Volume 5, Issue 3, March 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Unicast Topology Inference Scheme in Network Tomography

Supriya Gawade, Nikita Dusane, KavitaThange
BE (IT) GSMCOE, Pune,
M.S., India

Abstract- Network topology inference is a topic of on-going interest as researchers continue to identify the topology of internet. In this paper for improvement of the performance of the network, we are examining the network and its functionality; we studied which are blockages in our network, and also studied about the resources which are not used properly. For improvement of the performance of the network we have used the Trace-Route tool for getting the internal information of the network, this tool requires cooperation from the internal nodes in the network. In the wide public networks like the routers of the internet, it requires the cooperation of internal nodes in the network, this functionality does not exist in Trace route method for that reason we can't use trace route in the public network. In the following paper, we have proposed a method called as Advanced Sandwich Probe method. This method collects the required information from the network to generate the topology from the information which is collected. The following proposed method is the expansion of the sandwich probe method with the functionality of Trace-route method. Simulation and execution result of the proposed method after comparing with the previous method shows that the method improves the topology identification process conveniently and the proposed method also overcomes the functionality and working of existing method with variation of differences.

Keywords— Network tomography, Topology inference, Sandwich probe scheme, trace route, advanced sandwich probe.

I. INTRODUCTION

Network topology identification has been studied extensively because of great need for precise topology information in networks, distributed internet applications often need to know information about the characteristics of the network therefore in network monitoring, this tool can help network operators for obtaining routing information and the internal characteristics of network such as loss rate, delay etc. However, as the internet becoming more heterogeneous and unregulated, it is more challengeable to get such information. Even though the administrators of small networks can monitor the local traffic conditions and identify the congestion points and bottlenecks, very few networks are completely isolated. The user expected performance of a network thus depends on the performance of an internetwork, and monitoring this internetwork is very challenging part. [1] Presently most of the well-known network topology and behaviour identification tools such as trace route required help from the internal nodes of the network, such as routers because they send requests and messages to the internal nodes and receive feedback from those nodes. on the other hand, one cannot rely on the cooperation of nodes monitored by someone. For e.g., Trace route uses Internet Control Message Protocol (ICMP) messages sent by the routers. For the network like the Internet, routers are often configured in such a way that they should not respond to ICMP, and not to send these messages [2]. Depending on cooperation of internal nodes, and depending on specific settings and configuration details within them, makes this kind of approach very limited. Due to problems in the cooperation based method, some researchers have started to work on the problem of inferring internal network structure and statistics without depending on internal nodes.

Network tomography is the study of a network's internal characteristics using information derived from end point data. To get all needed information about network, first step is to infer network topology. [3] Our paper focus on inferring network topology by using information derived from end point data. Thus we focus on the case where we have one source and several destination receivers, and we use Unicast end-to-end messages to probe the network. Also we believe the routes in the network are fixed during the time these messages are sent.

II. RELATED WORK

In network tomography, we infer network information by using information derived from end points of network. So in network tomography, we initially gather information from end points in the network i.e. first is Data gathering phase and then after we use this gathered information to infer topology of the network i.e. Network topology inference.

A. Data Gathering Phase

Data collection methods for network topography depend on whether internal nodes in network are cooperating or not. If the network on which the tomography method is performed is a public network, and we do not have control of the internal nodes (routers, switches, etc.), we can only expect a limited amount of information from those nodes [2].If we have our own private network we can also program the internal nodes to collect and to report the information that we need. If internal nodes are cooperating then only it is easy to gather network information than when internal routers are not cooperating.

B. Active Data Gathering

Active methods send probe messages across the network to extract the information. Active methods can also use end-to-end probes so they do not need to rely on help from internal nodes. The problem of active methods is that they impose extra load on the network; these methods try to keep the extra load as low as possible so as not to affect the network's quality of service. The probes that are sent by the active methods can be multicast or can be unicast messages. Although not all networks can support multicast, some methods also requires multicast messages to collect the data they need [4]. Also there are some methods that try to optimize the size of the data sent with each probe using approaches like network coding .

C. Passive Data Gathering

Passive methods do not interfere with the network's normal traffic – they just act as monitors and try to reach a conclusion using the available characteristics of the traffic. This is possible if we are able to use the internal nodes for collecting data.

