Revisiting RTT models

Esteban Carisimo, Sebastián P. Grynberg, J.I Alvarez-Hamelin
INTECIN (UBA-CONICET), Universidad de Buenos Aires
Buenos Aires, Argentina
Email: *{carisimo,sgrynberg,ihameli}@cnet.fi.uba.ar

Among 56 studied cases we have observed that stable distributions fit better to the RTT than previous proposed models (i.e., Weibull distribution [1], Lognormal distribution [2]). In order to compare the accuracy of each model, we used the Total Variation Distance [3]. Finally, we show a simple active method, based on RTTs, which is useful to estimate link utilization.

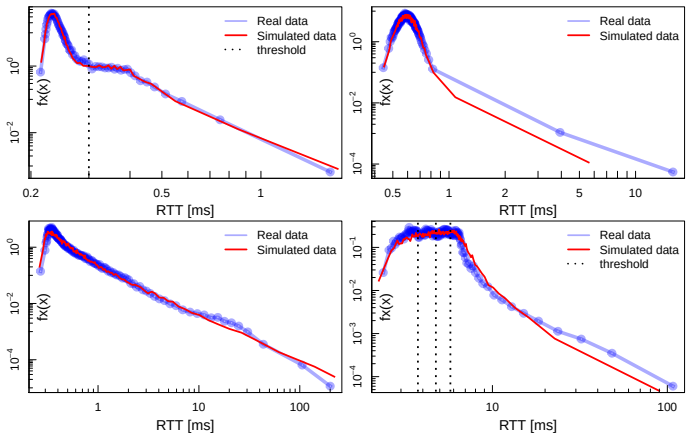

Fig. 1. From 14 Planet-lab nodes we performed a traceroute per second up to the 4^{th} hop, aiming to avoid backward path variation. These plots show histograms of RTTs (blue) for 4 different kinds of behaviors observed. Red curves correspond to histograms based on samples from either a stable distribution or a mixture of stable distributions that fit emipirical data. In some cases stable parameters we used Kogon-McCulloch method [4], [5].

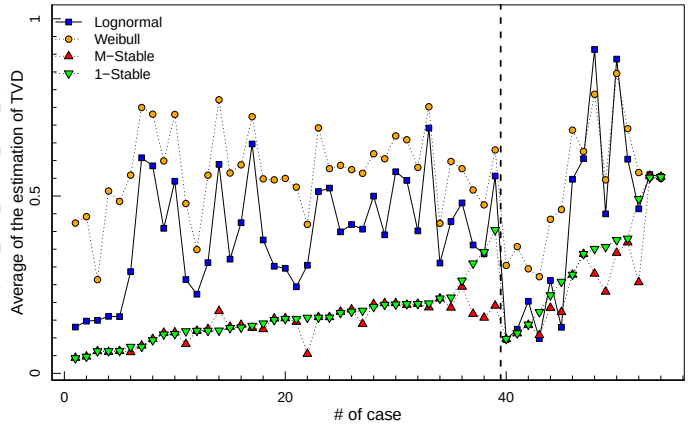

Fig. 2. We present the estimation of Total Variation Distance (TVD) for analyzed models: Weibull, Lognormal, 1-Stable (the dataset modeled with just one stable distribution) and M-Stable (the dataset modeled using a mixture of stable distributions). The dashed line distinguishes the cases where the Kogon-McCulloch method [4], [5]was used (to the left) from those where a different estimation method was employed (to the right of the dashed line).

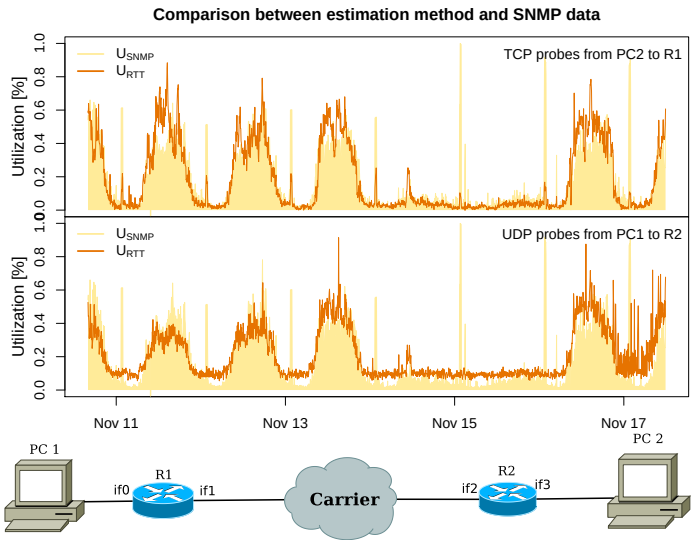

Fig. 3. Taking into account that RTT vary with traffic load [6], we developed an simple method to estimate link utilization. U_{RTT} was estimated as a ratio of RTTs over a threshold during a time window (300 packets in 5 minutes). We verified this method using U_{SNMP}: information obtained through polling the router's inferfaces by SNMP.

Figure 1 presents different patterns that we found on the 56 cases. Figure 2 shows that 1-Stable and M-Stable models reproduce empirical data more accurately than previous models. Indeed, 80% of the samples (TVD= 0.2) can be reproduced on 83% and 72% of the cases with M-Stable and 1-Stable respectively, while 19% and 0% with Lognormal and Weibull. We have observed that the tail of the stable distribution is related with traffic load. Therefore, if we set a threshold based on stable distribution parameters, we will be able to determinate which packet delayed on the router, leading to results in Figure 3.

References

[1] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, F. Tobagi, and C. Diot, "Analysis of measured single-hop delay from an operational backbone network," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2. IEEE, 2002, pp. 535–544.

[2] R. Fontugne, J. Mazel, and K. Fukuda, "An empirical mixture model for large-scale rtt measurements," in *INFOCOM 2015 IEEE*. IEEE, 2015, pp. 2470–2478.

[3] H. Thorisson, *Coupling, stationarity, and regeneration*. Springer New York, 2000.

[4] S. M. Kogon and D. B. Williams, "Characteristic function based estimation of stable distribution parameters," *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pp. 311–338, 1998.

[5] J. H. McCulloch, "Simple consistent estimators of stable distribution parameters," *Communications in Statistics-Simulation and Computation*, vol. 15, no. 4, pp. 1109–1136, 1986.

[6] M. Luckie, A. Dhamdhere, D. Clark, B. Huffaker *et al.*, "Challenges in inferring internet interdomain congestion," in *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, 2014, pp. 15–22.