Impact of the Inaccuracy of Distance Prediction Algorithms on Internet Applications: An Analytical and Comparative Study

R. Zhang, C. Tang, Y. C. Hu, S. Fahmy, and X. Lin
Purdue University, IBM T.J. Watson Research Center

Research Questions

- How well do network distance prediction (NDP) algorithms perform within application contexts?

- How do NDP algorithms' performance compare with on-demand measurement and optimal algorithms?

Background

- Landmark-based algorithms (triangulated heuristic)

- Coordinate-based algorithms (Global Network Positioning [GNP])

- Vector-based algorithms (Internet Distance Estimation Service [IDES])

- Server-based algorithms

Methodology

- Uses King, PlanetLab, and AMP datasets

- Compares *relative prediction error* and *directional relative prediction error*

- Accuracy predicting long and short links

Overlay Multicast

- Task: given a network of nodes, construct a peer-to-peer unicast tree to efficiently distribute content to multiple nodes

- Optimize on latency

- Evaluates using trees constructed by MST, ESM, and LGK algorithms

Evaluation

- Tree costs (MST, ESM): Sum of distances between tree nodes

- Delay stretch: Ratio of delay using tree vs. delay using direct unicast

- Results: Distance prediction algorithms produce poor trees, algorithm complexity does not help

