Computer Communications 34 (2011) 658–669

Contents lists available at ScienceDirect

Computer Communications

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

Palmtree: An IP alias resolution algorithm with linear probing complexity

M. Engin Tozal *, Kamil Sarac

The University of Texas at Dallas, Department of Computer Science Richardson, TX 75080, USA

ARTICLE INFO

Article history:
Available online 22 June 2010

Keywords:
Internet
Topology
Map
Alias
Router

ABSTRACT

Internet topology mapping studies utilize large scale topology maps to analyze various characteristics of the Internet. IP alias resolution, the task of mapping IP addresses to their corresponding routers, is an important task in building such topology maps. In this paper, we present a new probe-based IP alias resolution tool called `palmtree`. `Palmtree` can be used to complement the existing schemes in improving the overall success of alias resolution process during topology map construction. In addition, `palmtree` incurs a linear probing overhead to identify IP aliases. The experimental results obtained over Internet2 and GEANT networks as well as four major Internet Service Providers (ISPs) present quite promising results on the utility of `palmtree` in obtaining more accurate network topology maps.

1. Introduction

Internet consists of many networks operated by different organizations that do not necessarily publish their complete network topology maps. On the other hand, having an accurate topology map of the network helps us to understand its structural and operational characteristics such as reliability, connectivity, robustness, and efficiency; enhance current and future protocols; optimize networking structure; and improve synthetic network graph generators and other simulation tools.

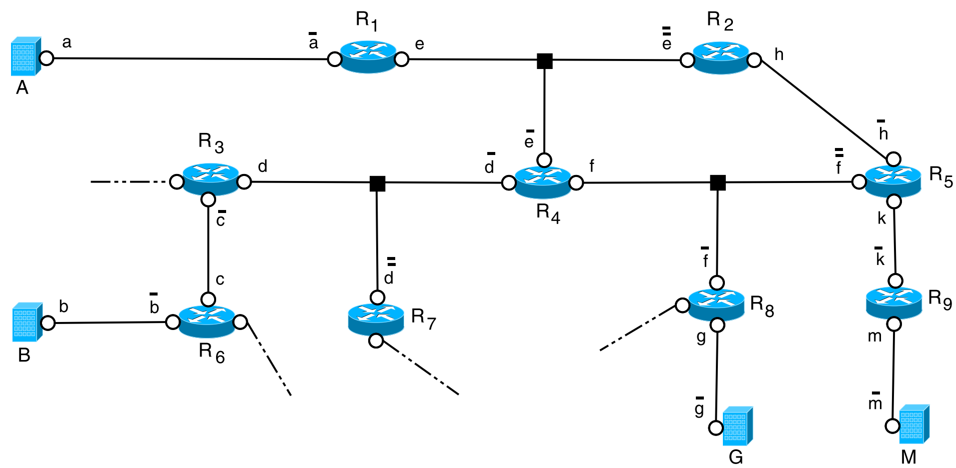
Many successful research projects and efforts attempting to derive an accurate and large scale topology map of the Internet have been introduced in the literature [14,16,20,13]. These efforts focus on different but correlated topology maps: *IP level* maps show IP addresses that are in use on the Internet; *router level* maps group the interfaces hosted by the same router into single units; *subnet level* maps enrich the router level maps with subnet level connectivity info; and *AS level* maps demonstrate the adjacency relationship between Autonomous Systems (ASes).

Obtaining router level topology maps in the Internet includes two steps: (1) collecting topology data from the target network and (2) processing this data to build a representative map corresponding to the target network. The common practice in data collection is to use the well-known `traceroute` [10,1] tool to collect path traces crossing over the target network. These traces are run from a set of vantage points towards a set of destinations such that each trace crosses over the target network and returns a partial topology of the underlying target network. Processing the collected data involves several tasks including subnet inference [7,18], star elimination [8], and IP alias resolution [4,19,5,3,18]. *IP alias resolution*, the task of identifying IP addresses that are accommodated by the same router in the target network, is the main focus of the work presented in this paper. IP alias resolution is an important task and inaccuracies in this task may significantly affect the accuracy of the resulting map with respect to the actual topology of the target network [6].

