

Network Distance Prediction for Enabling Service-Oriented Applications over Large-Scale Networks

Haojun Huang, Hao Yin, Geyong Min, Dapeng Oliver Wu, Yulei Wu, Tao Zuo, and Ke Li

Abstract—The knowledge of end-to-end network distances is essential to many service-oriented applications such as distributed content delivery and overlay network multicast, in which the clients are flexible to select their servers among a set of available ones based on network distance. However, due to high expenditure of global measurements in large-scale networks, it is infeasible to actively probe end-to-end network distances for all pairs. In order to address this issue, network distance prediction has been proposed by measuring a few pairs and then predicting the other ones without direct measurements, or splicing the path segments between each pair with the observation. It has been considered important to improve network performance, and enables service-oriented applications over large-scale networks. In this article, we first illustrate the basic ideas behind network distance prediction, and then categorize the current research work based on different criteria. We illustrate how different protocols work and discuss their merits and drawbacks. Finally, we summarize our findings and point out potential issues and future directions for further research.

I. INTRODUCTION

NETWORKS have become an indispensable part of our daily life as an information platform for communications. Over the past few decades, widespread distributed service-oriented applications, such as peer-to-peer file sharing services, overlay network multicast [1], have evolved considerably beyond the traditional client-server model, where an end-user (client) only communicates with one server. In contrast, in the current service-oriented applications, the end-users have much flexibility in choosing their servers, with little or no information about the potential performance of different communication paths towards them. In order to overcome this issue, Network Distance Prediction (NDP) has been proposed, which provides benefit for end-users to select intelligent paths based on network performance and constructs much more convenient networks. For instance, in content delivery networks, an end-user can conveniently obtain its desired Web resources from

H. Huang and H. Yin are with Department of Computer Science and Technology, Tsinghua University, Beijing, 100081 China (e-mail: hhj0704@hotmail.com; h-yin@mail.tsinghua.edu.cn.)

G. Min and Y. Wu are with the Department of Computer, University of Exeter, Exeter, EX4 4QF, UK (e-mail: g.min@exeter.ac.uk; Y.L.Wu@exeter.ac.uk).

D. Oliver Wu is with the Department of Electrical & Computer Engineering, University of Florida, Gainesville, Florida, 32611-6130, USA (e-mail: wu@ece.ufl.edu).

T. Zuo is with the Department of Electrical & Information Engineering, Wuhan University of Science and Technology, Wuhan, 430081, China (e-mail: zuotao@wust.edu.cn).

K. Li is with School of Information Science and Technology, Southwest Jiao Tong University, Chengdu, 610031 China (e-mail: keli@swjtu.edu.cn).

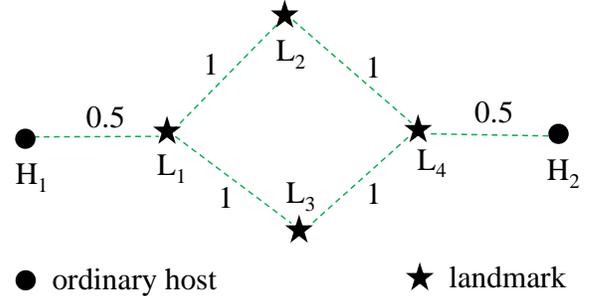


Fig. 1. Network distance prediction derived from real measurements.

a particular site with the knowledge of the predicted network distance.

Similar to [2-15], in this article the network distance between two nodes is defined as the communication delay or latency between them, in the form of either one-way delay or more often Round Trip Time (RTT). Obviously, it is infeasible to ceaselessly probe network distances among all pairwise nodes in large-scale networks because global accurate measurements are difficult and costly to achieve and maintain. A natural idea is to probe a small set of pairs and then to predict the distances between others without direct measurements or to splice the path segments between each pair, this is so-called NDP. This understanding has motivated a great deal of research to develop NDP [1]. Fig. 1 illustrates NDP operations by matrix factorization [9], with four landmarks and two ordinary hosts. Based on the knowledge of inter-landmark distances, the distance matrix among landmarks L_1 , L_2 , L_3 and L_4 can be factorized to the incoming and outgoing vectors for distance prediction among the other hosts. Given an ordinary host H_1 , which desires to know the distance to another host, for instance, H_2 , it first measures its distance vectors to and from landmarks L_1 , L_2 , L_3 and L_4 as $[0.5, 1.5, 1.5, 2.5]$. Then, it can calculate its outgoing and incoming vectors as $Y_{H_1} = [1.5, 0, 1]$ and $X_{H_1} = [-1.5, 0, 1]$ (the calculation shown in [9]), respectively. Similarly, H_2 can obtain its distance vector as $[2.5, 1.5, 1.5, 0.5]$, and its outgoing and incoming vectors: $X_{H_2} = [-1.5, 0, -1]$ and $Y_{H_2} = [-1.5, 0, 1]$. If H_1 learns the incoming vector of H_2 , the distance between them can then be predicted as $X_{H_1} \cdot Y_{H_2} = 3.25$ with a tolerable predicted error of 8.3%, instead of relying on direct measurements. This means that, as long as an acceptable predicted network distance can be obtained for host H_1 , the small measurement cost can be neglected and the remaining overhead is amortized over all distance predictions.