- Cause: Poor prediction of close distances

Nearest Neighbor
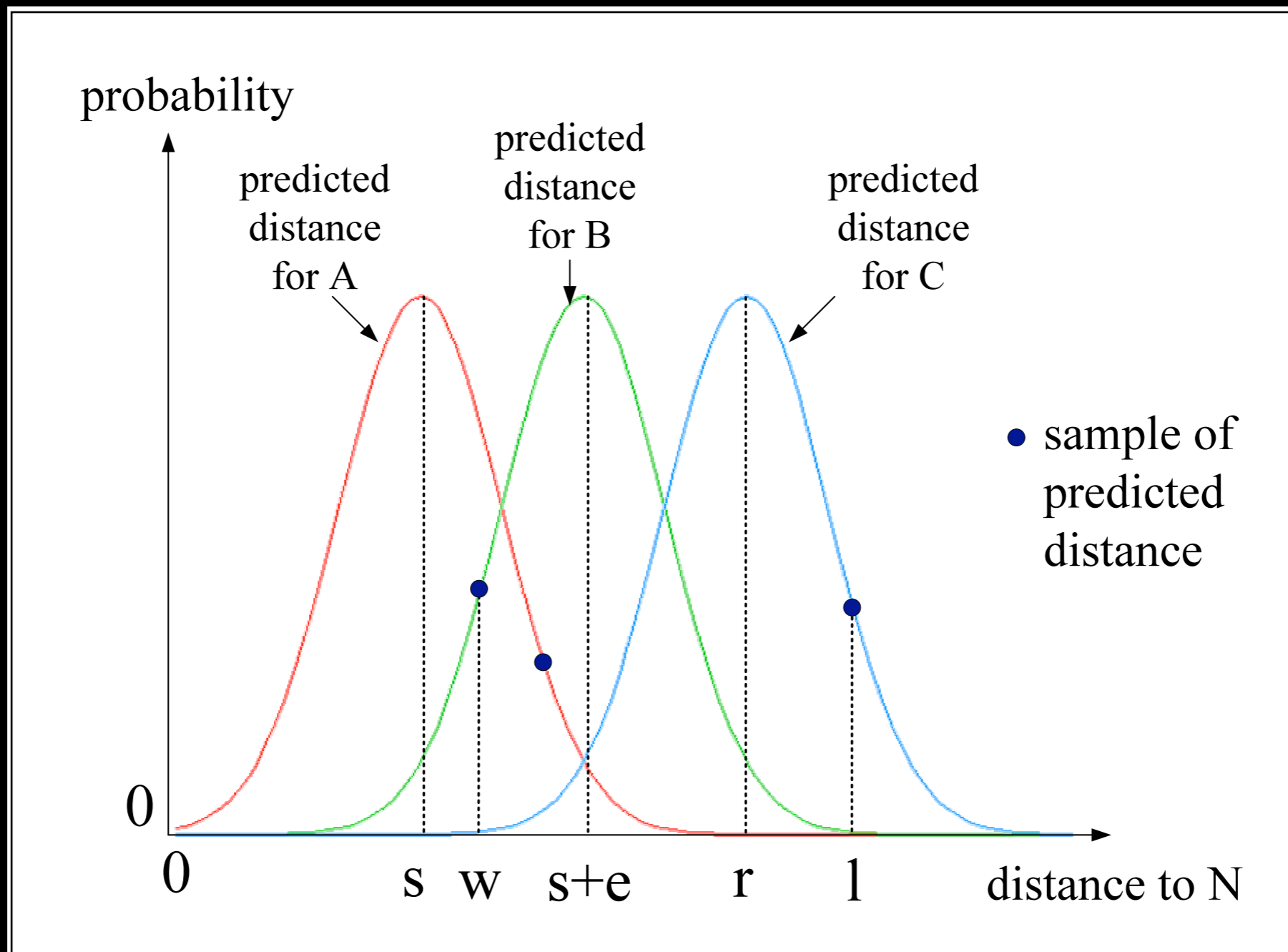
Directional Relative Error
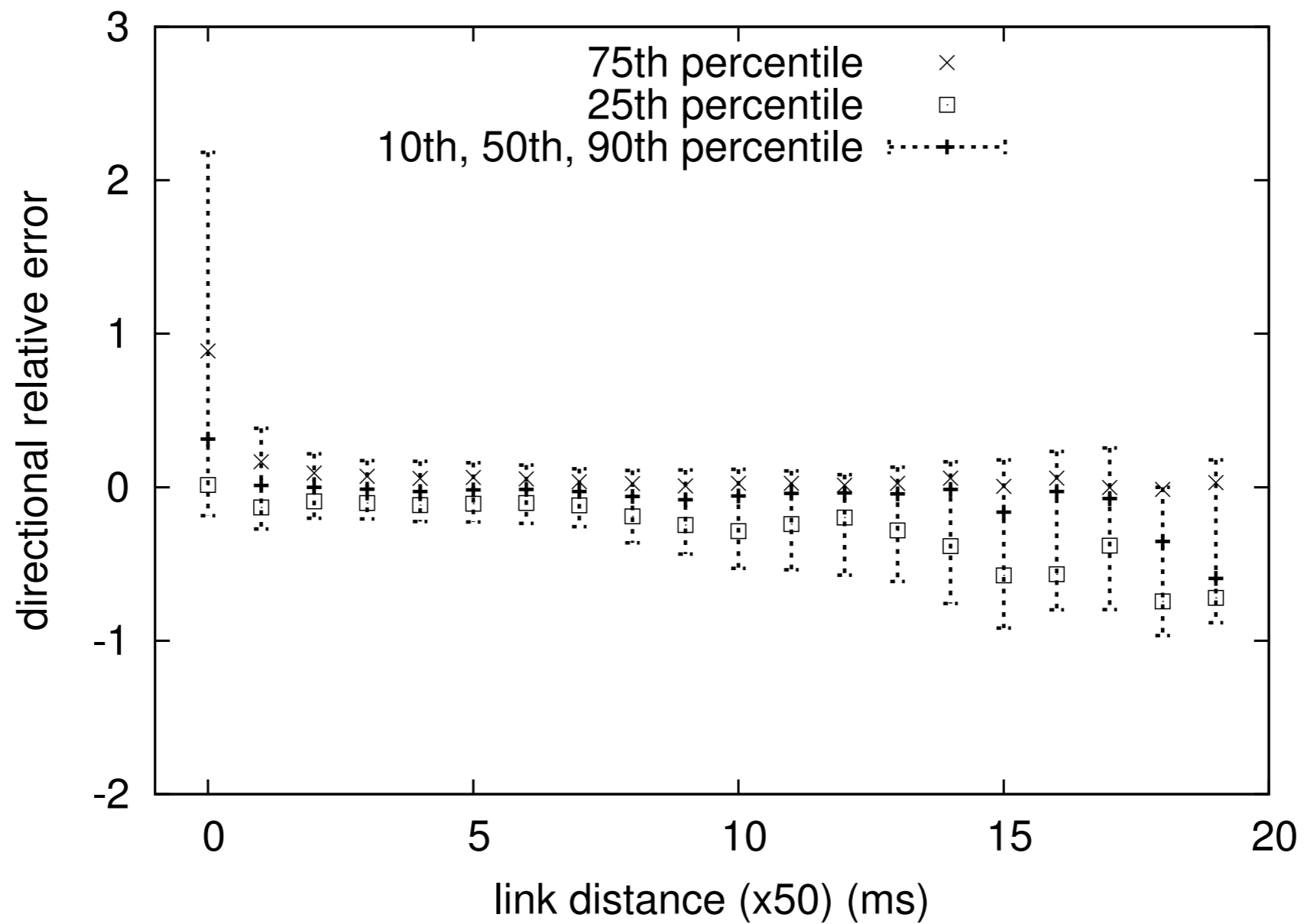
Other Analysis

