Dimension Reduction of Network Bottleneck Bandwidth Data Space

Peng Sun, Yang Chen*, Yibo Zhu,

Xiaoming Fu*, Beixing Deng, Xing Li

Department of Electronic Engineering, Tsinghua University, Beijing, China

*Institute of Computer Science, University of Goettingen, Germany

Motivation

For latency-aware applications, round-trip-time estimation has been studied extensively. But there are also lots of bandwidth-aware applications. Prediction of bottleneck bandwidth has received much less attention.

Therefore, we attempt to design a new system to predict bottleneck bandwidth, based on matrix factorization.

As a first step, we need to prove:

1) Low-rank nature of bottleneck bandwidth matrices

2) Feasibility of reducing dimension of bottleneck bandwidth data space

Matrix Factorization

The network bottleneck bandwidth data space can be modelled as square matrix ***B***. Apply Principle Component Analysis on ***B***:

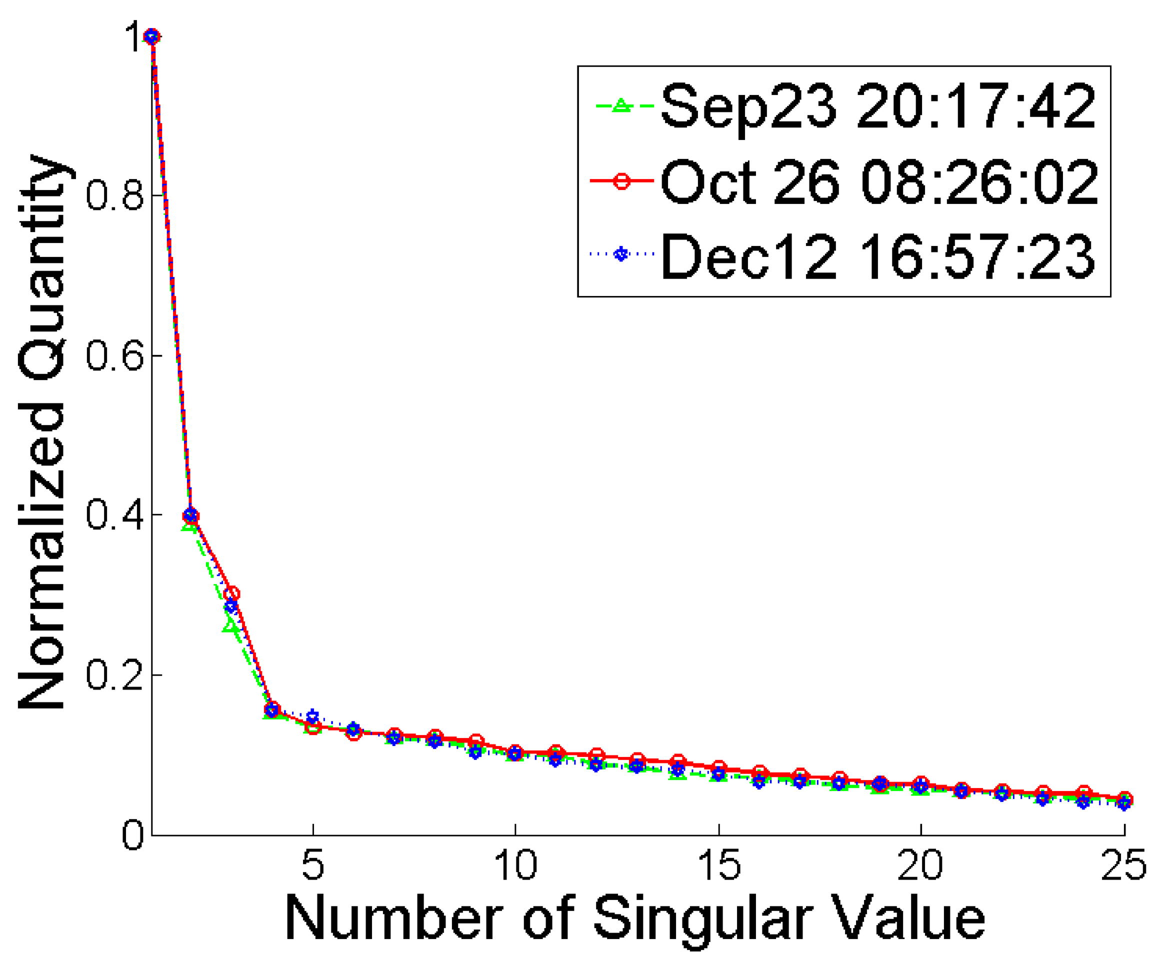
$$B = USV^T$$

$$= \begin{bmatrix} U_{11} & \cdots & U_{1k} & \cdots & U_{1n} \\ U_{21} & \ddots & & & \vdots \\ U_{31} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ U_{n1} & \cdots & U_{nk} & \cdots & U_{nn} \end{bmatrix} \begin{bmatrix} S_{11} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & S_{kk} & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & S_{nn} \end{bmatrix} \begin{bmatrix} V_{11} & \cdots & V_{1k} & \cdots & V_{1n} \\ V_{21} & \ddots & & & \vdots \\ V_{31} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ V_{n1} & \cdots & V_{nk} & \cdots & V_{nn} \end{bmatrix}^T$$

$$\approx U'S'V'^T = (U'\sqrt{S'})(V'\sqrt{S'})^T = XY^T$$

Principle Component Analysis

We attempt to analyze the magnitude of singular values of ***B***.



