Journal of Network and Computer Applications 45 (2014) 191–202

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Conditional clustered matrix factorization based network coordinate system

Rawya Rizk [a,*], Walaa Saber [a], Hany Harb [b]

[a] Electrical Engineering Department, Port Said University, 42523 Port Said, Egypt
[b] The Computers & Systems Engineering Department, Faculty of Engineering, Al-Azhar University, Cairo, Egypt

ARTICLE INFO

Article history:
Received 11 January 2014
Received in revised form
26 May 2014
Accepted 16 July 2014
Available online 1 August 2014

Keywords:
Clustering
Matrix factorization
Network coordinate
Prediction accuracy
Silhouette graph

ABSTRACT

Network coordinate (NC) systems that use Euclidean distances suffer from the existence of Triangle Inequality Violations (TIVs). Matrix factorization (MF) based NC system is an alternative approach towards better prediction accuracy and can remove TIV. However, the prediction accuracy for short links in these systems still suffers from low prediction accuracy compared with the overall prediction accuracy. Two-layer systems have been proposed to improve the prediction accuracy for short links. They divide the whole space into several location-based clusters and run NC systems on both global layer and local layer. However, these systems do not improve the prediction accuracy for short links in the clusters with a few hosts. In this paper, the Conditional Clustered Network Coordinate (*CCNC*) system is proposed. It divides the space into a number of clusters in a balanced, dynamic, and decentralized way. In the proposed system, any node can join or disjoin the system without affecting the system accuracy. The performance of the *CCNC* system is evaluated with the King data set and the PlanetLab data set to be compared against two well known NC systems: *Phoenix* and *Pancake*. The simulation results show that *CCNC* outperforms *Phoenix* and *Pancake* significantly in terms of estimation accuracy, expected time to construct the clusters, and the communication overhead. Moreover, *CCNC* converges very fast and it is simple, scalable, dynamic, and robust under different dimension values.

1. Introduction

The Network Coordinate (NC) system is an efficient and scalable mechanism for Internet distance prediction between any two Internet nodes without explicit measurements (Donnet et al., 2010). The NC system reduces the active probing overhead significantly, which is especially beneficial to large-scale distributed applications. To date, NC systems are widely used in many services, including locality-aware server selection (Fu et al., 2013; Zhang et al., 2014), which is beneficial to Bit Torrent-based file sharing (Pietzuch and Ledlie, 2006) and cloud services (Klein et al., 2012; Malik et al., 2014). Detour finding (Zhu et al., 2010; Lim et al., 2005) is another promising application based on NC systems, which is useful in overlay routing (Klein et al., 2013) and multi-player online gaming (Agarwal and Lorch, 2009). In the next generation network, NC systems are again useful in areas such as route selection for IPv6 multihomed sites (de Launois et al., 2005).

In most of the NC systems such as *Vivaldi* (Dabek et al., 2004), *GNP* (Ng and Zhang 2002), *PIC* (Costa et al., 2004), and *ICS* (Lim et al., 2005), each host is assigned a set of numbers called coordinates to represent its position in the Euclidean space, and the distance between any two hosts can be predicted by their coordinates using Euclidean distance. Unfortunately, the accuracy of predicted distances is largely hurt by the persistent occurrence of Triangle Inequality Violation (TIV) in measured Internet distances. On the other hand, Matrix Factorization (MF) based NC systems such as *IDES* (Mao et al., 2006), *DMF* (Liao et al., 2010), and *Phoenix* (Chen et al., 2011) provide an alternative approach towards better prediction accuracy and completely remove the TIV constraint. However, the experimental study found that the prediction accuracy for short links is significantly worse compared with the overall prediction accuracy for this approach (Chen et al., 2011; Zhang et al., 2006). Two-layer systems that divide the whole space into several location-based clusters and run NC system per cluster like *Pharos* (Chen et al., 2009), *Pancake* (Chen et al., 2010), and *Tarantula* (Chen et al., 2011) have been proposed to improve the prediction accuracy for short links. However, the clustering operation for these systems was based on partitioning the whole network space into clusters based on a predetermined set of a common landmark/anchor infrastructure. Such systems do not provide any flexibility or balancing in the clustering operation. In addition, the clusters with a few hosts have prediction accuracy lower than the other clusters.

* Corresponding author.
E-mail addresses: r.rizk@eng.psu.edu.eg (R. Rizk), walaa_saber@eng.psu.edu.eg (W. Saber), harbhany@yahoo.com (H. Harb).

http://dx.doi.org/10.1016/j.jnca.2014.07.027
1084-8045/

In this paper a new NC system is proposed whose goal is twofold. The first goal is constructing balanced clusters in an autonomous way. The second goal is providing a dynamic-managed clustering structure to overlay-based applications, in order to allow both topology awareness and scalability of these applications. The proposed system divides the space with a novel approach, forming several cluster-based systems. With this space dividing approach, the prediction accuracy for intra-cluster links for all clusters are improved.

The clustering operation for this system consists of three main phases. First, the initial clustering phase is based on the knowledge of a set of nodes where the system starts with a subset of nodes constructing minimum number of disjoint clusters that can contain maximum number of nodes and create the initial cluster heads (CHs) list. The second phase is the steady state phase which handles joining the new hosts to the system in an appropriate way. The distance prediction third phase applies local and global *Phoenix* system to predict the short and the long distances respectively.

The proposed system converges very fast reducing the expected time to construct clusters for the whole network. It also provides a more balanced clustering operation than other NC systems like *Pancake*. The clustering operation does not depend on geographical information which is hard to implement. It provides dynamic clustering that enhances the clustering of space periodically and then provides good prediction accuracy.

