

Simulation of self-similarity in network utilization patterns as a precursor to automated testing of intrusion detection systems

D. A. Nash and D. J. Ragsdale, United States Military Academy, West Point NY

Purpose

To examine the effect of wavelet compression on network traffic signals. Ultimately the goal is to design an efficient, automatic method for generating a traffic utilization model that may be used as input for network-based intrusion detection systems.

Method

A multiresolution wavelet decomposition is applied to collected real-world network traffic using an orthogonal filter bank. The resulting set of wavelet coefficients are filtered according to their magnitude, and a percentage of the smallest coefficients is discarded. The input signal is then reconstructed using only the retained coefficients, and the self-similarity of the result is measured by calculation of the associated Hurst parameter \hat{H} . This procedure is followed using various threshold values for the smallest coefficient magnitude.

Results

It was observed that substantial compression could be achieved using this method, while retaining reasonably high fidelity in the output signal. There was a roughly linear relationship between the threshold magnitude and the estimate of the associated Hurst parameter. Root mean squared errors in the range of 10 to 40 were observed.

New aspect of work

The effect of wavelet compression on the Hurst parameter of the reconstructed traffic has not been previously investigated. The discovery of the relationship between compression threshold and \hat{H} provides a baseline which may be used to construct models which trade off fidelity and efficiency as the simulation context requires.

Conclusions

This work establishes the foundation for further efforts, toward the ultimate goal of purely automatic traffic generation. We show that it is possible to reconstruct traffic patterns with reasonable fidelity in the RMSE sense, and that such reconstructions retain their characteristics of self-similarity in a predictable fashion. Future work will examine how the process of coefficient extraction may be applied to generate traffic with the required characteristics for arbitrary intervals.