The following paper posted here is not the official IEEE published version. The final published version of this paper can be found in the Proceedings of the IEEE International Conference on Communications (2005 : Seoul, Korea):pp. 143-147

Discover Multicast Network Internal Characteristics Based on Hamming Distance

Hui Tian and Hong Shen
Graduate School of Information Science
Japan Advanced Institute of Science and Technology
hui-t,shen@jaist.ac.jp

Abstract

One of the important techniques to monitor and control large-scale networks today is to implement only at the end. However, end-based control system needs to have the knowledge of network internal characteristics. The paper proposes a novel approach to discover network internal characteristics from end-to-end multicast traffic measurements, which requires no support from internal routers. Our approach is based on hamming distance of sequences on receipt/loss of probe packets maintained at each pair of nodes. As will discuss in this paper, our approach will mainly focus on identification of network internal characteristics of routing topology and loss performance. The simulation shows that hamming distance-based approach can discover the routing topology which is more accurate and efficient with a finite number of probe packets than before. The hamming distance matrix proposed in this paper can also effectively discover the loss performance of the network.

Key words: Multicast network, topology inference, loss performance, sequence, hamming distance.

1 Introduction

As the network grows explosively, tremendous traffic load and various applications drive people to develop techniques to model and control the network, analyze the performance and optimize the network. All these tasks require that the end-system have the knowledge of the network internal characteristics. Discovered internal information such as topology and localized loss performance plays an important role in resource management, loss recovery, congestion control and so on, as discussed in [1, 5, 4]. Therefore, different techniques are developed for identification of network internal characteristics.

Generally, the existing approaches to discover network internal characteristics are classified into three types: (i)

setting probers in the network, collecting statistics at internal nodes periodically and generate reports on topology or link-level performance; (ii) characterizing the network based on end-to-end behavior of point-to-point traffic such as that generated by TCP or UDP; (iii) inferring link-level loss behavior and topology from end-to-end measurements.

The first approach requires support from internal nodes. It takes long time and extra equipment or software to collect data and analyze them. Usually only authorized network administrator and developer can perform the task. The approach cannot update the network internal information in time, neither is it scalable. The second approach keeps the same problem as the first one. The only difference is that it is performed in the edge of the network. Compared with the former two approaches, the third one is the most intelligent approach. It infers network internal characteristics from end-to-end measurements, without internal network cooperation. This is called network tomography which is very hot in recent years because the end-based inference approach is much more practical and scalable than previous approaches.

The first crucial step in the tomography process is topology identification. There has been much research on topology inference from end-to-end measurements as showed in [2, 4, 6, 7, 8, 10]. Based on identified topology, more internal characteristics such as loss performance, link delay can be inferred. As used in [9, 8], we also use multicast probe traffic as end-to-end measurements in this paper. The key idea underlying this approach is that receivers sharing common paths on the multicast tree associated with a given source will see correlations in their packet losses or delays. The multicast topology and internal loss performance can thus be inferred based on the shared loss or delay statistics on transmitted probe packets. However, we find that the prevalent method to estimate correlation used in [2, 6, 7, 8, 12] for siblings identification may produce fault results in the cases discussed in Section 3. Therefore, we propose the approach of hamming distance-based classification in this paper which is based on hamming distance

of sequences on receipt/loss of probe packets maintained at each pair of sibling nodes. Moreover, we propose a hamming distance matrix approach to effectively identify the network internal loss performance. It is more simple and easier to implement by applying the approach we proposed in this paper than previous work as discussed in [3, 11]. In our simulation, the efficiency and effectiveness of the hamming distance based topology inference approach is validated.

The paper is organized as follows. In Section 2 the problem is stated. New contributions are given for overcoming existing problems in Section 3, and simulation is also given in this section. Section 4 proposes a hamming distance matrix for loss performance identification. Section 5 concludes the paper.

2 Problem Statement

As presented in [8], we model the physical multicast tree by a tree comprising actual network elements (the nodes) and communication links connecting them.