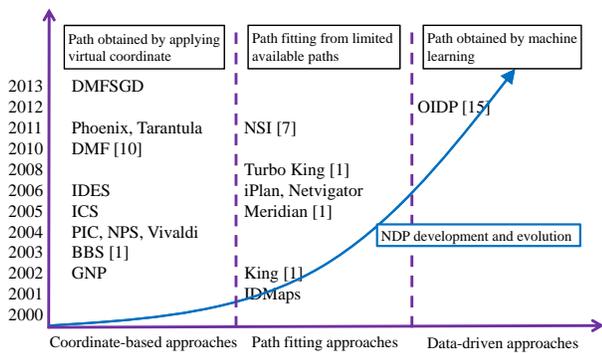


Fig. 2. Network distance prediction categories and historical developments.

The NDP has been considered important to improve the performance of many service-oriented applications and bridge the gap between the end-users and large-scale networks, and thus received increasing attention. However, the existing approaches have been proven to be a difficult task to use in deployed applications [1-3,5,8]. In this article, we survey the various NDP approaches reported in the current literature, elaborated in Fig. 2, as a reference for further research. We analyze their emerging challenges and discuss the future NDP developments.

The remainder of this article is organized as follows. The next section illustrates how different approaches work, and discusses their merits and drawbacks, followed by the existing evaluation metrics. Then, the emerging challenges behind NDP are highlighted, and the open issues and opportunities for further research are outlined in the following section. The final section draws the conclusions.

II. SURVEY ON NETWORK DISTANCE PREDICTION

Essentially, NDP resorts to the predicted distance without performing direct measurements, with the aid of infrastructures such as landmarks, tracers, DNS servers, and routers, to represent the actual distance for given host pairwise. It means that the predicted distances should be far close to the actual distance. Therefore, the NDP makes sense if and only if an acceptable predicted accuracy, with quantified evaluation metrics investigated in the next section, can be guaranteed. The major challenges to design NDP include symmetry, consistency, security, dynamics, cluster and Triangle Inequality Violations (TIVs), as elaborated in the following section. Currently, the NDP approaches can be classified into three categories: coordinate-based approaches, path fitting approaches and data-driven approaches. Fig. 2 depicts most classic NDP approaches following categories, their historical development and evolution. In this section, we explain their operations with special focus on how they work, and discuss their merits and drawbacks.

A. Coordinate-Based Approaches

The basic ideal of such approaches is to design a finite-dimension virtual metric space and embed the hosts into that space with the constraint that errors between predicted distances and measured distances are minimized. The network distance between two reachable hosts is then predicted as the distance between their coordinates.

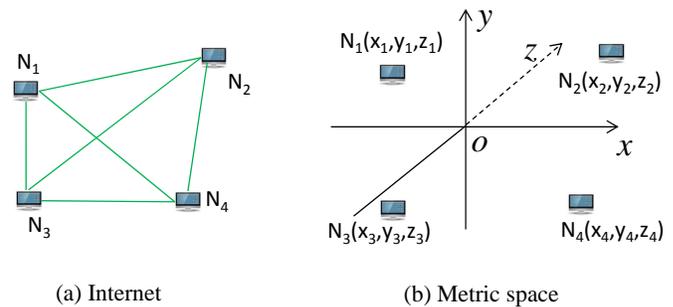


Fig. 3. Correspondence between the Internet and the metric space. The network distances are represented in green lines in Fig.3(a), and metric distances in Fig.3 (b).

Fig. 3 elaborates a matching between the large-scale networks such as the Internet and the metric space. The Internet network distances between four hosts are represented in green lines in Fig. 3(a). This distance can be, for instance, the RTT. Fig. 3(b) embeds the predicted network distances into a metric space. In such a space, the predicted distance is evaluated using the traditional distance metrics such as the relative error and stress [2,9-10]. In the following, we investigate the main coordinate-based NDP approaches.

- **GNP.** Global Network Positioning (GNP) [2] is the first network embedding system to predict network distance, which relies on a small number of fixed landmarks. It assigns the locations of all hosts in an n -dimensional Euclidean space, n being the number of landmarks. Given any two reachable hosts, GNP approximates the latency between them in the original networks as their distance in this space, with the assumption that the predicted distances among any three hosts satisfy the triangle inequality.

It starts by instructing the n landmarks to measure all the latencies among them. With these information, it computes all the landmark coordinates such that the distances among their coordinates in Euclidean space are as close as possible to their measured latencies in the original networks. Then, the ordinary hosts calculate their own coordinates with respect to the landmarks. In this way, any network distance among pairwise hosts can be computed. A distinct drawback of GNP is that it is vulnerable to landmark failures since hosts join and leave networks frequently.

- **PIC.** Practical Internet Coordinates (PIC) [3] is the first security-aware mechanisms to predict the Internet network distance. In PIC, an ordinary host selects any host whose coordinate has already been given as a landmark. This is similar to GNP, but GNP selects a fixed set of landmarks for the whole ordinary nodes as the reference nodes. On the contrary, it needs to probe the network distance to each landmark, and uses an active node discovery protocol to find out some nearby hosts used for computing coordinates. In this way, it obtains the coordinates of all landmarks and then uses a multidimensional global optimization method, e.g. Simplex DownHill, to compute its coordinates such that the errors in the predicted distances are minimized. In order to prevent network attacks, PIC explores reference select point techniques based on triangle inequality to detect independent malicious participants.

