A Probe Prediction Approach to Overlay Network Monitoring

Shun-an Wu, Qiao Yan, Xue-song Qiu, Yanjie Ren

State Key Laboratory of Networking and Switching Technology

Beijing University of Posts and Telecommunications

Beijing, 100876, China

Email: {sawu, qiaoyan, yjren}@metarnet.com, xsqiu@bupt.edu.cn

Abstract—**Nowadays the overlay network has greatly improved the performance of the Internet. The overlay network flexibly selects its communication paths and targets and thus can benefit from estimation of end-to-end network performances. For an overlay network with n end hosts, most of the existing systems have to send O (n^2) probes into the network and then they calculate the performances of all links. Although these systems to some extent can determine the performances of the links, they have to send plenty of probes into the network, which has generated great traffic and imposed extra overload in the network. In order to address the problem, we propose a new approach based on probe prediction method by which we only need to measure a few probes in the probe set and predict out the responses of the rest probes and then we revise the final prediction results and find out the suspected congested links set. The experiments have shown that we only need to send about 20% of the total probes and infer all the responses of these probes with higher accuracy than ever before.**

Keywords-monitor; probe prediction; Overlay; result revision; end-to-end performance

I. INTRODUCTION

The proposal of overlay provides a promising way to overcome the Internet ossification. The overlay system flexibly selects its communication paths and targets, and thus can benefit from estimation of end-to-end network performances such as latency and packet loss rate. So it is very important to monitor the network links performance. However, for an overlay network with n end hosts, most of the existing systems have to send O (n^2) probes to measure the end-to-end performances. Although these systems to some extent can determine the links' performances, still they have following drawbacks. On one hand they lack scalability; on the other hand too many probes have generated great traffic in the network, which may influent the network performance. So how to determine the links performances with least probes becomes the current focus of research.

The existing probe selection methods can be divided into two classes: scalable end-to-end performance estimation schemes [1] [2] [3] [4] [5] [6] and algebraic end-to-end performance inference schemes [9] [11]. The methods in the former category also can be categorized into clustering based methods [5] [6], coordinate based methods [3] [4] and collaborative prediction based methods [1] [2]. The clustering based ones and the coordinate based ones can only be used to estimate the latency, and cannot be used to detect congestions and faults, for that the end hosts in these systems may take different routers to remote end hosts. Although the collaborative based methods can be used to estimate both latency and path packet loss rate, they just randomly monitor a subset of the paths in the network without probe selection phase. However, the monitoring on some paths is unnecessary for that some paths can be calculated by a linear combination of other paths. Methods based on the algebraic end-to-end performance estimation schemes selectively monitor a basis set of k paths. And by monitoring the loss rates for these paths in the basis set, it can infer the loss rats for all end-to-end paths. This method can be used in scalability systems and usually the number of paths in the basis set (i.e. k) equals to O (n log n). However, under some special conditions such as the rank of the routing matrix that formed by the whole paths is approaching full rank, the number of paths in the basis set (i.e. k) can be very large.

In this paper, we propose a new network monitoring method which is based on collaborative prediction [8]. We just monitor a small subset of paths and infer out the packet loss rates of the rest paths. Besides, we propose a method to improve the prediction results. Meanwhile, our method can also be extended to other metrics, such as latency.

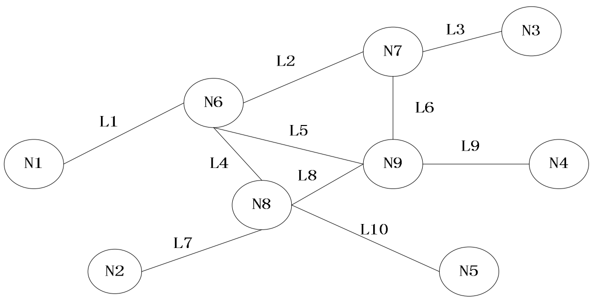
The rest of this paper is organized as follows. Section II presents a general description on the Overlay network monitoring problem. Section III presents the MMMF model and our probe selection method. Section IV revises the prediction results and gives the suspect congested link set. Section V evaluates the proposed algorithms and the paper concludes in section VI.

