Modeling Available Bandwidth for an Efficient QoS Characterization of a Network Path ⋆

Alexander Chobanyan[1], Matt W. Mutka[1], V.S. Mandrekar[2], and Ning Xi[3]

[1] Department of Computer Science and Engineering, Michigan State University
chobany1@cse.msu.edu, mutka@cse.msu.edu
[2] Department of Statistics and Probability, Michigan State University,
mandrekar@stt.msu.edu
[3] Department of Electrical and Computer Engineering, Michigan State University,
xin@egr.msu.edu

Abstract. Estimating the reliability of an end-to-end network path is critically important for applications that support remote real-time task execution. Available bandwidth, which is defined as a minimum spare capacity of links constituting a network path, is an important QoS characteristic of the path. In this work we demonstrate a new approach to modelling available bandwidth behavior from a time-series analysis prospective. In particular, we introduce a notion of *crossing probability* – the probability that available bandwidth drops below the QoS critical threshold for the period of time required for a real-time task execution. We estimate "crossing probability" by an application of the $ARCH^2$ (AutoRegressive Conditional Heteroscedasticity) model to available bandwidth behavior. We estimate model coefficients β_0 and β_1 to quickly output "crossing probability" for arbitrary values of threshold and length of the real-time task. The model was evaluated on real bandwidth measurements across multiple network paths.

Index terms–Network measurements, Statistics, Stochastic processes . . .

1 Introduction

Over the past several years the importance of Internet-related technologies have rapidly increased. The growth of the Internet has not only a quantitative nature but also affects the type of information that may be transferred and consequently the methods for real-time interaction between communicating parties. Many rapidly emerging network applications may benefit not only by transferring traditional media-types, such as video and audio, but also by remotely manipulating, touching and feeling a remote object. One may imagine the influence of new media types when new progressive technologies emerge, such as tele-medicine that allows physician to remotely touch and feel a patient, tele-education that opens wider possibilities to distant learning about various objects

⋆ This research is supported in part by NSF Grant No. 0334035.

via sensing, and tele-commerce that allows a customer to feel and manipulate an object before conducting a financial transaction.

Extending transferable media to new types, such as haptic and temperature data, give rise to many challenges such as channel synchronization, increased real-time requirements for tele-operated task execution, and others. In particular, the nature of real-time tasks related to new media types requires Quality-of-Service (QoS) guarantees to communicating parties over the potentially unreliable and uncontrolled Internet. Available bandwidth is an important QoS characteristic of the network path. It may be briefly defined as the maximum rate that a sender may put bits on a medium without affecting existing cross-traffic, or in other words, as a residual capacity of the network path. In this paper we present a statistical model for a time-series that represents the available end-to-end bandwidth. The proposed model provides an important *probabilistic* QoS channel specification in terms of coefficients that allow the rapid computation of *crossing probability*–the probability that the available bandwidth becomes less than an arbitrary pre-defined critical value over an arbitrary pre-defined time-frame. The aforementioned coefficients are evaluated based on observed available bandwidth behavior over a reasonable period of time and are updated at run-time reflecting the changing dynamics of a network environment. The critical threshold and the time-frame should be defined by the QoS specifications of a particular real-time task.

The available bandwidth time-series is directly related to the cross-traffic that is present in the path. It has been shown [1] that network traffic has self-similar properties, which imply a non-summable autocorrelation function and a long-range dependence between observations. Therefore observations of available bandwidth cannot be considered as independent, which makes statistical inference more challenging. All previous work [2] concentrated on obtaining the value of available bandwidth averaged over a reasonably small time-window without an attempt to provide a broader probabilistic picture that characterizes the network path. Figure 1 shows an example of how the available bandwidth between two nodes may behave over time. As depicted in the plot, the available bandwidth may instantaneously drop far below its mean value, thus causing degradation of service, such as distorted video or audio. Despite the fact that such "spikes" may last only a few seconds, they may be fatal for many real-time remotely-operated tasks. As illustrated in figure 1, the value of the available bandwidth averaged over an arbitrary large (or small) time-window cannot characterize the network path well in terms of its bandwidth QoS properties.

