# Worst Case Round Trip Time Prediction and Statistical Analysis Using Extreme Values Theory

Zhen Zhu, Albert M.K. Cheng
Real-Time System Laboratory, Department of Computer Science
University of Houston, Houston, Texas, 77204, USA
{zhu, cheng}@cs.uh.edu

**Abstract**

Scheduling packets using the Least-Laxity-First (LLF) strategy in a real-time network requires knowing the Worst Case Transferring Time (WCTT). In this paper, the Worst Case Round Trip Time (WCRTT) is analyzed and predicted by Extreme Values Theory for the first time. Two Extreme Values analysis methods are conducted. They are Block Maxima and Threshold Exceedances. The Block Maxima helps to compute the Return Level and the Threshold Exceedances supplies the WCRTT associated with a Confidence Level, which makes the WCRTT prediction more reliable and convincible. Two different link's data are compared and conclusions are drawn. It is shown that the Extreme Values Theory is successful in modeling the WCRTT and could be utilized in other network delay predictions as well.

## 1. Introduction

Today's Internet provides best-effort service only. In the coming decades, methods for QoS guarantees are going to bring more reliable and higher quality data service [13]. Transmitting packets in real-time is a challenging task since the missing of a delivery deadline may not be tolerated.

Widely studied scheduling schemes include naïve First-Come-First-Serve (FCFS), Early-Deadline-First (EDF) and Static-Priority (SP) [8]. However these methods do not consider the urgency of a packet. Recently Least-Laxity-First (LLF) [10] was also suggested for dealing with the network packet scheduling. It assumes that each packet is labeled with its delivery deadline $D$. In each node, we first estimate the delay $C$ from current node to the destination, then calculate laxity $L = D - C$. We discard the packet if $L < 0$, otherwise, add the packet into the ready queue. Here comes the question: how can we know the exact value of delay $C$ in the above scheme?

Generally, the exact $C$ will never be known ahead of the transmission because of extremely complicated and various network conditions leading to uncertainty for delay estimation. Hence, it becomes difficult to implement the LLF scheme in real life. We note that there are several factors affecting the value $C$, like the physical clock time, the geographical distance, the number of hops, and so on. To make life easier, here we only consider the situation with an invariant travel route. Therefore, the delay will only vary with the time and the packet size. Our experimental result shows that the relationship between $C$ and the packet size is given by a simple linear relation for a certain link. Hence, only the end-to-end delay of a fixed-sized packet will be investigated carefully here. We use the Round Trip Time (RTT) rather than the One Way Trip Time for simplicity. The Round Trip Time is roughly equal to double of One Way Trip Time in most cases. In near future, we are going to investigate the latter more carefully.

The Extreme Values (EV) statistics [6] will be used in this paper to analyze the RTT data. Extremes are new variables depending on the initial distribution and the sample size. Under certain circumstances, some results are distribution-free. In real-time scheduling, it has been used to compute the Worst Case Execution Time (WCET) for a particular program running on a superscalar architecture processor [4]. However this is the first time that the EV theory is used to analyze and predict the network packet delay bounds.

We argue that EV statistics is especially suitable for seeking the upper bound of end-to-end internet delay. The reasons are:

a) The temporal variant pattern of Internet data traffic has proven to be similar to that of the weather [9]. Hence, we can utilize similar analysis methods to investigate the network traffic.
b) It is easy to collect the maximum value from the data block since the network traffic data also shows definite periodical patterns coinciding to the calendar.
c) The stable network topology guarantees that the underlying distribution is comparably stable and leads to a reliable statistical analysis.

In this paper, we attempt to employ the EV statistics to analyze the RTT to achieve the WCRTT and then make a solid step toward analyzing the real network delay in the future. We introduce the EV statistics theory, and then show the prediction results. Next, we compare results from two different links. Finally, conclusions and further works are described.

## 2. Extreme Value Statistic

The objective of the Extreme Value Theory is to analyze the extreme values (largest or smallest). The extreme values are regarded as statistical variables which inherit their properties from the underlying distribution function. Two ways of identifying extreme values are considered,

- The first approach handles the extreme values happening in successive periods, for instance, days, months, or years. The maxima of the observed data sets consist of the sample. They are called *Block Maxima*.
- The second one focuses on the data exceeding a threshold. The threshold is either a constant or a value which is larger than a certain proportion of data. They are called *Threshold Exceedances*.

