Network Topology Tomography Under Multipath Routing

Shengli Pan, *Student Member, IEEE*, Yingjie Zhou, Fucai Yu, Feng Qian, and Guangmin Hu

***Abstract*—Network topology tomography can infer a tree topology for single-source networks using end-to-end measurements. However, multipath routing, which introduces multiple paths between end-hosts, violates the tree topology model. In this letter, we demonstrate that such nontree topologies are also identifiable. We employ graph cuts and show the nontree topology can be decomposed into two identifiable subtopologies. And by reconnecting the cut paths, the nontree topology can be recovered after obtaining these subtopologies. To detect the paths that share cuts, we propose a scheme based on measurements of the end-to-end packet arrival order. Simulation results show that our scheme achieves the desirable detection accuracy.**

***Index Terms*—Network tomography, topology identification, multipath routing, end-to-end measurement.**

I. Introduction

NETWORK topology tomography [1], [2], based on end-to-end measurements, is an appealing alternative for logical topology inference when methods that use direct measurements (e.g., traceroute [2]) are not tractable. Long-term interest has been given to the tree topology inference [3]–[9]. However, the tree model for single-source networks will be violated if multipath routing [10] gets involved.

In general, tree topology inference assumes single-path routing, where the routing path between any two end-hosts remains unique. Notice that in a tree topology, every two end-to-end paths construct an inverted "Y" structure [1]. Then through end-to-end measurements, e.g., the "back-to-back" probing [1], [2], the delay covariances between paths can be used to determine where their branching nodes are [5], [11]. Abstracting the delay covariance as a length metric for the shared paths, [7] generally shows that the tree topology is identifiable from the shared path lengths of every two of its end-to-end paths. Moreover, the multi-source topology (having the "Y" structures) has also been proved as identifiable by [12] when the lengths of both the end-to-end paths and their shared paths are known. The multi-source topology is shown reconstructable via merging the tree sub-topologies.

Although the above methods work well under single-path routing, the topology of a network will not be fully discovered when it employs multipath routing [10]. E.g., since two end-hosts could have more than one routing path, a single-source network actually gets a non-tree topology rather than its inferred tree topology. Besides the concerns about mapping the probing flows to their actual routing paths in [10] and [13], a more tricky issue is that the shared path of two paths can be nonconsecutive, making it unable to be used as usual to locate the branching node or the joining node directly [12].

In this letter, we demonstrate that it is possible to decompose the non-tree topology into a tree sub-topology and a multi-source sub-topology. To identify cut-sharing paths for such non-tree topology decomposition, we propose an algorithm based on the detection of the cut-sharing paths with end-to-end measurements of the packet arrival order. Simulation results validate the good accuracy of our scheme.

Manuscript received August 15, 2015; accepted January 6, 2016. Date of publication January 11, 2016; date of current version March 8, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61171091, Grant 61571094, and Grant 61201127, and in part by the Science and Technology on Communication Security Laboratory (no. 9140C110503140C11054). The associate editor coordinating the review of this paper and approving it for publication was G. Reali.

S. Pan, F. Yu, F. Qian, and G. Hu are with the School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail:hgm@uestc.edu.cn)

Y. Zhou is with the School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China, and also with the College of Computer Science, Sichuan University, Chengdu 610041, China.

Digital Object Identifier 10.1109/LCOMM.2016.2516531

II. Model Under Multipath Routing

We aim at the logical topology [2] of the single-source network, where multipath routing is introduced by per-flow load balancing [10]. Particularly, packets with the same five-tuple will be forwarded to the same next hop by the load balancer, and accordingly they will follow the same route. Let the directed acyclic graph Γ be the non-tree topology of the single-source network under multipath routing, and let σ be the unique source node. We denote the set of end-to-end paths from σ by $\mathcal{P} = \{p_1, \ldots, p_i, \ldots\}$, where p_i is the i-th path. When p_i has the destination node d, we say $p_i \to d$. If there are also other paths from σ to d, we say $d \in \mathbb{D}$; Otherwise, $d \in D$. For convenience, $p_i \to D$ (resp. $p_i \to \mathbb{D}$) is also used when p_i is destined for any node in D (resp. $\mathbb{D}$).