The key difference between our approach and recent work is that we apply random time-series statistical techniques to available bandwidth behavior. We create a probabilistic framework for network path QoS characterization. A number of various statistical models that consider dependence between observations may describe instantaneous spikes and reasonably predict the "crossing probability" based on the chosen model. Many successful time-series analysis techniques have been applied to similar applications in econometrics and natural sciences. We note, in particular, that the ARCH model used in econometrics is able to describe spikes and possesses self-similar properties, which is important for making inferences about the "crossing" probability.

The notion of "crossing probability" on its own gives rise to conceptually novel approach to available bandwidth estimation. Real time applications may benefit from this approach by reallocating part of their QoS-sensitive channels over different network paths when the *crossing probability* exceeds an acceptable threshold. Channel reallocation may happen before something fatal occurs for a real-time task but after the probability for this fatal event happens to exceed a reasonable threshold.

2 Background

A network path may be viewed as a set of store-and-forward links $L_i, i = 1, 2, ..., k$. The capacity of the i-th link C_i may be defined as the maximum bit rate that the sender may transfer information through this link when no other traffic is present. Assume that the link is partially used by a separate communication. Let $T_i(t)$ be defined as the number of bits transferred through this link by time t. The spare capacity of link i, or in other words, its available bandwidth $A_i(t)$ at time t may be defined as $A_i(t) \equiv C_i - \frac{dT_i}{dt}(t)$. The bandwidth $A(t)$ available at time t of a path may be defined as $A(t) \equiv \min_{i=1,...,k} A_i(t)$. From this definition it follows that the available bandwidth cannot be measured precisely and therefore needs to be approximated by the mean available bandwidth at time t for the interval τ: $A_\tau(t) = \min_{i=1,...,l}\{C_i - \frac{T_i(t+\tau)\quad T_i(t)}{\tau}\}$. With smaller τ, one may achieve better flexibility in describing the available bandwidth behavior. In general, the approximation of available bandwidth is QoS relevant if τ is less than the time that receiver needs to empty its buffer. Fluctuations of the available bandwidth should not cause degradation of quality at the receiver as long as those fluctuations do not affect the number of bits that the receiver needs to keep its buffer non-empty.

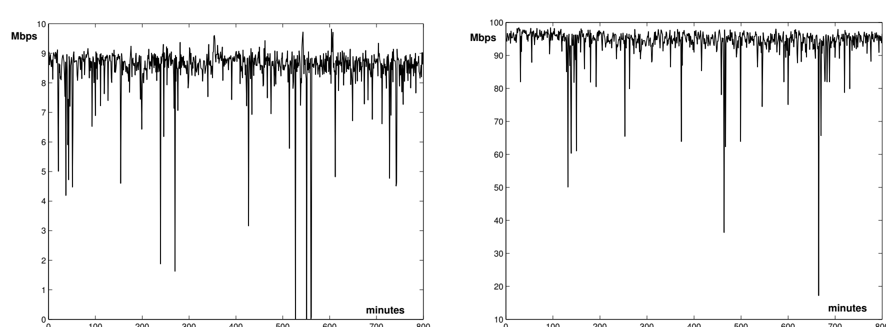


Fig. 1. Estimated available end-to-end bandwidth (Mbps) across different network paths. The time-series are sampled at a rate of one observation per minute.

Significant work conducted regarding mean available bandwidth estimation is well reflected by Strauss, et al. [3] and Hu, et al. [4]. Probe gap models used in Spruce [3] and Delphi [5] measure delay between probe packets and make further inference about available bandwidth based on that delay. As it was shown by Prasad et al. [6] context switches and enabled "interrupt coalescence" feature at receiver's side may distort measurements of a gap between packet pairs and therefore significantly influence probe-gap based available bandwidth measurements.

