Int. J. Communications, Network and System Sciences, 2009, 5, 400-406
doi:10.4236/ijcns.2009.25045 Published Online August 2009 (http://www.SciRP.org/journal/ijcns/).

Network Delay Model for Overlay Network Application

Tian JIN, Haiyan JIN
School of Electronics and Information Engineering, BeiHang University, Beijing, China
Department of Enterprise Management, North China Electric Power University, Beijing, China
Email: *jintian@buaa.edu.cn, jhy*99@263.*net*
Received May 6, 2009; *revised* June 11, 2009; *accepted July* 8, 2009

ABSTRACT

This paper presents a method to model network delay for overlay network application. The network topology measurement technology and network AS information is used to build up model of network delay via AS and geographic distance. Based on global Internet measurement result, we calculated the parameters of the model. Furthermore, the model verification is done by comparing on AS-MMI protocol and HMTP protocol.

Keywords: Network Model, Autonomous System, Overlay Network

1. Introduction

The network delay measurement in large scale overlay system become impossible due to the Internet expansion. So, in this paper we introduce a method of modeling network delay with AS and geographic distance. This method use network topology information to aid us know more information about the connectivity of Internet, and reduce the time and range of network delay measurement [1,2]. Contrast with network bandwidth, delay is more stable. Thus, the system can share a much more stable result of network topology measurement.

First, we will show how to measure the Internet topology. And then, we will show how to modeling the delay of network by the topology and geographic information. At last, we will demonstrate the protocol performance by using the model and ping result. That result shows our model of Internet can reflect network delay quite well.

2. Related Work

By knowing the location of a client host, an application, such as a Web service, could send the user location-based targeted information, classify users based on location, and improve the performance of overlay application.

Previous work on the measurement-based geolocation of Internet hosts uses the positions of reference hosts [3, 4], called landmarks, with well-known geographic location as the possible location estimates for the target host. These might limit the accuracy of the resulting location estimation, because the closest reference host may still be far from the target.

Some applications such as GeoTrack, GeoPing and GeoCluster have developed to map IP-to-location information, but none of them has detailed model. In this paper, we will show how the model could be measured.

3. Topology Measurement

Some researches report that there are 36888 routers and 42269 links in 10 major AS. And number of AS over Internet has exceeded 18000.

Due to the bandwidth inside an AS is much more than inter-connection of AS, we can use the AS relationship to describe the network connectivity. In multimedia communication, network delay always happens between AS inter-connections. So, the network delay can be measured by the AS count of packet passed. According to Oregon University OIX project, we can gather the router information all over the world, and summarize the Internet connectivity relationship.

Though research in topology measurement is very popular, but most of them do not take geography into consideration [5]. With the development of network, the inter-connections between AS will become more and more short. The delay of network will mostly depend on

Table 1. Major AS network router and links.

AS Num	Name	Place	Router	Link
1221	Telstra	(Australia)	2,796	3,000
1239	Sprintlink	(US)	8,355	9,500
1755	Ebone	(Europe)	596	500
2914	Verio	(US)	7,336	6,800
3257	Tiscali	(Europe)	865	700
3356	Level3	(US)	3,446	6,700
3967	Exodus	(US)	900	1,100
4755	VSNL	(India)	121	69
6461	Abovenet	(US)	2,259	1,400
7018	AT&T	(US)	10,214	12,500

geographic delay. So, the network delay can be modeled by AS and geographic distance. We will introduce the technique of deducing network delay by analyzing route information and generating network topology. The key issues contains, how to analyze AS information; how to generate AS connectivity and how to model geographic distance of AS.

The network topology measurement take four stages: The first is gathering Internet route table, analyzing BGP route and mapping between IP and AS. The second stage is analyzing BGP route and AS path. We can deduce AS connectivity at this stage and calculate shortest connectivity path of AS. The third stage is finding AS geographic information according to AS registry information. The last stage is calculating AS communication delay by AS and geographic information.

3.1. Gathering Route Information

BGP is an external gateway routing protocol of TCP/IP. It is designed for solving large scale network route problem. The BGP route is synchronized all over Internet, so it can reflect the topology of current Internet.