For two paths from D, it is common that they never rejoin once they split (e.g., in the tree topology). While they are from $\mathbb{D}$, they can join again. E.g., in Fig. 1, $p_{12} \to 6$ and $p_9 \to 5$ join $p_{10} \to 6$ individually at r_3 and r_1. This normally results in the nonconsecutive shared paths, i.e., the shared path between two paths will end at a branching node but start again at a joining node. However, it is reasonable to limit their joining times to one, especially within an autonomous domain [10]. Accordingly, let $p_i \Diamond p_j$ denote the (unique) joining node between p_i and p_j, and let $\mathcal{R}$ be the set of joining nodes in Γ. We say $p_i \Diamond p_j \in \varnothing$ when these two paths never rejoin, and denote their shared section by $p_i \wedge^0 p_j$. While $p_i \Diamond p_j \in \mathcal{R}$, they will get another shared section which starts from r and additionally, we denote it by $p_i \wedge^1 p_j$. Let $p_i \wedge p_j$ be the entire shared path. Then, we have $p_i \wedge p_j = p_i \wedge^0 p_j \bigcup p_i \wedge^1 p_j$, where $p_i \wedge^1 p_j = \varnothing$ if $p_i \Diamond p_j \in \varnothing$.

When a node v or a link l appears on path p_i, we say $v \in p_i$ or $l \in p_i$. We denote the length measure by ζ, e.g., $\zeta(p_i)$ is the length of p_i and $\zeta(l)$ represents the length of l. Then according

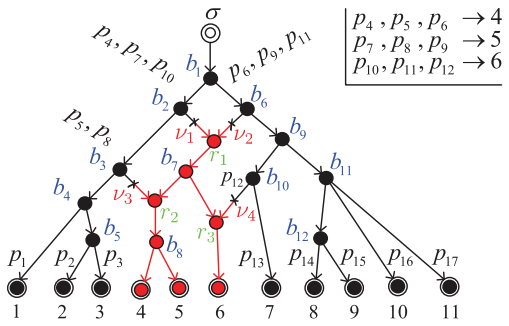

Fig. 1. The non-tree topology Γ composed of 17 end-to-end paths. Destination 4, 5 and 6 each get three active paths from σ, while there is a unique path from σ to each of the rest destinations. E.g., p_{17} is the unique path to destination 11 while p_4 (traversing b_1, b_2, r_1, b_7, r_2 and b_8), p_5 and p_6 are all destined for destination 4. We have $p_4 \Diamond p_5 = r_2$, $p_4 \Diamond p_6 = r_1$ and $p_5 \Diamond p_6 = r_2$. The set of joining nodes is $\mathcal{R} = \{r_1, r_2, r_3\}$. To decompose Γ, we cut $\mathbb{P} = \{p_4, p_5, \ldots, p_{12}\}$ at ν_1, ν_2, ν_3, and ν_4, and obtain Γ^{up} and Γ^{dn} (in red links), which are composed of $\{p_1, p_2, p_3, p_{13}, \ldots, p_{17}\} \bigcup \{p_4^{up}, \ldots, p_{12}^{up}\}$ and $\{p_4^{dn}, \ldots, p_{12}^{dn}\}$, respectively.

to [1], the linear additive model is used to express the length relationships between the paths and their constituent links,

$$\zeta(p_i) = \sum_{l \in p_i} \zeta(l). \tag{1}$$

In addition, we denote a path cut for p_i by $\mathcal{C}_{p_i} = \langle v, n \rangle$, where $v \in p_i$ and $n \in p_i$ are the two nodes connected by $l \in p_i$. Via cutting $l \in p_i$, p_i can be divided into two subpaths. Let $v \in p_i$ denote the break point. Then the subpath from σ to v is denoted by p_i^{up} and the other one from v is set by p_i^{dn}. E.g., in Fig. 1, $p_4 \to 4$ with $\mathcal{C}_{p_4} = \langle b_2, r_1 \rangle$ is cut into p_4^{up} and p_4^{dn} at ν_1. Obviously, we have $\zeta(p_4) = \zeta(p_4^{up}) + \zeta(p_4^{dn})$.

