

A Solution for Anonymous Routers Discovery Based on Source-Routing Traceroute

Jiang Du and Yu Li

Department of Computer Science and Technology
Chongqing University of Posts and Telecommunications
Chongqing, China
{clouddu & lyris.joy}@gmail.com

Abstract—Discovery of computer network's topology is always an important basement for network management. As IPv6 becomes the official standard, topology discovery methods should be changed as well. In IPv6 network, Anonymous Routers (AR) is one problem that we must deal with. AR is one kind of router we don't know its interface information, which makes it very difficult to find out their interconnections through common ways. This paper introduces a solution to deal with it, using source-routing traceroute command and node's degree to merge ARs or find out their relation, and finally get the topology graph.

Keywords- computer network; topology discovery; AR; source-routing traceroute

I. INTRODUCTION

Computer network topology reflects the real connections between network devices and has great significance for network management. Efficient and accurate topology algorithm is a research topic that people always working on. Anonymous Routers (AR) is one kind of routers hiding its IP addresses and other information in topology discovery which exists in IPv4 as well. If a router is configured not to respond the ICMP probe packages, or just cannot respond in time because of network congestion, it becomes an AR. In IPv6 networks router's interfaces may have configured no global unicast addresses that make the router an AR.

The existing solutions for AR are ignoring them[1] or treating every AR as an independent device [2], or merging continuous ARs as one real router [3]. A proper discovery method can show people the network topology accurately. But there is still no means can solve this problem very well. In this paper we use routing information collected from single-source detection and analyze the topology characters caused by AR. We improve the topology discovery methods for AR networks based on source-routing traceroute and use nodes' degree as a help to make the discovery more accurate.

Rest of the paper is organized as follows. Section II briefly introduced the existing topology discover methods for AR-networks as well as their advantages and disadvantages. Section III describes theories and network pattern we use in this paper. Section IV describes our methods in detail which is based on source routing traceroute. And finally the conclusion in section V.

II. RELETED WORK

In [4], Qian-fang Xu mentioned that source-routing traceroute can be used to get information of any two nodes in the network and then discover two or more ARs which are connected directly. But the author didn't give the detail algorithm. In paper [5], algorithm proposed by Huai-zhou Shi et al is to treat every address that responses with "timeout" as an anonymous element. If two anonymous elements' TTL' difference is less than 3, they should be one router in real network. This method is wide range applicable and high-efficiency. But the algorithm use only TTL as condition which is very easy to misjudge when there's more anonymous elements in a network. Gunes M. et al handle this problem using graph theory in their paper [6]. They first analyze the network topology that contains AR and then depart the abundant parts into kinds of math graphs for further process. For a typical AR network structure, the discovery accuracy is very well but the calculation is much too complex. And if there are two or more ARs directly interconnected with each other, this algorithm can't judge the connection correctly. In [7], Zhang Yong et al combined ARs while they have a same hop length by traceroute. In fact there are ARs having same hop but independent from each other, which may cause some mistakes.

Considering the accuracy and efficiency of network topology discovery, source-routing traceroute is still a proper method. Our algorithm is based on it and the detail will be given in the later section.

III. BACKGROUND AND SYSTEM MODEL

Traceroute can trace the route (path) that packets across IP network while source routing traceroute allows people to specify some nodes on the packets' route and trace the special path we need. According to the path we can get some information of this network and discover its topology.

A. Network Model

Our algorithm is based on the result of single-source traceroute and assuming that every router supports source routing and every traceroute packet passed through a same path each time, and all the connection has the same bandwidth. Here we ignore all hosts and remain only routers to make our algorithm be easy to describe. All other non-anonymous nodes have been merged or been discovered. The network model is as Fig. 1 shows, in which black ones represent the anonymous routers.

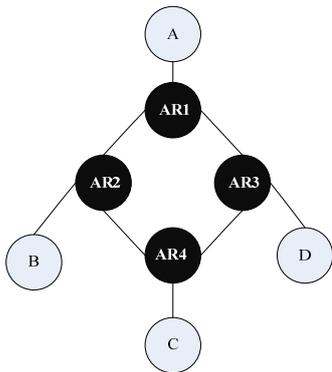


Figure 1. Network Model with anonymous routers

The result of single-source traceroute is a tree graph as Fig. 2 shows, in which the root and leaves are non-anonymous routers and other nodes are ARs. Our work is to find out which ARs belongs to a router in real network, and which two ARs are connected directly.