The probe rate model used in Pathload [2] use sequences or "trains" of packets sent at different rates. By detecting a trend in packet transmission times sent at a given rate Pathload decides whether the specified rate is higher or lower than the available bandwidth. Then the available bandwidth can be found by a simple binary search. Pathload is believed to be more intrusive than tools using the probe gap model. To our knowledge Pathload is, however, the only tool that considers effect of context switches and interrupt coalescence at the receiver side and is therefore the most stable and accurate among available bandwidth measurement tools to date.

From the definition of available bandwidth, it follows that it is closely related to cross-traffic in the network path. Cross-traffic, defined as $\frac{dT}{dt}(t)$, represents a random time-series that has been extensively analyzed by both statisticians and network engineers for its statistical properties. If the most tight link remains unchanged throughout measurements (which is usually the case), then the available bandwidth time-series may be evaluated by subtracting the cross-traffic from a tight link's unchanged capacity. It follows that in the aforementioned case the available bandwidth and cross-traffic are exactly the same time-series in terms of their statistical characteristics. Therefore we believe that the model chosen for available bandwidth characterization should incorporate at least basic properties of models that are presently used for Internet traffic modelling. Below we provide some background of such models.

All parametric traffic models assume weak stationarity of the traffic time-series, which means that the correlation between time-series observations remains the same under any shift in the time domain. Leland, et al. [1] showed in 1994 that network traffic possesses an important characteristic called self-similarity. Most important implication of self-similarity is "slowly" (slower than $1/h$) decreasing correlation between observations located at lag h apart implying long-range dependence.

Many recent works in internet traffic modelling [7–9] confirmed that self-similarity is one of key-importance properties of a present Internet traffic. Researchers tried to build models that visually look similar to observed time-series and have an autocorrelation structure similar to what has been observed. Model parameter estimation based on observed data should remain unchanged or vary slowly over time. Furthermore, the model should be relatively simple and amenable for the further statistical analysis. We follow the same strategy to unify recent work in Internet traffic analysis, available bandwidth estimation tools, and time-series studies with respect to the analysis of QoS characteristics of a network path.

3 Data collection and preprocessing

For this work we made continuous measurements of available bandwidth between a number of machines located in Michigan State, Berkeley, MIT, University of Sydney, and Uupsala University (Sweden), thus covering a broad range of geographical locations. All machines chosen for the experiment had very low processor utilization factor since the Pathload measurements are very sensitive to processor load of a machine responsible for receiving trains of measurement packets (context switch / interrupt coalescence effect). Figure 1 depicts the available bandwidth behavior for the period of 800 minutes of a few network paths. The distance between nodes varies from 4 to 20 hops. In each example in figure 1 we observe a small trend that is 2 Mbps around the mean value. Each example has bursts reflecting the heavy-tailed self-similar nature of Internet traffic-related series. For further fitting of stationary self-similar models we preprocess time-series of interest by subtracting the trend. For all examples considered in figure 1, trend-subtraction changes each observation for less than 2 Mbps, which is negligible with respect to spikes that are the primary cause for instantaneous unexpected degradation of service. Predicting the probability of such a spike to occur over a certain a QoS sensitive time-slice is our primary goal. For the aforementioned purpose we can neglect 2-3 Mbps trends and further consider the “adjusted” time-series. Figure 2 shows the original time-series representing available bandwidth between nodes located in MSU and Berkeley along with the corresponding “adjusted” time-series. Figure 2 shows that trend-removal, inversion and “lining-up” to zero level do not significantly change the behavior pattern of the original time-series of interest. For trend-removal, we used the Spencer 15-point moving average filter described by Brockwell, et al. [10], which skips polynomials of degree three and smoothes everything that has a larger degree.

It is also undesirable to have instantaneous spikes affecting trend estimation. Therefore a “trimming” procedure is applied before passing data through a Spencer filter. The whole procedure may be done at run-time. The size of the window that slides across the time-series observations is 19. We delete two maximal and two minimal observations, and pass the remaining 15 observations to a Spencer filter that outputs the estimated trend. Then trend is subtracted from the original time-series, the result is inverted, lined-up to zero, and thus finally the “adjusted” time series is obtained.