III. Topology Decomposition and Identifiability

The tree topology and the multi-source topology are shown to be identifiable using the lengths of the end-to-end paths and their shared paths [7], [12]. Their topology identifiability lies in that the location information of either the branching nodes or the joining nodes can be well revealed when comparing the related shared path lengths to the end-to-end path lengths. In these topologies, any two paths from the same source construct an inverted "Y" structure, while they form a "Y" when they are from different sources but destined for the identical destination node. According to [14], here we formally present the topology definition for the single-path routing networks.

Definition 1: Given a logical topology, we call it a single-path routing topology if and only if both following conditions hold: (C1) There exists only one path between any pair of end-hosts; (C2) Any two end-to-end paths never meet after they split. Moreover, if it is single-source, it is a tree topology; Otherwise, it is a multi-source topology. Any single-path routing topology is identifiable.

Given the single-source topology Γ, we let $\mathbb{P} \subseteq \mathcal{P}$ be the set of end-to-end paths that we have cut. And denote Γ^{up} and Γ^{dn} be the two sub-topologies composed of $(\mathcal{P} \setminus \mathbb{P}) \bigcup \{p_i^{up} | p_i \in \mathbb{P}\}$ and $\{p_i^{dn} | p_i \in \mathbb{P}\}$, respectively (e.g., Fig. 1). Then according to Definition 1, we need to carefully select $\{\mathcal{C}_{p_i} | p_i \in \mathbb{P}\}$ in order to ensure that either in Γ^{up} or Γ^{dn}, every two end-to-end paths get a consecutive shared path.

A. Decompose Topology Using Path Cuts

It is noted that $\forall p_s \to D$, there is no joining node on p_s, i.e., $\forall p_i \neq p_s$, (p_s, p_i) constructs an inverted "Y" structure and $p_s \wedge p_i = p_s \wedge^0 p_i$. Then based on Definition 1, we do not need to cut p_s and can leave $\{p_s \to D\}$ all in Γ^{up}. However, $\forall p_i \to d$ $(d \in \mathbb{D})$, there always exists a joining node on p_i, as we can find $p_i' \to d$ such that $p_i \Diamond p_i' \notin \varnothing$. Nevertheless, a joining node means a nonconsecutive shared path and the related paths need to be cut. Therefore, we have $\mathbb{P} = \{p_i \to \mathbb{D}\}$.

Moreover, $\forall p_i, p_j \in \mathbb{P}$ $(p_i \Diamond p_j \notin \varnothing)$, there must not exist $r \in \mathcal{R}$ such that $r \in p_i \wedge^0 p_j$. Otherwise, we can find $p_k \to \mathbb{D}$ which can join p_i or p_j for more than one time. Let $r^\star \in p_i$ be the nearest joining node to σ on p_i. We have,

Lemma 1: $\forall p_i \in \mathbb{P}$, the path cut is set by $\mathcal{C}_{p_i}^\star = \langle b, r^\star \rangle$. Then $\forall p_i, p_j \in \mathbb{P}$, if $p_i \Diamond p_j \notin \varnothing$, we have $p_i^{up} \wedge p_j = p_i^{up} \wedge p_j^{up} = p_i \wedge^0 p_j$ and $p_i^{dn} \wedge p_j = p_i^{dn} \wedge p_j^{dn} = p_i \wedge^1 p_j$.