The path information of two Autonomous Systems is recorded in route table of all core routers. We can get the AS connectivity by analyzing these information. If the information is gathering from difference routers over the world, the topology will be more accurate.

Oregon University's OIX project provided summary of some core router's BGP route table. We can extract mapping between IP and AS from that table. For instance, following mapping can be obtained by previous route Table 2:

Table 2. IP-AS mapping from route table.

IP Address	AS Num
6.1.0.0/16	668
6.2.0.0/22	668
6.3.0.0/18	668
6.4.0.0/16	668

Table 3. AS peers listed in route table.

AS Num	Peer AS Num
4538	9407
9407	7660
7660	11537
11537	668
19782	11537

3.2. AS Information Analyzing

In BGP route table, AS path is the route path at AS level. So, analyzing AS path information, we can deduce the connectivity of AS. For instance, we can summarize the following connectivity of AS from the route Table 3:

According to BGP route data of Oregon University at year 2004, there are 18431 AS and 39886 links in Internet. The data briefly described how the Internet is connected at that time.

3.3. Analyzing AS Connectivity

We can calculate the shortest path between two AS by using shortest tree algorithm. The route protocol (such as OSPF) has taken the shortest path into consideration. The result will reflect the theoretical minimal distance between two AS, which is similar as actual distance [6,7]. We use the number of AS passed in communication (also called as AS length) to represent the distance of AS. The BGP route in Internet might always change, but most BGP route change does not interfere with AS length. The length of AS path is stable in most time.

We can deduce AS length by previous route example:

The distance of AS can be calculated via following algorithm:

Define $Path(u)$ is, the set of AS directly connect with AS u.

Define $\mathrm{Len}(u,v)$ is, distance of AS u and AS v.

If "u v" or "v u" exist in AS path of BGP route Table 4,

Table 4. AS distance according to route table.

AS Num	AS Num	Dis	AS Num	AS Num	Dis
4538	9407	1	9407	19782	3
4538	7660	2	7660	11537	1
4538	11537	3	7660	668	2
4538	668	4	7660	19782	2
4538	19782	4	11537	668	1
9407	7660	1	11537	19782	2
9407	11537	2	668	19782	2
9407	668	3			

then $u \in Path(v)$.

If $u \in Path(v)$, then $\text{Len}(u,v) = 1$. If $u \notin Path(v)$, then $\text{Len}(u,v) = 0$.

If $w \in Path(v)$, and $\text{Len}(u,v) \neq 0$, and $\text{Len}(u,w) = 0$, then $\text{Len}(u,w) = \text{Len}(u,v) + 1$.

If $w \in Path(v)$, and $\text{Len}(u,v) \neq 0$, and $\text{Len}(u,w) \neq 0$, then $\text{Len}(u,w) = \min(\text{Len}(u,v)+1, Len(u,w))$.

According to BGP route data from Oregon University's OIX project at year 2004, the average distance of AS is 3.765305, AS distance of 18431 AS is shown in Table 5:

3.4. Using Geographic Information

With the development of Internet and speed, the geographic ratio in communication delay will be increased. So, we must take the geographic delay into account [8]. For simple calculation, we can get the AS number by the IP address, and then get country information by AS registry. And the geographic distance can be calculated by the longitude and latitude information of that country. For mote accurate calculation, we can use city information rather than country information for calculation.

The information of IP Address and city, AS number and country information can be retrieval from whois server. The latitude and longitude information can be retrieval from NetGeo and other projects [6,7]. With the earth's radius and following formula, we can calculate the theoretical distance of two nodes:

DISTANCE=R*ARCOS[SIN(A)SIN(C)
+COS(A)COS(C)COS(B-D)];

R is earth's mean radius: 6371km.

(A, B) is latitude and longitude of node 1.

(C, D) is latitude and longitude of node 2.

The network delay consists of geographic delay and AS communication delay. So, we can calculate the backbone link delay and geographic delay by topology measurement result. The following formula shows the network delay model:

RTT＝T(N) + P*D

The P is related with current network condition.

D is the geographic distance between two nodes.

T(N) is the delay of AS length N between two nodes.