D. Topology Inference

There are basically two major two types of methods for inferring network topology. One of them is maximum likelihood

E. Maximum-Likelihood Approach

Verdi was one of the first researchers working on the concept of network tomography. He was also working on problem of the estimating end-to-end path properties by using link measurements. In the year 1996 he suggested the following equation to model the problem of estimating the number of packets sent between each pair of end nodes [12]. Y = AX, (1.1) Where Y is a vector of the measured data, which is the number of packets passing through each direct link during the specific time period. So Y has element for each direct link of the network. X is a vector which contains the amount of traffic on each end to end path during that time period. A is a binary routing matrix. and let p be the no of end-to-end paths whose traffic is being estimated and |L| is the number of direct links in the network. A is a |L| × p binary matrix in which the element in row i and column j is 1 if and only if i-th end-to-end path containing the J-th direct link. Some researchers used this model for other network tomography problems. In the problem Vardi was doing work on path characteristics were preferred and link characteristics were given. In the other problems with the same model, link characteristics (such as delay, loss, etc.) or the routing matrix may be preferred. The measurements are mostly path-based or link-based. Y = (Y1, Yt) is a vector of the measured data, X = (X1, Xs) is the vector of parameters to be estimated and A is a |t × s| binary routing matrix. Usually Xi is a random variable taken parameterized by some Qi and Qs = (Q1... Qs) is what we need to estimate. the maximum-likelihood approach are relatively computationally complex, and they are suitable only for small routing trees .and therefore, there are some of the methods for dealing with the trade-off between complexity and the accuracy of these approaches. so one the ideas is to use the pseudo-likelihood models. This means dividing the problem which is described by the Eq. 1.1 into some several sub-problems, and solving each separately, and then combining the solutions to get to the final solution for it. For the topology inference problem it means finding the logical topology induced by some different groups of receivers and then merging the topologies for creating the complete logical topology.

F. Constructive Methods

The constructive algorithms require much less computation than the maximum-likelihood approaches, but usually give less accurate results [6]. There are several proposed constructive methods; most of them have a similar bottom-up approach. First, we find a group of leaves that seem to be neighbors. Then we join them as a single node, set proper estimated values for the new node, and continue until only a single node remains. Coates et al. [6] proposed such a method in 2004 named ALT. ALT uses unicast measurements and group's two nodes in each step so the resultant tree is always a binary tree. Therefore, this method is suited only for inferring binary topologies. Ratnasamy et al. [7] suggested a similar method which uses multicast packet loss measurements. In 2002, Duffield et al. [4] extended Ranasamy's algorithm to make it suitable for general, non-binary trees. More recently, Ni et al. [8] suggested a joining algorithm for general trees which uses unicast measurements. They also suggested a dynamic algorithm to add or delete a receiver from the network. Using their method, it is possible to add nodes one by one and construct the whole topology.

III. PROPOSED SYSTEM

We proposed a technique called as Advanced Sandwich Probe. This method collects the needed information from the network to create the topology from the collected information. The proposed method is the extension of the sandwich probe method with the functionality of Trace-route method. Simulation result of the proposed method as comparing with the previous method shows that the method improve the topology identification process. The proposed method overcomes the functionality of existing method with the variation of differences.

A. Data Gathering

In our scheme we first estimate all the Xi, j i.e. the delay of common path which is for all receivers in the network same as per sandwich probe method described above.

In second step, we send sandwich probe to single receiver(i) instead of two different nodes and initially set TTL value of large packet(q) is to certain value k, after k hops q will be drop, and then we can measure the delay in p1 and p2 i.e.

Yi,k. and After that we will also increment the value of TTL of q packet (suppose TTL=k+1) then q will dropped after k+1 hops and again we will measure delay in packets to reach i. This process of incrementing the TTL value of q and measuring delay in packets to reach i is continued until q reaches to i.

We also do above step for each receiver in the network and get Yi,k for each receiver Eg. of measuring Yi,k is shown in fig 3.As shown in figure we set TTL of q packet =2 i.e. after two hops packet q will be drop and delay in packets to reach final node is estimated. Now the value of k is incremented by 1 i.e. 3 and again the time delay is measured this process is continued till q reaches successfully to final node.

The time measurements are performed in the end node instead of the starting node. This method gives us one way delays to the nodes along the path, as opposed to the round trip delay times elicited by Traceroute. The ASP doesn't give us any information about the addresses of the nodes along the path. Delay varies as per the hop so we take multiple readings and takes average of it.

