Performance Evaluation 67 (2010) 1030–1043

Contents lists available at ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva

# Modeling TCP throughput: An elaborated large-deviations-based model and its empirical validation

Patrick Loiseau[a,*], Paulo Gonçalves[b], Julien Barral[c], Pascale Vicat-Blanc Primet[b]

[a] INRIA Paris-Rocquencourt, Le Chesnay, France
[b] INRIA Rhône-Alpes, Université de Lyon/École Normale Supérieure de Lyon, Lyon, France
[c] LAGA, Université Paris 13, Villetaneuse, France

ARTICLE INFO

*Article history:*
Available online 12 August 2010

*Keywords:*
TCP modeling
Large deviations
Network traffic
Performance

ABSTRACT

In today's Internet, a large part of the traffic is carried using the TCP transport protocol. Characterization of the variations of TCP traffic is thus an important issue, both for resource provisioning and Quality of Service purposes. However, most existing models are limited to the prediction of the (almost-sure) mean TCP throughput and are unable to characterize deviations from this value.

In this paper, we propose a method to describe the deviations of a long TCP flow's throughput from its almost-sure mean value. This method relies on an ergodic large-deviations result, which was recently proved to hold on almost every single realization for a large class of stochastic processes. Applying this result to a Markov chain modeling the congestion window's evolution of a long-lived TCP flow, we show that it is practically possible to quantify and to statistically bound the throughput's variations at different scales of interest for applications. Our Markov-chain model can take into account various network conditions and we demonstrate the accuracy of our method's prediction in different situations using simulations, experiments and real-world Internet traffic. In particular, in the classical case of Bernoulli losses, we demonstrate: (i) the consistency of our method with the widely-used square-root formula predicting the almost-sure mean throughput, and (ii) its ability to additionally predict finer properties reflecting the traffic's variability at different scales.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

### 1.1. Motivations

A deep understanding of the properties of network traffic is important for Internet Service Providers to optimally control the traffic, to dimension the hardware and software, and eventually to offer users the best possible Quality of Service. In this context, a lot of research has been recently focusing on the mathematical modeling of network traffic, especially from a statistical viewpoint. However, comprehensive modeling of the traffic is a very arduous problem because it encompasses several difficulties of different natures, such as the transport protocols and their associated control mechanisms, the flows

[*] Corresponding author.
*E-mail addresses:* Patrick.Loiseau@inria.fr (P. Loiseau), Paulo.Goncalves@ens-lyon.fr (P. Gonçalves), barral@math.univ-paris13.fr (J. Barral), Pascale.Primet@ens-lyon.fr (P. Vicat-Blanc Primet).

0166-5316/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.peva.2010.08.016

dependencies and the complexity due to the network's topology. The design of models, simple yet rich enough to reproduce essential characteristics observed in the traffic, is then an important challenge for resource optimization, traffic control or prediction, with a real impact on industrial applications.

In particular, great attention has been devoted to the statistical modeling of the TCP protocol, the most widely used transport protocol in today's Internet. Following Padhye's paper [1] and its well-known square-root formula, many models have appeared in the last decade to predict the mean throughput achieved by a long-lived TCP connection in various network conditions (see next section). Such models are of great interest and have been extensively used because long-lived connections accurately model long TCP flows, whose ratio is constantly increasing in Internet traffic.

While it provides very useful information, prediction of the sole mean throughput may be insufficient for some purposes. For example, information about the throughput's fluctuations can be needed to evaluate the risks of congestion. An important step towards a better comprehension of the traffic, to optimize resource utilization at several levels, is then to characterize the variations of TCP traffic around its mean.

In this paper, we consider the flow's throughput averaged at scale $n$, that is throughput's sample mean calculated over time windows of size $n \times$ RTTs, with $n$ significantly larger than one but significantly smaller than the total flow length. For different values of the scale $n$, we demonstrate that it is possible to analytically derive the statistical distribution of the corresponding throughput, and hence to compute the probability that it deviates from its global mean value. We also prove that in practice, the throughput deviations that are observable on a finite-size realization of a TCP flow are statistically bounded.