- **NPS.** Network Positioning System (NPS) [6] is a hierarchical NDP system, which maintains consistency while enabling decentralization to compute accurate and stable network locations. Essentially, it is a GNP extension aiming to overcome its landmarks failure by selecting any ordinary host that knows its own position to serve as a landmark for other hosts by a membership server. The role of the membership server is to initialize configuration parameters of ordinary hosts, such as identifying the landmarks. In order to ensure the consistency of hosts, NPS imposes a hierarchical position dependency among the hosts, by eliminating any landmark called malicious node that provides significant relative errors compared to the other landmarks.

- **ICS.** Internet Coordinate System (ICS) [8] is a low-dimensional NDP system, which retains as much topology information as possible. In ICS, the distances between a host and landmarks (called beacon nodes in ICS) are expressed as a distance vector, whose dimension is equal to the number of landmarks. For any host, it does not need to measure the distance from itself to all the landmarks, but rather to a set of landmarks and thus obtains its distance vector. The landmarks calculate their locations by multiplying the distance vector with a transformation matrix based on Principal Component Analysis (PCA), which aims at the distance dimension reduction by refining the variables that retain most of the original information. Such a linear transformation essentially extracts the topology information from delay measurements between landmarks and keeps it in a novel metric space.

- **IDES.** Internet Distance Estimation Service (IDES) [9] is the first system based on matrix factorization to predict distances for large-scale networks, which has the same landmark-based architecture as GNP while enduring TIVs. Its essential idea is to approximate a large distance matrix whose elements represent pairwise distances by the product of two smaller matrices. Such a model allows a representation of distances violating TIVs and asymmetric distances based on matrix factorization, and can be regarded as a form of dimensionality reduction with both singular value decomposition and non-negative matrix factorization algorithms. In particular, it addresses the questions of the impact of both landmark placement and measurement error on predicted performance.

- **Tarantula.** Tarantula [4] is an alternative hierarchical NDP system, which focuses on mitigating the impact of TIVs. It dynamically divides the whole networks into three clusters following the major actual clusters of the world networks—America, Asia-Pacific, and Europe, and integrates each two clusters into a subspace, meaning each host belongs to one cluster and two subspaces at the same time. It runs a Vivaldi [1] system on each cluster or subspace. Considering the inter-cluster links account for more TIVs than other links in the hierarchical NDP [1,6], Tarantula respectively uses the cluster-based system and subspace-based system to predict the intra-cluster and inter-cluster distances. In this way, it outperforms the existing hierarchical NDP greatly in predicted accuracy.

- **DMFSGD.** Decentralized Matrix Factorization by Stochastic Gradient Descent (DMFSGD) [10] is an IDES extension seeking to overcome the limitation generated by the failure of landmarks. Different from IDES, it only requires

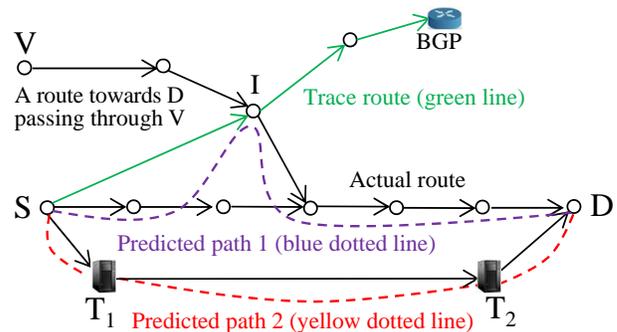


Fig. 4. The predicted network distance derived from the path fitting.

each host to measure local distances to and from a small set of neighbors and then predict the distances to the other hosts, without explicit matrix constructions and special hosts such as the fixed landmarks and central servers. DMFSGD is simple, decentralized, and scalable. With these features, it addresses many practical issues in NDP such as measurement dynamics and network churn with time in large-scale networks.

- **Phoenix.** Phoenix [5] is a weight-based NDP system by matrix factorization. Basically, it is an extension of DMFSGD and IEDS that aims at address the inaccurate coordinate impacts on distance prediction by introducing weights to reference coordinates. It assigns each landmark a graded weight in its coordinate based on its accuracy, and trusts the hosts with the higher weight values more than the others. Therefore, Phoenix can substantially mitigate the impact of the error propagation and improve the prediction accuracy over the existing approaches.

B. Path Fitting Approaches

Different from the coordinate-based NDP approaches that treat the network as a black-box, the path fitting approaches utilize the internal network structure such as the network topology, DNS servers, tracers and existing routing to predict the network distance, by splicing the path segments with the observation or approximating the path among the nodes closer to client-server as the network distance from the client to the server. Essentially, the path fitting approaches use direct approximative measurements [10] or reactive aggregations [11] instead of predictions to approximate the network distance for given client-servers.