II. PROBLEM FORMULATION

Fig.1 is an example of network model. Fig.1 (a) is the logical network topology of a managed network. The relationship between the probes and the links can be represented by the matrix in Fig.1 (b). The rows in the matrix E correspond to probes and columns correspond to links. E_{ij} in the matrix equals 1 if the probe passes the link , and $E_{ij} = 0$, otherwise. In the Overlay network, probe is sent from

This work was partly supported by t 973 Project of China (2007310703) , National S&T Major Project (2011ZX03003-002-01) and NSFC (60821001, 60802035, 60973108).

one end host to another one, and the number of probes equals the number of end host pairs, so we can represent the responses of the probes by the matrix in Fig.1 (c), where the rows correspond to the sources of the probes and the columns correspond to the destination of the probes. Each entry in the matrix represents a certain state of the corresponding probe, Here, the entries in the diagonal line of the matrix do not exist for that the sources and the destinations are the same. We add extra probes correspond to the entries in the diagonal line of the matrix which pass no links and the responses of these probes are all 0 (i.e. normal state).



(a) A logical network topology

	$L1$	$L2$	$L3$	$L4$	$L5$	$L6$	$L7$	$L8$	$L9$	$L10$
$P1$	1	0	0	1	0	0	0	0	0	1
$P2$	0	1	1	1	0	0	1	0	0	0
$P3$	1	0	0	1	0	0	1	0	0	0
$P4$	0	0	1	0	0	1	0	1	0	1
$P5$	0	0	1	0	0	1	0	0	1	0
$P6$	1	1	1	0	0	0	0	0	0	0
$P7$	0	0	0	0	0	0	1	1	1	0
$P8$	0	0	0	0	0	0	1	0	0	1
$P9$	0	0	0	0	0	0	0	1	1	1
$P10$	1	0	0	0	1	0	0	0	1	0

(b) The matrix E that represents the relationship between probes

	$N1$	$N2$	$N3$	$N4$	$N5$
$N1$	0	0	1	0	1
$N2$	0	0	1	0	0
$N3$	1	1	0	1	0
$N4$	0	0	1	0	1
$N5$	1	0	0	1	0

(c) The matrix that represents the responses of the probes

Figure 1. A network model

All the probes in Fig.1 (c) constitute of the probe set. Our goal is to determine all the paths' performances by a minimal subset of the total probe paths with maximum accuracy.

III. MMMF MODEL AND PROBE SELECTION

A. MMMF model

Most of the existent Overlay monitoring algorithms need to send a great part of the probes even all to measure the packet loss rates of the paths between each end host pair, which will generate too much traffic in the network. In order to save traffic we only send a small subset of the probe set to form a sparse matrix like Fig.1 (c) and use MMMF to predict out the responses of the rest probes. In MMMF prediction algorithm, the entries in the matrix need to be 1, -1 or 0 (1 and -1 represent OK and false state, and 0 represents unknown state), while the entries in the matrix in Fig.1 (c) that represents the responses of the probes are either 1 or 0 (1 for OK state and 0 for false state). So we need to transform the matrix in Fig.1 (c) into binary matrix, where 0 entries are transformed into 1 and 1 entries are transformed into -1 and the unsent probes are filled with 0. Fig.2 is an example of the transform from the matrix in Fig.1 (c) to binary matrix. Fig.2 (b) is the probe-response matrix with probes N1->N3, N3->N1 … N4->N5 unmeasured. Fig.2 (b) is the matrix transformed from the matrix in Fig.2 (a).After we select a small subset of the probe set to monitor we can get a sparse matrix. Then we use MMMF to get a full matrix approximate to the sparse matrix, which can predict the probe responses which have not been measured. The prediction result is shown in Fig.3 (a). As the entries in the approximate matrix are decimals from -1 to 1, we need to transform the decimals into -1 or 1. If the entry is greater than 0 it is transformed into 1, otherwise it is transformed into -1. Fig.3 (b) is the transformed result.

