Time-Series Analysis for Performance Monitoring and Anomaly Detection in Computer Networks

Marcos Portnoi, *Member, IEEE*
Department of Computer and Information Sciences
University of Delaware
mportnoi@udel.edu

Priscilla Santos Moraes
Department of Computer and Information Sciences
University of Delaware
pmoraes@udel.edu

Martin Swany
Department of Computer and Information Sciences
University of Delaware
swany@cis.udel.edu

ABSTRACT

We survey, in this work, applications for time-series statistical analysis of computer network data, specifically for performance and anomaly detection. In the realm of Quality of Service, network agents could control the fair distribution of resources based on historical behavior of applications, instead of on deterministic algorithms. Virtual circuits, for instance, can be allocated on demand for applications that exhibit a past of high utilization. Furthermore, in a network performance monitoring architecture, such as perfSONAR, services may benefit from time-series analysis of measurement data to trigger events, audit statistical behavior, or detect anomalies in the network. These anomalies might indicate performance or security issues. Finally, time-series analysis enables forecasting, that can be employed to predict future performance.

Categories and Subject Descriptors
G.3 [**Probability and Statistics**]: Multivariate statistics.

General Terms
Management, Measurement, Performance, Security.

Keywords
Time-series, anomaly, forecasting, performance, measurement, quality of service.

1. INTRODUCTION

In a computer network, applications usually compete for network resources. This can often result in applications receiving a fair share of the resources, but fairness as interpreted by the network and its protocols. The user might have requirements that are not completely or properly satisfied by this interpretation.

Technologies exist to administer guarantees of minimal Quality of Service to chosen applications, and/or manage a fair usage of the network according to the requirements of these applications. Mainly, these guarantees may be established:

- Previously, by service level agreements;
- Requested actively by the applications at runtime.

In our work, we propose allowing the *network* control the assignment of resources. To better address fairness in the point of view of the applications, this control is achieved based on a *previous history* of the applications' behavior, in order to better capture each application's requirements. The previous history is furnished by means of time-series data produced by monitoring metrics of choice in the network.

Essentially, our system employs a *probe*, which monitors network flows. A network flow is identified by five attributes: Source IP, Source Port, Destination IP, Destination Port, Protocol. Inside the probe, a *Statistical Engine* uses Finite Automata techniques to detect flow behavior. When a historical high demanding flow appears in the network, the engine recognizes it and takes action. For instance, it might trigger the creation of a Virtual Circuit that will transport that specific high demanding flow.

Currently, this Statistical Engine takes into consideration the duration of the flow for detection. Other possibilities include bandwidth usage, and specific source and destination. Also, we research other methods for statistical analysis, such as mean standard deviations (MSD) and cumulative distribution function (CDF).

The time-series analysis performed by the Statistical Engine may also be employed in a network performance monitoring architecture, such as perfSONAR, to provide services for event triggering, alarming, and statistical auditing. One such application is *anomaly detection*, which can be utilized for performance and security management. A version of the Statistical Engine is under our development for use in the Pegasus Workflow Management System, through the project *Synthesized Tools for Archiving, Monitoring Performance and Enhanced DEbugging* (STAMPEDE), where the applications generate scientific workflows.

Finally, *forecasting* is also a relevant exercise for future capabilities of the Statistical Engine, where the history of the network behavior and usage is exploited to predict future performance.

2. REFERENCES

Synthesized Tools for Archiving, Monitoring Performance and Enhanced Debugging (STAMPEDE). (2010). *Synthesized Tools for Archiving, Monitoring Performance and Enhanced Debugging (STAMPEDE)*.

Alarcon-Aquino, V., & Barria, J. (2001). Anomaly detection in communication networks using wavelets. *Communications, IEEE Proceedings-, 148*, 355-362.

Basseville, M., & Nikiforov, I. V. (1993). *Detection of abrupt changes: theory and application.* Prentice-Hall, Inc.

Chen, C., & Liu, L.-M. (1993). Forecasting time series with outliers. *Journal of Forecasting, 12*, 13-35.

Chen, D., Hu, H., Wang, Z., & Chen, J. (2008). A Novel Method for Network Anomaly Detection Using Superstatistics., (pp. 595-598).

Crone, S., & Dhawan, R. (2007). Forecasting Seasonal Time Series with Neural Networks: A Sensitivity Analysis of Architecture Parameters., (pp. 2099-2104).

Denning, D. E. (1987). An Intrusion-Detection Model. *IEEE Trans. Softw. Eng., 13*, 222-232.