### 2.1 Modeling Block Maxima by EV functions

For independent and identically distributed (iid) random variables $X_1, \ldots, X_n$ with a common distribution function $F$, for each variable, one may observe a largest occurrence among the sample. If the maxima distribution is governed by $F^n(x)$, an important result was given [5] as Theorem 1,

Theorem 1: If $F^n(b_n + a_n x)$ has a non-degenerate limiting distribution function for constants $b_n$, and $a_n > 0$, then $F^n(x) \rightarrow G(\mathrm{x})$ as $n \rightarrow \infty$, where $G$ is the limiting forms of the maxima.

We apply one-parameter representation suggested in [11],

$$G_\gamma = \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right), & \gamma \neq 0 \\ \exp\left(-e^{-x}\right) \quad , & \gamma = 0 \end{cases} \tag{1}$$

On adding location and scale parameters, the $x$ will be replaced by $(x-\mu)/\sigma$, where $\mu$ is the location parameter and $\sigma$ is the scale parameter.

### 2.2 Return Level of the Maxima

The Return Level [3] is a value expected to occur exactly once during the Return Period. The Return Level and Return Period are defined as following:

If there are iid random variables $X_1, \ldots, X_n$ with common distribution function $F$, then Return Level $R^T$ is

$$R^T = G_\gamma^{-1}(1 - 1/T) \tag{2}$$

where $T$ is the Return Period. This can be one of the WCRTT prediction method.

### 2.3 Modeling Threshold Exceedances by GP functions

The second primary aim is modeling the Threshold Exceedances. The exceedances are chosen as they exceed a pre-determined threshold. The Threshold Exceedances are modeled by another family of functions, the Generalized Pareto (GP) Distributions. Similarly, there is a limiting theorem for exceedances [2],

Theorem 2: If $F^{[n]}(b_u + a_u x)$ has a contiuous limiting distribution function, then $F^{[n]}(x) \rightarrow W_{\gamma,\mathrm{u},\sigma}(x)$ as $u$ goes to the right endpoint $\omega(F)$, where $W$ is the limiting form of the exccedances.

Again, the GP functions are unified model by taking one parameterization by $\gamma$,

$$W_\gamma(x) = \begin{cases} 1-(1+\gamma x)^{-1/\gamma}, & \gamma \neq 0 \\ 1-\exp(-x)\ , & \gamma = 0 \end{cases} \tag{3}$$

Please note that the distribution of negative $\gamma$ has a right endpoint at $1/|\gamma|$. Most of WCRTT models have $\gamma > 0$, which implies that it could be infinitely large statistically.

### 2.4 Confidence Level for WCRTT Evaluation

The purpose of modeling the Threshold Exceedances is estimating WCRTT with a required Confidence Level. In real life, a WCRTT does not guarantee not being exceeded forever. Therefore, a WCRTT associated with the Confidence Level makes more sense in scheduling applications. If $H$ is used to specify the Confidence Level, the WCRTT with that confidence level is defined by

$$Y^H = W_\gamma^{-1}(H) \tag{4}$$

where $W_\gamma$ is GP function.

The Confidence Level definition is based on the threshold selection. Essentially, when we choose WCRTT = Threshold, the Confidence Level is zero and for WCRTT = $\infty$, the Confidence Level is 1.

## 3. Collecting RTT Sample

The NLANR [14] Active Measurement Project (AMP) is a distributed network of more than 100 active monitors all over the world. There are more than 10,000 pairs from these monitors. RTT data for experiments are all downloaded from the NLANR website.

