Research Article

Adaptive Loss Inference Using Unicast End-to-End Measurements

Yan Qiao,^{1,2} Jun Jiao,¹ and Huimin Ma¹

¹School of Information and Computer, Anhui Agriculture University, Hefei 230061, China

²State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Yan Qiao; qiaoyan101@gmail.com

Received 8 July 2016; Revised 9 November 2016; Accepted 28 November 2016

Academic Editor: Mohammad D. Aliyu

Copyright © 2016 Yan Qiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We address the problem of inferring link loss rates from unicast end-to-end measurements on the basis of network tomography. Because measurement probes will incur additional traffic overheads, most tomography-based approaches perform the inference by collecting the measurements only on selected paths to reduce the overhead. However, all previous approaches select paths offline, which will inevitably miss many potential identifiable links, whose loss rates should be unbiasedly determined. Furthermore, if element failures exist, an appreciable number of the selected paths may become unavailable. In this paper, we creatively propose an adaptive loss inference approach in which the paths are selected sequentially depending on the previous measurement results. In each round, we compute the loss rates of links that can be unbiasedly determined based on the current measurement results and remove them from the system. Meanwhile, we locate the most possible failures based on the current measurement outcomes to avoid selecting unavailable paths in subsequent rounds. In this way, all identifiable and potential identifiable links can be determined unbiasedly using only 20% of all available end-to-end measurements. Compared with a previous classical approach through extensive simulations, the results strongly confirm the promising performance of our proposed approach.

1. Introduction

The robustness of communication networks is extremely important for both users and network service providers. However, as the network increases in size and diversity, it becomes extremely difficult to monitor the characteristics of the network interior, such as link loss rates and packet latency. The first reason is that general organizations only have administrative access to a small fraction of the network's internal nodes, whereas commercial factors often prevent organizations from sharing internal performance data. The second reason is that the servers and routers in the network are generally operated by businesses, and those businesses may be unwilling or unable to cooperate in collecting the network traffic measurements for network management. Thus, monitoring the network interior has to rely on end-to-end measurements.

Network performance tomography (or network tomography) is proposed to acquire the characteristics of the network

interior by efficiently probing only end-to-end paths [1–3], rather than by directly monitoring every network element. It formulates the problem of inferring link metrics from end-to-end measurement results as a large linear system. Link metrics can be calculated by solving the linear equations in the system. Because the end-to-end measurements inevitably impose additional traffic on the networks, it is important to appropriately select end-to-end paths such that the desired inference capability can be achieved with as few end-to-end measurements as possible.

Given all available paths, the state-of-the-art solutions in network tomography select a subset of available paths, determined by finding an arbitrary basis of the linear system [1, 4]. However, these methods typically assume a simple network model, in which all network elements are reliable. Tati et al. in [3] argue that failures of network elements are common events in modern networks and that the typical durations of link failures in IP networks are longer than the lengths of time windows for measurement collection. After

one or more links break down, measurements on the selected paths that cover these link failures will be unavailable.

In addition, all existing tomography techniques select measurement paths offline without considering the real-time states of the current system. These approaches can be quite inefficient: they need to run repeatedly an unnecessarily large set of measurement probes capable of determining all links' loss rates, many of which might in fact never lose packets. In practice, a small number of end-to-end measurements that cover these *good* links are definitely sufficient to determine their loss rates. In general, most links in communication networks are such *good* links. Thus, a more adaptive tomography technique is necessary.

To improve the performance of existing network tomography applications, in this paper, we, for the first time, propose an efficient loss inference method (named ALIA) that performs the path selection and link metrics inference in an adaptive manner. More specifically, it first selects and measures the most informative end-to-end path, depending on the previous measurement results. Then, it infers the particular metrics of links that can be unbiasedly determined currently. Finally, it removes all determined links from the system and returns to the first step. The path selection and metrics inference will be performed repeatedly until no informative measurements remain. In this manner, all identifiable links and potential identifiable links can be unbiasedly determined using only a small set of end-to-end measurements. Moreover, link failures can be detected in time to avoid more unavailable measurements. In this paper, we only specifically focus on inferring the metrics of link loss rates, although our approach is also applicable to other link metrics. In summary, this paper makes the following contributions to the field of network tomography applications.

- (1) We propose an adaptive loss inference engine motivated by our two observations. The selection of measurement paths is adaptive to the current system state, depending on the earlier measurement results. Early measurements can provide the following important information. On the one hand, *good* measurement outcomes can unbiasedly determine all links on these path (see Section 4.1). On the other hand, failed measurements can be used to locate the link failures. Furthermore, because the links that have been determined can be removed, the system scale will be reduced in each subsequent round.
- (2) We develop an efficient path selection method and an accurate fault localization method. In each selection round, we select the longest path that does not lie in the current system space, and we demonstrate that measuring this path can obtain the maximum information; once a new measurement fails, we add all links on the path into a suspected fault set. We also define a weight ω for each suspected link: if a link's ω exceeds a certain threshold, then we consider it to be a real link failure. Hence, all paths that transverse the real failures will not be selected in the subsequent selection round.

- (3) We confirm the benefits of our proposed method in comparison with the previous solution in realistic network scenarios through simulations. The results show that, in most cases, our new approach uses only half of the other solution's measurement cost and computational time, but it unbiasedly determines even more links than all available end-to-end measurements. Moreover, the accuracy of our fault localization is 95%, with less than 20% false positives. All of the results strongly demonstrate that the proposed approach significantly improves the performance of network tomography applications.

The remainder of this paper is organized as follows. We first survey the related works in Section 2. Then, we present the definitions and formally describe our problem in Section 3. In Section 4, we present two observations to explain the main problems in the existing tomography techniques. Then, we present our adaptive loss inference-based approach in Section 5. Finally, we evaluate our new approach on realistic topologies in Section 6 prior to concluding the paper in Section 7.

