

Landmark-based End-to-End Bandwidth Inference

Mohammad Malli, Chadi Barakat, Walid Dabbous
Projet Planète, INRIA-Sophia Antipolis, France

The end-to-end delay among a set of peers can be easily and scalably estimated using a landmark approach [2], [3], [4], [5]. For example, one can calculate coordinates for peers and infer the delay as being the Euclidean distance separating them. Peers' coordinates are deduced from delay measurements to a small number N of landmarks $L\{L_1, \dots, L_N\}$. We wonder whether it is feasible to use such distributed solution for bandwidth estimation. This requires that each peer determines its bandwidth vector by measuring the direct and reverse bandwidth on its path with each landmark. The bandwidth vectors of two peers is then used to estimate the bandwidth between them in both ways. If we arrive to design such solution, we will be able to infer the bandwidth between peers in a manner which is (i) scalable since the system overhead will be linear with the number of peers in the system, and (ii) easy to implement since peers will not need to know and probe each other; any node can estimate the bandwidth between any two peers based on their bandwidth vectors.

For a couple of peers, we denote by (i) *direct path* the network path that joins them directly using IP routing, and by (ii) *indirect path* the path that joins them via a landmark node. N indirect paths (N being the number of landmarks) are assigned to each direct path.

Our idea is to estimate the bandwidth of the direct path using those of the indirect paths. The indirect paths that have more links common with the direct path, are assigned more weight in the estimation function. We consider different estimation functions and we study the impact of the landmarks' locations on the accuracy of the estimations.

For a direct path joining two peers, we estimate its end-to-end bandwidth using the following class of linear functions:

$$EB = \sum_{i=1}^N P_i \cdot BB_i, \quad (1)$$

where BB_i is the bandwidth of the indirect path that passes by landmark L_i , and P_i is a normalized weight (i.e., $\sum_{i=1}^N P_i = 1$) assigned to this indirect path based on the location of its corresponding landmark with respect to the two peers. By varying the weight P_i , we are able to cover different policies for bandwidth estimation ranging from the one that gives the same priority to all landmarks to the one that privileges the landmark that we deem the most suitable for the direct path bandwidth inference.

In [5], the authors observe that 8 to 12 landmarks are enough for a good delay estimation at the scale of the Internet. We consider the same number of landmarks for bandwidth estimation. Therefore, we take 8 Planetlab nodes [1] selected

from different European countries as landmarks. We also take 14 Planetlab nodes completely distributed in Europe as peers. Each of these peers measures the round trip delay RTT and the direct and reverse available bandwidth ABw^1 to 34 Planetlab nodes distributed worldwide. This leads to 476 measured paths. Then, we infer the bandwidth of these paths ($ABw_{estimated}$) using Equation (1) and we compare the estimations with the measured values ($ABw_{measured}$). Furthermore, we study the correlation between the estimation accuracy and the landmarks' locations.

Our landmark nodes are chosen with the main concern to have a high bandwidth connectivity to the Internet. This is an important requirement since we want to avoid having the bottleneck of an indirect path determined by the region around the landmark. We want it to be rather determined by the regions around the peers. In fact, it is more probable that the latter regions are common with the direct path compared to that around the landmark.

We consider different forms of the weights P_i , and consequently of the end-to-end bandwidth estimation function. By doing that, we are able to study the correlation between the estimation accuracy and the locations of the landmarks. We divide the study into two main parts: (i) the estimation function depends on how much the direct path and the indirect paths' delays are close to each other, (ii) the estimation function depends on the delay closeness between the landmarks and the path end points.

One possibility is to estimate the end-to-end bandwidth of a direct path using that of the indirect path having the shortest delay. Even though we found satisfactory results, we believe this method is not sufficient for providing accurate estimation since direct IP routing may lead to an end-to-end delay larger than the one of the shortest indirect path, with both paths having different sets of links and hence different bottlenecks. The accuracy could improve by considering more than one indirect path in the estimation function while assigning more weight to those having shorter delays. This consideration is mainly recommended when there are more than one indirect path having delays on the order of that of the shortest one.

Thus, we consider all the N indirect paths in the bandwidth estimation function (Equation 1) with the following expression for the weight P_i :

$$P_i = \frac{C_i}{\sum_{i=1}^N C_i}, \quad \text{for } i = \{1, \dots, N\} \quad (2)$$

¹This is the remaining bandwidth left on a path between two nodes. It is determined by the residual bandwidth at the bottleneck link. In the rest of the paper, the term bandwidth is used to denote the available bandwidth.

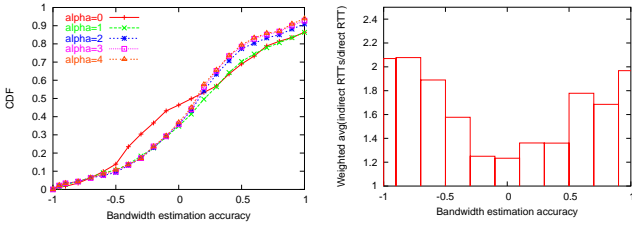


Fig. 1. ABw estimation accuracy

Fig. 2. ABw estimation accuracy

where,

$$C_i = \left(\frac{RR_{min}}{RR_i} \right)^\alpha, \quad (3)$$

RR_i is the round trip delay of the indirect path that passes by the landmark L_i , RR_{min} is that of the shortest indirect path among the N indirect paths, and α is a positive real number.