Let $T = (V, L)$ denote a multicast tree with node set V and link set L. The root node 0 is the source of probe packets, and $R \subset V$ denotes the set of leaf nodes representing the receivers. A link is said to be internal if neither of its endpoints is the root or a leaf node. Each non-leaf node k has a set of children node $d(k) = \{d_i(k) \mid 1 \leq i \leq n_k\}$, and each non-root node k has a parent $p(k)$. The link $(p(k), k) \in L$ is denoted by link k. Let $a(U)$ denote the nearest common ancestor of a node set $U \subset V$. Nodes in U are said to be siblings if they have the same parent, i.e., if $f(k) = a(U), \forall k \in U$. The subtree of T rooted at k is denoted by $T(k) = (V(k), L(k))$, and the receiver set $R(k) = R \cap V(k)$.

For each link an independent Bernoulli loss model is assumed with each probing packet being successfully transmitted across link k with probability p_k. Thus the progress of each probing packet down the tree is described by an independent copy of a stochastic process $X = (X_k)_{k \in V}$ as follows. $X_0 = 1$, $X_k = 1$ if the probing packet reaches node $k \in V$ and 0 otherwise. If $X_k = 0$, $X_j = 0, \forall j \in d(k)$. Otherwise, $P[X_j = 1 | X_k = 1] = p_j$ and $P[X_j = 0 | X_k = 1] = 1 - p_j = \alpha_j$. Define $p_0 = 1$. The pair (T, p) is called a loss tree. $P_{T,p}$ denotes the distribution of X on the loss tree (T, p). We use a boolean variable $X_k^{(i)}$ to denote the loss measurement of node k for ith probe packet. For n probe packets, the 0-1 sequence maintained on the node k is denoted by $\{X_k^{(i)}\}$, $1 \leq i \leq n$, $k \in V$.

Thus the problem is: given the measurements performed between the root node and receivers $X_R^{(i)}$, how to infer the multicast network topology $T = (V, L)$ and its internal loss performance.

3 Hamming Distance-based topology identification

We begin the section with discussing the main idea behind existing algorithms for topology inference and we point out a problem that affects them. We then introduce hamming distance approach to overcome this problem. Due to its simplicity and efficiency, hamming distance-based classification approach can play the role on identifying siblings in BLT algorithm [8].

When a probe packet is sent from the root of the multicast tree, several copies are generated at each router encountered. When a copy of probe packet i reaches node k, $X_k^{(i)}$ is set to 1. Otherwise, $X_k^{(i)}$ is set to 0, indicating that node k doesn't receive any copy of probe packet i. For those internal routers, $X_k^{(i)}$ is 1 if the ith probe packet reaches it and 0 otherwise. In practice, $X_k^{(i)}$ of internal nodes is obtained by $\vee_{l \in R(k)} X_l^{(i)}$ because the internal node is said to receive a probe packet surely if any receiver descended from it receives the probe packets. Thus, each node obtains a "0-1" sequence $\{X_k^{(i)}\}$, $0 < i < n$, $k \in V$. We can hence compare the correlation between the sequences so as to reconstruct the multicast network topology and infer the internal link characteristics.

In the previous work such as [8, 7, 6, 12], the product of the probabilities of successful transmission on all the links between the root 0 to node k, denoted by $A(k) = \Pi_{j \succeq k} p_j$, is widely used in topology inference from measured end-to-end loss. $A(k) = \Pi_{j \succeq k} p_j$ is the product of the probabilities of successful transmission on each link between k and the root 0. If the probability of one probe packet is successfully transmitted from the root to the nearest ancestor node of a receivers subset is minimized, the receivers are considered as siblings. This approach comes down to minimizing the following common form.

$$A(i, j) = \frac{P[\vee_{l \in R(i)} X_l = 1] P[\vee_{l \in R(j)} X_l = 1]}{P[\vee_{l \in R(i)} X_l = \vee_{l \in R(j)} X_j = 1]}. \quad (1)$$