Based on measurement result, value of P and T is P=20us/km, T(2)=10ms, T(3)=55ms, T(4)=78ms, T(5)= 92ms. The Internet is in a changing state, so previous parameter will also change with Internet's development. That value only reflects current Internet measurement result from Chapter 4.

Table 5. Summary of AS distance.

Dis	Count	Ratio	Dis	Count	Ratio
1	79772	0.0235%	8	133926	0.0394%
2	18585690	5.4715%	9	11467	0.0034%
3	121181350	35.6748%	10	792	0.0002%
4	135241350	39.8139%	11	65	0.0000%
5	52057273	15.3252%	12	8	0.0000%
6	10960601	0.32267%	13	1	0.0000%
7	1431035	0.4213%			

4. Network Delay Analyze

The topology measurement is related with reality network, so we can not setup the topology measurement test with simulation. We use PlanetLab [9] node as source of reality measurement, and use ping (or extended tcpping [10]) to measure reality network delay. The comparison will confirm the relationship between AS length and network delay.

Due to the ping firewall in Internet, some ping measurement will not reach to some host. To get the better measurement result, we use TCP instead of ICMP for node delay test [1,2,11], and use connection confirm time instead of RTT time to represent network delay.

We use ScriptRoute's packet data service to measure the network delay between node installed ScriptRoute. The TCP port 3355 is used for tcpping measurement instead of ICMP ping. The node from 83 AS and 129 nodes joined the test, they are distributed as Figure 1:

We select 14 nodes' result, and make detail analyses for network topology model in Table 6 and Table 7.

According to AS distance calculation algorithm at previous chapter, we can get the AS distance of the nodes in Table 8 and Table 9.

After we summarize AS distance, geographic distance and RTT information from previous table, we can calculate network delay data from 81 nodes. Among the result, most of AS distance is 2 to 4, which is 88.9% of all data. The distribution of AS distance is shown at Table 10.

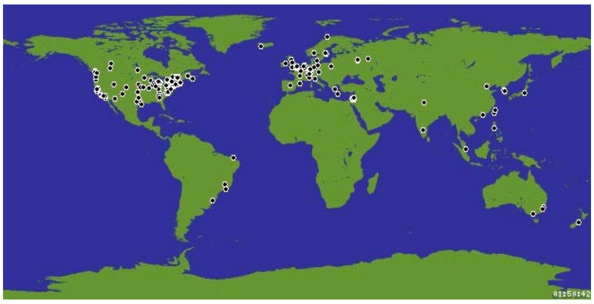

Figure 1. Topology measurement node distribution.

Table 6. Topology measurement node list.

	IP	AS	Area	Lat.	Long.
S1	128.208.4.155	73	us	47.65	-122.31
S2	129.97.75.240	549	ca	45.35	-72.52
S3	142.103.2.2	271	ca	49.26	-123.23
S4	212.192.241.155	2848	ru	55.65	37.5
S5	165.132.126.58	4665	kr	37.53	127
S6	140.109.17.180	9264	tw	25.02	121.37
S7	130.161.40.154	1128	nl	52.02	4.35
S8	132.72.23.10	378	il	31.5	34.75
S9	198.32.154.195	11537	us	40.72	-73.99
S10	195.37.16.101	680	de	48.58	13.47
S11	140.192.37.134	20130	us	41.88	-87.63
S12	206.117.37.5	226	us	33.98	-118.46
S13	128.83.143.153	18	us	30.28	-97.74
S14	202.141.62.35	23731	in	29.85	77.9

Table 7. The geographic distance (km).

