Performance Evaluation 91 (2015) 16–37

Contents lists available at ScienceDirect

Performance Evaluation

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

# On optimal monitor placement for localizing node failures via network tomography[☆]

Liang Ma[a,*], Ting He[a], Ananthram Swami[b], Don Towsley[c], Kin K. Leung[d]

[a] *IBM T. J. Watson Research Center, Yorktown, NY, USA*
[b] *Army Research Laboratory, Adelphi, MD, USA*
[c] *University of Massachusetts, Amherst, MA, USA*
[d] *Imperial College, London, UK*

ARTICLE INFO

*Article history:*
Available online 3 July 2015

*Keywords:*
Network tomography
Failure localization
Maximum node identifiability
Optimal monitor placement algorithm

ABSTRACT

We investigate the problem of placing monitors to localize node failures in a communication network from binary states (normal/failed) of end-to-end paths, under the assumption that a path is in normal state if and only if it contains no failed nodes. To uniquely localize failed nodes, the measurement paths must show different symptoms (path states) under different failure events. Our goal is to deploy the minimum set of monitors to satisfy this condition for a given probing mechanism. We consider three families of probing mechanisms, according to whether measurement paths are (i) arbitrarily controllable, (ii) controllable but cycle-free, or (iii) uncontrollable (i.e., determined by the default routing protocol). We first establish theoretical conditions that characterize network-wide failure identifiability through a per-node identifiability measure that can be efficiently evaluated for the above three probing mechanisms. Leveraging these results, we develop a generic monitor placement algorithm, applicable under any probing mechanism, that incrementally selects monitors to optimize the per-node measure. The proposed algorithm is shown to be optimal for probing mechanism (i), and provides upper and lower bounds on the minimum number of monitors required by the other probing mechanisms. In the special case of single-node failures, we develop an improved monitor placement algorithm that is optimal for probing mechanism (ii) and has linear time complexity. Using these algorithms, we study the impact of the probing mechanism on the number of monitors required for uniquely localizing node failures. Our results based on real network topologies show that although more complicated to implement, probing mechanisms that allow monitors to control measurement paths substantially reduce the required number of monitors.



☆ Research was sponsored by the US Army Research Laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the US Government, the UK Ministry of Defence or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

* Corresponding author.
*E-mail addresses:* maliang@us.ibm.com (L. Ma), the@us.ibm.com (T. He), ananthram.swami.civ@mail.mil (A. Swami), towsley@cs.umass.edu (D. Towsley), kin.leung@imperial.ac.uk (K.K. Leung).

http://dx.doi.org/10.1016/j.peva.2015.06.003

## 1. Introduction

Effective monitoring of network performance is essential for network operators to build a reliable communication network robust against service disruptions. In order to achieve this goal, the monitoring infrastructure must be able to detect network misbehaviors (e.g., unusually high loss/latency, unreachability) and localize the sources (e.g., malfunction of certain routers) of these misbehaviors in an accurate and timely manner. Knowledge of where problematic network elements reside in the network is particularly useful for fast service recovery, e.g., migration of affected services and/or rerouting of traffic. However, localizing network elements that cause a service disruption is challenging. The straightforward approach of directly monitoring the health of individual network elements incurs a high traffic overhead and is not always feasible due to access control or lack of protocol support at internal nodes. Moreover, built-in monitoring agents running on network elements cannot detect problems caused by unanticipated interactions between network layers, where end-to-end communication is disrupted but individual network elements along the path remain functioning (a.k.a. *silent failures*) [1]. These limitations call for a new approach to diagnose the health of network elements based on the health of end-to-end communications perceived between measurement points.

This different approach is generally known as *network tomography* [2], where a canonical application infers internal network characteristics by measuring end-to-end performance from a subset of nodes with monitoring capabilities, referred to as *monitors*. Unlike the approach of direct measurements that employ control packets, network tomography only relies on end-to-end performance (e.g., path connectivity) experienced by data packets, thus capable of reducing overhead, minimizing dependence on protocols, and detecting silent failures. In cases where the network characteristic of interest is binary (e.g., *normal* or *failed*), the problem is known as *Boolean network tomography* [3].

Given binary observations on paths (normal/failed), it is usually impossible to uniquely identify the states of individual network elements (nodes/links). For example, if two elements always appear together in any measurement path that contains one of them, then upon observing failures on these paths, we can at most infer that one (or both) of these elements has failed but not which one. Most existing works [2,4,1] address this uncertainty by focusing on a most probable solution that explains all path observations with the minimum number of failures. There is no guarantee, however, on how well this solution approximates the true failure locations.

Generally, to distinguish two possible sets of failures, there must exist a measurement path that traverses at least one element in one set and none of the elements in the other set. It is highly nontrivial to place monitors, such that this condition is satisfied with minimum cost, due to the large solution space (all combinations of monitor locations) and large number of constraints (all pairs of sets of failure locations). Several heuristics have been proposed to place monitors to uniquely localize a bounded number of link failures under specific probing mechanisms (e.g., traceroute) [5–7]. There is, however, a lack of understanding on the minimum number of monitors required for a generic probing mechanism and how this number varies for different probing mechanisms and different bounds on the number of failures.

In this paper, we apply Boolean network tomography to localize node failures. Node failures can be used to model failures of both physical nodes and links, with the latter represented as logical "nodes" connected to endpoints of the corresponding links. We consider the following problem: What is the minimum set of nodes that should be employed as monitors, such that *any* set of up to $k$ node failures can be uniquely localized from the states of paths that are measurable under a given probing mechanism? We study this problem in the context of three families of probing mechanisms: (i) *Controllable Arbitrary-path Probing (CAP)*, where monitors can measure arbitrary paths subject to connectivity, (ii) *Controllable Simple-path Probing (CSP)*, where measurements are limited to cycle-free paths between monitors, and (iii) *Uncontrollable Probing (UP)*, where only paths between monitors selected by the default routing protocol can be measured. These probing mechanisms assume different levels of control over the routing of probes and are feasible in different network scenarios (see Section 2.3). A comparison of the minimum numbers of monitors required to uniquely localize the same number of failures in the same topology under different probing mechanisms thus establishes a fundamental tradeoff between the controllability of probing mechanism and the cost of monitor deployment.

### 1.1. Related work

Existing works on tomography-based failure localization mostly consider link failures. Given path measurements, the standard approach is to find the minimum set of links whose failures explain all of the measurements, which gives the most probable set of failure locations under the assumption that failures are low-probability events. Using this approach, [2,4] propose solutions for networks with tree topologies, which are later extended to general topologies by [1]. In a Bayesian formulation, [8] proposes a two-staged solution that first estimates the failure probabilities of different links and then infers the most likely set of failed links using subsequent measurements. Augmenting path measurements with available control plane information (e.g., routing messages), [9,10] propose a greedy heuristic for troubleshooting network unreachability that has better accuracy than benchmarks using only path measurements.

Our work belongs to a complementary line of works that aim at *uniquely* localizing failures by strategically placing monitors. The optimal solution is known to be hard when monitors cannot control the routing of probes. Specifically, under round-trip probing (e.g., ping, traceroute) where only sources of probes need to be monitors, [5] shows that the optimal monitor placement is NP-hard and proposes a greedy approximation algorithm. Under the same probing mechanism, [6]