Poster: Neighbor Selection Based on TIV Severity Sort Model in Vivaldi Network Coordinate System

Peng Sun*, Yang Chen†, Beixing Deng* and Xing Li*
*Department of Electronic Engineering, Tsinghua University, China
Email: sp06@mails.tsinghua.edu.cn
†Institute of Computer Science, University of Göttingen, Germany
Email: yang.chen@cs.uni-goettingen.de

***Abstract*—Network Coordinate (NC) system is an efficient and scalable mechanism to estimate the distance between Internet hosts. However, the existence of Triangle Inequality Violation (TIV) decreases the accuracy of NC system. With focus on most widely used NC system, *Vivaldi*, we propose an effective mechanism of neighbor selection based on TIV Severity Sort to improve *Vivaldi* performance. By sorting existing hosts based on corresponding edges' TIV severity, the 90th percentile relative error(NPRE) of *Vivaldi* is decreased by 13.9%. The convergence rate is improved, and the final median prediction error is 7.9% smaller.**

I. Introduction

Accurate distance estimation between two hosts is crucial to many peer-to-peer applications on the Internet. Network Coordinate (NC) system is one promising approach to predict the distance (round trip time) without time-consuming end-to-end measurement. The basic idea of NC system is to assign the Internet host a coordinate to represent its position, and calculate the distance using Euclidean Distance.

Among the already proposed NC systems, *Vivaldi* [1] is regarded as the most representative and most widely-used one. It has been deployed in many P2P services, such as Bamboo DHT, Stream-Based Overlay Network and Azureus BitTorrent. Several studies [2] [3] have demonstrated that the accuracy of *Vivaldi* can be greatly degraded by the Triangle Inequality Violation(TIV) phenomenon between Internet hosts. Several methods have been proposed to evaluate the severity of TIV. [2] introduces the concept of TIV severity.

In this paper, we propose a new mechanism of neighbor selection based on TIV Severity Sort to improve *Vivaldi*'s accuracy, named as *TIV Severity Adjusted Vivaldi*. Instead of random selection in the original *Vivaldi*, for each host, the existing hosts are sorted based on the TIV severity of the edges within them. Then the least-severe ones are selected as that host's neighbors. In this *TIV Severity Adjusted Vivaldi*, each host has neighbors less TIV-severe for coordinate reference. It results in a more accurate distance estimation. The major contributions of our work are as follows. 1) Improve the accuracy: The 90th percentile relative error is decreased by 13.9%. 2) Improve the convergence rate: The convergence rate is increased and final median prediction error is decreased by 7.9%.

In the following sections, we first illustrate the proposed mechanism. The performance of *TIV Severity Adjusted Vivaldi* is evaluated and compared with the original system in Section III. We conclude this paper in Section IV.

II. System Design

A. TIV in Internet Hosts

Triangle Inequality Violation occurs when hosts A, B and C are in a relationship like $d(A,B) + d(B,C) < d(A,C)$, where $d(X,Y)$ represents the round trip time between host X and Y. *Vivaldi* is based on Euclidean space embedding. In Euclidean space triangle inequality has to be obeyed. Therefore, when facing TIV, *Vivaldi* forces edges to shrink or stretch in the embedding space to resolve the violation. It introduces inaccuracy into the calculation of coordinates.

The 90th percentile relative error (NPRE) is used to evaluate the accuracy of *Vivaldi*. It guarantees 90% of the hosts have lower relative error than the value of NPRE. Previous studies have shown that the more severe the TIV, the higher the NPRE of *Vivaldi*. 48.07% of all triples of hosts in the AMP [4] data set are violating triangle inequality (defined as TIV ratio), while TIV ratio in the P2PSim [4] data set is 97.33%. The NPRE of AMP is only 20.2%, compared to P2PSim's being 56.3%.

B. Neighbor Selection Based on TIV Severity

TIV severity [2] is defined as follows. Let S be the set of all hosts and $A, B, C \in S$. Set Γ is defined as $\{X | d(A,X) + d(X,C) < d(A,C), X \in S\}$:

$$TS_{A,C} = \frac{1}{|S|} \sum_{B \in \Gamma} \frac{d(A,C)}{d(A,B) + d(B,C)}$$

The concept is extended to measuring TIV severity of a host. The TIV severity of a host is defined as the average of TIV severity values of the edges between that host and all the others in S.

For host A

$$TS_A = \frac{1}{|S|} \sum_{B \in S} TS_{A,B}$$

Our mechanism of neighbor selection runs as follows. Suppose each host has L neighbors for reference. The first L joining-in hosts select themselves as each other's neighbors. When the number of existing hosts in the system (defined as K) is larger than L, for every new host (e.g. A), the existing

Algorithm : *TIV Severity Adjusted Vivaldi*

1: Neighbor Set Up

WHEN $K < L$
  select all existing hosts as i's neighbors
  incorporate i into S
WHEN $K \geq L$
  WHILE $j \in S$
    calculate TS_j and sort
  select smallest-TS L hosts as i's neighbors
  incorporate i into S