In Fig. 1, we get the path cut $\mathcal{C}_{p_5}^\star = \langle b_3, r_2 \rangle$ and $\mathcal{C}_{p_7}^\star = \langle b_2, r_1 \rangle$. We have $p_5^{up} \wedge p_7 = p_5^{up} \wedge p_7^{up} = p_5 \wedge^0 p_7$. Similarly, $p_5^{up} \wedge p_4 = p_5^{up} \wedge p_4^{up} = p_5 \wedge^0 p_4$. It is clear that if $p_i \Diamond p_j \in \varnothing$, (p_i, p_j) forms an inverted "Y" structure and $p_i \wedge^0 p_j = p_i \wedge p_j$. Further when $\nexists r \in \mathcal{R}$ such that $r \in p_i \wedge^0 p_j$, we will get $p_i^{up} \wedge p_j^{up} = p_i \wedge p_j$ and $\mathcal{C}_{p_i}^\star \neq \mathcal{C}_{p_j}^\star$. E.g., (p_4, p_{12}) forms the inverted "Y" structure and gets no joining node on their shared path. We have $p_4^{up} \wedge p_{12}^{up} = p_4 \wedge p_{12}$ and $\mathcal{C}_{p_4}^\star = \langle b_2, r_1 \rangle \neq \mathcal{C}_{p_{12}}^\star = \langle b_{10}, r_3 \rangle$. However, when $\exists r \in \mathcal{R}$ such that $r \in p_i \wedge^0 p_j$ (e.g., $r_1 \in p_4 \wedge^0 p_7$), then,

Lemma 2: $\forall p_i, p_j \in \mathbb{P}$, we have $\mathcal{C}_{p_i}^\star = \mathcal{C}_{p_j}^\star$ and $p_i^{up} = p_j^{up}$ if and only if $\exists r \in \mathcal{R}$ such that $r \in p_i \wedge^0 p_j$.

Since a path is cut before the joining node that is nearest to σ, any two paths will share the same cut as long as they never split before they traverse a joining node. And the same cut indicates the same break point. E.g., p_4 and p_7 share the cut $\langle b_2, r_1 \rangle$ and are both cut at ν_1. p_4^{up} and p_7^{up} follow the same path to ν_1, exactly where p_4^{dn} and p_7^{dn} start from. Obviously, in Fig. 1, both $p_4^{up} \wedge p_7^{up}$ and $p_4^{dn} \wedge p_7^{dn}$ remain consecutive, although their lengths depend on where we insert ν_1 in the link connecting b_2 to r_1. Also it is easy to find out that both Γ^{up} and Γ^{dn} meet Definition 1. Consequently, we have

Theorem 1: Given Γ, it can be divided into a tree topology and a multi-source topology. I.e., Γ^{up} is a tree topology while Γ^{dn} belongs to a multi-source topology.

The related proofs are somewhat straightforward and will be omitted regarding the space limitations.

B. Identify Topology from Sub-topologies

Theorem 1 indicates that Γ can be reconstructed via connecting the related paths in $\mathbb{P}$ once the two sub-topologies (Γ^{up} and Γ^{dn}) are identified. To this end, there could be two tasks: (i) One is how to determine the sets of end-to-end paths for the two sub-topologies. More specifically, how to detect the paths that share the same cut. As demonstrated by Lemma 2, paths with the same cut (e.g., p_4 and p_7 in Fig. 1) will be the same end-to-end path in Γ^{up}, and they will be further regarded as having the same source node in Γ^{dn}; (ii) According to [7] and [12], the other one is how to conduct end-to-end probing for the shared path length measurement $\{\zeta(p_i \wedge^0 p_k) | p_i \to D\} \bigcup \{\zeta(p_j \wedge^0 p_k) | p_j, p_k \to \mathbb{D};\ \mathcal{C}_{p_j}^\star \neq \mathcal{C}_{p_k}^\star\}$

and $\{\zeta(p_j^{dn})|p_j \to \mathbb{D}\} \bigcup \{\zeta(p_j^{dn} \wedge p_k^{dn})|\mathcal{C}^{\star}_{p_j} = \mathcal{C}^{\star}_{p_k}\} \bigcup \{\zeta(p_j \wedge^1 p'_j)|p_j, p'_j \to d \in \mathbb{D}\}$ to identify Γ^{up} and Γ^{dn}, respectively.