2. Related Work

Network tomography has been widely used in, but not limited to, the fields of inferring individual link characteristics [5], network topology inference [6], and estimating the complete set of end-to-end measurements from an incomplete set [7]. In this paper, we only specifically focus on the applications of inferring link loss rates. Existing works on link loss inference primarily focus on two problems. The first problem is how to select a set of minimum paths to reduce the traffic overhead while maximizing their performance. The second problem is how to unbiasedly determine the loss rates of most links using the measurement outcomes of the selected paths.

Chen et al. [1] first proposed selecting the independent paths by finding a basis of the linear system through QR decomposition. Ma et al. in [2] proposed STPC (spanning tree-based path construction) to construct linearly independent monitor-to-monitor paths that can uniquely determine most links under an environment in which all network routers support the source routing policy. Zheng and Cao [5] selected a minimum path set that can identify all *identifiable* links and cover all *unidentifiable* links. Tati et al. [3] considered the presence of link failures in current networks and proposed RoMe for tolerating link failures by selecting the path set with the maximum *expected rank*. Zhao et al. [4] proposed *LEND*, which can infer the loss rates of all identifiable links and minimal identifiable link sequences with the least bias.

All of the above methods select paths offline and infer the link metrics after the path selection stages. The adaptive path selection approach has already been used in the field of fault diagnosis in our previous works [8, 9]. However, the methods proposed in [8, 9] can only select the measurement paths capable of distinguishing the two states of network elements: normal or fault. In this paper, we aim to select the end-to-end

FIGURE 1: An example of a network system.

TABLE 1: Paths and their transmission rates in Figure 1.

Paths	Routes	ϕ_i
$P_1: H_1 \rightarrow H_2$	$e_1 \rightarrow e_2 \rightarrow e_9$	0.75
$P_2: H_1 \rightarrow H_3$	$e_1 \rightarrow e_2 \rightarrow e_{10}$	0.85
$P_3: H_1 \rightarrow H_4$	$e_1 \rightarrow e_3 \rightarrow e_7$	0.86
$P_4: H_1 \rightarrow H_5$	$e_1 \rightarrow e_2 \rightarrow e_6 \rightarrow e_8$	1.00
$P_5: H_2 \rightarrow H_3$	$e_9 \rightarrow e_{10}$	0.64
$P_6: H_2 \rightarrow H_4$	$e_9 \rightarrow e_4 \rightarrow e_7$	0.43
$P_7: H_2 \rightarrow H_5$	$e_9 \rightarrow e_6 \rightarrow e_8$	0.75
$P_8: H_3 \rightarrow H_4$	$e_{10} \rightarrow e_4 \rightarrow e_7$	0.48
$P_9: H_3 \rightarrow H_5$	$e_{10} \rightarrow e_6 \rightarrow e_8$	0.85
$P_{10}: H_4 \rightarrow H_5$	$e_7 \rightarrow e_5 \rightarrow e_8$	0.76

measurements, which can unbiasedly determine the loss rates of most links.

In our recent work [10], we proposed a path selection method named APSA. It divided the path selection into two steps: covering path selection to select the max-coverage paths and solving path selection to select the most important paths using the graph construction and decomposition method. However, APSA only focuses on overcoming the problem of path selection without considering the link loss inference problem. Furthermore, it cannot be applied to the networks that present link failures.

3. Definitions and Problem Formulation

We consider the Internet loss inference systems that consist of routers and communication links. Some routers in the network can be directly connected by end hosts, which can send and receive probing packets. For example, Figure 1 is a network system with 10 links and 9 routers, 5 of which are connected by end hosts. Because the link between an end system and a router is quite short and stable, we only count the performance of the links among routers.

Let $G = (\mathbf{v}, \mathbf{e})$ denote the network with a set of nodes (routers) $\mathbf{v}$ and links $\mathbf{e}$. The numbers of nodes and links are denoted by $|\mathbf{v}|$ and $|\mathbf{e}|$, respectively. We define a path P_i as a sequence of links starting from a source host and ending at a destination host. All paths in the network form the path set $\mathbf{P}$, which contains $|\mathbf{P}|$ paths. Table 1 lists the available end-to-end path set $\mathbf{P}$ in Figure 1.

For a given network $G = (\mathbf{v}, \mathbf{e})$ and a path set $\mathbf{P}$, we define the routing matrix $\mathbf{R}$ with dimension $M \times N$, where $M = |\mathbf{P}|$ and $N = |\mathbf{e}|$, as follows: each row of $\mathbf{R}$ represents a path in the network, and the columns represent links; $\mathbf{R}_{ij} = 1$ when path P_i traverses link e_j , and $\mathbf{R}_{ij} = 0$ otherwise. For example, the routing matrix of paths in Table 1 is as follows:

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}. \quad (1)$$

Let $\hat{\phi}_i$ be a random variable given the fraction of the number of probe packets that arrive correctly at the destination monitor in the current measurement. Let $\hat{\phi}_{e_j}$ be the fraction of packets from all paths passing through link e_j that have not been lost at that link. For any path P_i , we define its transmission rate as $\phi_i = \mathbb{E}(\hat{\phi}_i)$. Similarly, the transmission rate of link e_j can be defined as $\phi_{e_j} = \mathbb{E}(\hat{\phi}_{e_j})$.

Given the routing matrix $\mathbf{R}$, the relationship between the transmission rates of paths in $\mathbf{P}$ and the transmission rates of links in $\mathbf{e}$ can be formulated as follows:

$$\phi_i = \prod_{j=1}^N \phi_{e_j}^{(\mathbf{R}_{ij})}. \quad (2)$$

Taking the logarithms on both sides of (2), we can rewrite it as

$$\log \phi_i = \sum_{j=1}^N \mathbf{R}_{ij} \log \phi_{e_j}. \quad (3)$$

Let $X_j = \log \phi_{e_j}$ and $Y_i = \log \phi_i$, which are grouped in vectors $\mathbf{X} = \{X_1, \dots, X_N\}$ and $\mathbf{Y} = \{Y_1, \dots, Y_M\}$, respectively. Equation (3) is equivalent to

$$\mathbf{R}\mathbf{X} = \mathbf{Y}. \quad (4)$$

The above formulations are similar to those in the other tomography-based works, including our previous study [10].

