Journal of Network and Computer Applications 56 (2015) 77–87

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

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

# Network tomography application in mobile ad-hoc network using stitching algorithm

M.S.S. Khan [a], Anup Kumar [b], Bin Xie [c], Prasanna K. Sahoo [d]

[a] Department of Electrical Engineering and Computer Science, Texas A&M University, Kingsville, TX 78363, USA
[b] Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA
[c] InfoBeyond LLC, Louisville, KY, USA
[d] Department of Mathematics, University of Louisville, Louisville, KY 40292, USA

ARTICLE INFO

*Article history:*
Received 30 July 2014
Received in revised form
9 May 2015
Accepted 4 July 2015
Available online 14 July 2015

*Keywords:*
Mobile network
Ad-hoc network
EM
MLE
Stitching algorithm

ABSTRACT

In this paper, a network tomography (NT) approach is proposed to study network performance in ad-hoc networks. An analysis of network performance is presented in a dynamic MANET. An expectation–maximization (EM) algorithm is used in NT to estimate the network performance parameter in accordance with network performance observations. Over the dynamic network, a new algorithm has been proposed, called the stitching algorithm, to aggregate dynamic performance values for network links. Specifically, the stitching algorithm gives local and global performance parameters such as link delays over specified time periods. Therefore, the network behavior as well as the corresponding local network performance in a mobile network can be derived over a continuous period.


## 1. Introduction

In the last decade, the Internet has become more sophisticated with the explosion in the number of heterogeneous network devices such as servers, routers, and gateways.

Moreover, mobile ad-hoc networks (Agrawal and Xie, 2009) represent a growing percentage of the devices connected to the Internet. Measuring the network performance of traditional network nodes is fairly straightforward because of their static nature. A mobile ad-hoc network (MANET) can be characterized as a dynamic network topology with low bandwidth at each network node. Consequently, performance monitoring for MANETs poses additional challenges compared to those of static nodes on the Internet. Unlike the Internet, traffic in MANET is forwarded to the destination node by hop-by-hop forwarding with the help of intermediate nodes between the source and the destination; hence, an ad-hoc node can be classified as a terminal and a router as well. With increased mobility, route changes are more frequent compared to Internet traffic. The Internet enjoys more bandwidth than the ad-hoc network, which has limited bandwidth which restricts MANET performance; hence, it is very important to monitor its performance. Moreover, MANETs traditionally use internal measurement techniques to evaluate network performance. However, this does not provide to the nodes a global perspective of the network topology. This paper investigates MANET network performance using NT (Castro et al., 2004), which has provided a new way to evaluate network performance.

NT was first proposed by Vardi (1996) who used reducing hierarchical statistical models such as moment- and likelihood-based estimation and Markov chain Monte Carlo algorithms. NT for the Internet is an approach in which network performance can be inferred from incomplete traffic observations; NT uses statistical methodologies to estimate network performance accurately in real time. NT attempts to characterize internal performance (e.g., link delay) from available measurements only. Large-scale network performance inference involves estimating performance parameters from a limited subset of network nodes. This estimation is carried out by various statistical inference algorithms that determine performance attributes that cannot be directly observed.

This paper presents a mathematical NT model, which can be classified as a statistical inverse problem (Lawrence et al., 2007) that probabilistically estimates network traffic performance in a dynamic network environment. In the proposed model, a MANET consists of number of subsequent steady states. In each steady state, the MANET has stable network topology for traffic flows. The MANET moves to the next steady state through a change in network topology (e.g., when a node moves relative to the position of other nodes). The proposed model uses pseudo log- likelihood

*E-mail address:* adhoc.khan@gmail.com (M.S.S. Khan).

http://dx.doi.org/10.1016/j.jnca.2015.07.003
1084-8045/

estimation (PLE) algorithms to evaluate network traffic flow performance with the evolution of the MANET topology. In this work, PLE is an approach to reveal the distribution of network performance parameters such as route delays, link performance, and connectivity. In each steady MANET state, the PLE algorithm evaluates the network performance parameters. The stitching algorithm is then used to "stitch" the network performance parameters over different steady states to provide a global performance analysis of the MANET or a more fine-grained analysis such as an edge or path performance analysis over a period of time. The proposed MANET NT approach includes the following technical contributions:

- *PLE algorithm*: The proposed PLE algorithm is developed based on the expectation–maximization (EM) (Dempster et al., 1977; Meng and Dyk, 1997) model. A path matrix *A* can be a full-ranked or not depending on the topology of the network. This idea will greatly influence, how the *X* are calculated. If the path matrix is a full-ranked, then we have unique solution and if path matrix is not full-ranked then we have to estimated *X*. The proposed PLE algorithm is robust enough to estimates network performance parameters from various observation instances , i.e., a path matrix, which are of non-trivial dimension. Based on the evaluation of multiple observation instances, the expected network performance can be computed for a given steady state.
- *Stitching algorithm*: The stitching algorithm meaningfully aggregates network performance parameters estimated from multiple steady states. It evaluates the performance parameters over the time domain by following changes in the network topology.
- *Algorithm validation*: The correctness of the proposed PLE algorithms was validated and the analytical MANET network performance demonstrated in terms of time and geographical domains. In the proposed implementation, NS-2 (Mccanne et al., 1996-1997) was first used to simulate a number of network scenarios and to collect data over different time and node-mobility scenarios. By using the PLE and stitching algorithm on the collected data, the estimated network performance was graphically plotted in the time and geographical domains. In this way, network performance can be observed in terms of MANET nodes and routes at various time instances.