It is known that through the end-to-end probing such as the "back-to-back" probing, $p_i \wedge^0 p_k$ (and $p_j \wedge^0 p_k$) can be effectively measured. Without loss of generality, for $p_j \to d \in \mathbb{D}$, we set $\zeta(p_j^{up}) = \max\{\zeta(p_j \wedge^0 p_k), \zeta(p_j \wedge^0 p'_j)|p_k \to D; p'_j \to d\}$, which means the break point is right inserted besides the branching node of the path cut. E.g., in Fig. 1, we will obtain $\zeta(p_4^{up}) = \zeta(p_1 \wedge^0 p_4)$ and $\zeta(p_4^{dn}) = \zeta(p_4) - \zeta(p_4^{up})$ accordingly . Then for Γ^{dn}, we get $\zeta(p_j^{dn} \wedge p_k^{dn}) = \zeta(p_j \wedge^0 p_k) - \zeta(p_j^{up})$ when $\mathcal{C}^{\star}_{p_j} = \mathcal{C}^{\star}_{p_k}$. However, it is quite challenging to measure $\zeta(p_j \wedge^1 p'_j)$ directly. Although the loss-based metric [14] might seem to be a promising solution for $\zeta(p_j \wedge p'_j)$, it is generally not compatible with the "back-to-back" probing which uses the delay-based metric for $\zeta(p_j \wedge^0 p'_j)$.

Thus in this letter, we only focus our interest on the task (i) above. For $p_i \to d_1 \in \mathbb{D}$, $p_j \to d_2 \in \mathbb{D}$, if $\mathcal{C}^{\star}_{p_i} = \mathcal{C}^{\star}_{p_j}$, we will have $r \in \mathcal{R}$ and $r \in p_i \wedge^0 p_j$ according to Lemma 2. Such joining node on the shared path generally indicates that we can always find $p'_i \to d_1$ and $p'_j \to d_2$ such that $p'_i \Diamond p_i = p'_j \Diamond p_j = r$ and $\mathcal{C}^{\star}_{p'_i} = \mathcal{C}^{\star}_{p'_j}$. Referring the structure of (p_i, p_j; p'_i, p'_j) as the "rocket" depicted in Fig. 2(b), we actually need to detect such rocket structures in Γ instead. E.g., in Fig. 1, (p_4, p_7; p_6, p_9) is a rocket. Then p_4 and p_7 (resp. p_6 and p_9) are shown cut-sharing. Just like p_4 and p_7, p_6 and p_9 also appear in Γ^{up} as a single path and later enter Γ^{dn} from their break point (v_2). Let `rocket`(p_i, p_j; p'_i, p'_j) refer to the related detection procedure of rocket structures, and let $\mathbb{C}$ be the classification result of cut-sharing paths. We present the following algorithm to identify cut-sharing paths.

Algorithm 1. Identify Cut-Sharing Paths by Detecting Rockets

Input: $\mathbb{D}, \mathbb{P}$
  $\forall p_k \in \mathbb{P}$ and $n \leqslant |\mathbb{P}|$, initial $\mathbb{C}(v_n) \Leftarrow \{p_k\}$;
  for each $d_1, d_2 \in \mathbb{D}$ **do**
    for each $p_i, p'_i \to d_1$; $p_j, p'_j \to d_2$ **do**
      if `rocket`(p_i, p_j; p'_i, p'_j) is true **then**
        find $\mathbb{C}(v_{n1}) \ni p_i$ and $\mathbb{C}(v_{n2}) \ni p_j$;
        if $n1 \neq n2$ **then**
          $\mathbb{C}(v_{\min\{n1,n2\}}) \Leftarrow \mathbb{C}(v_{n1}) \bigcup \mathbb{C}(v_{n2})$;
          $\mathbb{C}(v_{\max\{n1,n2\}}) \Leftarrow \varnothing$;
        end if
        find $\mathbb{C}(v_{n3}) \ni p'_i$ and $\mathbb{C}(v_{n4}) \ni p'_j$;
        if $n3 \neq n4$ **then**
          $\mathbb{C}(v_{\min\{n3,n4\}}) \Leftarrow \mathbb{C}(v_{n3}) \bigcup \mathbb{C}(v_{n4})$;
          $\mathbb{C}(v_{\max\{n3,n4\}}) \Leftarrow \varnothing$;
        end if
      end if
    end for
  end for
Output: $\mathbb{C}$