To identify the loss rates (loss rate = 1 - transmission rate) of individual links, it is necessary to solve the linear equations (4). Normally, the number of rows in $\mathbf{R}$ is considerably larger than the number of columns. Unfortunately, in most cases, $\mathbf{R}$ is still column deficient. Consequently, we cannot obtain the unique solution of (4) without additional information of the system. However, some of the links in (4) can be unbiasedly determined, which we call *identifiable links*. For example, in (1),

e_9 and e_{10} are identifiable links because $X_9 = (Y_1 - Y_2 + Y_5)/2$ and $X_{10} = (Y_2 - Y_1 + Y_5)/2$, and the remainder of the links are all *unidentifiable links*.

Normally, the number of end-to-end paths $|P|$ is on the order of $O(|V|^2)$. For relatively large networks, probing on all paths will cost considerable probing time as well as large traffic overhead. Therefore, it is necessary to carefully select the probing paths that are the most informative for the inference. In this paper, our goal is to select the fewest probing paths to unbiasedly determine the most links.

4. Observations

In this section, we present two observations to highlight the critical problems that prevent the former loss inference approaches from achieving the desired performance.

4.1. Observation 1: Good Paths. If one path has a loss rate of zero (or near zero), we define this path as a *good path*. *Good paths* indicate that all links on them do not lose any packets. Therefore, links that are classified as *unidentifiable* can also be unbiasedly determined if they are lying on *good paths*. Former methods that select an arbitrary basis of paths without considering the *good paths* will inevitably leave out an appreciable number of such *good links* that should be unbiasedly determined. For example, the rank of (1) is 7, which means that the number of paths in the basis is 7. If we select $P_1, P_2, P_3, P_5, P_6, P_7$, and P_{10} as the basic path set, for which the measurement results are listed in Table 1, only e_9 and e_{10} can be uniquely determined because they are *identifiable links*. However, if we measure one additional path, such as $P_4 = \{e_1, e_2, e_6, e_8\}$, and observe that the loss rate of P_4 is 0, then there are 4 more links that can be unbiasedly determined: e_1, e_2, e_6 , and e_8 . In some cases removing these *good links* (e_1, e_2, e_6 , and e_8) from the system may generate additional *identifiable links*. In the rest part of this paper, we define all links such as e_1, e_2, e_6 , and e_8 and any other additional identifiable links (if possible) that should be unbiasedly determined as *potential identifiable links*.

In fact, there is a considerable number of paths that present 0 (or near 0) loss rates in most networks, which means that an appreciable number of *potential identifiable links* can be missed by the previous loss inference methods. We draw the cumulative distribution of loss rates on paths under different fractions of lossy links in Figure 2. In this figure, we use the realistic topology AS1239 from the Rocketfuel Project [11]. The detailed settings are provided in Section 6. As shown in Figure 2, more than 32% of paths are *good paths* when the network has 30% links that lose packets. This proportion is up to 98% when there are only 1% of links that lose packets. In the remainder of this paper, we consider a path to be *good* if the path loss rate is under 2%.

4.2. Observation 2: Failure Links. Existing works assume a simple network model, in which all network elements are reliable. However, failures of network elements are common events in modern networks due to maintenance procedures, hardware malfunctions, energy outages, or disasters [12]. The typical durations of link failures in IP networks are longer

FIGURE 2: The cumulative distribution of loss rates on paths.

FIGURE 3: Rank of a basis under failures (adopted from [3]).

than the lengths of time windows for measurement collection in network tomography [13]. Hence, the link failures may prevent the collection of some measurements. For example, suppose that $P_1, P_2, P_3, P_5, P_6, P_7$, and P_{10} are selected for inference and that the rank of the system is 7. If e_1 is now in a failure state, only paths P_5, P_6, P_7 , and P_{10} can be successfully probed. Hence, the provided rank is reduced to 4, and none of the links can be unbiasedly determined.

We plot the average ranks provided by two arbitrary bases and by all paths as we increase the number of link failures in Figure 3. The results indicate that link failures significantly degrade the quality of the selected paths.

5. Adaptive Loss Inference Algorithm

In this section, we design an adaptive loss inference algorithm (named ALIA), which has the following main advantages. On the one hand, it can unbiasedly determine all identifiable and potential identifiable links using even fewer end-to-end measurements than the system basis. On the other hand, it can locate the link failures during the inference process to avoid probing on the unavailable paths.

5.1. Overview. The structure of our approach is outlined in Figure 4. We first select the path $P_{\max}$ that is the most informative to determine the system links, and then we probe

FIGURE 4: The structure of ALIA.

the path to obtain the path transmission rate $\phi(P_{\max})$. If $\phi(P_{\max}) > 0$, we know that all links on the paths are not in failure states, and these links can be used to filter the suspected link failures. Subsequently, the new measurement outcome $\phi(P_{\max})$ is used to infer the link loss rates. All links that can currently be unbiasedly determined will be removed from the system to reduce the system scale; otherwise, if $\phi(P_{\max}) = 0$ (we use $\phi(P_{\max}) = 0$ to denote the failed path), there is at least one link on $P_{\max}$ that is (are) in failure states. Therefore, we temporarily add all links on $P_{\max}$ to the suspected failure set, and then we check whether each suspected failure meets a certain condition. If so, we mark it as a real failure. In the next round, we will select the most informative path $P_{\max}$ in the reduced system that will not transverse the real failures. The above steps will be repeated until there is no informative path remaining in the system.

5.2. Path Selection. The information of one path is considered from two aspects. The first is the number of links that the path includes. The second is whether the path can increase the rank of the current system.