The following discussion will be related to analysis of the “adjusted” time-series. Note that in most cases the “adjustment” procedure does not significantly distort data. Even if it does, we may detect a bad on-going trend, report the potential QoS problem to the communicating parties, and still learn at the same time about the pattern of spikes.

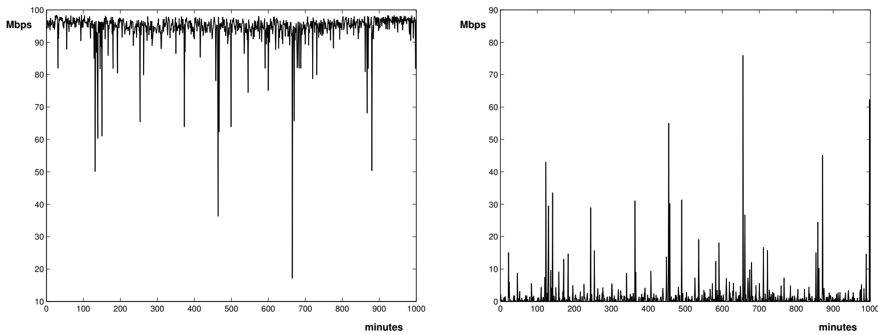


Fig. 2. Estimated available end-to-end bandwidth (Mbps). Trend and inversion effect.

To our knowledge rare cases fall beyond the scope of this work, such as when there exists long-lasting constantly present trend exceeding 5 Mbps or when routers on the network path frequently change routes for packets thus making available bandwidth oscillating around multiple different levels.

4 Modeling Available Bandwidth

4.1 \Crossing Probability"

Assume that an operator is about to decide whether to perform a real-time QoS critical task over a potentially unreliable network path. The most important question may be stated as follows. What is the probability of experiencing QoS critical degradation of service during the time required for a given real-time task execution? An example is the probability that a picture transferred to the operator from a remotely-controlled walking robot distorts to a point that the operator is not able to understand where the robot is located with respect to an approaching wall.

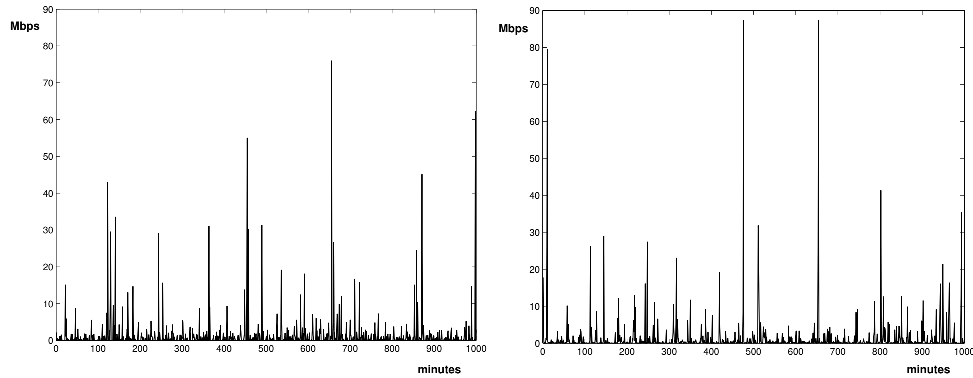
Let us slightly formalize the problem. The natural assumption is that the operator should be aware of the real-time task duration, minimal allowed transmission rate and the receiver's buffer size. These three metrics may be considered as basic QoS characteristics of a given task. QoS requirements are violated if the available bandwidth averaged over time $\tau = \frac{receiver's\ buffer\ size}{minimal\ allowed\ trasmission\ rate}$ drops below the threshold of minimal allowed transmission rate during the task's execution time. Note that the task's execution time T corresponds to a sequence of observations of mean available bandwidth time-series of a length $N = \frac{T}{\tau}$. Now rephrase the main QoS question as follows: What is the probability that the minimal value over the next N consecutive observations of available bandwidth time-series drops below a given threshold where both threshold and N are specified by real-time task constraints? For the "adjusted" and inverted mean available bandwidth time-series, the aforementioned probability may be re-defined as *the probability that the maximum over the next consecutive N observations of the time-series crosses the pre-defined QoS critical threshold*. We believe that finding this probability should be an ultimate goal and statistically correct problem statement for the QoS characteristics of a network path. We here-and-after refer to this probability as to a "crossing probability". Even finding a reasonable upper bounds for such probability is a worthwhile contribution to the QoS characterization of a network path. Below we build a model that describes the observed available bandwidth data. The model is self-similar and amenable to analysis with respect to the aforementioned "crossing probability."