C. Detection of Rocket Structures

"`rocket`(p_i, p_j; p'_i, p'_j)" is implemented in this section. We assume the cooperation of end-hosts during measurement [1] and require internal nodes do not reorder packets that traverse them. As in Fig. 2(a), if the packet on p'_i arrives at the joining node r earlier than the packet on p_i, it will also leave r to d_1 earlier. The arrival order of the two packets at d_1 then will be exactly the same with their arrival order at r. Moreover, for the rocket in Fig. 2(b), when we send two closely spaced packets on p_i and p'_i, and two extra closely spaced packets for p_j and p'_j, the packet arrival order for p_i and p'_i at d_1 will be the same with the packet arrival order for p_j and p'_j at d_2.

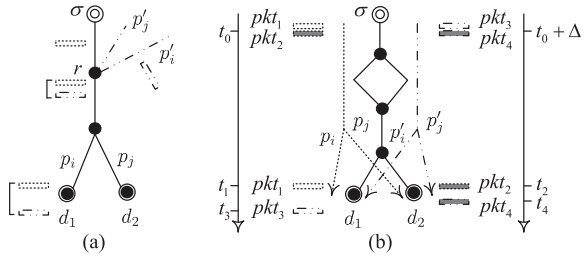

Fig. 2. The packet arrival order and the "rocket" structure.

Denote the four individual packets to p_i, p_j, p'_i and p'_j by pkt_1, pkt_2, pkt_3 and pkt_4. As depicted in Fig. 2(b), pkt_1 and pkt_2 depart at t_0 while pkt_3 and pkt_4 are sent at $t_0 + \Delta$ ($\Delta \in [-\phi, \phi]$). Accordingly, their arrival times are t_1, t_2, t_3 and t_4. Let $\Pr(t_1 < t_3)$ be the probability that pkt_1 reaches d_1 earlier than pkt_3. Similarly, $\Pr(t_2 < t_4)$ is the probability that pkt_2 arrives at d_2 before pkt_4. We denote $\pi = 1$ if at d_1 and d_2 we observe the same arrival order, i.e., $t_1 < t_3$ & $t_2 < t_4$, or $t_1 > t_3$ & $t_2 > t_4$. Otherwise, $\pi = 0$. Therefore,

$$\begin{aligned}\Pr(\pi = 1) = {} & \Pr(t_1 < t_3)\Pr(t_2 < t_4|t_1 < t_3) \\ & + \Pr(t_1 > t_3)\Pr(t_2 > t_4|t_1 > t_3). \end{aligned} \tag{2}$$

When p_i and p_j take the same route from σ to r, it is believed that pkt_1 and pkt_2 will retain closely spaced as they arrive at r. If (p_i, p_j; p'_i, p'_j) constructs a rocket, pkt_3 and pkt_4 will also remain close to each other at r since the path from σ to r for p'_i and p'_j are identical, too. Consequently, both $\Pr(t_2 < t_4|t_1 < t_3)$ and $\Pr(t_2 > t_4|t_1 > t_3)$ should be equal to 1. While (p_i, p_j; p'_i, p'_j) is a non-rocket, it either implies that p'_i and p'_j (or p_i and p_j) can take different routes from σ to $p_i \Diamond p'_i$, or simply $p_i \Diamond p'_i \neq p_j \Diamond p'_j$, making both $\Pr(t_2 < t_4|t_1 < t_3)$ and $\Pr(t_2 > t_4|t_1 > t_3)$ smaller than 1. Consequently we have,

$$\begin{aligned}\Pr(\pi = 1|\text{rocket}) &= \Pr(t_1 < t_3) * 1 + \Pr(t_1 > t_3) * 1 \\ &= 1 > \Pr(\pi = 1|\text{non-rocket}). \end{aligned} \tag{3}$$

Let τ be the threshold to identify the rocket from the non-rocket. We have $\tau = 1$ in theory according to (3). However, there will always be some measurement noise during the back-to-back probing, making τ smaller than 1 in practice.