2: Coordinate Iteration

WHILE $i \in S$
  WHILE j is i's neighbors
    $w = e_i/(e_i + e_j)$
    $e_s = |\|x_i - x_j\| - rtt_{ij}|/rtt_{ij}$
    $e_i = e_s \times c_e \times w + e_i \times (1 - c_e \times w)$
    $\delta = c_c \times w$
    $x_i = x_i + \delta \times (rtt_{ij} - \|x_i - x_j\|) \times u(x_i - x_j)$

K: Number of Existing Hosts, L: Number of Neighbors
S: Set of Existing Hosts

hosts' TIV severity are calculated based on the distance information within those K existing hosts and sorted in ascending order. Using only existing hosts' information guarantees the efficiency of *Vivaldi*. Since smaller TIV severity means less TIV-severe, the first L hosts of the sequence are selected as A's neighbors.

The pseudo-code algorithm of our *TIV Severity Adjusted Vivaldi* is shown above. Host i is a newly joining-in host. It has coordinate x_i and local error e_i. x_i is adjusted based on observation of the distance between i and its neighbor j (rtt_{ij}). In the algorithm, c_e and c_c are tunable parameters. For each host i, the initial value of e_i is set to 1.

III. Performance Evaluation

In our experiment, we compare the original *Vivaldi* and *TIV Severity Adjusted Vivaldi*. The NC system uses a 10-D Euclidean space and sets 32 neighbors for each host.

We use three typical data sets from real Internet measurement to evaluate the performance of our proposed mechanism. The data sets are AMP, P2PSim and Meridian. Simulations are run on all three data sets and similar conclusions are drawn. Due to space limitation, we only show the results of P2PSim. 256 nodes are randomly extracted out of this 1143-node data set to give a smaller data set (P2P-256) for our experiments.

First, we evaluate the performance of relative error. Fig.1 shows the comparison of the proposed mechanism and the original *Vivaldi*. The result of P2P-256 shows that the NPRE of *TIV Severity Adjusted Vivaldi* is 45.7%, 13.9% smaller than that of the original *Vivaldi*-53.1%.

Next, we evaluate the numbers of rounds required for convergence under a flash-crowd scenario, meaning all hosts join simultaneously. The median prediction error is plotted as a function of *Vivaldi* update rounds used per host in Fig.2. The result shows *TIV Severity Adjusted Vivaldi* converges faster than the original one. In the 10th round, the median prediction

Fig. 1. Relative Error Rate

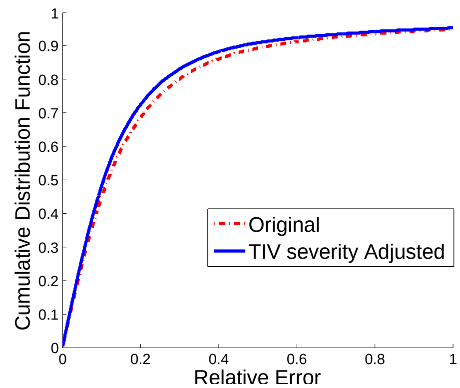
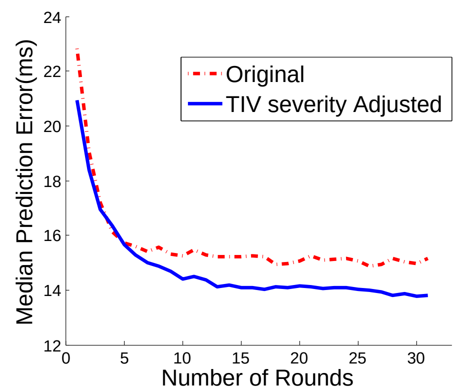


Fig. 2. Convergence Behavior



error of *TIV Severity Adjusted Vivaldi* converges to 14.4ms, while that of the original is 15.24ms. What's more, the final error of our mechanism is smaller. In P2P-256, it is 13.78ms, 7.9% smaller than the original one's 14.96ms.

IV. Conclusion

In this paper, we propose an effective mechanism of neighbor selection in *Vivaldi*. In the selection of a new host's neighbors, *TIV Severity Adjusted Vivaldi* sorts existing hosts based on TIV severity of the edges within them, and the least TIV-severe ones are picked as the neighbors. Evaluation results show that compared with the original *Vivaldi*, our proposed mechanism not only achieves lower relative error, but also converges faster and has lower final median prediction error.

Acknowledgment

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

References

[1] F. Dabek, R. Cox, F. Kaashoek and R. Morris. *Vivaldi: A Decentralized Network Coordinate System*. In Proc. of ACM SIGCOMM, 2004.
[2] G. Wang, B. Zhang, T.S.E. Ng. *Towards Network Triangle Inequality Violation Aware Distributed Systems*. In Proc. of ACM IMC, 2007.
[3] S. Lee, Z.-L. Zhang, S. Sahu, and D. Saha. *On Suitability of Euclidean Embedding of Internet Hosts*. In Proc. of ACM SIGMETRICS, 2006.
[4] Y. Mao, L. Saul, J.M. Smith. *IDES: An Internet Distance Estimation Service for Large Network*. IEEE Journal on Selected Areas in Communications (JSAC), 2006.