In practice, $A(i, j)$ is estimated by $A^{(n)}(i, j)$ as follows:

$$A^{(n)}(i, j) = \frac{\sum_{m=1}^{n} X_i^{(m)} \cdot \sum_{m=1}^{n} X_j^{(m)}}{n \cdot \sum_{m=1}^{n} X_i^{(m)} \cdot X_j^{(m)}}, \quad (2)$$

where $X_k^{(m)} = \vee_{l \in R(k)} X_l^{(m)}$. Each probability in Equation (1) is estimated by the observation from n probe packets, eg., $P[\vee_{l \in R(i)} X_l = 1]$ is estimated by $\frac{\sum_{m=1}^{n} X_i^{(m)}}{n}$. As n goes to infinity, $A^{(n)}(i, j)$ is consistent with $A(i, j)$, $i, j \in V$. However, with a finite number of probe packets, the estimation may cause obvious mistakes as shown in Figure 1 when determining whether two nodes are siblings.
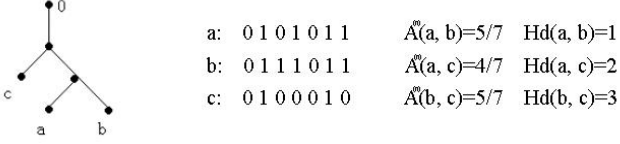
a:	0 1 0 1 0 1 1	$A^{(n)}$(a, b)=5/7	Hd(a, b)=1
b:	0 1 1 1 0 1 1	$A^{(n)}$(a, c)=4/7	Hd(a, c)=2
c:	0 1 0 0 0 1 0	$A^{(n)}$(b, c)=5/7	Hd(b, c)=3

Figure 1. Comparison of Hamming distance($H_d(\cdot, \cdot)$) and the previous estimated function($A^{(n)}(\cdot, \cdot)$) for each pair node

According to estimated value $A^{(n)}(i, j)$, if receiver j loses a lot of probe packets, the value of $A^{(n)}(i, j)$ will be very small, even smaller than the value between i and its actual sibling, e.g., $A^{(n)}(a, c) < A^{(n)}(a, b)$ because of great loss at receiver c. Even when the number of probe packet increases, such bias cannot be eliminated completely. Because the estimated probability in Equation (2) doesn't equal to the true probability unless the number of probe packets is infinite. In order to reduce the bias with a finite number of probe packets, we propose our hamming distance classification approach.

The hamming distance between nodes p and q is defined as the number of different bits between their sequences, which is given in the following equation, where "$\oplus$" is the exclusive-OR operator.

$$H_d(u, v) = \sum_{m=1}^{n} X_u^{(m)} \oplus X_v^{(m)}. \qquad (3)$$

For instance, the hamming distances of each pair sequences in Figure 2 are computed. $H_d(a, c) > H_d(a, b)$ is congruent with what the figure shows obviously because node a and node b are siblings in the real network which failed to be inferred by the value of $A^{(n)}(a, b)$ and $A^{(n)}(a, c)$. According to the real network, the node pair of b and a are siblings, b and c are non-siblings. The hamming distance approach distinguishes the sibling and non-sibling with different values successfully. However, $A^{(n)}(i, j)$ cannot mark out the differences between the pair sequences with 7 probe packets in Figure 1.

In the multicast network, the nearer two nodes are located, the more similar two "0-1" sequences they maintained are. The reason for such phenomenon is that the probe packets from the root to the receivers may pass many common links. If the receivers are siblings, the paths the probe packets pass from the root to their parent nodes are the same. Therefore, the correlation of two "0-1" sequences between a node and its siblings is greater than that of all other pairs of sequences between the node and its non-sibling node.

In order to infer the topology of the multicast network, all the siblings need to be identified correctly. That is, we should find out the similarity of each pair of sequences. The problem of identifying siblings in the multicast network can be stated as marking out the similarity and dissimilarity of all pairs of bit sequences. For a bit sequence, hamming distance is the simplest and most efficient method to identify the similarity and dissimilarity among different sequences.

Therefore, we use hamming distance to identify siblings for multicast network topology inference. The previous BLT algorithm can thus use hamming distance-based classification resulting in a more accurate topology with higher efficiency. The new algorithm based on hamming distance, we call HDLT, follows a similar strategy to the BLT algorithm with the following key modifications. The receiver pairs are selected by finding i, j such that

$$\sum_{m=1}^{n} X_i^{(m)} \oplus X_j^{(m)}$$

is minimized. This selection criterion is superior to the previous one not only in efficiency, but also in accuracy which is obvious in Figure 2.
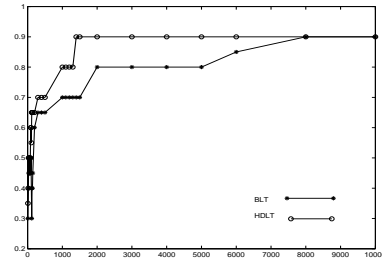

Figure 2. Comparison of BLT and HDLT. With a small number of probe packets, HDLT can infer a more accurate topology than BLT.