## 2. Related work

The authors in Duan et al. (2009), Chen et al. (2010), Eriksson et al. (2010) used NT to discover the network topology of the Internet, which is generally referred to as network topology identification. On the Internet, the network topology is hidden from users because the Internet is a complex heterogeneous network. These networks are regulated by different Internet Service Providers (ISPs), and collaboration among them cannot be assumed. NT infers a logical network topology without the cooperation of the nodes. The authors in Duan et al. (2009) proposed a two-by-two component identification methodology based on bidirectional measurements and an emerging logical topology algorithm. Furthermore, Eriksson et al. (2010) advocated the practical use of tomographic inference for accurate router-level topology discovery. The authors developed a depth-first search (DFS) ordering algorithm that clusters end-host targets based on shared infrastructure. DFS further enables discovery of the logical tree topology of a network using the round-trip time (RTT). Furthermore, Chen et al. (2010) investigated network delay tomography in relation to topology. These authors defined a characteristic function that is the Fourier transform of the distribution and developed a Fourier domain inference algorithm based on flexible mixture models of link delays. Unlike this work, Liang and Yu (2003) developed a maximum pseudo-likelihood estimation algorithm to mitigate computational complexity because NT is targeted for analysis on the large-scale Internet.

In addition to topology, the NT approach as described in Caceres et al. (1999), Lo Presti et al. (2002), Yao and Cai (2010), Tsang et al. (2003) and in Zhu (2011), Singhal and Michailidis (2011), Yao et al. (2010), Firooz and Roy (2010), Ghita et al. (2010), Gui et al. (2011) has been extensively used to evaluate aspects of network performance such as the loss rate of links and the delay distribution of a link. NT makes it possible to evaluate packet loss ratios and link delays only from end-to-end observations, without knowing the internal network structure. Caceres et al. (1999), for example, proposed an algorithm to characterize internal packet loss on the Internet. Lo Presti et al. (2002) studied internal delays using a multicast-based inference NT algorithm. Zhu (2011) designed an explicit estimator based on the Law of Large Numbers to evaluate the loss rate. The estimator in this work found the maximum likelihood estimate (MLE) of a link or path without using iterative approximation. Singhal and Michailidis (2011) presented a structural model to estimate the characteristics of source–destination flows based on aggregated link measurements. The work presented by Yao et al. (2010) was a study of NT in the presence of network failures, in particular adversarial or random errors and adversarial or random erasures. In the presence of network failures, these authors developed algorithms that not only estimated the network topology, but also located random errors and adversarial erasures. Firooz and Roy (2010) studied NT using combinatorial compressed sensing. These authors assumed that the most likely used links incur a high delay in packet delivery and therefore the focus of NT is to identify these highly used links and provide a best estimate of link delays. Ghita et al. (2010) used NT to identify correlated links in a network, i.e., situations where the performance of one link may depend on the performance of other links. NT for correlated links assumes that links are no longer statistically independent and that the performance of one link can affect the performance of others. Gui et al. (2011) proposed a linear algebraic network-scale tomography framework to estimate by active inference link loss rates on mesh topologies using NT. Demir and Khan (2011) developed a searchlight localization algorithm with NT to locate the sources of distributed denial of service (DDoS) from sequences of DDoS attacks. Yu et al. (2010) presented cognitive radio NT (CRNT) to enhance network spectrum-sensing capabilities. These authors developed CRN tomography to infer spectrum status at both the link and network levels. Specifically, CRN tomography is a way to combine statistical measurement, processing, and inference techniques to provide spectrum parameters and traffic/interference patterns at both the link and network levels. It is important to notice that none of these research contribution has used the NT approach on ad-hoc networks.

## 3. Mathematical model for NT

The matrix equation in (1) is used to model the NT of an ad-hoc network as an inverse matrix problem:

$$\mathbf{Y} = A\mathbf{X}. \tag{1}$$

Given $\mathbf{Y}$ and $A$, $\mathbf{X}$ can be evaluated by this equation. In the equation, $\mathbf{Y}$ is a vector of measurements (e.g., end-to-end delays) observed over different time instants at a number of different network nodes, $A$ is the routing matrix, and $\mathbf{X}$ is the estimated time-independent performance vector (e.g., the mean delay vector). By varying the $\mathbf{Y}$ vectors, NT achieves link-level network