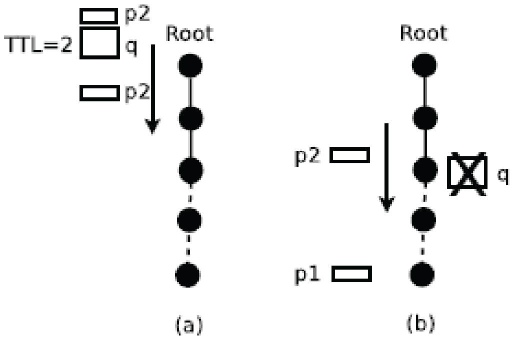


Fig. 1 Step2 in Advanced Sandwich Probe

B. Topology Inference

When Data gathering is completed we have values if Xi,j and Yi,k for all receivers in the network. Xi,j contains common segment delay for each receiver pair in the network. Yi,k is the set of delay difference for packets for each receiver. We use constructive method for inferring topology of the network. Then we start with a tree containing only the root, and try to add leaves to the tree one by one. Assume we have a tree, which is a partial tree for the whole network. Assume we want to add a new leaf n to the tree.

First we find the leaf *n'* in the tree that maximizes *Xn', n*. The paths from the nodes *n* and *n'* to the root have a common segment, and *Xn', n* is our estimate of the delay along that segment. We need to find out at which node in the path from *n'* to the root these two paths separate. The delay from the separation node to the root has to be close to *Xn', n*. So we find the number *k* which minimizes |Xn', n – Yn', k|, Which means the *k*-th node in the path of the root to *n'* has the closest estimated delay to *Xn', n*. From above we got common segment in n and n' and where they are separating, so we can add n to the network. Same like this we have to add node one by one and infer network topology.

Figure 2 shows the network topology that we are studying. The input of the inference algorithm includes the sandwich probes' output i.e., pair wise common path estimations or X = (xi, j) and Y = (Yi, k). Table 1 and Table 2 show X and Y calculated from Figure 4. In Table 1 the element in row i and column j shows xi, j. In Table 2 each row represents the measurements for one leaf. In row i the first element is yi, 1, and the second one is yi, 2, etc.

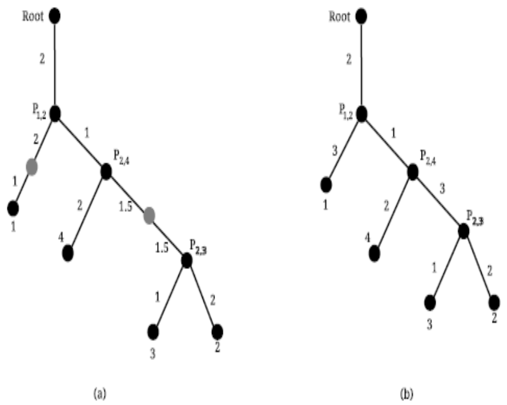


Fig. 2 Actual Tree

Table 1 Actual x values of the tree

Xi,j	1	2	3	4
1	0	2	2	2
2	2	0	6	3
3	2	6	0	3
4	2	3	3	3

Table 2 Actual Y values of the tree

Xi,j	1	2	3	4
1	0	1.90	2.00	2.05
2	1.90	0	6.10	2.90
3	2.00	6.10	0.0	3.10
4	2.05	2.90	3.10	0.0

Table 3 Estimated X values of tree

YI,k	
1	1.90 3.8 4.90
2	1.90 0 6.10 5.90 7.85
3	2.00 2.90 4.45 6.10 6.90
4	2.10 3.10 4.90

We start by adding a single node as the root and the leaves are added one by one to the tree. Inserting the first node gives us the tree in Figure 2-a, the number of the nodes and the link delays come from the first row in Table 4. The length of the first link is 1.9, which is the first element of the row. The length of the second link is the difference between the second and the first elements i.e., 3.8−1.9 = 1.9 and the length of the third one is the difference between the fourth and the third elements. and To insert the second leaf we first look into Table 3 and then find X (1, 2) which is 1.9. Now we should find the node in the path from leaf 1 to the root whose distance to the root is closest to the 1.9. This is the first node from the root and is called P (1, 2) in Figure 2-b. This means the paths of leaf 1 and 2 are separate at this node. Now using the second row of Table 4 we create the rest of the path of leaf 2 and we get the tree in Figure 5-b. The separation point is P (1, 2) whose distance to the root is 1.9. The distance from next node in the path of the leaf 2 to the root is 3.05, so the length of the next link is 3.05 − 1.9 = 1.15 and the rest of the path is created with the same process for the first leaf.

Finally to add the last leaf, the procedure is the same as for the third one. First we find the node n in Table 3 that maximizes x (4, n) which is leaf 3. The node with the closest distance to x (4, 3) is the node called P (2, 4) in Figure 5-d; x (4, 3) = 3.1 and the root's distance to P (2, 4) is 3.05. According to Table 4 the path of the leaf 4 contains one more link and we create that using the fourth row in Table 4 The final result of the algorithm is shown in Figure 2-d.