Fig. 4 illustrates how path fitting approaches predict the network distance. Given nodes S and D , the network distance between them can be obtained by either splicing a short path segment from S to an intersection I_1 from which a path towards D has been observed, or approximating the distance between S and its nearest Tracer (or DNS) T_1 , plus the distance between D and its nearest Tracer T_2 , plus the distance between T_1 and T_2 , or even approximately equaling to the distance from T_1 to T_2 . Three classic examples of those approaches are IDMaps [10], iPlane [11], and Netvigator [12].

- **IDMaps.** Internet Distance Map Service (IDMaps) is the first Internet distance prediction system and has been considered as the predecessor of NDP approaches. It starts by building a simplified overlay topology map of the Internet performed by special hosts (called Tracers in IDMaps) based

TABLE I
CHARACTERISTICS OF NETWORK DISTANCE PREDICTIONS AVAILABLE IN THE CURRENT LITERATURE

Approaches	Measurement overhead	Prediction prerequisites	Embedding model	Churn recovery	Infrastructure dependability	Scalability	Prediction Accuracy
GNP	$O(m^2 + mn)$	A small fixed set of landmarks	Euclidean	No	No	Medium	General
PIC	$O(m^2 + mn)$	A small set of landmarks, P2P substrate	Euclidean	No	No	High	General
NPS	$O(m^2 + mn)$	A small set of landmarks	Euclidean	No	No	Medium	General
Tarantula	$O(m^2 + mn)$	Three clusters and subspaces	Euclidean	Yes	No	High	High
ICS	$O(m^2 + mn)$	A small set of landmarks	Lipschiz	No	No	High	General
IDES	$O(m^2 + mn)$	A small fixed set of landmarks	Matrix factorization	No	No	Medium	High
DMFSGD	$O(m^2 + mn)$	A small set of landmarks	Matrix factorization	Yes	No	High	High
Phoenix	$O(m^2 + mn)$	A small set of landmarks	Matrix factorization	Yes	No	High	High
IDMaps	$O(r^2 + k)$	A small fixed set of tracers	-	Yes	Yes	High	High
iPlane	$O(q^2 + n)$	A small set of vantage hosts mapping the Internet topology	-	Yes	Yes	High	High
Netvigator	$O(mn + pn)$	A small set of landmarks and milestones	-	Yes	Yes	Medium	General

Note: m : the number of landmarks; n : the number of hosts; r : the number of tracers; k : the number of clustered address prefixes; p : the number of milestones; q : the number of vantage hosts.

on network measurements. Then, it performs the shortest path routing on this map, such that the routing can be used as the predicted network distance for any two reachable hosts with valid IP addresses. Given two hosts x and y , the predicted distance between them is expressed as the sum of the distance from x to its nearest Tracer T_1 , the distance from y to its nearest Tracer T_2 , and the shortest routing distance from T_1 to T_2 over this map. IDMaps provides general distance query such that the service-oriented applications can easily obtain the network distances among the end-uses and services.

- **iPlane.** Information Plane (iPlane) is a scalable service that aims at providing the accurate Internet path performance predictions for application-level overlay networks. It continuously requires vantage hosts that locate in different geographically regions to map the Internet topology such that they obtain the observed paths with a rich set of link and router attributes such as latency, available bandwidth, and loss rate. With such information, it can predict paths for arbitrary two reachable nodes by being combined with the measured performance of path segments in the Internet. In order to reduce measurement overhead, it clusters IP prefixes into border gateway protocol atoms such that it generates the target list. Compared to the current NDP systems, iPlane not only provides latency prediction between two reachable nodes, but also automatically infers important network behavior information such as loss rate, capacity, available bandwidth, and isolated anomalies.

- **Netvigator.** Network Navigator (Netvigator) is an efficient NDP system that focuses on proximity estimation and distance prediction. Initially, it requires each host to send and receive probe packets to and from the landmarks and milestones (also called intermediate routers in Netvigator) if any, and then constructs a landmark vector including the distances to all the landmarks and the milestones that the probe packets pass through and sends it to a central server with the global information table. Once receiving a query from a client host, the central server applies the clustering algorithm to identify k closest candidates for its proximity distance. In this way, the proximity distance from a client host to some server host can be predicted. A significant merit of Netvigator is that it can avoid false clustering by introducing the milestones used for

refinement, and therefore achieves great accurate predictions of the proximity estimation.

C. Data-Driven Approaches

Nowadays, we are in the age transitioned from a hypothesis-driven world to a data-driven world brought by big data. Big data has changed our most basic understanding of how to comprehend network behavior and explore the Internet. For instance, assuming algorithms A and B can solve the same problem, the result of A is clearly better than B in the small-scale data world, i.e. algorithm A can achieve the better performance. However, such situation may not exist as the amount of data increases. There are too much such similar phenomena in the current networks. These findings have brought computer science and its deuterogenic disciplines a landmark revelation: the data itself (rather than the approaches and models used by data analysis) can guarantee the validity of the data analysis results with the increase of the available data. Even though the approaches or models are lack of precision, it can make the conclusion closer to the truth as long as owning enough data.

