errors.
- Apply these findings to improve existing alpha-stable based network anomaly detectors.
- Observe related long-term trends and characteristics of networks.
- Extend these results to wireless networks.