First, let us observe an RTT spectrum over a week. In Figure 1, the data shows the RTT from March 30, 2003 to April 5, 2003. The pinger is in the Baylor College of Medicine, Houston, TX and the pingee is in the Virginia Tech, Blacksburg, VA. As we have expected, the RTT value shows significant temporal changes during a week. The period is approximately a day. In the daytime, especially in the afternoon and the evening, more traffic is expected and leads to large RTT values. Thus, the RTT spectrums are similar between most links. Based on the periodical pattern of the RTT spectrum, an assumption is made in this paper:

***<u>Assumption</u>: The distribution function of RTT for a certain link varies with the time. The period of that variation is 24 hours.***

Next, the analysis will be implemented for different time intervals of a day. Since the temporal variation over a network delay is very slow [1], we assume that the distribution for a continuous two hours is roughly the same. Therefore, we have twelve models to be computed. They represent 0-2AM, 2-4AM, … and so on…

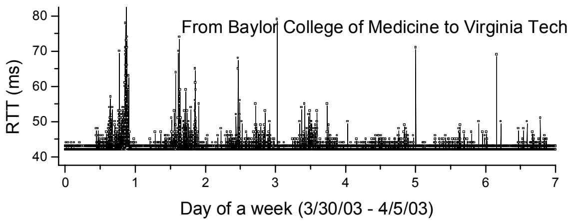



Figure 1. A week long RTT display for link from Baylor College of Medicine to Virginia Tech

## 4. Predicting Return Levels of the Block Maxima

The example being presented for predictions uses the RTT data pinging from the Baylor College of Medicine (BCM) to the Virginia Polytechnic Institute (VTech). The function we choose is the EV model unified by $\gamma$ expressed by formula (1). The Maximum Likelihood Estimator (MLE) is applied to compute the best parameters fitting the function to the sample data. For twelve time intervals, we have twelve sets of parameters $\gamma$, $\mu$ and $\sigma$ values estimated. Then, applying equation (2) can give us a return level for a certain return period. In this paper, we specify five different return periods, 2, 7, 30, 90, and 200 days for these twelve different models. The results are shown in Figure 2.

Figure 2 gives the statistical result that shows how frequently the maxima will come back in the future. From this we see that larger return periods result in a larger magnitude of Return Levels and a larger deviation of Return Levels as well. This is confirmed by our impression that the larger values show more uncertainties upon the occurring probability. We also compute the confidence intervals of our results by the bootstrap method [12]. For a reason of limited space, we omit this part in this paper.

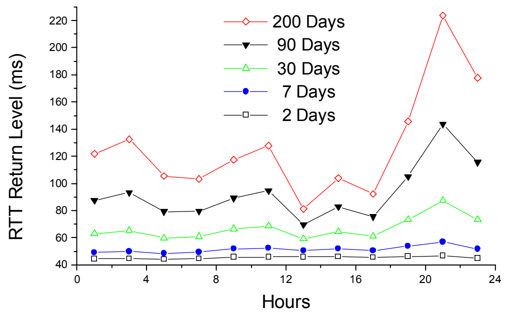



Figure 2. Return Levels for five different Return Periods

## 5. Predicting WCRTT with Different Confidence Levels

The same data as in section 4 are applied for GP modeling to predict the WCRTT. In this paper, we use an intuitive method to select the threshold value: the threshold is such an amount that the number of exceedances is limited to less than one thousand.

Similarly, we will apply the GP model (3) (4) to forecast the WCRTT with a different confidence level requirement. The confidence level will be stated as a percentage, like 90% or 95%. Here, we consider five confidence levels and see the WCRTT for the twelve models in Figure 3.

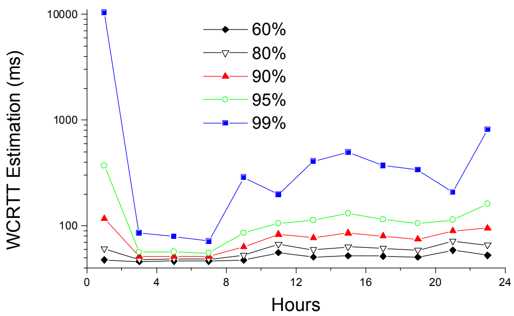



Figure 3. WCRTT evaluations for five different confidence levels

In Figure 3, it is apparent that the higher confidence level requirements lead to a larger evaluation of WCRTT. When the confidence level is 99%, the WCRTT for 0-2AM is more than 10,000 ms. We argue that the error is caused by a few unusual data appearing between 500 and 700 ms. We also note that the 95% confidence level's curve is much lower than 99%'s. Therefore, we might decrease our confidence level and save a lot in scheduler resources.