In the near future, our understanding on network designs will be driven more by the abundance of data rather than hypotheses, described as the fourth paradigm of scientific exploration. We argue that data-driven idea will transfer the design philosophy of future Internet in all aspects including architecture design, resource management, task scheduling, and network distance prediction [14]. However, the existing researches have not taken full advantage of data-driven thought to steer the design of NDP. The quasi prototype of the data-driven NDP approaches can be found in [15], where an Internet NDP approach seeks to capture geographical characteristics between Internet host pairs by machine learning, instead of relying on direct measurements. Although without explicitly exposing data-driven thought, it still reveals a novel and efficient solution to the NDP.

We think that the NDP can be executed by employing data-driven approaches. Given client u and server v , the network distance between them, denoted by $P(u, v)$, can be obtained through machine learning, by analyzing large amounts of user

behavior data. Let $P(t_{i,j})$ be the path from node i to node j with the network distance of $t_{i,j}$. The network distance $P(u,v)$ can be expressed as

$$\begin{cases} P(u,v) = \sum P(t_{u,i_1}) + P(t_{i_1,i_2}) + \dots + P(t_{i_n,v}), \\ \text{s.t. } \min(t_{u,i_1} + t_{i_1,i_2} + \dots + t_{i_n,v}), t_{i_j,i_{j+1}} > 0. \end{cases} \quad (1)$$

This achievement depends on the enough available data and our ability to harness big data. It will be the inevitable trend that many data-driven NDP approaches will be proposed in the future to provide intelligent server selection suggestions for clients. In order to facilitate the understanding of the existing NDP approaches comprehensively, Table I summarizes the above discussed approaches together with the performance criterias, including measurement overhead, prediction prerequisites, embedding model, churn recovery, infrastructure dependability, scalability, and prediction accuracy.

III. EVALUATION METRICS

In order to quantify the magnitude of the differences between predicted distances and original distances, some evaluation metrics have been proposed. Let $d(i,j)$ denote the measured distance, $\bar{d}(i,j)$ be the predicted distance computed from some function, ϕ be the metric space, and $\phi(x)$ be the coordinate of node x in ϕ . The current evaluation metrics can be summarized as follows.

- **Relative Error.** The relative error [2,9-10], denoted by e_r , is defined as $e_r = \frac{d(i,j) - \bar{d}(i,j)}{\min(d(i,j), \bar{d}(i,j))}$, or $\frac{|d(i,j) - \bar{d}(i,j)|}{\min(d(i,j), \bar{d}(i,j))}$. This metric was proposed to evaluate the accuracy of the distance prediction.

- **Stress.** The stress [10] is given by $\sqrt{\frac{\sum_{i \neq j} (d(i,j) - \bar{d}(i,j))^2}{\sum_{i \neq j} d(i,j)^2}}$, which measures the overall fitness of the embedding and is used to illustrate the convergence of the proposed NDP schemes.

- **Median Absolute Estimation Error (MAEE).** The MAEE [8] is given by $\text{median}_{i,j}(|d(i,j) - \bar{d}(i,j)|)$. This metric is designed to evaluate the absolute prediction error between predicted distances and measured distances for any pair of reachable nodes.

- **Distortion.** Let $r(\phi, x, y) = \frac{\phi(x) - \phi(y)}{d(x,y)}$, $\text{exp}(\phi) = \max r(\phi, x, y)$, and $\text{con}(\phi) = \min r(\phi, x, y)$ as x and y range. The distortion of ϕ is defined as the ratio of $\frac{\text{exp}(\phi)}{\text{con}(\phi)}$. The distortion is a worst-case measure of the quality of an embedding, and used to measure the worst-case change in the relative distances of the embedding.

- **Local Relative Rank Loss.** Let $p(z) = \{(x,y) | x=y \text{ and } \text{swapped}(z,x,y)\}$, and $s = \frac{(|N|-1)(|N|-2)}{2}$. The local relative rank loss, denoted by $\text{rrl}(\phi, x)$, is defined as $\text{rrl}(\phi, x) = \frac{|p(z)|}{s}$, where $|N|$ is the set of nodes, (x,y) are elements of $N \times N$, and $\text{swapped}(z,x,y)$ is true if the relative relationship of z to x and y is different in the original networks and the embedding space. This metric is designed to reflect the probability that the relationship between any two nodes in the original networks will have a different relative order in embedding space.

- **Closest Neighbors Loss.** Given node x , the closest neighbors loss, denoted by $\text{cnl}(\phi, x)$, is defined as 0 if the nodes

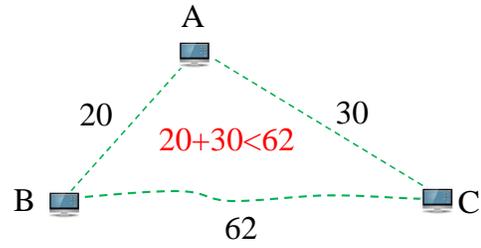


Fig. 5. Triangle inequality violations derived from real measurements.

closest to x remain closest in the embedding space ϕ , and 1 otherwise. For n nodes, the average closest neighbors loss is defined by $\frac{\sum_{i=1}^n \text{cnl}(\phi, x_i)}{n}$. This metric is designed to reflect the average percentage of nodes whose closest neighbors are not preserved in the embedding space.