For the first aspect, we select the path (denoted by $P_{\max}$) that includes the most links over all candidate paths. Such path can provide more information than others because once we observe that the transmission rate of path $P_{\max}$ equals 1 (i.e., $\phi(P_{\max}) = 1$), we can determine all links on $P_{\max}$ as *good* links. Although $1 < \phi(P_{\max}) < 0$, we can also use all links on the path to filter the suspected failures. Moreover, although there are relatively more links that will be considered as suspected failures if the measurement on $P_{\max}$ is failed (i.e., $\phi(P_{\max}) = 0$), this problem can be postponed or even ignored because the probability of $\phi(P_{\max}) = 0$ is typically very small.

For the second aspect, the rank of the system is proportional to the number of identifiable links [1]. Therefore, we select the path that is not lying in the current system space. Suppose that Q is an orthonormal basis for the current path space; path P^* is not lying in the space if and only if $\|P_{\max}\| \neq \|P_{\max}Q\|$ [4]. Adding this path to the current system will increase the rank of the system by 1. Note that because the determined links will be removed in each round, the space of the system also changes round by round. Therefore, we need to compute the orthonormal basis of the current path space every round prior to the path selection.

In our previous work [10], we also proposed a path selection algorithm named APSA, which divides the process of path selection into two steps: selecting the covering paths which can cover all links and selecting solving paths which

can determine the most links using the *graph construction and decomposition method*. However, link loss rates cannot be inferred during the path selection in APSA, and also, APSA cannot handle the link failures in the network. Thus, the *graph construction and decomposition method* has not been adopted in this paper.

5.3. Loss Inference. The loss rates of links can be unbiasedly determined in two ways. The first is from the good paths whose transmission rates are 1. The other is through solving the system equations in (4).

For the first way, once we observe that the new selected path has a loss rate of 0, all links on this path also have loss rates of 0. In such a case, we directly remove these links from the system. For the second way, we determine all *identifiable links* in current system through Theorem 1 proposed in our previous work [9], as follows.

Assume that $\mathbf{R}$ is an m by n matrix and that the rank of $\mathbf{R}$ is r (here, let $r < n$ because if $r = n$, all links are identifiable). Let $N(\mathbf{R})$ denote the null space of $\mathbf{R}$, that is, for any vector $\eta \in N(\mathbf{R})$, $\mathbf{R}\eta = 0$. $\{\eta_1, \eta_2, \dots, \eta_{n-r}\}$ represents an arbitrary basis for $N(\mathbf{R})$, where $\eta_i = \{\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}\}$, $1 \leq i \leq n-r$.

Theorem 1. Link e_j , which is represented by the j th column of $\mathbf{R}$, can be uniquely identified if and only if for all η_i , $1 \leq i \leq n-r$, $\alpha_{ij} = 0$.

After determining all *identifiable links*, it is necessary to compute their loss rates. This is performed by finding an arbitrary solution $\bar{\mathbf{X}}$ through $\bar{\mathbf{X}} = \text{pinv}(\mathbf{A})\mathbf{Y}$, where $\bar{\mathbf{X}} = \{\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m\}$ and $\text{pinv}(\mathbf{A})$ is the pseudoinverse matrix of $\mathbf{A}$. The loss rate of identifiable link e_i is equivalent to $1 - \exp(\bar{X}_i)$. We also adjust the value of $\mathbf{Y}$ in the system by $\mathbf{Y} = \mathbf{Y} - \bar{\mathbf{X}}_i$ when e_i is removed from the corresponding path.

Note that some paths in the system may become *good* when the *identifiable links* have been removed. In such a case, we consider all links on these paths as *good* links and remove them from the system, and then we repeat the above steps until there is no link in the system that can be unbiasedly determined.

5.4. Fault Localization. In ALIA, link failures are located during the loss inference process. We define $\mathcal{F}$ as a suspected failure set, which consists of link combinations. Suppose that $P^* = \{e_1^*, e_2^*, \dots, e_k^*\}$ is the new selected path. If its transmission rate $\phi(P^*) > 0$, then we filter the links in $\mathcal{F}$

from the path $P^* = \{e_1^*, e_2^*, \dots, e_k^*\}$ and set the states of all links on P^* to be *normal*. Otherwise, if $\phi(P^*) = 0$, we pick up links from $e_1^*, e_2^*, \dots, e_k^*$ whose states have not been set to be *normal* and add them to the suspected fault set $\tilde{\mathcal{F}}$. For each link e_k in $\tilde{\mathcal{F}}$, we define a weight ω_k , as follows:

$$\omega_k = \frac{|\{P_j \mid P_j \in \mathbf{P}_F, e_k \in P_j\}|}{|\{P_i \mid P_i \in \mathbf{P}_F\}|}. \quad (5)$$

Here, $\mathbf{P}_F$ is the set of currently failed paths. The physical meaning of (5) is the fraction of failed paths that currently cover link e_k in all failed paths. We define two thresholds, λ and β , as follows: the weights of suspected links will not be computed until the number of failed paths reaches the threshold λ . For each link e_k in $\tilde{\mathcal{F}}$, if $\omega_k > \beta$, we mark e_k as a real failure and add it to the real failure set $\mathcal{F}$. Once the real failure $\mathcal{F}$ is not empty, paths that include the real failures will not be selected in the further selection round. If the real failure set $\mathcal{F}$ is still empty after we have finished the path selection, we simply put all links in the suspected set $\tilde{\mathcal{F}}$ into the real failure set $\mathcal{F}$.

In our previous work [10], loss inference and fault localization were not taken into account.