4.2 ARCH model

We propose a parametric model that allows efficient computation of the "crossing probability" for any given number of future observations N and the threshold M based on the estimated model parameters. We used the "$ARCH^2(1)$" model [11] for description of the mean available bandwidth data. Therefore we first provide background on ARCH models and describe reasons why this model has been chosen as a candidate for the mean available bandwidth behavior description.

The ARCH(p) (Auto-Regressive Conditional Heteroscedasticity) time-series model was introduced by Engle [11] for applications of financial events. The ARCH(p) model may be defined as an *autocorrelated* Gaussian noise with dependency between observations described as follows. $X_t = \sqrt{\beta_0 + \sum_{i=1}^{p} \beta_i X_{t-i}^2} Z_t$ where Z_t are independent observations following a standard normal distribution. The simplest partial case is the $ARCH(1)$ time-series defined respectively as $X_t = \sqrt{\beta_0 + \beta_1 X_{t-1}^2} Z_t$. Finally, the $ARCH^2(1)$ time-series will be defined as $Y_t = (\beta_0 + \beta_1 Y_{t-1}) Z_t^2$. From the definition of $ARCH^2(1)$, it follows that the model is completely defined by a pair of coefficients $[\beta_0, \beta_1]$. Both $ARCH(1)$ and $ARCH^2(1)$ models were extensively applied to finance-related time-series that at least visually have the similarity to the "adjusted" mean available bandwidth time-series pattern. Before going any further, we provide a picture with "adjusted" mean available bandwidth on the left and a simulated $ARCH^2(1)$ time-series on the right. As shown in figure 3, the patterns of both time-series are the same. Another encouraging factor is that it has been shown by Embrechts, et al. [12], that the $ARCH^2(1)$ time-series possesses self-similar characteristics.

In order to apply a parametric model to "crossing probability" estimation, one needs to ensure that the model parameters may be efficiently estimated at run-time. The model needs to be mathematically analyzed in order to find an efficient method to compute the "crossing probability" based on the estimated model parameters. In addition, one needs to build confidence intervals around the estimated probability in order to be aware of the model's accuracy. Then the model needs to be empirically verified. We therefore provide below our work on unification of the results in $ARCH^2(1)$ parameters estimation and "crossing probability" computation together with the scheme of confidence interval construction.



(a) Measured available end-to-end bandwidth (Mbps)

(b) Simulated $ARCH^2(1)$ observations with $\beta_0 = \beta_1 = 0.25$

Fig. 3. Simulated versus observed series

The basic work in ARCH parameter estimation has been conducted by Weiss [13]. Weiss [13] proposes to estimate ARCH parameters by maximizing Quasi Log Likelihood Function evaluated as: $L(\beta_0, \beta_1) = \sum_{i=1}^{n-1}[\log(\beta_0 + \beta_1 X_i^2) + \frac{X_{i+1}^2}{\beta_0+\beta_1 X_i^2}]$ Here n is the number of observations in the window that is used for parameters estimation. The parameter n is very important. We discuss its proper adjustment separately. The advantage of the QMLE (Quasi Maximum Likelihood) estimator of $[\beta_0, \beta_1]$ is that it is asymptotically unbiased. Also, it is known [14] that $[\hat{\beta}_0, \hat{\beta}_1]$ is normally distributed with mean $[\beta_{0true}, \beta_{1true}]$ and the covariance matrix is called Fisher Information. The normality of estimators gives background for computing the confidence interval of any level around the estimated values of model parameters.