- **k -Closest Preservation.** Given node x , let $\text{cn}(k, x)$ denote its k closest neighbors in the original networks, and $\text{cn}(\phi, k, x)$ be its k closest neighbors in the metric space ϕ . The k -closest preservation is defined by $\frac{|\text{cn}(k, x) \cap \text{cn}(\phi, k, x)|}{k}$. This metric is used to reflect the ability of node x to keep the first k closest neighbors in the embedding space.

IV. EMERGING CHALLENGES AND OPEN ISSUES

A. Emerging Challenges in NDP

This subsection highlights the important challenges in NDP.

- **Symmetry.** Given any nodes i and j , the distance between them is given by the function $d(i,j)$. The symmetry requires that for i and j , $d(i,j) = d(j,i)$. However, the network distance between any two reachable nodes is not necessarily symmetric due to the network structure and routing policy [1,16]. The existing researches [16] have measured the proportion of asymmetric routing: more than 20% of the links have borne the asymmetric flow/packet/byte and more than 14% of the flows in the Internet have shown autonomous system asymmetry. These challenge accurate predictions of network distances.

- **TIV.** Most NDP approaches, such as [2-4] and others, are based on the embedding of host positions in a finite-dimensional space, commonly the Euclidean coordinate system, where Euclidean distance is used to form the desired estimate. However, in such a system the predicted distance always violates the triangle inequality due to routing policies or path inflation. The existing studies show that TIVs are widespread and persistent [6,16]. A TIV occurs among a triple of nodes in the Internet when the latencies between them cannot form a valid triangle, and changes with time. Fig. 5 elaborates such a scenario derived from real measurements [16]. A TIV represents a real network path that there exists a closer route to a host through an intermediate host than the direct route, but can not hold for metric space. Thus, the TIV inevitably yields inaccurate predictions of network distances.

- **Consistency.** Let $d(i,j)$ and $\bar{d}(i,j)$ denote the measured distance and the predicted distance from node i to node j , respectively. Given any reachable pairwise nodes i and j , and nodes u and v , the consistency requires that the path among them $\bar{d}(i,j) > \bar{d}(u,v)$ if and only if $d(i,j) > d(u,v)$, and vice versa. In practice, there is a significant disparity in NDP in term of the prediction accuracy for different distance ranges

TABLE II
COMMON CHALLENGES AFFECTING NETWORK DISTANCE PREDICTIONS

Approaches	Common Challenges					
	Symmetry	TIV	Consistency	Security	Cluster	Dynamics
GNP	Low	Medium	Low	High	No	Medium
PIC	Low	Medium	Low	Low	Low	Low
NPS	Low	Medium	Low	Low	Low	Low
Tarantula	Low	Medium	Low	High	Low	Low
ICS	No	No	Low	Medium	No	Medium
IDES	No	No	Low	High	No	Medium
DMFSGD	No	No	Low	High	No	Low
Phoenix	No	No	Low	High	No	Low
IDMaps	No	No	No	Medium	No	Medium
iPlane	No	No	No	Medium	No	Medium
Netvigator	No	Medium	No	Medium	No	Medium

due to many factors such as dynamic topology/paths, time-varying traffic, congestion, metric space [1], thus lowering the validity and scalability of NDP. How to achieve significant consistency in NDP is a common challenge for the current approaches.

- **Security.** It has been shown that NDP methods are rather primitive [1] and cannot defend against all types of attacks including disorder attack, repulsion attack and isolation attack. Such attacks make NDP very susceptible to the malicious nodes either from inside or outside of the network. The current solution is to test the TIV [8] or eliminate the landmarks that provide the significant relative errors [6], which has been considered suboptimized to fulfill security requirement. The security has become a great challenge in NDP.

- **Cluster.** In order to achieve comparable prediction quality, the hierarchical NDP approaches divided the networks into several clusters in a distributed way. In such approaches, each host keeps different sets of neighbors and coordinates in different layers, such that it can predict the intra-cluster distance and inter-cluster distance based on the local coordinates and the global coordinates, respectively. Two significant issues of these approaches are that there are still no consensuses on how to divide the networks into clusters and how many clusters the networks should be divided into [6,8], which also challenges the predicted quality.

- **Dynamics.** The NDP schemes should be resilient against the network dynamics mainly due to host dynamics and time-varying traffic, including host failures/joining, temporary network partitioning, and traffic change/overload. In particular, there exist the distance predicted schemes, such as GNP and IDES, suffering from landmark failures and overloading, and furthermore, all the predicted distances originated from these schemes are partially decided by direct measurements. However, such measurement results fluctuate frequently with time because of time-varying traffic, which has a slight impact on the direct measurements. Such dynamics prevent the system from building models of network distances and potentially degrade the predicted accuracy of NDP.

Table II summarizes how the main approaches proposed previously is influenced by each of these challenges.

B. Open Issues and Future Directions in NDP

In this subsection, we list some potential issues and future directions for further research on NDP.

- **Embedding Models.** Strictly speaking, the network distance does not match the notion of metric space due to

the inherent network characteristics such as violations of symmetry and TIVs. Thus, we can infer that any predicted distance obtained from NDP is an approximate network distance. Currently, there are many embedding models such as Euclidean space, hyperbolic space and hybrid space used to predict network distance, while these models can not very accurately draw out the characteristics of the real networks. For instance, the widespread TIVs in NDP cannot represent networks with complex routing policies such as sub-optimal routing or asymmetric routing, and thus yield inaccurate predictions of network distances. In order to address such issues, we should focus on how to design embedding models imposing the network features.