4 Hamming Distance Matrix for Loss Performance Analysis

In this section, a hamming distance matrix is defined at first. Based on the matrix, we will discuss how to analyze the network internal loss performance.

As we have given above, each receiver maintains a 0-1 sequence which is denoted by $\{X_r^{(m)}\}$, $1 \le m \le n$, $r \in R$. We assume the number of receivers in a multicast network to be l. For simplicity, let d_{ij} denote the hamming distance between receiver i and j, i.e., $d_{ij} = H_d(i, j)$. Denote D to be the hamming distance matrix. Then D is defined as follows:

$$
\mathbf{D} = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1l} \\ d_{21} & d_{22} & \dots & d_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ d_{l1} & d_{l2} & \dots & d_{ll} \end{pmatrix}.
$$

Obviously, the matrix is a symmetric matrix, that is, $d_{ij} = d_{ji}$ for $i \neq j$. And it is also easy to know $d_{ij} = 0$ if $i = j$. Apart from these properties, the matrix implies more information. For instance, if receiver i and j are siblings, d_{ij} is supposed to be smaller than d_{ik}, $k \neq j$ and $k \neq i$. More generally, the nearer two nodes are located, the smaller their hamming distance is supposed to be because their sequences are more similar due to more shared common link condition which we have discussed in Section 3. Therefore, such a hamming distance matrix shows a lot of information on internal topology in the condition that those internal links have different congestion and loss situation.

Now let's see what this matrix can do for loss performance analysis and even identification. Since we have inferred the multicast network topology in Section 3, how to discover the loss performance based on the topology is now concerned. For clearly clarifying the applications of the hamming distance matrix in loss performance analysis, we firstly consider a network where only one link causing severe loss as Figure 3. Link s in Figure 3 loses a lot of packets. All other links work in good condition.
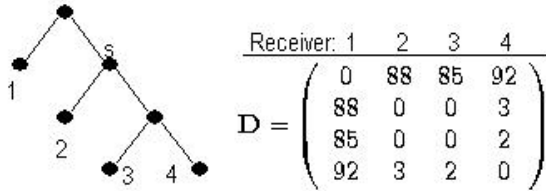

Figure 3. A simple hamming distance matrix.

According to the hamming distance matrix in Figure 3, we can easily find the hamming distances between 1 and any of the rest receivers are very great, while the hamming distances among receiver 2, 3 and 4 are quite similar and small. Then we can infer that the nearest common link of receiver 2, 3 and 4 work in poor condition.

As for a network where internal links condition may be very complex, we need to transform the matrix into several blocks according to different values of components in the matrix. Usually we classify the components by an experience difference which depends on how many probe packets are sent in total. With the experience difference, we can do elemental transformation on the hamming distance matrix and obtain a matrix in several regular blocks. The receivers

are accordingly classified into several groups. Thus we can infer all the possible bad links in the topology according to those classified receivers by blocks easily. For example the network topology, the hamming distance matrix and the transformed matrix are given in Figure 4, we aim to analyze the link loss performance and locate the ill-performing link easily.
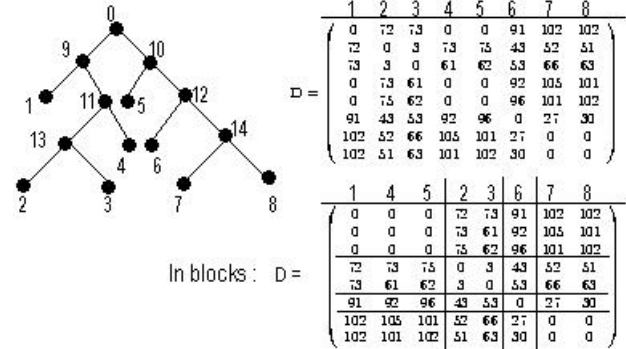

Figure 4. Hamming distance matrix and matrix in blocks

There is a group including receiver 1, 4 and 5 which receive almost all the probe packets. This group is called base group which can be easily identified by the 0-1 sequence maintained by any receiver in this group. The links in the path from the source to the receiver in this group all works in good condition. We find all other groups have the common property, that is, the hamming distance between any receiver in these group and any receiver in base group is very great while the hamming distance between the receivers in each group is very small. These groups are called loss group. We can then decide the link connecting to the nearest common ancestor node of each loss group is one of the links who are causing severe loss. In the topology of Figure 4, link 13, 6 and 14 are links we aim to identify. While all other links work well in this network.

