

# Synergy: Blending Heterogeneous Measurement Elements for Effective Network Monitoring

Awais Ahmed Awan

Department of Computer Science  
Queen Mary University of London  
awais.awan@dcs.qmul.ac.uk

Andrew Moore

Department of Computer Science  
Queen Mary University of London  
andrew.moore@dcs.qmul.ac.uk

## ABSTRACT

Network traffic matrices are important for various network planning and management operations. Previous work for estimation of traffic matrices is based on either link load records obtained from SNMP or flow level sampled data available from Net-flow records. Sampled flow-level data provides us good approximations for traffic matrices, however, static sampling rates are not desired. People have proposed sampling solutions with dynamic sampling rates which adapt according to the changing behavior of the network. SNMP link level reports, on the other hand are computationally less extensive but do not provide optimum solutions for network traffic estimation. This work addresses the problem of measurement in a wider scope and we propose to study merging together data from multiple sources and using it effectively for network monitoring and measurement tasks. We address this problem in both time and spatial domain and intend to develop a framework for dynamically choosing monitoring points in network and blend the existing and readily information from different sources for network measurement. This information mix from multiple sources can be used for more promising and generalized anomaly detection models. We address this problem in three dimensional space of time, space and application level details.

## Categories and Subject Descriptors

C.2.3 [Computer-Communication Networks]: Network Operations – *Network Monitoring* C.2.5 [Computer-Communication Networks]: Local and Wide Area Networks - *Internet*

## General Terms

Measurement, Performance, Reliability, Security

## Keywords

Network Traffic Monitoring, Spatial Domain Monitoring, Anomaly Detection, Distributed Measurement

## 1. INTRODUCTION AND MOTIVATION

Traffic measurement remains an active research area because of the changing trends in normal and abnormal Internet traffic. Various active and passive measurement techniques have been used for measurement tasks in past, however, there is still growing need to perform the measurement task in more accurate and reliable way. Flow level measurements, like Cisco net-flow, have

been effectively used for estimation of traffic matrices. However, carrying out measurement tasks on today's large IP network is a challenging task due to the large number of OD pairs present in the network. Net-flow sampled data has been used successfully in past [1][3][5] for inferring traffic statistics. Dynamically adapting sampling rates for specific measurement task [11] have gained importance over static sampling rates due to their improved accuracy in the presence of congestion in the network. However, only a little work [2][6][7] has been conducted to address the problem of sampling in spatial domain where we indicate that how to locate monitors in the network-wide domain and adjust sampling rates accordingly.

SNMP traffic data, easily available, gives us volumes of data which are easily manageable and there is no significant impact on router performance. However, there are inherent limitations in SNMP data which include poor data quality due to ambiguous and missing data and also irregular sampling, obtaining traffic matrix and types is difficult, difficulties in detecting DoS attacks, etc.

Recent advancements in measurement domain work include utilization of data from multiple sources. We have seen that link level measurements and flow level measurement have both their own advantages and disadvantages. Nick Duffield et. al. [9] have proposed to use multiple sampling methods in spatial domain and combine them in an optimum way. They use measurements taken from multiple routers for estimating network level flow rates. Similarly Gion Reto et. al. [6] have tried to address the same problem. They have proposed framework for selecting network nodes where monitors should be placed and have devised an automatic adjustment of sampling rates for these geographically dispersed monitors based on characteristics of traffic. Roughan et. al. [8] have proposed tomo-gravity model which effectively combines tomographic and gravity models for estimating traffic matrix. Their work primarily augments link load measurements with the available network and routing configuration information.

Motivated from the above work, we are currently investigating the concept of heterogeneous measurement elements for effective network monitoring. Idea of heterogeneous measurement elements is derived from the need of merging multiple data sources so that they give us robust traffic matrix estimations. We believe that certain disadvantages of different data sources can be reduced when combined with other sources. A spatial knowledge of data is important in this context and network wide data is required. Making decisions that which nodes will be participating for monitoring, which measurement strategy will be used for specific

location, what level of dynamism will be there for choosing measurement scheme, how to dynamically change measurement parameters with the change in traffic volumes, are important motivation for this work.