	R1	R2	R3	R4	R5	R6	R7
S1	0	3754	192	8375	8322	9740	7831
S2	3754	0	3776	7022	10594	12065	5424
S3	192	3776	0	8213	8156	9581	7714
S4	8375	7022	8213	0	6621	7359	2190
S5	8322	10594	8156	6621	0	1490	8615
S6	9740	12065	9581	7359	1490	0	9497
S7	7831	5424	7714	2190	8615	9497	0
S8	10920	8765	10774	2694	8121	8289	3356
S9	3864	528	3914	7511	11056	12535	5840
S10	8497	6175	8372	1807	8415	9158	751
S11	2787	1274	2861	8003	10512	12001	6595
S12	1554	4069	1744	9790	9578	10910	8958
S13	2848	2757	3006	9559	11156	12587	8158
S14	11163	11135	10972	4271	4570	4299	6347

	R8	R9	R10	R11	R12	R13	R14
S1	10920	3864	8497	2787	1554	2848	11163
S2	8765	528	6175	1274	4069	2757	11135
S3	10774	3914	8372	2861	1744	3006	10972
S4	2694	7511	1807	8003	9790	9559	4271
S5	8121	11056	8415	10512	9578	11156	4570
S6	8289	12535	9158	12001	10910	12587	4299
S7	3356	5840	751	6595	8958	8158	6347
S8	0	9156	2607	9951	12212	11511	4104
S9	9156	0	6587	1145	3958	2432	11651
S10	2607	6587	0	7346	9673	8910	5740
S11	9951	1145	7346	0	2825	1575	11902
S12	12212	3958	9673	2825	0	1990	12712
S13	11511	2432	8910	1575	1990	0	13313
S14	4104	11651	5740	11902	12712	13313	0

Distance calculated by longitude and latitude (km)

Table 8. Topology experiment AS distance.

	R1	R2	R3	R4	R5	R6	R7
S1	0	4	3	4	4	3	4
S2	4	0	3	5	5	3	5
S3	3	3	0	5	4	2	5
S4	4	5	5	0	4	3	4
S5	4	5	4	4	0	3	4
S6	3	3	2	3	3	0	3
S7	4	5	5	4	4	3	0
S8	4	5	3	3	4	3	3
S9	2	4	2	2	3	1	2
S10	3	4	2	3	3	2	3
S11	4	4	3	4	4	2	3
S12	3	3	2	3	3	2	3
S13	4	5	4	5	4	3	5
S14	4	4	3	4	4	3	4

	R8	R9	R10	R11	R12	R13	R14
S1	4	2	3	4	3	4	4
S2	5	4	4	4	3	5	4
S3	3	2	2	3	2	4	3
S4	3	2	3	4	3	5	4
S5	4	3	3	4	3	4	3
S6	3	1	2	2	2	3	2
S7	3	2	3	3	3	5	4
S8	0	2	2	4	3	4	5
S9	2	0	2	2	2	2	3
S10	2	2	0	3	3	4	3
S11	4	2	3	0	3	4	3
S12	3	2	3	3	0	3	3
S13	4	2	4	4	3	0	5
S14	5	3	3	3	3	5	0

Table 9. RTT time (ms) measured by tcpping.

R1	R2	R3	R4	R5	R6	R7
0	71.349	30.971	199.615	148.035	246.747	158.807
71.349	0	66.913	139.948	197.421	310.892	E
30.976	66.913	0	194.833	162.055	255.681	140.498
198.063	139.948	191.213	0	309.407	316.383	65.828
152.266	197.421	145.103	309.407	0	120.777	284.129
246.775	310.892	255.564	347.639	135.715	0	300.238
158.807	E	140.498	65.828	284.129	300.238	0
223.482	117.406	213.658	105.605	253.891	270.334	70.366
56.205	E	86.534	169.715	E	225.518	E
184.487	117.094	171.863	63.994	324.602	331.154	28.366
46.925	12.234	31.679	138.64	248.451	186.586	102.759
26.13	91.096	51.684	194.72	173.54	157.166	169.848
60.169	E	90.399	166.525	189.145	229.427	E
743.486	770.898	887.795	912.26	E	938.597	909.354
R8	**R9**	**R10**	**R11**	**R12**	**R13**	**R14**
223.543	56.205	184.542	50.386	26.056	60.169	742.947
117.406	E	117.094	12.234	91.096	E	770.898
202.403	86.534	172.014	49.707	56.315	90.399	917.678
104.881	169.715	65.103	133.095	185.149	166.525	922.6
368.949	E	319.806	178.883	176.751	189.145	896.655
372.332	225.518	331.148	197.726	196.177	229.427	938.417
70.366	E	28.366	102.759	169.848	E	909.354
0	195.149	83.075	136.026	218.953	150.46	921.899
195.149	0	E	29.099	32.648	E	E
83.398	E	0	135.035	187.613	157.074	912.963
174.036	29.099	126.183	0	55.733	28.134	728.829
164.449	32.648	165.351	45.122	0	35.5	684.068
150.46	E	157.074	28.134	35.5	0	776.616
936.419	E	907.139	803.945	751.519	776.616	0

Table 10. Topology experiment result summary.