We draw in Figure 1 the CDF of the estimation accuracy, which is calculated as $(ABw_{estimated} - ABw_{measured})/ABw_{measured}$ for different values of α . The figure shows that when the α parameter increases, the estimation accuracy improves. This is expected since when $\alpha = 0$, the bandwidth contribution of all indirect paths gets the same weight, and when α becomes large, the indirect paths having shorter delays, and hence better representation of the direct path, get more weight than those having larger delays. For $\alpha > 3$, we observe that the results become steady. This can be explained by the fact that the estimation becomes only dependent on the indirect paths having a delay on the order of that of the shortest indirect path. For $\alpha = 4$, the figure shows that 39.63% of the estimations are accurate within 25% and 68.36% of the estimations are accurate within 50%.

To show the correlation between the estimation accuracy and the difference in the delay between the direct and the indirect paths, we plot Figure 2 for the case $\alpha = 4$. For an estimation accuracy interval (on the x axis) of length 0.2, the y axis shows the sum $\sum_{i=1}^N P_i \cdot (RR_i/R_d)$, which is a weighted average of the ratio of the indirect paths' delays and the delay of the direct path (R_d). This sum is averaged over all estimations laying with the 0.2 intervals. The figure shows a clear correlation between the two entities plotted on the x and y axis. This means that when some landmarks are located such that the delay of their correspondent indirect paths is close to that of the direct path, the estimation accuracy becomes better.

Now, instead of relying our estimation on the end-to-end delay of the indirect paths, we focus on how close landmarks are to the direct path end points. Thus, for each pair of peers, we consider the N indirect paths in the bandwidth estimation model after assigning more weight for those going through landmarks that are closer to the two peers under consideration. Basically, we want to check if these latter indirect paths are more representative of the direct path than the ones having smaller end-to-end delays. We express the coefficients C_i as:

$$C_i = \left(\frac{R_{min}}{R_i} \right)^\alpha, \quad (4)$$

where,

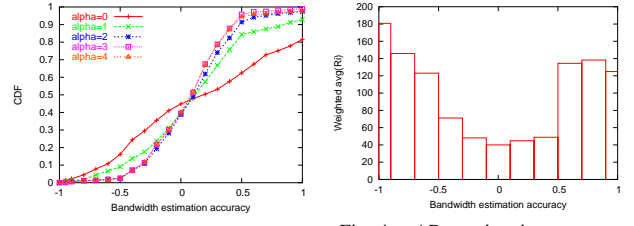


Fig. 3. ABw estimation accuracy

Fig. 4. ABw estimation accuracy

$$R_i = \min(R_{xi}, R_{yi}), \quad (5)$$

R_{xi} represents the round trip delay between the peer x and the landmark L_i , and

$$R_{min} = \min_{i=1..N} R_i. \quad (6)$$

We recalculate the P_i function (Equation 2) and subsequently the estimation function (Equation 1) with these new coefficients C_i . Then, we plot in Figure 3 the CDF of the accuracy function for all the bandwidth estimations and for different values of α . As before, when α increases, the indirect paths having landmarks close to one of the two peers get more weight. Again, the figure shows that the results become stationary for $\alpha > 3$. Furthermore, one can see clearly that the results are better than the previous case; 56.54% of the estimations are accurate within 25% and 92.62% of the estimations are accurate within 50%.

Figure 4 shows the correlation between the estimation accuracy and the landmarks' closeness to the two peers for the case $\alpha = 4$. For an estimation accuracy interval (on the x axis) of length 0.2, the y axis shows $\sum_{i=1}^N P_i \cdot R_i$ averaged over the estimations laying inside the interval. We can observe a clear correlation between the two entities on the x and y axis. This means that when some landmarks (among the N) are close to the path extremities, the estimation accuracy improves.

Thus, one can conclude that the bandwidth estimation model should attribute more weight to the indirect path having its landmark close to one of the two peers more than to the one having small end-to-end delay. Our interpretation is that the route provided by IP and those passing by the landmarks could be disjunct even if their delays are close to each other. On the other hand, the indirect paths going through landmarks that are close to peers are more expected to provide better representation of the direct path. Our future research aims at investigating more this problem and testing the validity of our solution in a real application as Peer-to-Peer file sharing.

REFERENCES

- [1] An open, distributed platform for developing, deploying and accessing planetary-scale network services, see <http://www.planet-lab.org/>.
- [2] E. Ng and H. Zhang: Predicting Internet network distance with coordinates-based approaches, IEEE Infocom, 2002.
- [3] L. Tang, and M. Crovella: Virtual Landmarks for the Internet, IMC'03, 2003.
- [4] Y. Shavitt, and T. Tankel: Big-bang simulation for embedding network distances in euclidean space, IEEE Infocom, 2004.
- [5] S. Ratnasamy and M. Handly and R. Karp and S. Shenker: Topologically-Aware Overlay Construction and Server Selection, IEEE Infocom, 2002.