Table 4 Estimated Y values of the tree

Yi,k	
1	2 4 5
2	2 3 4 6 8
3	2 3 4 6 7
4	2 3 5

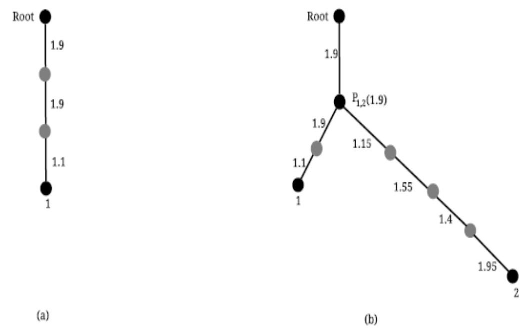
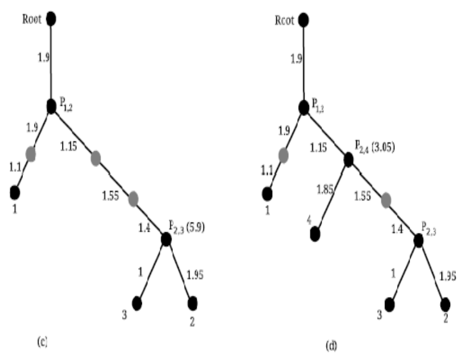


Fig. 2 (a,b) Example of topology inference



IV. CONCLUSION AND FUTURE WORK

Utility tools like trace route to get the internal information of the system. We have proposed the new technique called Advanced sandwich probe to gather important evaluation of the network. The performance result also shows that our proposed work overcomes the existing work with variation of differences

The information gained from TSP probing scheme can be used to find out other characteristics of the network, as well as the logical topology that is discussed in this paper. For instance our topology inference algorithm can be used to find out the physical topology of the network.

Another way to explore the work is to study the overhead of this method, more closely and try to minimize it. Running the probing scheme on some large real networks,e.g., the Internet, is another interesting field for future work and it can also help a lot with studying the overhead of network

REFERENCES

[1] Network Tomography: Recent Developments Rui Castro, Mark Coates, Gang Liang, Robert Nowak and bin Yu.J. Clerk Maxwell, a Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] B. Yao, R. Viswanathan, F. Chang, and D. Waddington, “Topology inference in the presence of anonymous routers,” in INFOCOM 2003.Twenty-Second Annual Joint Conferences of the IEEE Computer and Communications.IEEE Societies, **vol**. 1. IEEE, 2003, pp. 353–363K. Elissa, “Title of paper if known,” unpublished.

[3] N. Duffield, J. Horowitz, F. Lo Presti, and D. Towsley, “Multicast topology inference from measured end-to-end loss,” Information Theory, IEEE Transactions on, vol. 48, no. 1, pp. 26–45, 2002.

[4] N Duffield, J. Horowitz, F. Lo Presti, and D. Towsley, “Multicast topology inference from measured end-to-end loss,” Information Theory, IEEE Transactions on, vol. 48, no. 1, pp. 26–45, 2002.

[4] M. Gjoka, C. Fragouli, P. Sattari, and A. Markopoulou, “Loss tomography in general topologies with network coding,”in Global Telecommunications Conference, 2007.GLOBECOM’ 07.IEEE. IEEE, 2007, pp. 381–386.

[5] C. Fragouli and A. Markopoulou, “A network coding approach to network monitoring,” 2005.

[6] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, “Network tomography: recent developments,” Statistical Science, pp. 499–517, 2004

[7] S. Ratnasamy and S. McCanne, “Inference of multicast routing trees and bottleneck bandwidths uses end-to-end measurements,” in INFOCOM’99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies Proceedings IEEE, vol. 1 IEEE, 1999, pp. 353–360

[8] J. Ni, H. Xie, S. Tatikonda, and Y. Yang, “Efficient and dynamic routing topology inference from end-to-end measurements,” IEEE/ACM Transactions on Networking (TON), vol. 18, no. 1, pp. 123–135, 2010.

[9] M. Coates, R. Castro, R. Nowak, M. Gadhiok, R. King, and Y. Tsang, “Maximum likelihood network topology identification from edge-based unicast measurements,” ACM SIGMETRICS Performance Evaluation Review, vol. 30, no. 1, pp. 11–20, 2002.