- **QoS-based Network Distance Prediction.** The NDP designed at the beginning mainly aims to provide predicted distances without performing direct measurements for end-users to greatly benefit from intelligent path selection based on network performance. However, in most cases, they cannot provide QoS-based guaranteed distances except iPlane. For instance, the predicted distances fluctuate frequently and are not convincing in some ways. There exists a performance gap between them and the ideal NDP. We argue that it is desirable to design QoS-based NDP approaches in the future. To achieve this goal, we can focus on time-varying measurements by exploiting DSCP (Differentiated Services Code Point or ToS (Type of Service) bits to probe the network distance, and designing QoS-based predicted metrics to estimate the predicted accuracy.

- **Multi-Metric Network Distance Prediction.** Currently, many related work provides only a limited subset of the metrics of interest, commonly latency between a pair of nodes. In reality, however, the latency is just one of many metrics such as available bandwidth and packet loss rate that affect the performance of service-oriented applications. Compared to the latency as an additional parameter, the available bandwidth is a concave parameter, and the packet loss rate is a multiplicative parameter. If we simply embed these parameters into Euclidean space like the current approaches, there must cause prediction distance to be arbitrarily wrong. Therefore, designing multi-metric embedding models based on different performance metrics is an efficient approach to achieve the desired predicted performance and needs to be further investigated.

- **Predicted Errors.** The NDP inevitably generates the predicted errors caused by various factors such as landmarks

failure, embedding metric spaces, and evaluated metrics. The predicted errors directly determine the QoS for end-users. In order to provide better service to end-users and bridge the gap between service-oriented applications and large-scale networks, we need to investigate the impact of the predicted errors on the service-oriented applications and identify the reasons behind them, then to design the high-precision NDP. In addition, we should evaluate the predicted performance under the practical measurement platforms such as PlanetLab, DIMES, OneLab, EmuLab using the multi-metrics, and create a system that leverages both theoretical approaches and actual distance prediction in the network, thus catalyzing the evolution of the NDP into a service-oriented architecture.

• **Security-aware Network Distance Prediction.** The most recent predicted mechanisms assume that the participating nodes can be trusted. Unfortunately, it has been proven that NDP methods are rather primitive and cannot prevent a variety of attacks such as disorder, repulsion, isolation, and system control attacks, providing a potentially attractive fertile ground for the disruption or collapse of the many applications and overlays that would use these services. To the best of our knowledge, currently there are only a few simple methods in NDP to defend against malicious behaviors. For example, PIC uses a test based on the triangle inequality to detect malicious nodes, and NPS regards a landmark that provides significant relative errors compared to the reference nodes as a malicious node. Such security mechanisms designed to NDP approaches have shown that they are still susceptible to the intrusions.

The security-aware NDP studies are still in their infancy. How to prevent network attacks determines the QoS of distance prediction. It is desirable to design more security-aware NDP approaches and therefor they can prevent from various malicious behaviors.

V. CONCLUSIONS

The network distance prediction has been considered important to improve performance of service-oriented applications and bridge the gap between the end-users and large-scale networks, and thus has received increasing attention. In this article, we have investigated the important existing NDP approaches, and categorize the current research work based on different criteria. We provide general information on the behaviors of NDP approaches and discuss their merits and drawbacks. Finally, we point out potential issues and future directions for further research.

ACKNOWLEDGMENT

This work has been partially supported by the National Basic Research Program of China (No. 2011CB302601 and No. 2012CB315801), the National Natural Science Foundation of China (No. 61402343, No. 61222213, No. 61170290, and 61371126), the EU FP7 CLIMBER Project (No. PIRSES-GA-2012-318939), the Natural Science Foundation of Hubei Province (No. 2013CFB332), and the Jiangsu International Cooperation Program of Science and Technology (No. BZ2013018).