- Prediction algorithms add a cost to the tree

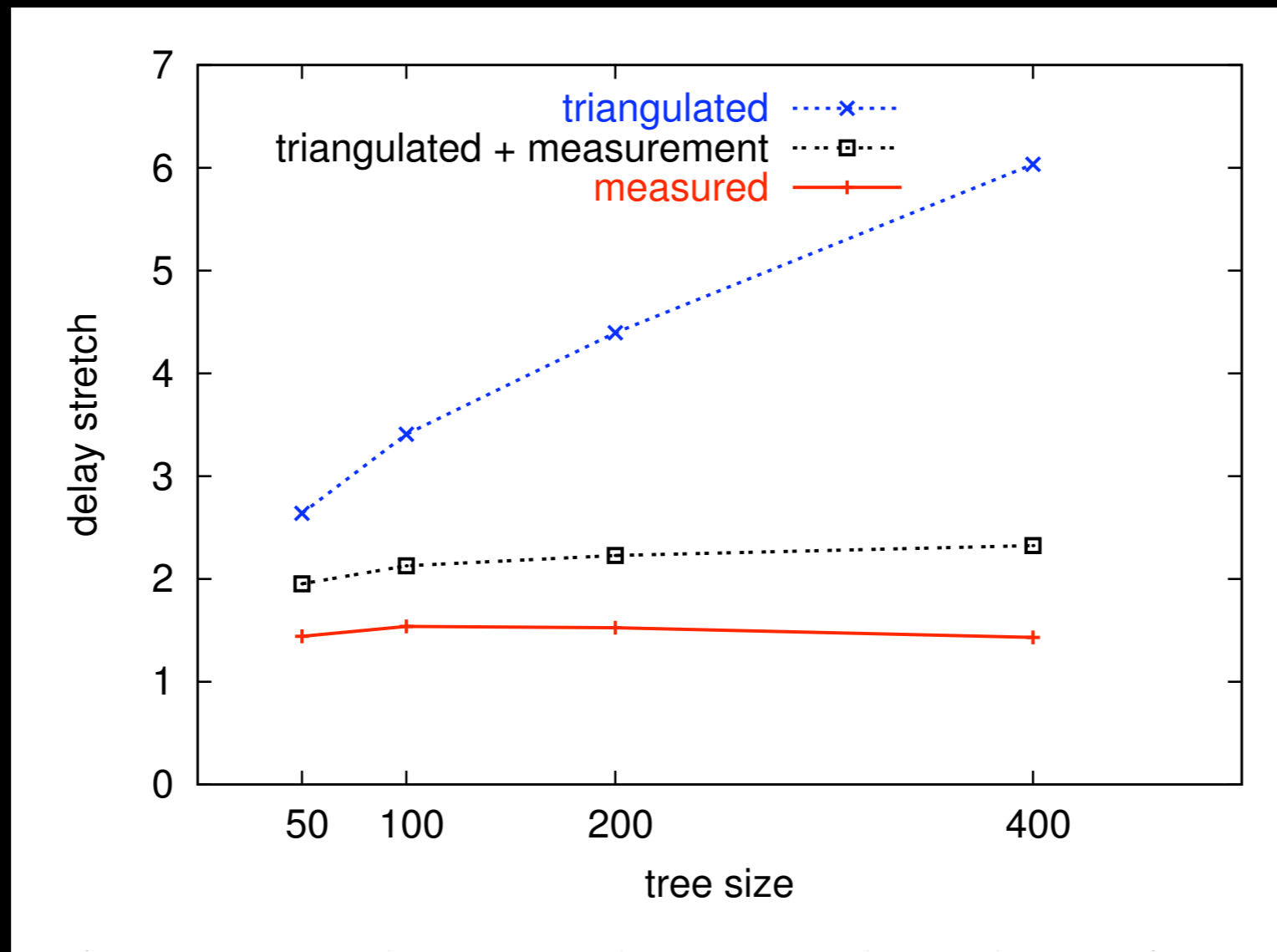- Network distances are close to normally distributed