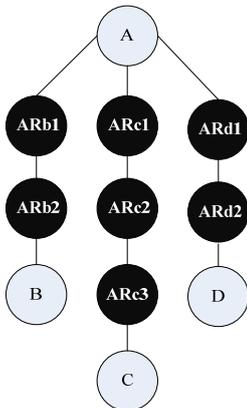


Figure 2. Topology after single source traceroute

Let R_{ij} be the non-anonymous router and AR_{ij} be anonymous routers, both of which are not the source node. i presents router R_i , the end node of one path. And j presents the hop number from source node. For instance, path A-B can be expressed as A-ARb1-ARb2-Rb3. Then the paths can be represent as TABLE I

Name	Path
A-Rb	A→ARb1→ARb2→Rb3
A-Rc	A→ARc1→ARc2→ARc3→Rc4
A-Rd	A→ARd1→ARd2→Rd3

The purpose of our algorithm is to find out relations between all the AR_{ij} . Using single-source traceroute and source routing traceroute it is easy to get hop length between any two leaf nodes, based on what we calculate the intersection or connection of two paths and finish topology discovery. And as mentioned before, nodes' degree is used to help finding out the ARs that is directly connected to a non-anonymous router.

B. Lemmas for Topology Discovery

For convenience of description, we use H_i to present the hop number from source to R_i (the path length). H_{ji} stands for the path length from source to R_i through R_j . As the routing strategy always makes the routing path a best way, there should be $H_i < H_{ji}$. Let $s = H_{ji} - H_i = 2n + k$, ($k = \{0, 1\}$, $n = \{0, 1, 2, \dots, n\}$). Obviously when s is an odd, $k = 1$ and $n = (H_{ji} - H_i - 1) / 2$. On the contrary, $k = 0$ and $n = (H_{ji} - H_i) / 2$. Let $z = H_i - n$.

When $k = 1$, it means that except for the twice additional hop from source routing address, there is one more hop between the two paths. And there must have no intersection. From a viewpoint of mathematic, it may be some kind of connection like $AR_i(z-1)$ connects to $AR_j(z+1)$ or $AR_i(z+2)$ connecting with $AR_j(z-2)$ and so on. But according to the routing strategy, traceroute packet should route through the shortest way. As shown in Fig. 3, if AR_{b1} is connected to AR_{c3} , the route path from A to C may change as A- AR_{b1} - AR_{c3} -Rc. Therefore situations mentioned above cannot appear in real network. So we think AR_{iz} and AR_{jz} are directly connected when $k = 1$ (Fig. 4).

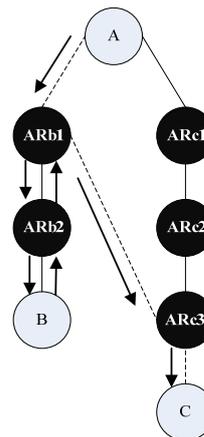


Figure 3. Wrong connection that may conflict with routing strategy

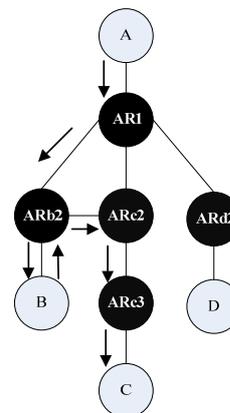


Figure 4. The connection between paths while $k = 1$

When $k = 0$, it means the additional hop caused by source routing traceroute is even, there are two cases at this time:

1) *There is an intersection in the two paths*

It means there is an AR in each path belongs to a router in real network. The extra path length ($H_{ji}-H_i$) caused by source routing is twice of the length from R_j to the intersection. As shown in Fig. 5, $H_{bc}-H_c=2$.

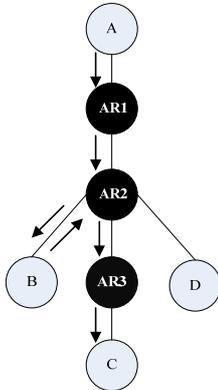


Figure 5. Path A-B and A-C has an intersection while $k=0$

2) *There is a connection between the two paths*

Assume that the two interconnected routers are AR_{xi} and AR_{yj} . Because of connection there is one additional hop for source routing traceroute. And as $k=0$, $|i-j|\%2=1$. According to routing strategy, $|i-j|<2$, or there must be a better way different from previous routing path. Therefore $|i-j|=1$, AR_{xi} is connected to $AR_{y(i+1)}$, or AR_{xi} connects with $AR_{y(i-1)}$, as shown in Fig. 6, also $H_{bc}-H_c=2$.