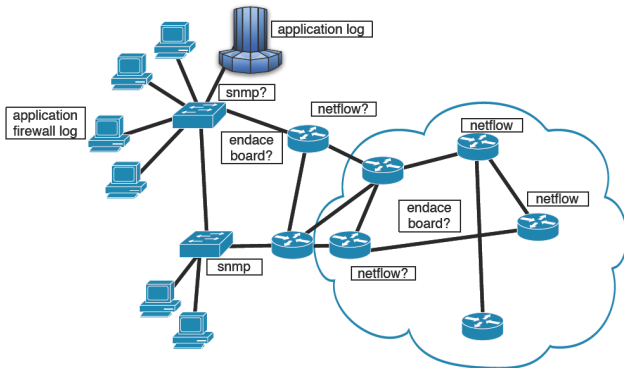


Figure 1: Multiple monitoring vantage points

## 2. PROPOSED METHODOLOGY

We intend to propose a framework for combining heterogeneous measurement elements. Past work show that fusing multiple data sources together and dynamically adjusting of the measurement parameters greatly improve the traffic matrix estimations. We Plan to organize this in a formal methodology. Based on such a framework, we can assess and compare various measurement and monitoring techniques.

## 3. APPLICATIONS

The proposed work opens new research directives in the measurement domain. Network wide data when combined in an optimal way, that is, enough data from multiple sources with minimum overhead, is an important asset and can be used in various network planning, management and security tasks. Our proposed work can be used to obtain anomaly matrices formulated using network wide information. Based on the information available, we can model the behavior of Internet traffic and study anomalies and variances. We believe that spatial domain data is essential for planning overall network activities and measurement tasks.

## 4. CONCLUSION

We have outlined an ambitious plan to examine the problem of network monitor placement. We intend examination of the data

from diverse sources to provide an insight into the accuracy of each. We then intend an examination of the merging of these sources to provide improved sources of information. Finally, we plan to express the spatial location problem as a cost-optimization trading density/sophistication of monitoring systems against accuracy of result.

## 5. REFERENCES

- [1] Yin Zhang, Mathew Roughan, Carsten Lund, David Donoho. An information theoretic approach to traffic matrix estimation. In *ACM SIGCOMM*, Germany August 2003.
- [2] Anukool Lakhina, Mark Crovella, Christophe Diot. Detecting Mining anomalies using traffic feature distribution. In *ACM SIGCOMM*, Philadelphia August 2005.
- [3] N. G. Duffield and M. Grossglauser, "Trajectory Sampling for Direct Traffic Observation", *IEEE/ACM Transactions on Networking*, vol. 9, pp. 280-292, 2001.
- [4] N.G. Duffield, C. Lund, M. Thorup, "Charging from sampled network usage," *ACM SIGCOMM Internet Measurement Workshop 2001*, San Francisco, CA, November 1-2, 2001.
- [5] N.G. Duffield, C. Lund, M. Thorup, "Learn more, sample less: control of volume and variance in network measurements", *IEEE Trans. of Information Theory*, vol. 51, pp. 1756-1775, 2005.
- [6] Gion Reto Cantieni, Gianluca Iannaccone, Chadi Barakat, Christophe Diot, Patrick Thiran. Reformulating the Monitor Placement Problem: optimal Network-Wide Sampling ??
- [7] A. Lakhina, M. Crovella, and C. Diot. Diagnosing Network-Wide Traffic Anomalies. In *ACM SIGCOMM*, Portland, August 2004.
- [8] Yin Zhang, Matthew Roughan, Nick Duffield, Albert Greenberg. Fast accurate Computations of Large-Scale IP Traffic Matrices from Link Loads. In *ACM SIGMETRICS*, 2003
- [9] Nick Duffield, Carsten Lund, Mikkel Thorup. Optimal Combination of Sampled Network Measurements. In *IMC 2005*
- [10] J. Cao, D. Davis, S. Vander Wiel, and B. Yu. Time-varying network tomography:router link data. *Journal of American Statistics Association*, pages 1063.1075, 2000.
- [11] Baek-Young Choi, Jaesung Park, Zhi-Li Zhang. Adaptive Random Sampling for Traffic Load Measurements. In *IEEE ICC 2003*.