	$N1$	$N2$	$N3$	$N4$	$N5$
$N1$	0	0		0	1
$N2$	0	0	1		0
$N3$		1	0	1	
$N4$	0	0	1	0	
$N5$	1	0			0

(a) The probe-response matrix with some probes unmeasured

	$N1$	$N2$	$N3$	$N4$	$N5$
$N1$	1	1	0	1	−1
$N2$	1	1	−1	0	1
$N3$	0	−1	1	−1	0
$N4$	1	1	−1	1	0
$N5$	−1	1	0	0	1

(b) The matrix transformed from the matrix

Figure 2. An example of transform from a sparse matrix to a binary matrix

	$N1$	$N2$	$N3$	$N4$	$N5$
$N1$	1	1	−0.868	1	−1
$N2$	1	1	−1	0.857	1
$N3$	−0.795	−1	1	−1	−0.149
$N4$	1	1	−1	1	0.007
$N5$	−1	1	−0.279	0.028	1

(a) The prediction result

	$N1$	$N2$	$N3$	$N4$	$N5$
$N1$	1	1	−1	1	−1
$N2$	1	1	−1	1	1
$N3$	−1	−1	1	−1	−1
$N4$	1	1	−1	1	1
$N5$	−1	1	−1	1	1

(b) The matrix transformed from the prediction result

Figure 3. The prediction result

However, the accuracy for primary prediction may not be satisfactory. So there turns up a problem of improving the prediction accuracy. There exists a relationship between the absolute value of the matrix entry and the prediction certainty, the entry with large absolute value is quite sure for being predicted correctly, but a entry with small absolute value (i.e. it is close to 0) is not so sure for that we are uncertain to separate it to be OK or false. In Fig.3 (a) we select the probe corresponds to 0.007. With this idea we provide a new method different from the randomly selecting method. We firstly select a small subset of the probe set to measure, and then use MMMF to predict the responses of the rest probes. After that, we select a few probes with the smallest absolute value to send, and use MMMF again to get a more approximate matrix. Repeat this step for several times until the prediction accuracy is up to a certain level.

B. Probe selection

In the general MMMF model, the subset of the probe set is selected randomly. In order to reduce the total number of probes sent into the network, we study another method to select a subset of the prior probe set. We also randomly select a small subset of the probe set but instead of monitoring them all we just monitor the basis of the subset. Therefore, we can get all the packet loss rate of the paths in the subset.

The path P_i is a vector $v \in \{0,1\}^s$ where s is the number of links and the jth entry v_j is 1 if the path P_i passes the link L_j and 0 otherwise. Let p_i denote the packet loss rate of path P_i and l_j denote the packet loss rate of link L_j and assume that the packet loss rate is independent among links we can have following equation.

$$1 - p_i = \prod_{j=1}^{s}(1 - l_j)^{v_j} \quad (1)$$

Define a vector $x \in R^s$ with elements $x_j = \log(1 - l_j)$, we have the following equation, where v^T is the transpose of v.

$$\log(1 - p_i) = \sum_{j=1}^{s} v_j \log(1 - l_j) = \sum_{j=1}^{s} v_j x_j = v^T x \quad (2)$$

Define a vector $b \in R^s$ with elements $b_i = \log(1 - p_i)$ and with the matrix E in Fig.1 (c) we have

$$E^T x = b \quad (3)$$

We decompose x into $x = x_E + x_N$, where x_E is a projection on the paths space and x_N is a projection on the null space (i.e. $E^T x = 0$). As the paths can be written as a unique combination of the max linearly independent paths, so the packet loss rate of the rest paths in the subset can be inferred from these linearly independent paths.

IV. PREDICTION RESULTS REVISION

Due to the fact that the prediction results cannot be 100% correct, usually contradictory predictions exist. Even though some probes are predicted incorrectly, for a link in normal state, there are often only a few probes that pass it would be in negative state. For example, if a link in congested state would cause three probes to be negative, it is a small probability event that all these three probes are predicted to be negative when the link is in normal state. Similarly, for a real congested link, it is highly unlikely that all probes that pass it are predicted to be positive. Based on this fact, we propose a simple but effective approach to deal with the incorrect prediction results. Experiments in section V show that it could efficiently reduce the incorrect predictions.