Since estimators of $[\beta_0, \beta_1]$ depend on past observations and therefore have a random nature, we also provide a way for rapid run-time computation of the confidence region of any level (by default 95% CR). The 95% confidence region for $[\beta_0, \beta_1]$ is defined as a random region in R^2 that contains the point $[\beta_0, \beta_1]$ with probability ≥ 0.95. This region for the QMLE estimator has an elliptical form defined by the equation obtained by Hotelling [15] $(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'\hat{\Sigma}^{-1}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})(n-2)/2 = F_{2,n-2}(0.95)$ where $F_{2,n-2}(0.95)$ is the 95% quantile of the F-distribution with 2 and $n-2$ degrees of freedom, and $\hat{\Sigma}$ is the estimated Fisher information. In order to compute a confidence region quickly at run-time, we tabulated $\hat{\Sigma}$ for a fixed $n = 720$ and a grid of β_0 and β_1. Fisher information was calculated by simulating 1000 independent $ARCH^2(1)$ time-series for each fixed pair of β_0, β_1 and finding the difference between estimated and true values of the aforementioned vector. β_0 and β_1 then were varied within the region $(0.5, 3.5)$ and $(0, 0.6)$, respectively. This region covers all potentially acceptable values that β_0 and β_1 might take while describing the available bandwidth behavior within the range of 200 Mbps. It is known that the estimated covariance matrix is proportional to n^{-1}. Therefore, when having available tables for $n = 720$, one can recompute $\hat{\Sigma}$ for an arbitrary n in less than a second.

4.3 Results from extremal values theory

Below we show mathematical results that allow computation of "crossing probability" for any specified threshold M, number of observations N, and estimated model parameters β_0, β_1. These results allow us to construct a 95% confidence interval for a "crossing probability" by considering the fact that the "crossing probability" is a monotonically increasing function of both β_0 and β_1 and that we find sufficiently many "crossing probability" points related to the 95% ellipse in the β_0, β_1 space. Embrechts, et al. [12] summarize recent work in extreme values by providing formulas for finding "crossing probability" for various time-series models. In particular, for $ARCH^2(1)$ model Embrechts, et al. have $\lim_{N \to \infty} P(N^{-\frac{1}{\kappa}} \max\{Y_1, ..., Y_N\} \leq M) = \exp^{-C_1 M^{-\kappa}}$ that allows us to rewrite this result for "crossing probability" P and large N as follows $P = 1 - \exp^{-C_1 M^{-\kappa} N}$ where κ is a unique root of the equation $\frac{(2\beta_1)^u}{\sqrt{\pi}} \Gamma(u + \frac{1}{2}) = 1$. Although Embrechts et al. [12] give an expression for C_1, it is more efficient to estimate it empirically off-line for different β_0 and β_1. For that purpose we simulated 800,000 observations for each pair of β_0 and β_1 coming from a grid with range $(0.5, 3.5)$ and $(0, 0.6)$, respectively. Then, based on the probability expression above, this tabulated result may be quickly generalized on-line to arbitrary N and M.

Calculating the estimated probability together with its confidence interval gives an opportunity to judge the size of the window n used for prediction of β_0 and β_1 coefficients. On one hand we want n to be as small as possible because we assume that model parameters remain unchanged for the time-period related to n. On the other hand, with larger n, accuracy of the prediction increases with corresponding shrinking of the 95% confidence region. In our experiments, for example, we put the following condition to the 95% confidence interval for the probability: the difference between the 95% upper bound and the predicted value should not exceed 0.2 for an arbitrary predicted probability value from the range [0;0.5]. Then for $N \in \{1, ..., 100\}$ and $M \in [60; 90]$ the lower boundary for n will be evaluated to 120. Note that dependence on N is natural: with increasing the period of time for which the prediction is made (duration of real-time task execution), the period of time used for prediction also needs to be correspondingly increased.