It shows that singular values decrease very fast. Considering the 'Oct 26' line, the 4th singular value (0.156) is the first one that is smaller than 0.2.

Methodology

Based on HP Scalable Sensing Service (S3), 250 interconnected hosts are extracted out for our evaluation. From September 23 to December 23 2009, we collect bottleneck bandwidth data every four hours. Finally we have ***491 datasets across 3 months*** for evaluation.

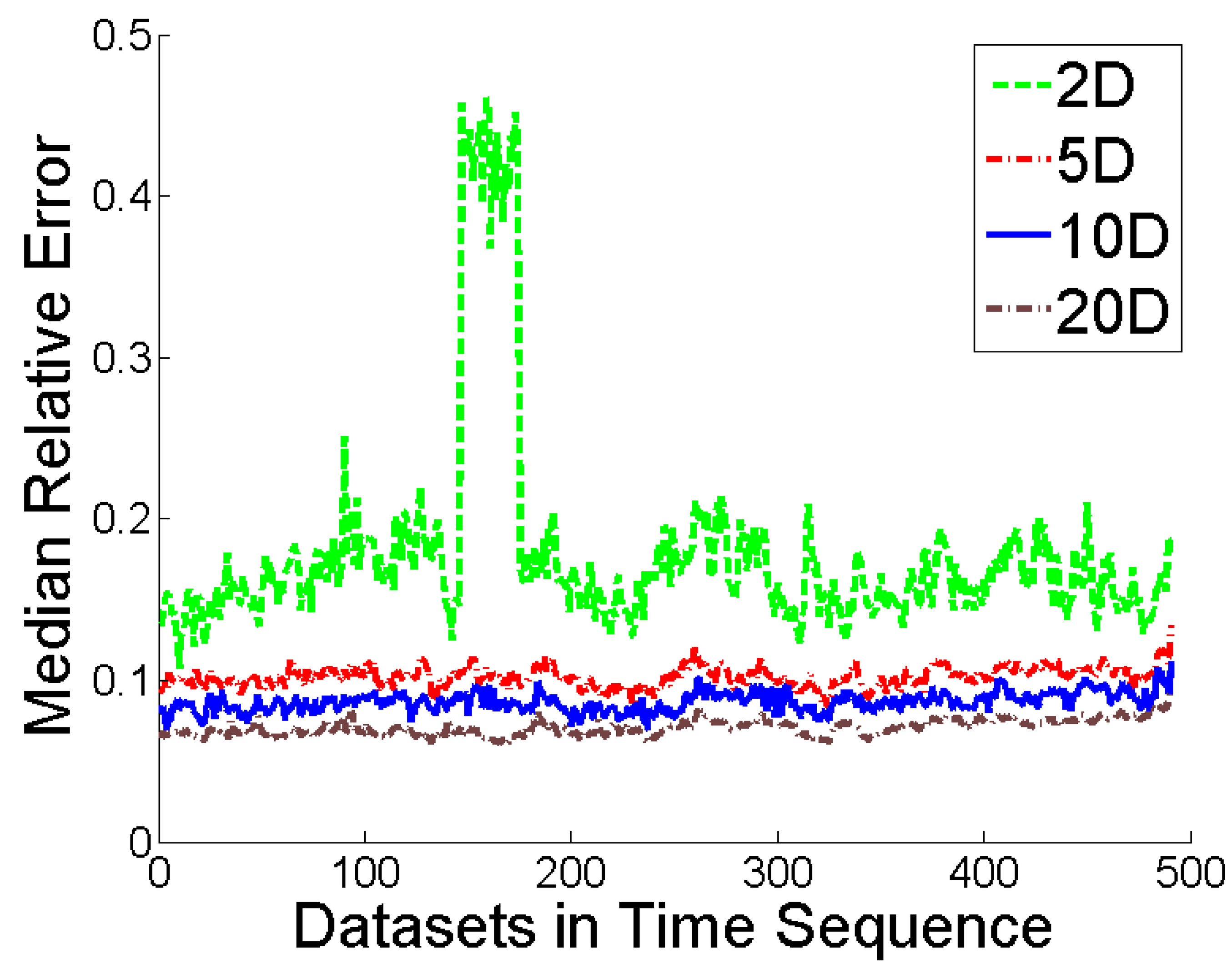
We compare the approximated matrix with the original one for evaluation. Relative error is defined as follows:

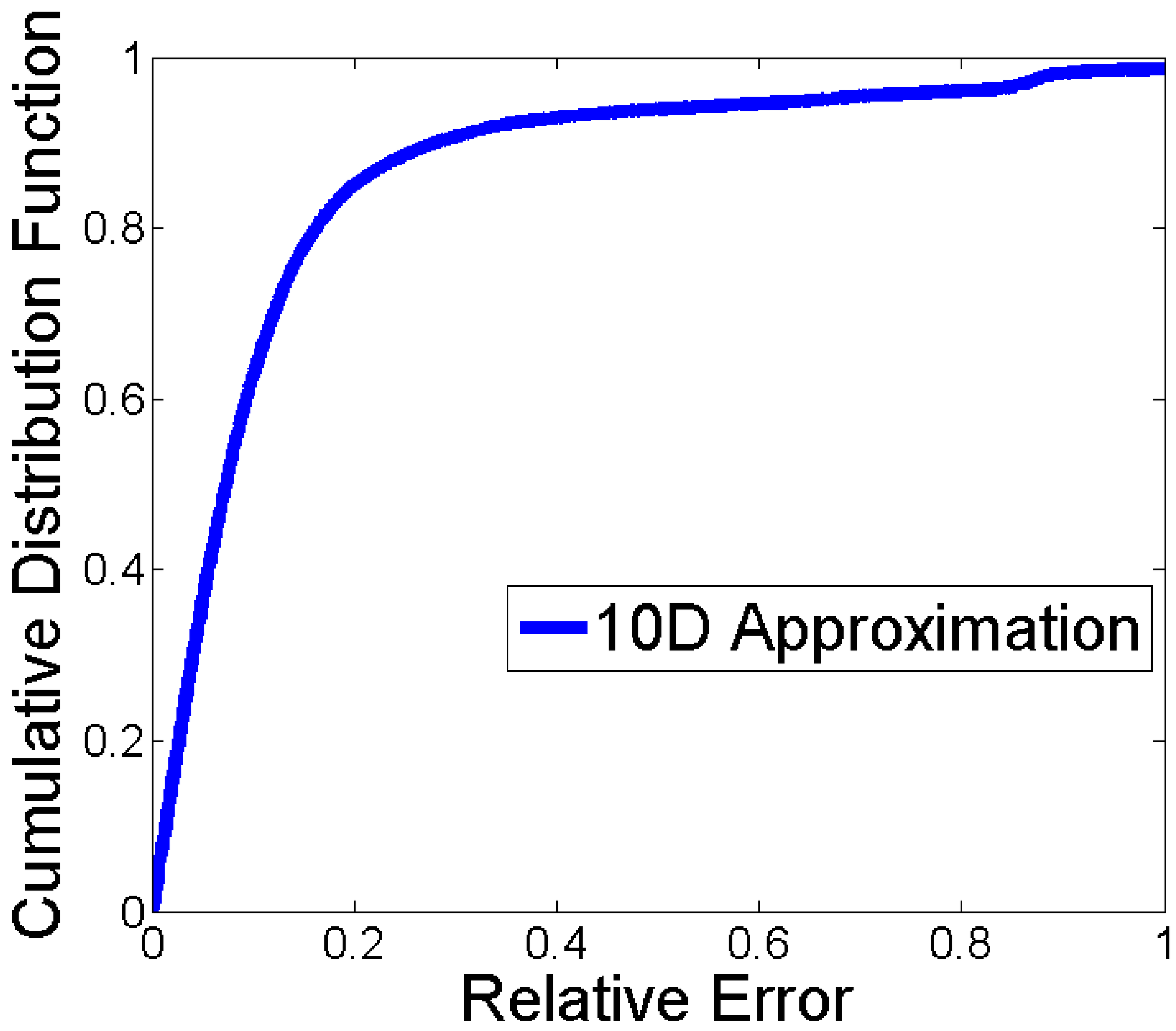
If $(i,j) \in \{(m,n) \mid b_{mn} \neq -1\}$, $\quad relative\ error_{ij} = \dfrac{|b_{ij}' - b_{ij}|}{b_{ij}}$

Evaluation of Dimension Reduction

The figure right shows the median relative error when the dimension of all the 491 datasets are reduced to 2D, 5D, 10D and 20D.

The average of median relative error for 10D approximation is only 8.65% among all the 491 datasets.





Considering the tradeoff between computation complexity and target dimension of reduction, a 10D approximation is carried out to show the cumulative distribution function of relative error in figure left.

The 90th percentile relative error is only 0.281, meaning that 90% of the data have lower relative error than 0.281.

Conclusion

1. Dimension of bottleneck bandwidth data space can be reduced ***from 250D to 10D***
2. The average of median relative error for approximation is ***only 8.65%*** among ***491 datasets***.
3. The 90th percentile relative error of 10D approximation is ***only 0.281***

Acknowledgement

This work is supported by National Science Foundation of China (No.60850003).

Future work

We would design a scalable bottleneck bandwidth prediction system based on matrix factorization, utilizing the low-rank nature and low relative error in dimension reduction.