5.5. The Algorithm Details. The details of the algorithm are shown in Algorithm 1. It inputs the candidate path matrix $\mathbf{R}$, the measurement module $\mathbf{M}$, and the thresholds β and λ and outputs the links $\mathbf{e}_1$ which can be unbiasedly determined, their transmission rates $\phi_{\mathbf{e}_1}$, and the real failures $\mathcal{F}$. It first initializes the real failure as null (line (1)) and then does the while loop until there is no informative path left in the matrix $\mathbf{R}$ (lines (2)~(24)). In the while loop, it first selects the informative path $P_{\max}$ from matrix $\mathbf{R}$ and then measures the transmission rate of this path (lines (3)~(4)). If the measurement is not failed it uses the links on the path to filter the failures in the suspected failure set $\tilde{\mathcal{F}}$ (line (6)) and further judges whether its transmission rate exceeds 0.98. If so, it assigns the transmission rates of links on the path to 1 (lines (8)~(9)); otherwise, it adds the path and its transmission rate into the current matrix $\mathbf{A}$ and the vector $\mathbf{Y}$, respectively, and then finds the identifiable links in current system (line (11)~line (13)) and computes their transmission rates. Finally, it adds the determined links and their transmission rates into $\mathbf{e}_1$ and $\phi_{\mathbf{e}_1}$, respectively, and removes them from the system (lines (14)~(16)). If the measurement on path $P_{\max}$ is failed, it adds the links on the paths to $\tilde{\mathcal{F}}$ and adds the path to the failed path set $\mathbf{P}_F$ (lines (18)~(19)). Once the number of paths in $\mathbf{P}_F$ exceeds λ , it checks each link in $\tilde{\mathcal{F}}$ using (5) to find out the real failures and then adds them into the set $\mathcal{F}$ (lines (21)~(24)).

The most time-consuming step in Algorithm 1 is to find out the identifiable links in current system (line (13)) based on Theorem 1. The time complexity is the order of $O(m \times n)$, where m and n are the dimensions of the current matrix $\mathbf{A}$. Hence, the total time complexity of Algorithm 1 is the order of $O(M \times m \times n)$, where M is the number of finally selected paths. The dimension of $\mathbf{A}$ changes in each round but is at most equal to $M \times N$, where N is the number of network links. And

```

Input:  $\mathbf{R}, \mathbf{M}, \beta, \lambda$ 
Output:  $\mathbf{e}_1, \phi_{\mathbf{e}_1}, \mathcal{F}$ 
(1)  $\mathcal{F} \leftarrow \text{NULL}$ 
(2) while nonzero( $\mathbf{R}$ ) do
(3)    $P_{\max} \leftarrow \text{maxCover}(\mathbf{R}, \mathcal{F})$ 
(4)    $\phi(P_{\max}) \leftarrow M(P_{\max})$ 
(5)   if  $\phi(P_{\max}) > 0$  then
(6)      $\tilde{\mathcal{F}} \leftarrow \text{filterFailures}(\tilde{\mathcal{F}}, P_{\max})$ 
(7)     if  $\phi(P_{\max}) \geq 0.98$  then
(8)        $e^* \leftarrow \text{getLinks}(P_{\max})$ 
(9)        $\phi_{e^*} \leftarrow 1$ 
(10)    else
(11)       $\mathbf{A} \leftarrow [\mathbf{A}; P_{\max}]$ 
(12)       $\mathbf{Y} \leftarrow [\mathbf{Y}; \phi(P_{\max})]$ 
(13)       $[e^*, \phi_{e^*}] \leftarrow \text{computeIdenLinks}(\mathbf{A}, \mathbf{Y})$ 
(14)       $\mathbf{e}_1 \leftarrow [e_1, e^*]$ 
(15)       $\phi_{\mathbf{e}_1} \leftarrow [\phi_{\mathbf{e}_1}, \phi_{e^*}]$ 
(16)       $\text{reduceSystem}(\mathbf{R}, \mathbf{A}, \mathbf{Y}, e^*)$ 
(17)    else
(18)       $\tilde{\mathcal{F}} \leftarrow \text{addFailures}(\tilde{\mathcal{F}}, P_{\max})$ 
(19)       $\mathbf{P}_F \leftarrow [\mathbf{P}_F, P_{\max}]$ 
(20)      if  $|\mathbf{P}_F| > \lambda$  then
(21)        forall the  $e_k \in \tilde{\mathcal{F}}$  do
(22)           $\omega_k \leftarrow \text{computeWeight}(\mathbf{P}_F)$ 
(23)          if  $\omega_k \geq \beta$  then
(24)             $\mathcal{F} \leftarrow \text{realFailures}(\mathcal{F}, e_k)$ 
(25) return  $\mathbf{e}_1, \phi_{\mathbf{e}_1}, \mathcal{F}$ 

```

ALGORITHM 1: Adaptive loss inference algorithm.

the number of paths that ALIA finally selected M is much smaller than the rank r of candidate path matrix $\mathbf{R}$ in most cases, according to our experimental results in Section 6.3. Therefore, the time complexity of Algorithm 1 is at most the order of $O(r^2 \times N)$.

5.6. A Working Example. In Table 1, suppose that e_7 is now in a failure state. We set $\lambda = 3$ and $\beta = 0.8$. ALIA finally selects 6 paths and unbiasedly determines 6 links during the 6 rounds, as follows:

- (i) In the first round, $P_4 = \{e_1, e_2, e_6, e_8\}$ is selected because it covers the most links. Then, we measure the transmission rate P_4 and obtain $\phi(P_4) = 1$ through probing on the path. Subsequently, we record the transmission rates of e_1, e_2, e_6 , and e_8 as 1 and remove them from the system. The candidate path matrix of the current system R_1 is shown in Figure 5(a).
- (ii) We select path $P_6 = \{e_4, e_7, e_9\}$ in the second round and obtain $\phi(P_6) = 0$. Therefore, e_4, e_7 , and e_9 are added to the suspected failure set $\tilde{\mathcal{F}}$.
- (iii) In the third round, $P_8 = \{e_4, e_7, e_{10}\}$ is selected, and it also has $\phi(P_8) = 0$. Currently, $\tilde{\mathcal{F}} = \{e_4, e_7, e_9, e_{10}\}$.
- (iv) In the fourth round, $P_3 = \{e_3, e_7\}$ is selected, and $\phi(P_3) = 0$. Now, $\tilde{\mathcal{F}} = \{e_3, e_4, e_7, e_9, e_{10}\}$. Because the number of failed paths reaches $\lambda = 3$, we begin to compute the weights for each suspected failure as

$$\begin{aligned}
R_1 &= \begin{array}{c} \begin{array}{cccccc|l} e_3 & e_4 & e_5 & e_7 & e_9 & e_{10} & \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & P_1 \rightarrow \text{vi} \\ 0 & 0 & 0 & 0 & 0 & 1 & P_2 \\ 1 & 0 & 0 & 1 & 0 & 0 & P_3 \rightarrow \text{iv} \\ 0 & 0 & 0 & 0 & 1 & 1 & P_5 \rightarrow \text{v} \\ 0 & 1 & 0 & 1 & 1 & 0 & P_6 \rightarrow \text{ii} \\ 0 & 0 & 0 & 0 & 1 & 0 & P_7 \\ 0 & 1 & 0 & 1 & 0 & 1 & P_8 \rightarrow \text{iii} \\ 0 & 0 & 0 & 0 & 0 & 1 & P_9 \\ 0 & 0 & 1 & 1 & 0 & 0 & P_{10} \rightarrow \text{X} \end{array} \\ (a) \end{array}
\end{aligned}$$

$$\begin{aligned}
R_6 &= \begin{array}{c} \begin{array}{cccc|l} e_3 & e_4 & e_5 & e_7 & \\ \hline 0 & 0 & 0 & 0 & P_2 \\ 0 & 0 & 0 & 0 & P_7 \\ 0 & 0 & 0 & 0 & P_9 \end{array} \\ (b) \end{array}
\end{aligned}$$

FIGURE 5: The candidate path matrix in current system.

$$\begin{aligned}
A_1 &= \begin{array}{c} \begin{array}{cccccc|l} e_3 & e_4 & e_5 & e_7 & e_9 & e_{10} & \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & P_5 \end{array} \quad Y_5 = \log(0.64) \\ (a) \end{array}
\end{aligned}$$

$$\begin{aligned}
A_2 &= \begin{array}{c} \begin{array}{cccccc|l} e_3 & e_4 & e_5 & e_7 & e_9 & e_{10} & \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & P_5 \quad Y_5 = \log(0.64) \\ 0 & 0 & 0 & 0 & 1 & 0 & P_1 \quad Y_1 = \log(0.75) \end{array} \\ (b) \end{array}
\end{aligned}$$

$$\begin{aligned}
Q_1 &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -0.7071 \\ -0.7071 \end{bmatrix} \\ \eta_2 &= \begin{bmatrix} 0 & 0 & -0.7071 & -0.7071 & e_3 \\ 0 & 0 & -0.7071 & -0.7071 & e_4 \\ 1 & 0 & 0 & 0 & e_5 \\ 0 & 1 & 0 & 0 & e_7 \\ 0 & 0 & 0 & 0 & e_9 \\ 0 & 0 & 0 & 0 & e_{10} \end{bmatrix}
\end{aligned}$$

FIGURE 6: The path matrix of selected paths in current system.

follows: $\omega_3 = 1/3 = 0.33$, $\omega_4 = 2/3 = 0.67$, $\omega_7 = 3/3 = 1$, $\omega_9 = 1/3 = 0.33$, and $\omega_{10} = 1/3 = 0.33$. Because $\omega_7 > \beta$, we consider e_7 as a real failure and add it to $\mathcal{F}$.

(v) In the fifth round, $P_5 = \{e_9, e_{10}\}$ is selected, and $\phi(P_5) = 0.64$. Hence, the matrix of selected paths A_1 and its orthonormal basis of the path space Q_1 is shown in Figure 6(a).

(vi) In the sixth round, we select path $P_1 = \{e_9\}$ and check that $\|P_1\| \neq \|P_1 Q_1\|$. P_{10} is skipped because $P_{10} = \{e_5, e_7\}$ covers the real failure e_7 . Then, we obtain $\phi(P_9) = 0.75$, and the selected paths' matrix of the current system is shown as A_2 in Figure 6(b). Through the basis of the null space η_2 and Theorem 1, we know that e_9 and e_{10} are identifiable links. We compute an arbitrary solution of the system through $\bar{X} = \text{pinv}(A_2)Y$ and obtain $\phi(e_9) = \exp(\bar{X}_9) = 0.75$ and $\phi(e_{10}) = \exp(\bar{X}_{10}) = 0.8533$.

After removing e_9 and e_{10} from the system, the candidate path matrix R_6 is as shown in Figure 5(b), where there are no informative paths remaining. ALIA now returns the loss rates of all determined links e_1, e_2, e_6, e_8, e_9 , and e_{10} , as well as the links failure e_7 .

6. Evaluation

6.1. Evaluation Setup

Topologies. We conduct our experiments on the realistic ISP topologies from the Rocketfuel Project [11]. We select the topologies of two autonomous systems with labels AS6461 and AS3356, which are representatives for small and large

TABLE 2: The details of topologies.

Topologies	AS6461	AS3356
Monitors	40	60
Candidate paths	1560	3540
Nodes	182	624
Links	214	5298
Covered links	142	1266

topologies, respectively. The numbers of nodes and links in the topologies are presented in Table 2.

Candidate Paths. We randomly select 40 and 60 nodes as the monitors that can both initiate and receive probes, respectively. The candidate paths are generated between each monitor pair, and all topologies in Table 2 adopt the shortest path routing policy. Links that cannot be covered by any paths are removed from the system. The numbers of candidate paths and the links covered by these paths are also listed in Table 2.