5 Experimental evaluation

In this section we describe our method for testing the accuracy of our approach. The final target of modelling is prediction of a *probability.* The only way, therefore, to empirically verify the model is to apply the ergodic theorem, i.e., to have B groups with N observations in each and then count the number L of groups where the maximum over N observations within the group crosses a pre-defined threshold. Then the ratio L/B may be considered as an empirically computed value of a "crossing probability." We considered as our design accuracy the following criteria: within the frames of the model the difference between the predicted probability and its 95% upper confidence limit should not exceed 0.2 at any circumstances. Our simulations of $ARCH^2(1)$ series show that such "within-the-model" accuracy on a probability scale for $N \leq 90$ dictates the lower bound on the number of observations needed for model parameters estimation as $n = 120$. Higher accuracy leads to higher values of a lower bound for n.

In real-life, however, it is impossible to achieve such an ideal design of the experiment because of the following factors. First, one obtains B groups of measurements in a real scenario that belong to the same time-series (instead of having multiple independently generated series in a simulation scenario) and therefore are somehow dependent. Second, even if we assume that the model parameters remain the same over the estimation window of length n (that is reasonable if n correspond to 2 hours) it is too naive to assume that these parameters will remain the same over the whole time-period needed for collection of B groups of size N. Presently available bandwidth measurement tools allow consistently stable measurements at a rate of 1 sample/minute. Therefore, $B = 100$ reasonably separated groups of $N = 100$ observations in a group correspond to time needed for collection of 20,000 consecutive observations. With presently available sampling rate this time is evaluated to 20,000 minutes or 2 weeks of consistent measurements.

The efficiency of the upper bound on the "crossing probability" may be, nevertheless, evaluated as follows. We considered a number of network paths that have visually unchanged over time their daily behavior pattern. Below we show as an example the analysis of one such path between nodes in MSU and Berkeley. We consider first 1440 observations (1 day) as our training-validation dataset. The rest 18,560 observations constitute testing dataset. Training observations were used for evaluation of the width of the strip where the estimated probability resides. The testing dataset was used for computation of the empirical probability L/B. We expected an average probability estimated with the testing data to reside in a 95% confidence strip defined by a first day training observations.

Figure 4 shows two surfaces where the upper surface represents the 95% upper bound estimated by the model and the lower surface represents the empirically computed probability. "Crossing probability" was computed for different values of N and M varying within the range 15-100 observations and 60-90 Mbps, respectively. History size n was taken as it was mentioned above equal to 120. As was expected, for all thresholds M and sizes N the empirically computed probability reside below the estimated bound. Figure 4 shows, in particular, that two surfaces are nearly parallel to each other, i.e., the model reflects scales of both N and M in reasonable agreement with what was empirically observed. Also figure 4 shows that the estimation becomes less accurate when N approaches to n, which is understandable.

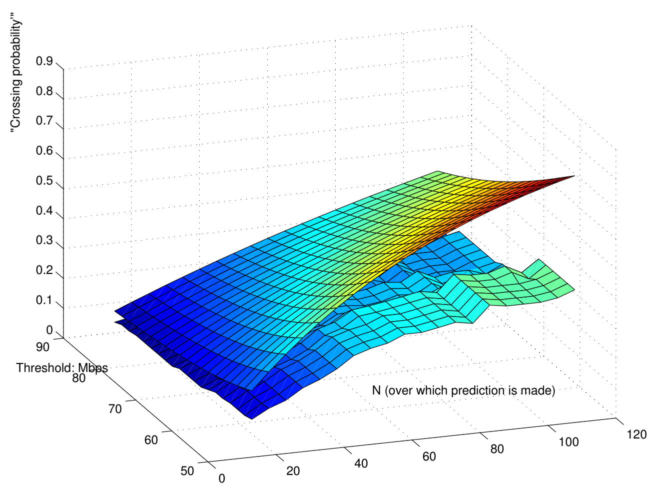


Fig. 4. Empirically computed "Crossing" probability within the frames of estimated 95% upper bound

We conclude that the model produced verifiable and an accurate prediction. We also note that such prediction is our suggested form of characterizing a present state of the network path with respect to its QoS properties.