Selected Measurements

- Since close neighbor prediction is problematic, substitute measurement for close nodes

- Algorithm: Take the set of nodes with smallest predicted distances and replace predictions with measurements

MST + GNP
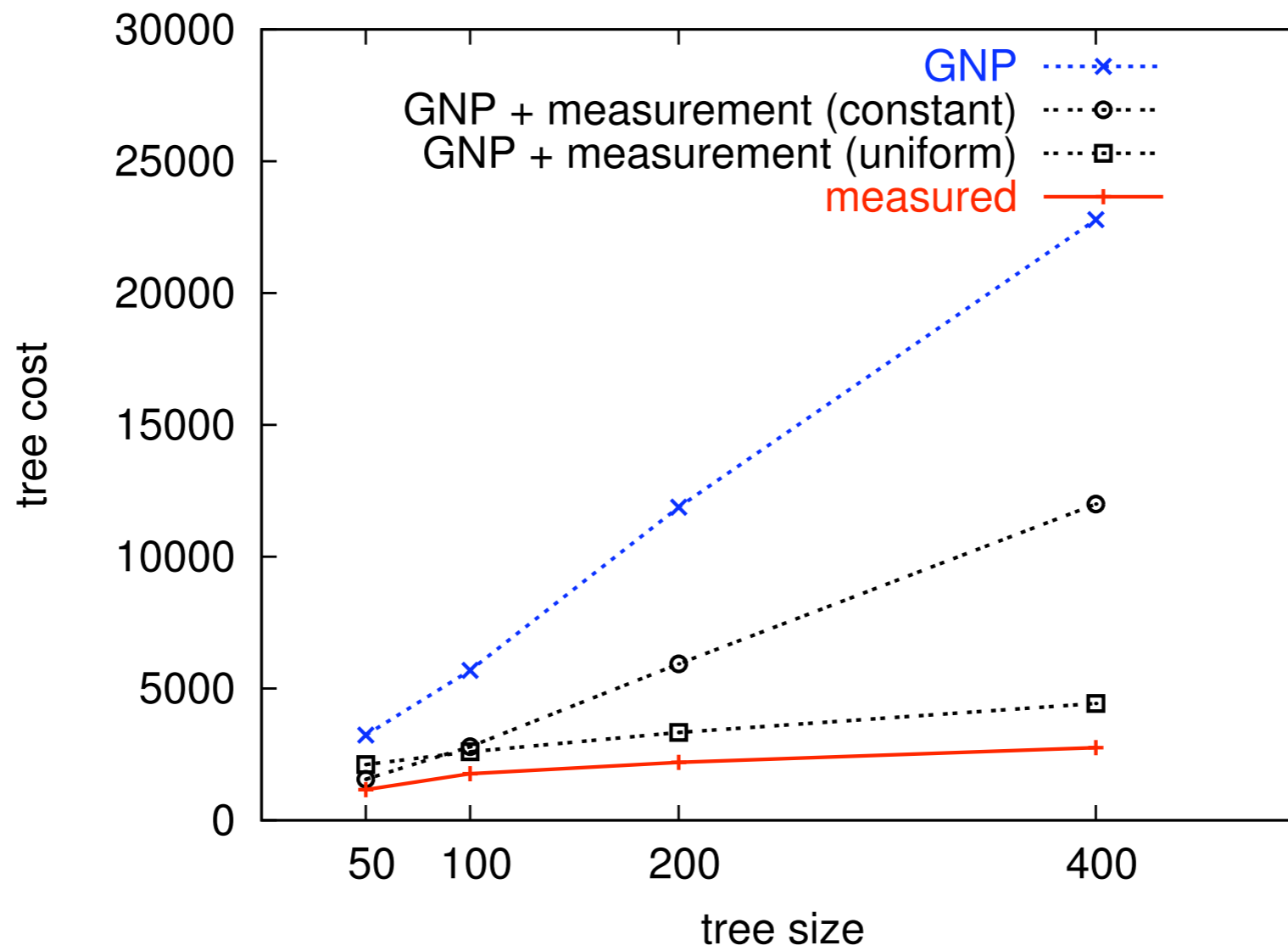
Triangulated Heuristic + LGK

Conclusions

- Measurement gains are similar for server selection and network overlay applications

- Complexity of algorithm does not produce significant gains for applications

- Differences between algorithms are smaller than differences between prediction and prediction + measurement