## 6. Comparisons of Two Different Links

We have performed the extreme values analysis for the link from BCM to VTech. To compare the experiment's result, another link's RTT is collected and analyzed. The second link is pinging from Korea to the Univ of Oregon. We want to see how different the intercontinental links are from shorter inland links.

The Korean link data is from Sept 9, 2002 to May 2, 2003, a total of 147 days' data available. We only compare the three parameters' result for these two links. The comparison between the two models can give us a sense of how different the data is distributed from the extreme theory point of view. The results are presented in Table 1.

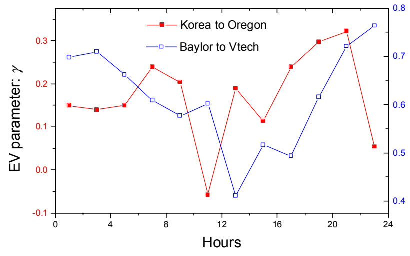
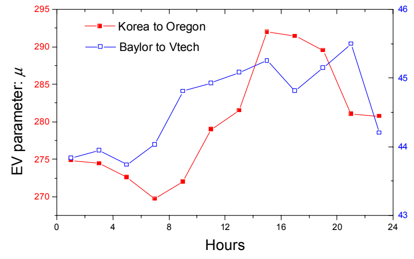
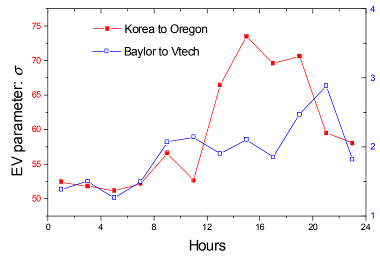
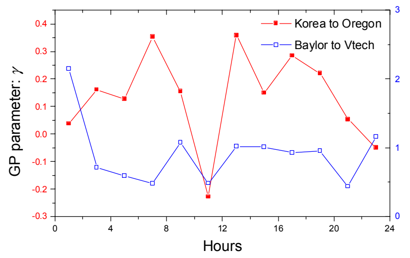
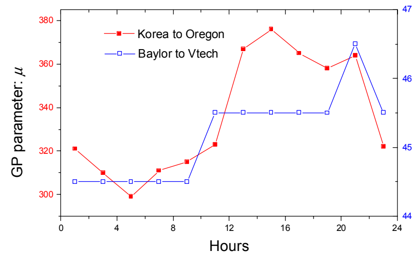
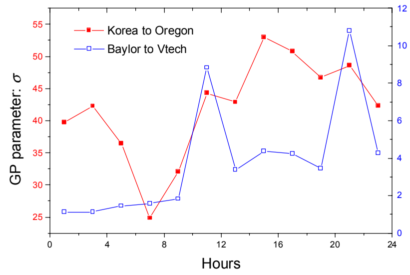
In Table 1, two curves were drawn overlapping in one plane in order to compare them easily. Solid curves are for the link from Korea to Oregon; its value is always illustrated with the left axis, the right axis otherwise. We summarize this table as followings:

1) $\gamma$ represents the degree of being skewed for the density curve in the EV models. The above table shows both $\gamma$ curves change with the time in one day and have almost the same temporal changing tendency, but the $\gamma$ of BCM link is double that of the Korean link. In the GP models, $\gamma$ represents the degree of "long tails". Unfortunately, we do not see much similarity in the tendency between these two links. They are varying around different levels, and the Korean link's $\mu$ is almost ten times that of the BCM link's, meaning that the former has much greater long tails distribution than the latter.
2) $\mu$ represents the fixed probability point in the density function for both models. In EV, it shows the probability equal to 0.36788. In GP, it is the left end point of the curves. We found that the $\mu$ curves are all similar for these links. They are increasing from the early morning, greatest in the evening, and decreasing through midnight. The difference is the absolute RTT since the geographical distance from Korea to Oregon is much longer than that from BCM to Vtech. In the GP analysis, the $\mu$ is the threshold we choose. Since the number of exceedances in our experiments is roughly the same, larger RTT data will lead to a larger threshold. This draws the same conclusion as in the EV models, that is, $\mu$ is related with the network traffic volume.
3) $\sigma$ is the scale parameters representing how evenly the data is distributed in the density distribution. In EV, both show quite similar tendencies in the $\sigma$ curves. Since $\mu$ is increasing in the evening, $\sigma$ is increasing also. Together this proves the fact that RTTs are becoming larger and more sparsely distributed in comparison to an earlier time in a particular day. However, we note that the Korean link's scale parameter is much larger than the BCM link's insinuating that an RTT with a long traveling route has more uncertainties. In GP, we ascertain few valuable conclusions. We can only say that both have a tendency of increasing. The evening $\sigma$'s are always larger than the mornings'. Moreover, once again the $\sigma$ for Korea's link is much larger than BCM's link.