AS Distance	Count	Ratio	Avg. Geographic Distance (km)	Avg. Network Delay (ms)
1	1	1.23%	12534	225.518
2	13	16.05%	6455	138.98
3	33	40.74%	6799	206.03
4	26	32.10%	7058	257.67
5	8	9.88%	8660	333.25

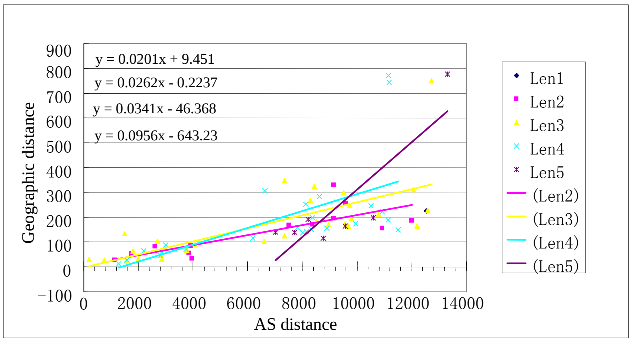

Figure 2. Network delay model.

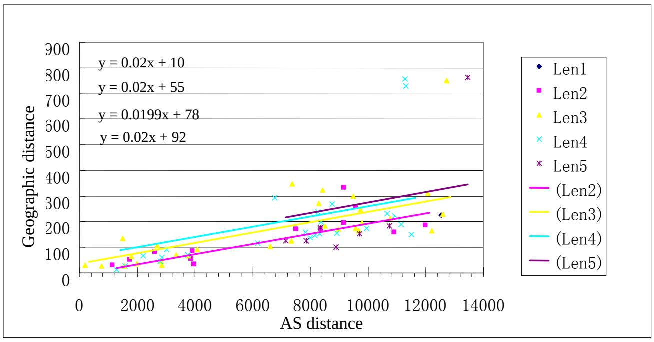

Figure 3. Normalized network delay model.

We can find from summary that the average geographic distance and network delay are dramatically increased when AS distance is increased. And the ratio of AS distance is similar as global value, in which ratio of AS distance from 1~5 is (0.0235%, 5.4715%, 35.6748%, 39.8139%, 15.3252%). From the result, we can draw the conclusion that there is relationship among AS distance, geographic distance and network delay.

Then we divided the measurement result into several groups based on AS distance. We can get the model of geographic distance and network delay in different AS distance by using linear regression method.

Due to the network delay cause by geographic distance is direct proportion with geographic distance, the different Linear Regression line should have same slope. So, we slightly adjust all the Linear Regression line's slope into same value -0.02. The procedure is called as normalization. The model after normalization is:

From the result, the network delay can be calculated via following formula:

RTT＝T(N) + P*D

Inside, P is the parameter reflect geographic distance, P=20us/km.

D is the geographic distance of two nodes.

T(N) is the network delay caused by AS distance N, where T(2)=10ms, T(3)=55ms, T(4)=78ms, T(5)=92ms.

The result shows network delay seemed to be random, but it is related with geographic distance and AS distance. Commonly speaking, average network RTT delay will increase 0.02ms when geographic distance increase 1 kilometer; average network RTT delay will increase 11~45ms when AS distance increase 1.

5. Model Verification

We use the model of network delay in overlay network protocol design to test it [12]. We design an AS-MMI protocol based on MMI protocol for large scale multimedia communication. The gateway selection algorithm of AS-MMI protocol will use the model. By contrast, we use HMTP [13,14] for comparison, which will use partly measured RTT delay from previous table. The two protocols consider Shortest Path Tree (SPT) algorithm as best solution.