IV. Simulation and Evaluation

A. Simulation Setup

We evaluate with ns-2 simulations. We employ Fig. 1 as our simulated network and assume the even load balancing [10] is deployed. In the simulation network, the delays for the internal links range within $[20ms, 40ms]$ while their bandwidths are

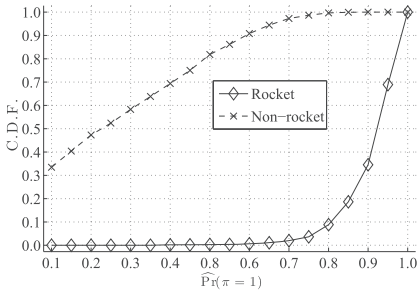

Fig. 3. C.D.F.s of the empirical $\Pr(\pi = 1)$ when $\mathcal{K} = 150$.

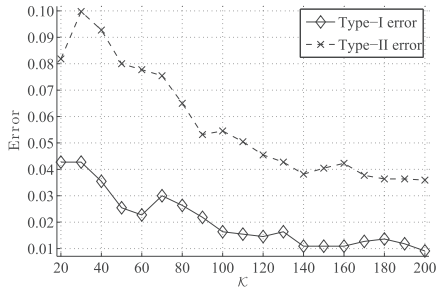

Fig. 4. Errors of Type-I and Type-II when $\tau = 0.70$.

set between $10Mbps$ and $20Mbps$. The edge links have a fixed delay of $5ms$ and a a fixed bandwidth of $5Mbps$. The cross traffic is composed of Pareto on/off flows and FTP flows and gets the network moderately loaded [6] with link bandwidth utilization and link loss rate ranging in $(35\%, 75\%)$ and $[0, 0.1\%)$, respectively. For every $(p_i, p_j; p_i', p_j')$, we employ the back-to-back probes (using UDP, 46 bytes each packet) to probe both (p_i, p_j) and (p_i', p_j') (as illustrated in Fig. 2(b)). Δ is uniformly selected from $[-\phi, \phi]$, where ϕ is set by the immanent delay difference between p_i and p_j. A measurement of π is considered as validated when no sent packet is dropped during the probing snapshot. We probe every other $100ms$ until we receive $\mathcal{K}$ validated measurements for π. The simulation is repeated for 100 times.

In Fig. 1, there are 27 different combinations for $(p_i, p_j; p_i', p_j')$, among which there are 5 rockets, i.e., $(p_5, p_8; p_4, p_7)$, $(p_5, p_8; p_6, p_9)$, $(p_4, p_7; p_6, p_9)$, $(p_4, p_{10}; p_6, p_{11})$ and $(p_7, p_{10}; p_9, p_{11})$. In each simulation, $(p_i, p_j; p_i', p_j')$ is said to be a rocket only if its empirical $\widehat{\Pr}(\pi = 1) = \frac{\sum^{\mathcal{K}} \pi = 1}{\mathcal{K}}$ is greater than a given threshold τ. We get a "Type-I" error (false negative) when $(p_i, p_j; p_i', p_j')$ truly is a rocket but we identify it as a non-rocket. While we receive a "Type-II" error (false positive) if $(p_i, p_j; p_i', p_j')$ is identified as a rocket but actually it is not.