Du, P., Abe, S., Ji, Y., Sato, S., & Ishiguro, M. (2008). Detecting and Tracing Traffic Volume Anomalies in SINET3 Backbone Network., (pp. 5833-5837).

Giorgi, G., & Narduzzi, C. (2008). Detection of Anomalous Behaviors in Networks From Traffic Measurements. *Instrumentation and Measurement, IEEE Transactions on, 57*, 2782-2791.

Gu, Y., McCallum, A., & Towsley, D. (2005). Detecting anomalies in network traffic using maximum entropy estimation. (pp. 32-32). USENIX Association.

Hanemann, A., Boote, J., Boyd, E., Durand, J., Kudarimoti, L., Lapacz, R., et al. (2005). PerfSONAR: A Service Oriented Architecture for Multi-Domain Network Monitoring., (pp. 241-254).

Hussain, F., Kalim, U., Latif, N., & Khayam, S. A. (n.d.). A Decision-Theoretic Approach to Detect Anomalies in Internet Paths. *A Decision-Theoretic Approach to Detect Anomalies in Internet Paths*.

Kai, H., Zhengwei, Q., & Bo, L. (2009). Network Anomaly Detection Based on Statistical Approach and Time Series Analysis., (pp. 205-211).

Kawahara, R., Kamiyama, N., Harada, S., Hasegawa, H., & Asano, S. (2008). Identifying Anomalous Traffic Sources Using Flow Statistics., (pp. 1-5).

Khanna, R., & Liu, H. (2008). Control theoretic approach to intrusion detection using a distributed hidden Markov model. *IEEE Wireless Communications, 15*, 24-33.

Kim, M., Kang, H., Hung, S., Chung, S., & Hong, J. (2004). A flow-based method for abnormal network traffic detection.

Kim, S. S., & Reddy, A. (2008). Statistical Techniques for Detecting Traffic Anomalies Through Packet Header Data. *Networking, IEEE/ACM Transactions on, 16*, 562-575.

Krishnamurthy, B., Sen, S., Zhang, Y., & Chen, Y. (2003). Sketch-based change detection: Methods, evaluation, and applications. *ACM New York, NY, USA*, (pp. 234-247).

Lall, A., Ogihara, M., & Xu, J. (2009). An Efficient Algorithm for Measuring Medium- to Large-Sized Flows in Network Traffic., (pp. 2711-2715).

Li, X., Bian, F., Crovella, M., Diot, C., Govindan, R., Iannaccone, G., et al. (2006). Detection and identification of network anomalies using sketch subspaces. (pp. 147-152). ACM.

Pelham, G. E., & Jenkins, G. M. (1994). *Time Series Analysis: Forecasting and Control.* Prentice Hall PTR.

Quinson, M. (2002). Dynamic performance forecasting for network-enabled servers in a metacomputing environment., *2.*

Ramasubramanian, P., & Kannan, A. (2004). Intelligent multi-agent based back-propagation neural network forecasting model for statistical database anomaly prevention system., (pp. 108-113).

Schweller, R., Li, Z., Chen, Y., Gao, Y., Gupta, A., Zhang, Y., et al. (2007). Reversible Sketches: Enabling Monitoring and Analysis Over High-Speed Data Streams. *Networking, IEEE/ACM Transactions on, 15*, 1059-1072.

Swany, M., & Wolski, R. (2002). Multivariate resource performance forecasting in the network weather service., (pp. 11-11).

Swany, M., & Wolski, R. (2002). Representing dynamic performance information in grid environments with the network weather service.

Thottan, M., & Ji, C. (1998). Proactive anomaly detection using distributed intelligent agents . *Network, IEEE, 12*, 21-27.

Wolski, R. (1997). Forecasting network performance to support dynamic scheduling usingthe network weather service., (pp. 316-325).

Wolski, R. (1998). Dynamically forecasting network performance using the network weather service. *Cluster Computing, 1*, 119-132.

Wolski, R., Spring, N., & Hayes, J. (1999). The network weather service: A distributed resource performance forecasting service for metacomputing. *Future Generation Computer Systems, 15*, 757-768.

Wu, Q., & Shao, Z. (2005). Network Anomaly Detection Using Time Series Analysis., (pp. 42-42).

Yasami, Y., Mozaffari, S., & Khorsandi, S. (2008). Stochastic learning automata-based time series analysis for network anomaly detection., (pp. 1-6).

Zonglin, L., Guangmin, H., Xingmiao, Y., & Dan, Y. (2009). Detecting distributed network traffic anomaly with network-wide correlation analysis. *EURASIP J. Adv. Signal Process, 2009*, 1-11.