6 Results Assessment and Future Work

We presented a "probabilistic" approach to a description of the available bandwidth behavior. We showed that the available bandwidth may be described by $ARCH^2(1)$ parametric model. We have studied and unified known results in $ARCH^2(1)$, tabulated relevant characteristics and provided an efficient method for computation of a "crossing probability". We tested the accuracy of our model on a channel with a relatively stable long-term behavior and proved that the approach has a range of applicability.

When describing available bandwidth from a time-series analysis prospective, it is important to have the time τ over which measurements are averaged as small as possible. Presently we are modifying and testing Pathload to increase its sampling rate to 10 samples/minute. With such a measurement tool, we may significantly decrease the total measurement time necessary to obtain the number of observations sufficient for the further statistical analysis. Recall that τ should be of the order of the receiver's buffer time in order to allow observations to be relevant with respect to a particular real-time task execution. Therefore, the real importance of the results provided in this work as well as the real range of their applicability may be extended after Pathload's proper modification.

Despite present challenges in available bandwidth measurement and determining the range of applicability of a particular parametric model, we see the main contribution of our approach in the correct "probabilistic" context of the problem statement. For real-time network applications, it is much more important to have at least tentative knowledge about the "crossing probability" rather than being aware of a single number representing the mean available bandwidth averaged over an inexactly defined time.

The notion of "crossing probability" opens research directions to predictive dynamic channel reallocation and resource planning for real-time network applications.

The aforementioned problems are subject of our future research. We, however, believe that this work is a significant step towards building a reliable QoS infrastructure over potentially unreliable media in order to facilitate the performance of many present real-time network applications.

References

1. W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, 1994.
2. M. Jain and C. Dovrolis, "End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput," in *ACM SIGCOMM*, 2002.
3. J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," in *IMC*, 2003.
4. N. Hu and P. Steenkiste, "Evaluation and characterization of available bandwidth techniques," in *IEEE JSAC Special Issue in Internet and WWW Measurement, Mapping, and Modeling*, 2003.
5. V. Ribeiro, M. Coates, R. Riedi, S.Sarvotham, and R. Baraniuk, "Multifractal cross traffic estimation," in *ITC specialist seminar on IP traffic Measurement*, 2000.
6. R. Prasad, M. Jain, and C. Dovrolis, "Effects of interrupt coalescence on network measurements," in *Passive and Active Measurement Workshop (PAM2004)*, 2004.
7. M. Zukerman, T. D. Neame, and R. G. Addie, "Internet traffic modeling and future technology implications," in *IEEE INFOCOM*, 2003.
8. M. Li, M. Claypool, R. Kinicki, and J. Nichols, "Characteristics of streaming media stored on the internet," Tech. Rep. WPI-CS-TR-03-18, Computer Science Department, Worcester Polytechnic Institute, 100 Institute Road, Worcester, Massachusetts 01609-2280, May 2003.
9. J. Aracil, R. Edell, and P. Varaiya, "A phenomenological approach to internet traffic self-similarity," in *35th Annual Allerton Conference on Communication, Control and Computing*, 1997.
10. P. J. Brockwell and R. Davis, *Time Series: Theory and Methods*. New York: Springer, 1991.
11. R. Engle, "Autoregressive conditional heteroscedasticity and estimates of variance of UK inflation," *Econometrica*, no. 50, pp. 987–1008, 1982.
12. P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*. Berlin Heidelberg: Springer, 1997.
13. A. Weiss, "Asymptotic theory for ARCH models: Estimation and testing," *Econometric Theory*, no. 2, p. 31, 1986.
14. J. A. Rice, *Mathematical Statistics and Data Analysis*. Belmont, California: Duxbury Press, 1995.
15. H. Hotelling, *Multivariate Quality Control. Techniques of Statistical Analysis*. McGraw-Hill, New York: Eisenhart C, Hastay MW, Wallis WA, 1947.