The rest of this paper is organized as follows. Section 2 presents the NC system background and related works. In Section 3, the design of the proposed NC system is introduced. In Section 4, the simulation results of the proposed system are presented and compared with the existing NC systems. Section 5 presents comparison between real and estimated data clustering. Finally, Section 6 presents the main conclusion and future work.

2. Background and related work

2.1. Definition of network coordinates

Assume that there are N Internet hosts in the network and M represents the $N \times N$ Internet distance matrix among these N nodes and $M(i,j)$ is the distance between host i and host j. With the coordinates of host i and host j, the distance between them can be calculated instead of performing direct end-to-end measurements. $M^E(i,j)$ denotes the predicted distance between host i and host j. NC systems seek to minimize the following objective function (F) in a decentralized way (Dabek et al., 2004).

$$F = \sum_i \sum_j (M(i,j) - M^E(i,j))^2 \quad (1)$$

To serve thousands of nodes effectively, the NC system should be scalable. Each node only does restrict measurements to calculate its NC. Thus, the time complexity of measurements is O(N). This total measurement overhead is much lower than the O(N^2) measurements required for a full mesh of N nodes.

The prediction accuracy of an NC system is often denoted by the relative error (RE) of the predicted distance over the real latency measured on Internet. RE between nodes i and j is defined as (Dabek et al., 2004; Zhang et al., 2006)

$$RE = \frac{|M^E(i,j) - M(i,j)|}{\min(M^E(i,j), M(i,j))} \quad (2)$$

Smaller RE indicates higher prediction accuracy. When measured latency is equal to predicted latency, the RE value will be zero.

2.2. Related network coordinate systems

Figure 1 shows the most widely used NC systems. Most of the existing NC systems use the Euclidean distance model such as *Vivaldi* (Dabek et al., 2004), *GNP* (Ng and Zhang 2002), *PIC* (Costa et al., 2004), and *ICS* (Lim et al., 2005). In this model, all the N hosts are embedded into d-dimensional Euclidean space R^d ($d \ll N$). *Vivaldi* becomes the most widely used one because of its simplicity and decentralized feature. The Euclidean distance model is mainly implemented in the NC systems in spite of its assumption that the predicted distances among any three hosts satisfy the triangle inequality. Unfortunately, TIV persistently occurs in the real Internet (Wang et al., 2007, 2013; Zhang and Zhang, 2012; Yamamoto and Yamazaki, 2013). As a result, the Internet distance cannot be predicted accurately with the Euclidean distance-based NC systems. Two-layered systems like *Pharos* (Chen et al., 2009), *Tarantula* (Chen et al., 2011), and hierarchical *Vivaldi* (Kaafar et al., 2008) have been proposed to remedy the impact of TIVs. They

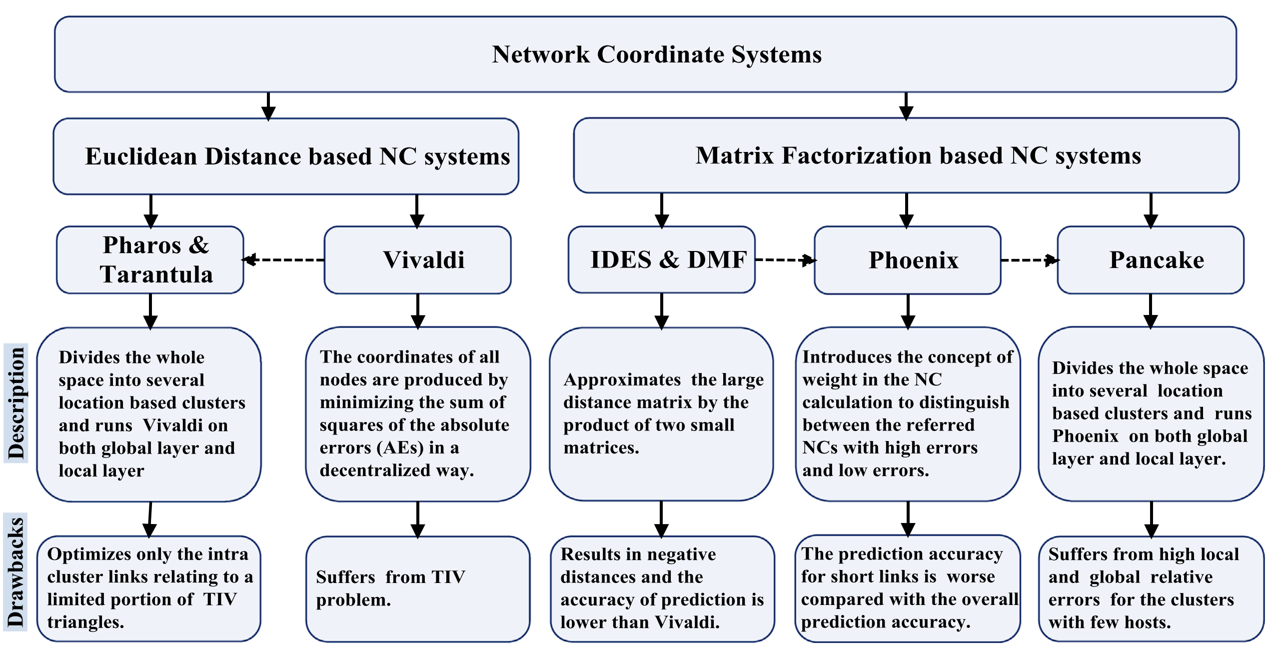

Fig. 1. Related network coordinate systems.