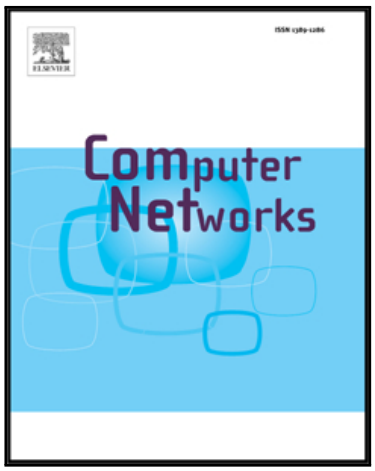

A BasisEvolution Framework for Network Traffic Anomaly Detection

Hui Xia, Bin Fang, Matthew Roughan, Kenjiro Cho, Paul Tune

PII: S1389-1286(18)30033-1
DOI: 10.1016/j.comnet.2018.01.025
Reference: COMPNW 6373

To appear in: *Computer Networks*

Received date: 13 January 2017
Revised date: 16 October 2017
Accepted date: 17 January 2018

Please cite this article as: Hui Xia, Bin Fang, Matthew Roughan, Kenjiro Cho, Paul Tune, A BasisEvolution Framework for Network Traffic Anomaly Detection, *Computer Networks* (2018), doi: 10.1016/j.comnet.2018.01.025

A BasisEvolution Framework for Network Traffic Anomaly Detection

Hui Xia[1], Bin Fang[2], Matthew Roughan[3], Kenjiro Cho[4], Paul Tune[5]

Abstract

Traffic anomalies arise from network problems, and so detection and diagnosis are useful tools for network managers. A great deal of progress has been made on this problem so far, but most approaches can be thought of as forcing the data to fit a single mould. Existing anomaly detection methods largely work by separating traffic signals into "normal" and "anomalous" types using historical data, but do so inflexibly, either requiring a long description for "normal" traffic, or a short, but inaccurate description. In essence, preconceived "basis" functions limit the ability to fit data, and the static nature of many algorithms prevents true adaptivity despite the fact that real Internet traffic evolves over time. In our approach we allow a very general class of functions to represent traffic data, limiting them only by invariant properties of network traffic such as diurnal and weekly cycles. This representation is designed to evolve so as to adapt to changing traffic over time. Our anomaly detection uses thresholding approximation residual error, combined with a generic clustering technique to report a group of anomalous points as a single anomaly event. We evaluate our method with orthogonal matching pursuit, principal component analysis, robust

[1]H. Xia is with the School of Accounting, Chongqing University of Technology, Chongqing, 400044, China. E-mail: summertulip@126.com. The majority of this work was conducted while H. Xia was visiting the University of Adelaide.

[2]Corresponding Author. B. Fang is with the School of Computer Science, Chongqing University, Chongqing, 400044, China. E-mail: fb@cqu.edu.cn

[3]M. Roughan is with ARC Centre of Excellence for Mathematical & Statistical Frontiers (ACEMS) at the School of Mathematical Science, and the University of Adelaide, Adelaide 5005, SA, Australia. E-mail: matthew.roughan@adelaide.edu.au

[4]Kenjiro Cho is at IIJ. E-mail: kjc@wide.ad.jp

[5]P. Tune is at Image Intelligence, E-mail: paul@imageintelligence.com