Link Loss. We allow each link to be congested with a probability p . Because p affects the selection result in our experiments, we vary p to evaluate the performances of the two algorithms. We use two different loss rate models, LLRD1 and LLRD2 of [14] (which are also used in [13, 15–17]), for assigning loss rates to links. In the LLRD1 model, congested links have loss rates uniformly distributed in $[0.05, 0.2]$, and good links have loss rates in $[0, 0.002]$. In the LLRD2 model, the loss rate ranges for congested and good links are $[0.002, 1]$ and $[0, 0.002]$, respectively. Because there is little difference between the two models, we only present our results for the

FIGURE 7: The number of paths selected by the two algorithms.

LLRD1 model. After assigning each link a loss rate, the actual losses on each link follow a Gilbert process. The link in the Gilbert model fluctuates between good and congested states. Links do not drop any packets when in the good state, and they drop all packets when in the congested state.

6.2. Baseline and Metrics. We compared our new algorithm (APSA) with the state-of-the-art path selection approach for network tomography called SelectPath [1]. Although several recent methods have been proposed to address the tomographic problems [3, 13, 15, 16, 18, 19], most of them are not comparable with our algorithm. For example, [13, 15, 16] do not select paths before they perform tomography, while selecting paths is a vital step in our algorithm; [18, 19] select monitoring paths to detect or locate the failures, but our algorithm aims to determine the loss rates of links. We choose SelectPath as the baseline not only because it works on problems similar to those of our algorithm but also because it is one of the most representative path selection algorithms that has been widely approved of in the research community.

SelectPath selects an arbitrary maximal set of linearly independent paths using QR decomposition with column pivoting [20]. Because there are also works that measure all available paths without path selection [13, 15, 16], we also present the performance of the entire candidate paths given their measurement results in our figures (marked as “All”).

The performances of the approaches are evaluated using the following three metrics: (i) probing cost, the number of selected paths; (ii) path quality, the number of links that can be unbiasedly determined from the selected paths given their loss rates; (iii) computing time, the period between the time they input the routing matrix and the time they return the selected paths and all determined links.

We also evaluate the performance of our algorithm on fault localization using two metrics: accuracy and false positive. Let $\mathbf{F}$ be the link set of real failures and $\mathcal{F}$ be the failures inferred by ALIA. $\bar{\mathbf{F}}$ is the set of links that are not in $\mathbf{F}$.

$$\text{accuracy} = \frac{|\mathbf{F} \cap \mathcal{F}|}{|\mathbf{F}|},$$

$$\text{False - positive} = \frac{|\bar{\mathbf{F}} \cap \mathcal{F}|}{|\mathcal{F}|}. \quad (6)$$

Due to space limitations, we only present the results on the varied fraction of failures and β 's, and we fix the threshold λ to 5. All the figures present results averaged over 20 runs.

6.3. Results

6.3.1. The Number of Paths Selected. In Figure 7, we plot the number of paths selected by ALIA and SelectPath as we vary the fraction of lossy links. The probability of failure is fixed to 0.1, and the threshold β is 0.4 (the same as in Figures 8 and 9). In both topologies, ALIA selects considerably fewer paths than the SelectPath algorithm, particularly when there are fewer lossy links in the network. Moreover, the advantage of ALIA becomes even more obvious in the relatively large networks (AS3356). As the fraction of lossy links increases, the number of paths selected by ALIA increases because the number of *good* paths decreases and our algorithm requires more paths to determine the system. Because the SelectPath algorithm always selects an arbitrary basis of the system, the curve of SelectPath gently fluctuates.

6.3.2. The Number of Links Determined. Next, we evaluate the link identifiability. Figure 8 shows the number of links that can be unbiasedly determined by the two algorithms and all available paths. As shown in this figure, ALIA can determine the most links among the approaches. In other words, ALIA uses the fewest measurements to unbiasedly determine the most links. In the figure, all three scenarios take the *good* measurement results into account (i.e., links on the selected paths whose loss rates are 0 will be considered as determined links). Thus, all of the curves decrease as the fraction of lossy links increases. The gap between ALIA and SelectPath indicates the potential identifiable links missed by SelectPath. ALIA performs even better than all available paths because it repeatedly removes the determined links from the system in each round, and additional identifiable links may emerge. For example, when we remove one lossy link, the lossy paths that

FIGURE 8: The number of links determined by the two algorithms.

FIGURE 9: The computing times of the two algorithms.

include this link may become *good*, and all other links on such paths can also be determined.

6.3.3. Computational Times. Figure 9 compares the computing times of the two algorithms. In both figures, ALIA runs considerably faster than SelectPath, because the inference system is reduced round by round as ALIA removes the determined links in every round. However, as the fraction of lossy links increases, the curve of ALIA goes up. The first reason is that the links that can be determined from the *good* paths will be reduced as the number of lossy links increases. The other reason is that ALIA needs to select more paths to determine the system with more lossy links. Nevertheless, for most communication networks, the fraction of lossy links is generally less than 10%. In other words, ALIA can reduce the computing time of SelectPath by more than 50% in most cases.

6.3.4. Accuracy on Fault Localization. Figure 10 shows the performance of ALIA on fault localization averaging over all topologies. We first vary the fraction of links failures, and we plot the results in Figure 10(a). Here, the threshold β and the fraction of lossy links are fixed to 0.4 and 0.1, respectively. In the figure, the accuracy curve slightly decreases while the

false positive curve slowly increases when the fraction of failures increases. Nevertheless, ALIA can locate more than 95% of failures even when 30% of the links are in failure states. In Figure 10(b), we vary the threshold β and fix the probability of failure to 0.2. From the figure, we can observe a valley in each of the curves. This result occurs because as β increases, fewer links in $\tilde{\mathcal{F}}$ will be considered as real failures. Consequently, the accuracy and false positive curves both decrease. When β increases to a certain value, there is no link that satisfies the real failure condition. In such a case, ALIA will place all links in the final $\tilde{\mathcal{F}}$ into the set $\mathcal{F}$, leading to relatively high accuracy and false positive. This is also the reason why we assign β to 0.4 in most of our experiments.