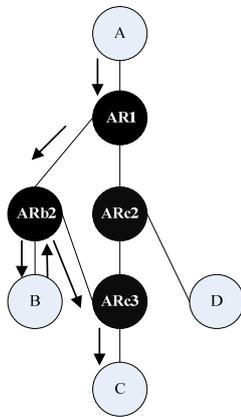


Figure 6. Connection between paths while $k=0$

Both situations above exist in real network. Considering hardware configuration, our algorithm firstly makes ARs an intersection and if the merge conflicts from a knowing condition, we choose to connect two ARs then.

After the analysis we get two important conclusions which are basis lemmas for our algorithm as well:

Lemma 1, when $k=(H_{ij}-H_i)\%2=1$, AR_{iz} is directly connected to AR_{jz} , where $z=H_i-(H_{ij}-H_i-1)/2$.

Lemma 2, When $k=(H_{ij}-H_i)\%2=0$, AR_{iz} and AR_{jz} is one router in real network and should be merged as one. If there is conflict then choose to connect AR_{iz} to $AR_{j(z+1)}$, or connect $AR_{i(z+1)}$ to AR_{jz} , where $z=H_i-(H_{ij}-H_i)/2$.

As mentioned before, we also use node's degree to help topology discovery, finding out topology of ARs which is directly connect to a non-anonymous router. As non-anonymous routers' information can be got by SNMP, including their active interface number, from which we will know how many devices are connected to them. Like network pattern in Fig. 1, through SNMP we can know router A has one interface. But after the single-source traceroute we find node A has 3 children (Fig. 2). According to the contrast we know AR_{b1} , AR_{c1} and AR_{d1} are actually one router, topology becomes as Fig. 7 shows.

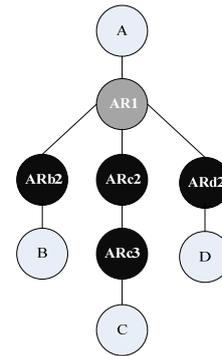


Figure 7. Topology discovered by node degree judgement

...lify the following calculation. An actual situation would be more complex so that node degree judgement should be called whenever it was needed.

IV. TOPOLOGY DISCOVERY ALGORITHM

In order to discover the topology of network with ARs, firstly we need the path length (hops) H from source node to destination node without source routing and order them by size. In proper order our algorithm uses the corresponding router to do source routing tracement. That means for source routing path $A-R_i-R_j$, there are $H_i \leq H_j$. After traceroute we'll get H_{ij} , and calculate $k=(H_{ij}-H_j)\%2$. Then use lemma 1 and lemma 2 to merge or connect ARs. Steps of the algorithm are as follow:

- 1) Use single source tracroute to get every path's length H_i and Order non-anonymous routers by H_i 's size, mark all the needed source routing paths as untreated;
- 2) See if there's some ARs can be merged according to non-anonymous routers' nodes degree;
- 3) To do source routing tracement with non-anonymous routers in proper order and get the path length of $A \rightarrow R_i \rightarrow R_j$, record as H_{ij} . Mark path $A \rightarrow R_i \rightarrow R_j$ as processed;
- 4) Calculate $k=(H_{ij}-H_j)\%2$;
- 5) If $k=1$, connect AR_{iz} to AR_{jz} , where $z=H_i-(H_{ij}-H_i-1)/2$; If $k=0$, merge AR_{iz} and AR_{jz} as one router; If there is conflict then choose to connect AR_{iz} to $AR_{j(z+1)}$, or connect $AR_{i(z+1)}$ to AR_{jz} , where $z=H_i-(H_{ij}-H_i)/2$;
- 6) If there are untreated source route path, repeat from step 2.

Take network model in Fig. 1 as an example. First we

know:

$H_b=3, H_c=4$ and $H_d=3$.

And order them by size:

$H_b=H_d < H_c$.

Then use R_b and R_d as source routing addresses to do traceroute in sequence. The paths are $A-R_b-R_d$, $A-R_b-R_c$ and $A-R_d-R_c$. In the beginning these three paths are all untreated. Next we use nodes degree as a judgment to find AR_{b1} , AR_{c1} and AR_{d1} can be merged as one like Fig. 7 shows.

Then the algorithm traces path $A-R_b-R_d$ and gets information as follow:

$H_{bd}=7$,

$k=(H_{bd}-H_c)\%2=0$,

$z=H_b-(H_{bd}-H_c)/2=1$.