## 7. Concluding Remarks and Future Works

We have seen that the Extremes Value models are successful in describing the Worst Case Round Trip Time. The EV statistics leads to a more precise and complete WCRTT prediction. We point out the following facts:

Table 1 Two models parameters comparison between two different links

| | $\gamma$ | $\mu$ | $\sigma$ |
|---|---|---|---|
| EV |  |  |  |
| GP |  |  |  |

1) Comparing the two different methods in this paper, we argue that they are similar but conceptually different. Mathematically, the two methods generate identical results under ideal circumstances.
2) Our assumption is challenged by the unusual data invoked by network failure. If the underlying distribution does not follow certain rules, there is no way to predict future behavior precisely. In [7], ten different RTT measurement cases are described. We are planning to design a dynamic model to tackle this problem.
3) We do not consider the size of the sample carefully. In our ongoing research, we want to collect more RTT data to check the correctness of our models. Also, more links' data can be compared and predicted.
4) RTT measurement and analysis is a step to support LLF scheduling. We attempt to use this method to analyze the One Way Trip Time in our ongoing work. Meanwhile, we are interested in the TCP packet delay prediction since most of the network traffic is based on that.
5) We present a long-term traffic prediction in this paper. Next time, we will derive the prediction of short-term traffic, which is assumed to be more stable since there is less possibility of changing network topology changing.

Overall, this statistical model will be studied more carefully in the near future to support network scheduling and data flow control.

**References**


[1] A. Acharya and J. Saltz, "A study of internet round-trip delay", *Univ of Maryland, Tech Rep. CS-TR-3736*, 1996

[2] A.A. Balkema and L. de Haan, "Residual life time at great age", *Ann. Probab.2* pp 792-804, 1974

[3] J. Danielsson and C. de Vries, "Value-at-Risk and Extreme Returns", *working paper*. 2001.

[4] S. Edgar and A. Burns, "Statistical Analysis of WCET for Scheduling", in *Proc. IEEE RTSS* 2001.

[5] R A Fisher and L.H.C. Tippett. "Limiting forms of the frequency distribution of the largest or smallest member of a sample." In *Proc. the Cambridge Philosophical Society*. Vol 24 pp180-190, 1928.

[6] E. J. Gumbel, "Statistics of Extremes", Columbia University Press, New York, 1958

[7] T. Hansen, J Otero, T McGregor and HW Braun "Active measurements data analysis techniques", website of NLANR, 2000

[8] C Li, R. Bettaiti and W.Zhao, "Static Priority Scheduling Priority Scheduling for ATM Networks", *IEEE Proc. RTSS*, 1997.

[9] H. Ohsaki, M. Morita, and M. Murata, "On modeling round-trip time dynamics of the Internet using system identification," in *Proc. ICOIN-16*, Jan. 2002

[10] S. Rao and A. M. K. Cheng, "Real-Time Multimedia Traffic Scheduling and Routing in Packet-Switched Networks," *J. VLSI Signal Processing*, Kluwer Academic Publishers, Vol. 34 No.1-2, 2003

[11] R. D. Reiss and M. Thomas, "Statistical Analysis of Extreme Values", 2nd edition, Birkhäuser Verlag, Berlin, 2000

[12] N Tajvidi, "Confidence intervals and accuracy estimation for heavy tailed generalized pareto distribution", *Thesis article*, Charlmers University of Technology. 1996

[13] X. Xiao and L. M. Ni, "Internet QoS: A Big Picture", *IEEE Network Mag.*, vol.13, pp. 8-18, Mar./Apr.1999

[14] http://www.nlanr.net