Therefore, if we can transform a hamming distance matrix into blocks according to the value of components, the receivers are classified into several groups. Then we can determine a base group which include those receivers who receive almost all probe packets, i.e., there isn't any links causing big loss in the path from the source to these receivers. And apart from this group, there are many loss groups. The link connecting to the nearest common ancestor node of each loss group are identified as those links where severe loss is caused. By this means, a hamming distance matrix D of a network can help to analyze and identify the network internal loss performance. Moreover, a phenomenon which values our attention is that there are usually only a few links who are in poor condition in a large scale

network in practice. Therefore, the components in hamming distance matrix usually appear to be different obviously and can be classified very easily.

5   Conclusion

We have discussed how to discover multicast network internal characteristics in this paper. We apply hamming distance of sequences on receipt/loss of probe packets maintained at each pair of nodes to topology and loss performance inference from end-to-end measurements. With a finite number of probe packets, the topology inferred by the new approach has been found to be more accurate and efficient than those inferred by the previous algorithms in our simulation. As to internal loss performance analysis, we defined a hamming distance matrix, which has shown potential usage and benefit to internal loss identification.

Acknowledgement

References

[1] B.N.Levine, S.Paul, and J.J.GARCIA-Luna-Aceves. Organizing multicast receivers deterministically according to packet-loss correlation. In *Proceedings of ACM Multimedia'98*, 1998.

[2] R. Caceres, N. G. Duffield, J. Horowitz, F. L. Presti, and D. Towsley. Loss based inference of multicast network topology. In *IEEE Conference on Decision and Control, Phoenix, AZ*, 1999.

[3] R. Caceres, N. G. Duffield, J. Horowitz, and D. Towsley. Multicast-based inference of network-internal loss characteristics. *IEEE Trans. Information Theory*, 1999.

[4] R. Castro, M. Coates, and R. Nowak. Maximum likelihood identification from end-to-end measurements. Technical report, DIMACS Workshop on Internet Measurement, Mapping and Modeling, DIMACS Center, Rutgers University, Piscataway, New Jersey, February 2002.

[5] D.Li and D.R.Cheriton. OtersOn-Tree Efficient Recovery using Subcasting:a reliable multicas protocol. In *Proceedings of IEEE ICNP'98*, 1998.

[6] N. Duffield, J. Horowitz, F. Presti, and D. Towsley. Multicast topology inference from end-to-end measurements. Technical report, ITC Seminar on IP Traffic Measurement, Modeling and Management, Monterey, CA, Sept. 2000.

[7] N. Duffield, J. Horowitz, and F. L. Presti. Adaptive multicast topology inference. In *IEEE Infocom 2001, Anchorage, Alaska*, April 22-26, 2001.

[8] N. G. Duffield, J. Horowitz, F. L. Presti, and D. Towsley. Multicast topology inference from measured end-to-end loss. *IEEE Trans. Information Theory*, 2002.

[9] N. G. Duffield and F. F. Presti. Multicast inference of packet delay variance at interior network links. In *Pro. IEEE Infocom 2000, Tel Aviv*, 2002.

[10] S. Ratanasamy and S. McCanne. Inference of multicast routing tree topologies and bottleneck bandwidths using end-to-end measurements. In *Proc. IEEE Inforcom'99, New York, NY*, 1999.

[11] H. Tian and H. Shen. Multicast-based inference of network internal loss performance. In *Proceedings of ISPAN'04*, May. 2004.

[12] H. Tian and H. Shen. An improved algorithm of multicast topology inference from end-to-end measurements. In *Proceedings of the 5th Int'l Symp. on High Performance Computing ISHPC-V*, Oct. 2003.