Let S_i be the probe that is predicted to be negative; let be the set of all probes that could be possibly predicted to be negative ($S_i \in S$); let S_n be the set of probes that are actually predicted to be negative. Let L_j be the congested links, and be the set of all congested links ($L_j \in L$). Then we use a metric as follows.

$$\alpha_{L_j} = \frac{\sum_{S_i \epsilon S_n} P(S_i|L_j)}{\sum_{S_i \epsilon S} P(S_i|L_j)}, \quad 0 \le \alpha_{F_j} \le 1 \quad (4)$$

For a particular link L_i, if $\alpha_{L_j} < \alpha$ (is a constant value between 0 and 1, such as 0.5), L_i should be regarded as a congested link.

Each link L_j has a corresponding α_{L_j}. If α_{L_j} exceeds the threshold (α), link L_j is considered as a suspected congested link. For each negative probe $S_i \epsilon S_n$, only if it could be explained by at least one suspected congested link, its result can be considered as correct, otherwise it should be revised to be positive.

V. EXPERIMENTAL EVALUATION

In this section we present our evaluation metrics, experiment methodology and the experiment results.

A. Experiment Methodology

The data source is gathered from NS2. The prediction and localization methods are implemented by Matlab. We measure all the path packet loss rates of every leaf-node pair. If the packet loss rate is greater than 0.05 we denote it by 1, else we denote it by -1. Thus we get a binary matrix to represent the responses of the probes. We consider the following dimensions for our experiment.

- Network topology: Firstly we generate 5 networks with the number of nodes from 100 to 500 by a popular topology generator, GT-ITM. The nodes include both internal nodes and end nodes (leaf nodes).30% of the nodes are leaf nodes and the rest are internal nodes.
- Link loss rate distribution: 90% of the links are selected to be in normal state and the rest are in congested state. The links in normal state are assigned with the loss rate at 0.001 and the congested links are assigned with the loss rate at 0.9.
- Loss model: The loss model is achieved by the error model in NS2. We set up an ftp between every leaf node pair and send 1000 packets at every leaf node. Then we count the packets that received at the other nodes. We assume that the packets are all received or drop (i.e. There is no packet that is still being transported on the network when we stop the NS2 simulation).

B. The Metrics

The metrics include prediction accuracy, and the final revision accuracy.

Let Ne be the number of the total probes, N the number of probes that have been send into the network and CN the number of probes whose responses have been predicted out correctly.

$$\text{Prediction Accuracy} = \frac{\text{CN}}{\text{Ne} - \text{N}}$$

Let P^* be the set of real congested links, $\overline{P}$ be the set of links in normal state and hypothesis be the set of suspected congested links. So the detection accuracy of the suspected congested links and the false positive accuracy can be written as follows.

$$\text{Detection Accuracy} = \frac{|P \cap P^*|}{|P^*|}$$

$$\text{False Positive} = \frac{|P \cap \overline{P}|}{|P^*|}$$

C. The Experiment Results

1) Monitoring all selected probes VS Monitoring the linearly independent group of the selected probes

In this experiment, we compare the prediction accuracy of the method of monitoring all selected probes with the method of only monitoring the linearly independent group of the selected probes. The number of nodes of the network topology is set to be 100. Firstly, we select two groups of probes in the probe set. The number of probes in both groups is 10% of the total probes, and the probes are selected randomly. We monitor all the probes in the first group while in the second group we only monitor the linearly independent group of the probes at first and then we infer out the responses of the rest probes in the second group. Then we use MMMF algorithm to predict the responses of the unselected probes. The result is shown in Fig.4.