REFERENCES

- [1] B. Donnet, B. Gueye, and M. Kaafar, "A Survey on Network Coordinates Systems, Design, and Security," *IEEE Commun. Surv. Tutor.*, 2010, 12(4):488-503.
- [2] T. S. E. Ng and H. Zhang, "Predicting Internet Network Distance with Coordinates-based Approaches," *Proc. IEEE INFOCOM 2002*, NY, USA, Jun. 2002, pp.170-179.
- [3] M. Costa, M. Castro, A. Rowstron, and P. Key, "PIC: Practical Internet Coordinates for Distance Estimation," *Proc. IEEE ICDCS 2004*, Tokyo, Japan, Mar. 2004, pp.178-187.
- [4] Z. Chen, Y. Chen, Y. Zhu, et al., "Tarantula: Towards an Accurate Network Coordinate System by Handling Major Portion of TIVs," *Proc. IEEE Globecom 2011*, Houston, USA, Dec. 2011, pp.5-9.
- [5] Y. Chen, X. Wang, C. Shi, et al., "Phoenix: A Weight-based Network Coordinate System using Matrix Factorization," *IEEE Trans. Net. Serv. Manage.*, 2011, 8(4):334-347.
- [6] T. S. E. Ng and H. Zhang, "A Network Positioning System for the Internet," *Proc. USENIX ATEC 2004*, Boston, USA, 2004, pp.141-154.
- [7] H. Oktay, A. S. Balkir, I. Foster, and D. Jensen, "Distance Estimation for Very Large Networks using MapReduce and Network Structure Indices," *Proc. WIN 2011 Workshop*, NY, USA, 2011, pp.1-6.
- [8] H. Lim, J. Hou, and C. Choi, "Constructing an Internet Coordinate System Based on Delay Measurement," *IEEE/ACM Trans. Netw.*, 2005, 13(3):513-525.
- [9] Y. Mao, L. Saul, and J. M. Smith, "IDES: An Internet Distance Estimation Service for Large Networks," *IEEE J. Sel. Areas Commun.*, Dec. 2006, 24(12):2273-2284.
- [10] Y. Liao, W. Du, P. Geurts, and G. Leduc, "DMFSGD: A Decentralized Matrix Factorization Algorithm for Network Distance Prediction," *IEEE/ACM Trans. Netw.*, 2013, 21(5):1511-1524.
- [11] P. Francis, S. Jamin, C. Jin, et al., "IDMaps: A Global Internet Host Distance Estimation Service," *IEEE/ACM Trans. Netw.*, 2001, 9(5):525-540.
- [12] H. Madhyastha, T. Isdal, M. Piatek, et al., "iPlane: An Information Plane for Distributed Services," *Proc. USENIX OSDI 2006*, Seattle, USA, Nov. 2006, pp.367-380.
- [13] P. Sharma, Z. Xu, S. Banerjee, and S. Lee, "Estimating Network Proximity and Latency," *ACM SIGCOMM Comp. Comm. Review*, 2006, 36(3):39-50.
- [14] H. Yin, Y. Jiang, C. Lin, et al., "Big Data: Transforming the Design Philosophy of Future Internet," *IEEE Net. Mag.*, 2014, 28(4):14-19.
- [15] A. Jain and J. Pasquale, "Internet Distance Prediction Using Node-Pair Geography," *Proc. IEEE NCA 2012*, Cambridge, USA, 2012, pp.71-78.
- [16] Y. Zhu, Y. Chen, et al., "Taming the Triangle Inequality Violations with Network Coordinate System on Real Internet," *Proc. ACM ReArch 2010*, Philadelphia, USA, 2010, pp.1-6.

BIOGRAPHIES

HAOJUN HUANG (hhj0704@hotmail.com) is a postdoctoral fellow in the RIIT at Tsinghua University. He received his B.S. degree in School of computer science and technology from the Wuhan University of Technology in 2005, and his PhD degree in School of Communication and Information Engineering from University of Electronic Science and Technology of China (UESTC) in 2012. He has published 20 papers in international journals and conferences, such as IEEE Commun. Mag. and IEEE Network Mag. His research interests focus on big data, mobile internet, wireless communication and ad hoc networks.

HAO YIN (h-yin@tsinghua.edu.cn) is a professor in the Research Institute of Information Technology (RIIT) at Tsinghua University. His research interests span broad aspects of multimedia networks, future networks, and big-datadriven network science and engineering. Some of his research results have been widely used in industry and adopted by industry standards. He has published over 100 papers in refereed journals and conferences.

GEYONG MIN (g.min@exeter.ac.uk) is a Professor of High Performance Computing and Networking in the Department of Mathematics and Computer Science within the College of Engineering, Mathematics and Physical Sciences at the University of Exeter, United Kingdom. He received the PhD degree in Computing Science from the University of Glasgow, United Kingdom, in 2003, and the B.Sc. degree in Computer Science from Huazhong University of Science and Technology, China, in 1995. His research interests include Future Internet, Computer Networks, Wireless Communications, Multimedia Systems, Information Security, High Performance Computing, Ubiquitous Computing, Modelling and Performance Engineering.

DAPENG OLIVER WU (wu@ece.ufl.edu) is a professor at the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL. He received B.E. in Electrical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 1990, M.E. in Electrical Engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1997, and Ph.D. in Electrical and Computer Engineering from Carnegie Mellon University, Pittsburgh, PA, in 2003. His research interests are in the areas of networking, communications, video coding, image processing, computer vision, signal processing, and machine learning.

YULEI WU (Y.L.Wu@exeter.ac.uk) is a Lecturer in Computer Science at the University of Exeter. He received his Ph.D. degree in Computing and Mathematics and B.Sc. (First Class Hons) degree in Computer Science from the University of Bradford, UK, in 2010 and 2006, respectively. His recent research focuses on Future Network Architecture and Protocols, Wireless Networks and Mobile Computing, Cloud Computing, and Performance Modelling and Analysis.

TAO ZUO (zuotao@wust.edu.cn) is an associated professor in College of Information Science and Engineering, Wuhan University of Science and Technology. He received his Ph.D. degree in Electronics and Information Engineering from Wuhan University in 2011. His research interests include photoelectric tracking, predictive filtering and robot network control.

KE LI (keli@swjtu.edu.cn) is a lecturer in the School of Information Science and Technology Southwest Jiao Tong University. Her major interests are network optimization, network design, and next generation networks. She has published over 10 papers in refereed journals and conferences.