

TULIP Distance Variation with Alpha

We have been analyzing the delay measurements collected from different landmarks in different regions of the world. Our basic motive was to get intra-regional and inter-regional value of alpha (a ratio between delay and distance). In order to get that we analyzed delay to distance ratio from world to landmark and landmark to whole world categorized in regions. In our discussion, we came to a conclusion that it is logically incorrect, why? The first question that we should ask: Is it possible that any landmark in Australia is selected as one of the three or four final landmarks to estimate the geographic coordinates of a host in Africa? This can never happen because the final landmarks that are chosen to be used in the geo-location algorithm are the ones that are located close by (have the shortest RTTs). So, if no landmark is going to be selected to find any host outside a certain radius there is not point of finding the correlation between distance and RTT (i.e. the alpha value) outside that radius for that landmark, because this makes things more difficult but still irrelevant. So, first we'll have to find that radius around a landmark, for which it is sensible to consider the landmark will be useful, and we need to focus on intra-locality instead of inter/intra regional aspect. Then, we have to report the distance to delay correlation within that radius and this will be the justification of Tiering as well.

For the last few weeks, we were trying to identify a good correlation between delay and distance across the globe. Our findings are that this correlation is poor (i.e. a scatter plot of distance vs RTT for multiple pairs of hosts does not closely follow a straight line, an exponential, or a power law etc).. So, right now instead of spending more and more time investigating that correlation let's just state that it is a weak correlation and show results confirming this. We will also provide justification why the inter and intraregional correlations between distance and RTT are often related to the geographical Internet path connectivity between regions. This will give a justification of that weak correlation.

As a parallel step, we are going to analyze the delay as a function of time over many years based upon the available historical data of different regions. We shall try to find some pattern, if there are any, during that analysis. We may expect some correlation of alpha and time between given pairs of hosts, given the exponential improvements in internet connectivity and reductions in RTT over time. If we do not find such a correlation, that is also interesting though surprising. As a last step, we shall get the results from other Geolocation techniques in hand and shall justify our findings based upon those results. We are going to write a 5 page paper for GLOBCOMM 2010 and the last date of submission to this conference is March 15.