7. Conclusion

In this paper, we present an adaptive loss inference approach (named ALIA) for network tomography applications. We first present two observations to argue that existing tomography-based approaches are far from perfect. On the one hand, selecting end-to-end measurements offline will inevitably miss an appreciable number of potential identifiable links. On the other hand, link failures will cause many measurements to be unavailable. Our proposed approach performs the path

FIGURE 10: Accuracy on fault localization.

selection and loss inference round by round based on the earlier measurement results until there are no informative measurements remaining. In this way, both the overall probing cost and the computing time can be significantly reduced. In addition, ALIA can determine all potential identifiable links and locate the link failures with high accuracy. Through extensive simulations on realistic ISP topologies, the results strongly confirm the promising performance of ALIA.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61402013, no. 31671589) and the Open Foundation of State Key Laboratory of Networking and Switching Technology (SKLNST-2016-1-02).

References

- [1] Y. Chen, D. Bindel, H. Song, and R. H. Katz, "An algebraic approach to practical and scalable overlay network monitoring," in *Proceedings of the ACM SIGCOMM 2004: Conference on Computer Communications*, pp. 55–66, Portland, Ore, USA, September 2004.
- [2] L. Ma, T. He, K. K. Leung, D. Towsley, and A. Swami, "Efficient identification of additive link metrics via network tomography," in *Proceedings of the IEEE 33rd International Conference on Distributed Computing Systems (ICDCS '13)*, pp. 581–590, July 2013.
- [3] S. Tati, S. Silvestri, T. He, and T. La Porta, "Robust network tomography in the presence of failures," in *Proceedings of the IEEE 34th International Conference on Distributed Computing Systems (ICDCS '14)*, pp. 481–492, IEEE, Madrid, Spain, July 2014.
- [4] Y. Zhao, Y. Chen, and D. Bindel, "Towards unbiased end-to-end network diagnosis," *IEEE/ACM Transactions on Networking*, vol. 17, no. 6, pp. 1724–1737, 2009.
- [5] Q. Zheng and G. Cao, "Minimizing probing cost and achieving identifiability in probe-based network link monitoring," *IEEE Transactions on Computers*, vol. 62, no. 3, pp. 510–523, 2013.
- [6] M. Coates, R. Castro, R. Nowak, M. Gadhiok, R. King, and Y. Tsang, "Maximum likelihood network topology identification from edge-based unicast measurements," in *Proceedings of the ACM (SIGMETRICS '02) International Conference on Measurement and Modeling of Computer Systems*, pp. 11–20, New York, NY, USA, June 2002.
- [7] Y. Zhang, M. Roughan, W. Willinger, and L. Qiu, "Spatio-temporal compressive sensing and internet traffic matrices," in *Proceedings of the ACM SIGCOMM Conference on Data Communication (SIGCOMM '09)*, pp. 267–278, ACM, Barcelona, Spain, August 2009.
- [8] L. Cheng, X. Qiu, L. Meng, Y. Qiao, and R. Boutaba, "Efficient active probing for fault diagnosis in large scale and noisy networks," in *Proceedings of the IEEE (INFOCOM '10)*, March 2010.
- [9] Y. Qiao, X. Qiu, L. Meng, and R. Gu, "Efficient loss inference algorithm using unicast end-to-end measurements," *Journal of Network and Systems Management*, vol. 21, no. 2, pp. 169–193, 2013.
- [10] Y. Qiao, J. Jiao, Y. Rao, and H. Ma, "Adaptive path selection for link loss inference in network tomography applications," *PLoS ONE*, vol. 11, no. 10, Article ID e0163706, 2016.
- [11] "Rocketfuel project: internet topologies," <http://www.cs.washington.edu/research/networking/rocketfuel/>.
- [12] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot, "Characterization of failures in an IP backbone," in *Proceedings of the Conference on Computer Communications - 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM '04)*, pp. 2307–2317, Hong Kong, China, March 2004.
- [13] H. X. Nguyen and P. Thiran, "Network loss inference with second order statistics of end-to-end flows," in *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement (IMC '07)*, pp. 227–240, San Diego, Calif, USA, October 2007.
- [14] V. N. Padmanabhan, L. Qiu, and H. J. Wang, "Server-based inference of internet performance," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM '03)*, vol. 1, pp. 145–155, Orlando, Fla, USA, April 2003.
- [15] H. X. Nguyen and P. Thiran, "The boolean solution to the congested IP link location problem: theory and practice," in *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, pp. 2117–2125, IEEE, Anchorage, Alaska, USA, May 2007.

- [16] D. Ghita, H. Nguyen, M. Kuran, K. Argyraki, and P. Thiran, "Netscope: practical network loss tomography," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM '10)*, pp. 1–9, IEEE, San Diego, Calif, USA, March 2010.
- [17] M. Malboubi, C. Vu, C.-N. Chuah, and P. Sharma, "Compressive sensing network inference with multiple-description fusion estimation," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM '13)*, pp. 1557–1563, Atlanta, Ga, USA, December 2013.
- [18] D. Jeswani, M. Natu, and R. K. Ghosh, "Adaptive monitoring: application of probing to adapt passive monitoring," *Journal of Network and Systems Management*, vol. 23, no. 4, pp. 950–977, 2015.
- [19] E. Cohen, A. Hassidim, H. Kaplan, Y. Mansour, D. Raz, and Y. Tzur, "Probe scheduling for efficient detection of silent failures," *Performance Evaluation*, vol. 79, no. 3, pp. 73–89, 2014.
- [20] G. H. Golub and V. C. F. Loan, "Matrix computations," *Mathematical Gazette*, vol. 47, no. 5, pp. 392–396, 1983.

Hindawi

Submit your manuscripts at
<http://www.hindawi.com>