B. Performance Evaluation

As the cut-sharing paths can be identified directly from rockets according to Algorithm 1, we here solely focus on the performance of our proposed scheme of rocket detection. We present the Cumulative Distribution Function (C.D.F.) of the empirical $\Pr(\pi = 1)$ in Fig. 3. The rocket and the non-rocket are depicted by the solid line with the diamond marker and the dashed line with the cross marker, respectively. As shown in Fig. 3, when we set $\mathcal{K} = 150$, almost 95% of $\widehat{\Pr}(\pi = 1)$ from the non-rocket is smaller than 0.70, while it is only about 5% for the rocket. Such a distinct difference not only validates the efficiency of using the arrival order measurements to detect the rocket, but also suggests a threshold of $\tau = 0.70$.

We investigate the errors in Fig. 4, where the Type-I error is shown in the dashed line and the Type-II error is drawn by the dashed line. It is not surprising that both the Type-I error and the Type-II error drop as $\mathcal{K}$ increases. Clearly, a larger $\mathcal{K}$ basically results in a more accurate $\widehat{\Pr}(\pi = 1)$, which favors preciser identifications of both the rocket and the non-rocket.

V. Conclusion and Future Works

Multipath routing introduces a non-tree topology to the single-source network. By showing the non-tree topology is still identifiable via decomposing it into two sub-topologies: a tree topology and a multi-source topology, a tomographic scheme of detecting the rocket structures is proposed for such topology decomposition. Simulation results demonstrate that our scheme is able to achieve a detection rate of around 95% with only about 150 packets per path. Future attention will be paid to modeling $\Pr(\pi = 1)$ with real-world parameters and to releasing the limitation of the joining times between paths.

References

[1] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, "Network tomography: Recent developments," *Stat. Sci.*, vol. 19, no. 3, pp. 499–517, 2004.

[2] X. Zhang and C. Phillips, "A survey on selective routing topology inference through active probing," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 1129–1141, Oct. 2012.

[3] N. G. Duffield, J. Horowitz, F. L. Presti, and D. Towsley, "Multicast topology inference from end-to-end measurements," *Adv. Perform. Anal.*, vol. 3, pp. 207–226, 2000.

[4] M. Coates, R. Castro, R. Nowak, M. Gadhiok, R. King, and Y. Tsang, "Maximum likelihood network topology identification from edge-based unicast measurements," in *Proc. ACM SIGMETRICS Perform. Eval. Rev.*, 2002, vol. 30, pp. 11–20.

[5] Y. Tsang, M. Yildiz, P. Barford, and R. Nowak, "Network radar: Tomography from round trip time measurements," in *Proc. 4th ACM SIGCOMM Conf. Internet Meas.*, 2004, pp. 175–180.

[6] M. F. Shih and A. O. Hero, "Hierarchical inference of unicast network topologies based on end-to-end measurements," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1708–1718, May 2007.

[7] J. Ni and S. Tatikonda, "Network tomography based on additive metrics," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 7798–7809, Dec. 2011.

[8] H. Nguyen and R. Zheng, "A binary independent component analysis approach to tree topology inference," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3071–3080, Jun. 2013.

[9] R. Zhang, Y. Li, and X. Li, "Topology inference with network tomography based on t-test," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 921–924, Jun. 2014.

[10] B. Augustin, T. Friedman, and R. Teixeira, "Measuring multipath routing in the Internet," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 830–840, Jun. 2011.

[11] N. G. Duffield and F. L. Presti, "Network tomography from measured end-to-end delay covariance," *IEEE/ACM Trans. Netw.*, vol. 12, no. 6, pp. 978–992, Dec. 2004.

[12] M. G. Rabbat, M. J. Coates, and R. D. Nowak, "Multiple-source Internet tomography," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 12, pp. 2221–2234, Dec. 2006.

[13] S. Pan, Z. Zhang, F. Yu, and G. Hu, "End-to-end measurements for network tomography under multipath routing," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 881–884, May 2014.

[14] H. X. Nguyen and P. Thiran, "Network loss inference with second order statistics of end-to-end flows," in *Proc. 7th ACM SIGCOMM Conf. Internet Meas.*, 2007, pp. 227–240.