According to lemma 2, we should merge AR_{b1} and AR_{d1} as one, which has been done by node degree judgment. Then trace path $A-R_b-R_c$:

$H_{bc}=6$,

$k=(H_{bc}-H_c)\%2=0$,

$z=H_b-(H_{bc}-H_c)/2=2$.

According to lemma 2, merge AR_{b2} and AR_{c2} . The result comes without conflict (Fig. 8).

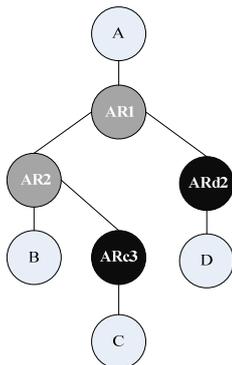


Figure 8. Topology after merge AR_{b2} and AR_{c2}

At last trace path $A-R_d-R_c$:

$H_{dc}=6$,

$k=(H_{dc}-H_c)\%2=0$,

$z=H_d-(H_{dc}-H_c)/2=2$.

According to lemma 2, we can treat AR_{d2} and AR_{c2} as one. But check conflict we see that AR_{c2} is AR_{b2} . Merge AR_{d2} and AR_{c2} means merge AR_{d2} and AR_{b2} . As path $A-R_b-R_d$ has been processed before and the result is that path $A-R_b$ and $A-R_d$ has an intersection at AR_{b1} (AR_{d1}), which makes it impossible to merge AR_{b2} and AR_{d2} again. So the algorithm chooses to connect AR_{d2} and AR_{c3} , which is the proper result (Fig. 9).

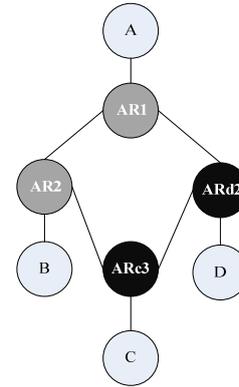


Figure 9. The final topology

Experiments show that for IPv6 networks our algorithm is efficient and performs well in IPv4 networks while devices support source routing. Limitation of this algorithm is that for networks like Fig. 10 shows, AR_3 would be merge into AR_2 or AR_4 no matter it is an independent router or not. In the real computer network environment, anonymous routers are not much and so as this topology. ARs are generally below 10% in common situation.

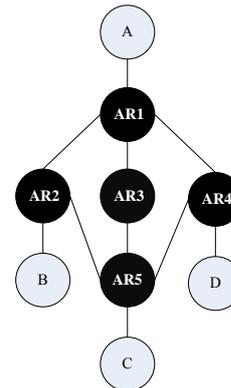


Figure 10. A more complex AR network topology

V. CONCLUSION

In this paper we have addressed the problem of discovering computer network topology that contains anonymous routers. We give an algorithm for performing topology discovery using source routing traceroute method, which is efficient to solve the topology discover problem caused by anonymous routers.

REFERENCES

- [1] S. Bilir, K. Sarac and T. Korkmaz, "Intersection Characteristics of End-to-end Internet Paths and Trees", in Proc. 13th IEEE International Conf. Network Protocols, Boston, 2005, pp. 6-9.
- [2] A. Broido and K. Claffy, "Internet Topology: Connectivity of IP Graphs", in Scalability and Traffic Control in IP Networks, Denver, 2001, pp. 19-24.
- [3] C. Bill, B. Hal and B. Steve, "Mapping and Visualizing the Internet", in USENIX Annual 2000 Technical Conference, San Diego, 2000, pp. 18-23.

- [4] Q. F. Xu, "Topology Discovery in IPv4/v6 Network". *Computer Engineering and Design*, 27, 23, pp. 4507-4509, 12. 2006.
- [5] H. Z. Shi, P. D. Zhu, "A new solution for Anonymous Routers Resolution", *Netinfo Security*, 11, pp. 45-46, 2008.
- [6] M. Gunes and K. Sarac, "Resolving Anonymous Routers in Internet Topology Measurement Studies", in *27th IEEE Conference on Computer Communications*, Phoenix, 2008, pp. 1750-1758.
- [7] Y. Zhang, S. Guo and W. Niu, Y. Yang, "Model Computing and Merging Algorithm for Anonymous Router in Network Topology", in *2010 First International Conference on Pervasive Computing, Signal Processing and Applications*, Harbin, 2010, pp. 247-250.