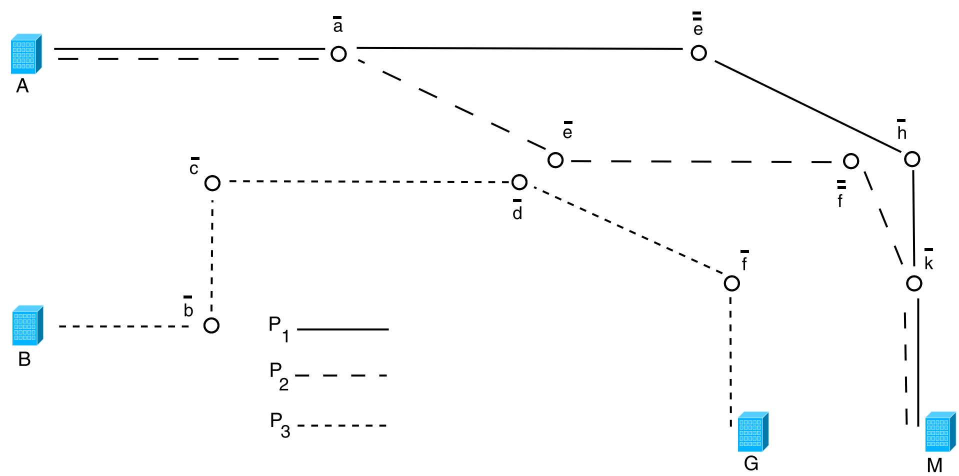
As an example, consider the use of `traceroute` in designing resilient overlay network systems where the goal is to use the obtained traces to identify node disjoint overlay paths. Fig. 1a shows the physical topology of a network that includes several routers as well as point-to-point and multi-access links. In the figure routers and hosts are represented with ovals and boxes, respectively and they are labeled with upper-case letters. Interfaces are shown with small circles attached to routers and named with lower-case letters. Additionally, we used bars (e.g., $\bar{i}$) to keep the picture lucid and to discretely label the interfaces (or IP addresses) being hosted on the same subnet instead of using more lower-case letters. Assume that our goal is to identify node disjoint paths between A and M and between B and G in this network. To keep it simple, let us say that the routers in the figure are responsive to `traceroute` probe packets and they report back the IP address of the incoming interface through which the packet has reached to the router. Fig. 1b shows the network topology collected by `traceroute` without using IP alias resolution where $P_1 = \{A, \bar{a}, \bar{\bar{e}}, \bar{h}, \bar{k}, M\}$ and $P_2 = \left\{A, \bar{a}, \bar{e}, \bar{\bar{f}}, \bar{k}, M\right\}$ are the two shortest paths between A and M and $P_3 = \{B, \bar{b}, \bar{c}, \bar{d}, \bar{f}, G\}$ is the only shortest path between B and G without IP alias resolution. Based on this topology map one would infer that the use of P_2 for A to M path along with the use of P_3 for B to G path would satisfy the node disjointness requirement. Yet, this would be an inaccurate conclusion as the second IP address ($\bar{e}$) of

* Corresponding author. Tel.: +1 4692358136.
E-mail addresses: engintozal@utdallas.edu (M. Engin Tozal), ksarac@utdallas.edu (K. Sarac).

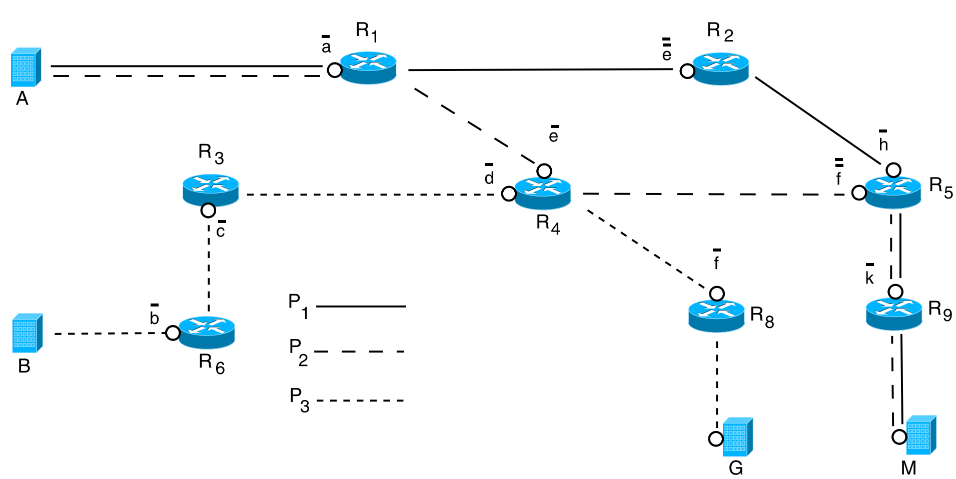
0140-3664/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2010.06.017



(a) A Network topology section between hosts A, B, G, and M with equally weighted links



(b) Traceroute view of the paths without IP alias resolution. Acquired two shortest paths between hosts A and M are $P_1 = \{A, \bar{a}, \bar{\bar{e}}, h, \bar{k}, M\}$ and $P_2 = \{A, \bar{a}, \bar{e}, \bar{\bar{f}}, \bar{k}, M\}$, respectively and a single shortest path between hosts B and G is $P_3 = \{B, \bar{b}, \bar{c}, \bar{d}, \bar{f}, G\}$



(c) Traceroute view of the same paths P_1, P_2, and P_3 with IP alias resolution

Fig. 1. *IP Alias Resolution Motivating Example.* An example network segment among hosts *A*, *B*, *G*, and *M* and traceroute views of the two paths between *A* and *M* as well as a single path between *B* and *G* with and without IP alias resolution, respectively.

path P_2 and the third IP address ($\bar{d}$) of path P_3 belong to the same router R_4. However, without IP alias resolution this fact is hidden in the topology map in Fig. 1b. As a result, these two paths are not really node disjoint. On the other hand, Fig. 1c showing the