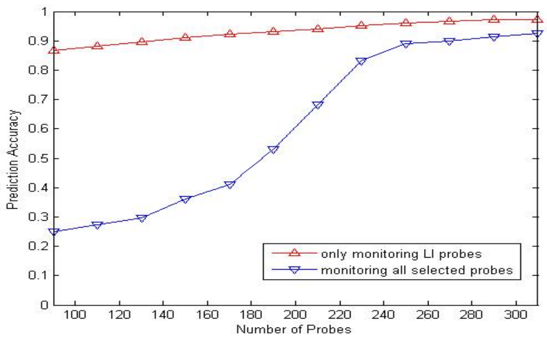


Figure 4. The prediction accuracy by the two methods

Fig. 4 clearly shows that to reach the prediction accuracy level at 0.9 the method of only monitoring linearly independent group of the selected probe only can save as many as 140 probes (about 16% of the total probes) to monitor compared with the method of monitoring all the selected probes. As the selected probes can be inferred from their linearly independent group there is no need to monitor them all, which can save a lot of traffic and cost generated by monitoring probes.

2) The number of probes needed to achieve certain prediction accuracy level

In this experiment, we find out the number of probes needed to achieve certain prediction accuracy level in the 5 networks with the number of nodes from 100 to 500. We first set the prediction accuracy level to be 0.90. Then from the experiment in (1) we randomly select 10% probes in the probe set and find the linearly independent group of them to monitor. After that we infer out the responses of all the selected probes from the monitored probes. Then we use MMMF to predict the responses of the unselected probes. We stop the prediction until the prediction accuracy is up to 0.9. Then we count the number of probes that monitored in each network. The result is shown below in Tab.1.

TABLE I. THE NUMBER OF PROBES NEEDED

Number of Nodes	100	200	300	400	500
Number of Probes	23.0%	22.0%	20.9%	20.0%	19.9%

From Tab. 1 we can clearly see that the number of probes that needed to monitor is only about 20% and the number decreases as the size of network grows. The reason is that the larger the network is, the more hidden factors lie in it and more accurate prediction we can achieve and less number of probes need to monitor. So, significant traffic and cost can be saved while monitoring the network.

3) The detection accuracy of the suspected congested links and false positive accuracy

In this experiment, the data is from the above experiment (i.e. we use MMMF to predict the results of the probes and stop when the prediction accuracy is up to 0.90). Then we find the probes whose results are 1 and calculate α_{L_j} for every link and then find out the suspected congested links set. With the metrics we can get the detection accuracy and false positive accuracy, as shown in table.2

TABLE II. THE DETECTION ACCURACY AND FALSE POSITIVE

Number of Nodes	100	200	300	400	500
Detection accuracy	99.0%	99.6%	99.8%	100%	99.9%
False positive	10.2%	8.3%	4.6%	2.4%	0.4%

From Tab. 2 we can see that the suspected congested links can cover almost 100% of the real congested links and the false positive accuracy is very low, which means in the future congested localization we only need to do the diagnosis in the suspected congested links set.

4) The final revision accuracy

After we find the suspected congested links set in experiment (3) we can revise the prediction results. If the probes predicted to be negative (i.e. 1) passes no suspected congested links we revise it to be positive (i.e. -1). The revision accuracy metric is the same with the prediction accuracy and the final revision accuracy is shown in Tab.3.

TABLE III. THE REVISION ACCURACY

Number of Nodes	100	200	300	400	500
Revision Accuracy	95.5%	95.7%	96.3%	97.1%	97.3%

In Tab.3 we can see that the final revision accuracy greatly improves the prediction accuracy (about 5%~7%) and can be up to 97 % or higher which is quite acceptable for future congested links localizations. Besides, with the size of the network grows the revision accuracy increases. One possible reason is that the larger the network is, the more paths it has, and the α_{L_j} will be more accurate.

VI. CONCLUSIONS

In this paper we propose an algorithm to monitor the overlay networks based on MMMF prediction method. By selecting a small subset of the probes set and finding their maximum linearly independent group to monitor and inferring out the responses of the rest selected probe, then by using MMMF algorithm we can only send 20% of the whole probes to get the responses of all probes. Then we propose a method to find out the suspected congested links set and revise the prediction results and the experiment shows that the final revision can be up to 97% or higher and the suspected congested links can cover almost 100% of the real congested links, which is quite acceptable for future congested links localization for overlay network.

REFERENCES

[1] Irina Rish and Gerald Tesauro, "Estimating End-to-End Performance by Collaborative Prediction with Active Sampling", Integrated Network Management, 2007, pp. 294-303, June 2007.