The performance experiment takes SPT algorithm for reference. And there are four factors to evaluate the time and quality of spanning tree.

Tree Cost means the cost of a tree is the sum of delays on the tree's links. Tree cost is a convenient metric to measure total network resource consumption of a tree.

Tree Delay means the delay from one member to another along the tree. The ratio between tree delay and unicast shortest path delay is delay ratio.

Tree Time means the total time used to build the span-

Table 11. Spanning tree result.

	HMTP	*AS-MMI*
Tree Cost	1074.317	1009.897
Tree Cost Ratio	1.67	1.57
Tree Delay	563.677	712.624
Tree Delay Ratio	1.18	1.49
Tree Time	2817.704	1009.897
Hit Node	4	5
Hit Ratio	33.3%	41.7%

ning tree. The time can reflect the effectiveness of protocol.

Hit Ratio: If the delay of the node and its parent node is no more than 10% of delay in SPT algorithm. We considerate this node as a Hit node, which means the node's delay is quite low. The ratio of hit node reflects how many nodes delay is acceptable.

Verification result is shown in Table 11:

We can find from previous table, AS-MMI protocol use much more short time in build up spanning tree than HMTP protocol. In spanning tree cost and hit ratio, the result of two protocols is similar. But the AS-MMI protocol will cause a little higher tree delay. That means the network model used by AS-MMI protocol can reflect most node's network delay, and tree build speed is much faster. But the AS distance is only theoretical distance which might be different as actual distance, some nodes' network delay might have errors. These nodes' delay caused tree delay increase dramatically, but does not have much effect interfere with tree cost and hit ratio.

Overall, the network model used by AS-MMI protocol can provide a fast method to check the network delay between different nodes. And the result shows it is accurate for most nodes of AS-MMI protocol.

6. Acknowledgment

Many thanks to the network and research support from State Key Laboratory of Software Development Environment, BeiHang University.

7. References

[1] M. Jain and C. Dovrolis, "End to end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput," Technical Report, University of Delaware, February 2002.

[2] M. Jain and C. Dovrolis, "Pathload: A measurement tool for end-to-end available bandwidth," In Proceedings of Passive and Active Measurements (PAM) Workshop, March 2002.

[3] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Improving the accuracy of measurement-based geographic location of Internet hosts," Computer Networks, Vol. 47, No. 4, pp. 503–523, March 2005.

[4] V. N. Padmanabhan and L. Subramanian, "An investigation of geographic mapping techniques for internet hosts," in Proceedings ACM SIGCOMM, San Diego, CA, pp. 173–185, August 2001.

[5] M. Kwon and S. Fahmy, "Topology-aware overlay networks for group communication," In ACM NOSSDAV, 2002.

[6] P. Francis, S. Jamin, V. Paxon, L. Zhang, D. Gryniewicz, and Y. Jin, "An architecture for a global internet host distance estimation service," Proceedings 18th IEEE Infocom'99, 1999.

[7] P. Francis, S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang, "Idmaps: A global internet host distance estimation service," IEEE/ACM Transactions on Networking, October 2001.

[8] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson, "Inferring link weights using end-to-end measurements," In ACM SIGCOMM Internet Measurement Workshop, 2002.

[9] Planet Lab Project, http://www.planet-lab.org/.

[10] ScriptRoute Project, http://www.scriptroute.org/.

[11] C. Dovrolis, P. Ramanathan, and D. Moore. "What do packet dispersion techniques measure?" Proceedings of IEEE INFOCOM, April 2001.

[12] S. Shi, "Design of overlay networks for internet multicast," Ph. D. dissertation, Washington University, August 2002.

[13] B. Zhang, S. Jamin, and L. Zhang, "Host multicast: A framework delivering multicast to end users," Proceedings of IEEE INFOCOM '02, 2002.

[14] B. Zhang, W. Wang, S. Jamin, D. Massey, and L. Zhang, "Universal IP multicast delivery," Computer Networks: The International Journal of Computer and Telecommunications Networking, pp. 781–806, April 2006.

[15] T. Cicic, "Networklevel multicast deployment and recovery," Ph. D. dissertation, October 2001.

