



INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student : AWNISH KUMAR

Roll Number : 11620104

Programme of Study : Ph.D.

Thesis Title: Estimation, Reconstruction, and Analysis of Internet Traffic Data using Decomposition Techniques

Name of Thesis Supervisor(s) : Dr. Vijaya V. Saradhi
Dr. T. Venkatesh

Thesis Submitted to the Department/ Center : Computer Science and Engineering Department

Date of completion of Thesis Viva-Voce Exam :
March 05, 2021

Key words for description of Thesis Work : Internet Traffic Data, Estimation, Compressive Sensing, Structural Analysis, Volume Anomaly Analysis, Multi-view Subspace Learning, Matrix Decomposition, Tensor Decomposition

SHORT ABSTRACT

With the ever-increasing demand for Internet services, the knowledge of Internet traffic data is very essential for addressing various network-wide applications within an Internet Protocol (IP) backbone network. However, the high-dimensional multivariate structure of Internet traffic data proves to be a major hurdle in addressing these applications. Decomposition techniques based on Principal Component Analysis (PCA), which provides a low-dimensional representation of Internet traffic data, have garnered widespread popularity among network researchers in the last few decades. Despite widespread applicability, PCA suffers from fundamental limitations stemmed in its assumption. This paves way for the research motivated towards exploring more apt decomposition techniques. In this thesis, we primarily focus on three applications, namely estimation, reconstruction, and analysis of Internet traffic data. The first contribution of this thesis is to propose the use of matrix-CUR decomposition and a multi-view learning technique based on CCA for traffic matrix estimation to avoid the overheads associated with direct measurement of Internet traffic data. The second contribution of the thesis is to propose the use of matrix-CUR decomposition for the reconstruction of missing values in the traffic matrix. For the reconstruction of missing values in traffic tensor, we introduce relative-error bound tensor-CUR (TCUR-REB) decomposition. Both these techniques are computationally inexpensive and TCUR-REB alleviates the limitation of having apriori knowledge of tensor rank. The third contribution of the thesis is to propose the interpretable decomposition, namely correspondence analysis and matrix-CUR decomposition, for structural analysis and volume anomaly analysis of traffic matrices. Both these techniques alleviate the limitations of PCA associated with the lack of interpretability and assumption of continuous random variables. We have performed extensive

experimentation on real and synthetic Internet traffic data to demonstrate the efficacy of the proposed techniques. Results show that these proposed techniques outperform the state-of-the-art techniques in terms of their corresponding evaluation measures.