[2] Jason D. M. Rennie and Nathan Srebro, "Fast Maximum Margin Matrix Factorization for Collaborative Prediction", In Proc of the 22nd ICML, pp. 713-719, 2005.

[3] T. S. E. Ng and H. Zhang, "Predicting Internet Network Distance with Coordinates-based Approaches". INFOCOM, 2002 Proceedings, IEEE.

[4] Choffnes D.R, Sanchez M.A, Bustamante F.E, "Network Positioning from the Edge-An Empirical Study of the Effectiveness of Network Positioning in P2P Systems",INFOCOM, 2010 Proceedings, IEEE, pp,1-5.

[5] Lijiao Liu, Fuke Shen, "An Approach based on locality-awareness and interest-focusing for web content distribution in P2P network", Sarnoff ICCSNA, September 2010.

[6] Francis P, Jamin S, Cheng Jin, Yixin Jin, Raz D, Shavitt Y, Zhang, L, "IDMaps: a global Internet host distance estimation service". Networking, IEEE/ACM, 2001.

[7] Jocelyne Elias, Fabio Martignon and Giuliana Carello, "Very Large-Scale neighborhood search algorithm for design of service overlay networks", Telcommunication System, June 2010.

[8] Nathan Srebro and Tommi S. Jaakola, "Maxmum-Margin Matrix Factorization", NIPS, 2005.

[9] Yan Chen, David Bindel and Randy H. Katz, "Tomography-based overlay network monitoring", Proceedings IMC' 03, ACM SIGCOMM, 2003.

[10] Tian Bu, Nick Duffield, Francesco Lo Presti and Don Towsley, "Network tomography on general topologies", Proceeding SIGMETRICS' 02, ACM, 2002.

[11] PYan Chen, PDavid Bindel, Hanhee Song, PRandy H. Katz, "An algebraic approach to practical and scalable overlay network monitoring", in Proc. ACM SIGCOMM,2004,pp. 55-66, August 2004

[12] J. Pearl, "Probabilistic reasoning in intelligent systems: network of plausible inference", Morgan Kaufmann, San mateo, California, 1988

[13] I. Rish, M.Brodie, Sheng Ma, N. Odinstova, A. Beygelzimer, G. Grabarmik and K. Hernandez, "Adaptive Diagnosis in Distribute Systems", Neural Networks, IEEE Transactions, pp. 1088-1109, September 2005.

[14] A. F. Atiya, Sung Goo Yoo, Kil To Chong, Hyongsuk Kim, "Packet Loss Rate Prediction Using the Sparse Basis Prediction Model", Neural Networks, IEEE, pp. 950-954, May 2007.

[15] C. Boutilier, R.Zemel, and B. Marlin, "Active Collaborative Filtering", In Proc. OF UAI, pp. 98-106,2003.

[16] Lu Cheng, Xue-song Qiu, Luoming Meng and Yan qiao, "Efficient Active Probing for Fault diagnosis in Large Scale and Noisy Networks". INFOCOM, 2010 Proceedings, IEEE, pp. 1-9, March 2010.

[17] Rongmei Zhang, Chunqiang Tang, Y. C. Hu, S. Fahmy and Xiaojun Lin, "Impact of the Inaccuracy of Distance Prediction Algorithm on Internet Applications-an Analytical and Comparative Study", INFOCOM 2006, pp.1-12, April 2006.

[18] Srikanth Kandula and Dina Katabi, "Shrink: A Tool for Failure Diagnosis in IP Networks", In proc. ACM SIGCOMM MineNet Workshop, August 2005

[19] Bo Sun and Zhenghao Zhang, "Probabilistic Diagnosis of Link Loss Using End-to-End Path Measurements and Maximum Likelihood Estimation".ICC, 2009, p:1-5

[20] Hongjie Sun, "An Active Congestion Detection Mechanism Based on Network Tomography". ICIME, 2010

[21] Dennis DeCoste, "Collaborative Prediction Using Ensembles of Maximum Margin Matrix Factorizations", ICML 06 Proceedings, ACM,2006.