These results complete the scope of Padhye's result in many different situations. For instance, it can be the case when the scale of interest is imposed by a buffer – recall that the condition $n \gg 1$ precludes network-level buffers – and we look for the maximum traffic variations at this scale. Then, our method allows determining the probability of congestion or starvation, depending on the problem at stake. Conversely, if we assume that the maximum acceptable deviation of the averaged throughput is now fixed, we may want to know at which minimum scale the traffic should be averaged to remain within this tolerable range. As an example, consider a generic server sequentially dealing with several clients. Given the elapsed time between two consecutive services of a same client, our method allows dimensioning the minimum size of the client buffer to ensure zero data loss. More generally, characterizing traffic variations beyond the mean throughput presents clear advantages in congestion and starvation problems. Yet, we do not elaborate any further on specific practical uses, and we focus on the technical aspects of the proposed method.

Our approach relies on a large-deviations principle that was recently proved to hold for any particular realization of a wide class of stochastic processes [2]. Here, we adapt this general result to the specific case of one long-lived TCP flow, and show that it allows predicting the deviations of the flow throughput around its mean value.

More precisely, it is the so-called *large-deviations spectrum*, a scale-invariant function, that conveys information about the distribution of the throughput measured over a given time-scale horizon. The large-deviations spectrum is reminiscent of a peculiar scaling law exclusively generated by the Markov property of the AIMD mechanism. As a result, it is important to stress that the present scale invariance must not be confused with long-range dependence, another scaling law whose impact on performance has already been extensively studied. Moreover, while long-range dependence was shown to be independent of the used protocol [3], this new scaling directly originates in TCP's control mechanisms, and the related large-deviations spectrum really is a straightforward fingerprint of the performance achieved by a TCP connection.

### *1.2. Summary of our contributions*

The present contribution is the methodological exploitation of a mathematical result we derived in [2], applied to the specific context of a TCP connection. We show that, in addition to the almost-sure mean value predicted by the existing models (such as the square-root formula), it is also possible to fully (statistically) characterize the deviation of the averaged throughput from its most frequent value.

More precisely, we show how to empirically estimate, on a TCP Reno connection, the so-called large-deviations spectrum of the throughput at different scales. In addition, we prove that a theoretical spectrum can analytically and consistently be derived from a Markov-chain model fitting the TCP time series. Practical validity of the proposed method is tested: (i) on simulations in the case of Bernoulli losses where an extensive comparison of our results with the results of Padhye's square-root formula is performed; and (ii) on real TCP traffic from controlled experiments. Then, we experimentally demonstrate that our model remains valid under less restrictive assumptions. Notably, we treat the case of softened Markov conditions, we confront our model to real uncontrolled Internet traffic, and we address examples of other TCP variants.

### *1.3. Organization of the paper*

The rest of the paper is organized as follows. In Section 2, we briefly review the works most closely related to our approach. We expose useful theoretical results in Section 3, in particular those of [2] on which the proposed method to describe the TCP throughput's variability relies. In Section 4, we present illustrations of the method in controlled situations, in particular in the Bernoulli case, using simulations and experiments. In Section 5, we present experimental validations of the method in more complex situations, and in particular on real-word Internet traffic. We also show an extension of the method to TCP traffic using a variant other than Reno.

## 2. Related work

Markov chains and similar tools have been extensively used in the last decade to model TCP traffic. All these models basically obtain steady-state information on the throughput, most often the first-order statistics (the mean), under various loss assumptions and including various complex TCP mechanisms (*e.g.*, timeout). The following is a non-exhaustive list of models that explicitly rely on the AIMD mechanism and are then conceptually the most related to our approach (bear in mind that we focus on long-lived TCP sources; papers interested in the case of multiple ON/OFF sources like [4–7] will thus not be reviewed here).

A first series of papers assumes Bernoulli losses (*i.e.* independent packet losses with probability $p_{\text{pkt}}$) [8,9,1], and use a simple Markov chain to model the congestion window's evolution. They obtain the well-known square-root formula for the mean throughput:

$$\mathbb{E}\{\text{throughput}\} \underset{p_{\text{pkt}}\to 0}{\sim} \sqrt{\frac{3}{2p_{\text{pkt}}}}, \quad \text{in packets per RTT.} \tag{1}$$

This famous formula has been extended to many different situations including the timeout mechanism [1] or the slow-start phase [10], maximal advertisement windows, fast retransmit and ack threshold [11], TCP variants other than Reno [12,13], to cite but a few. Yet, it is worth noticing that these generalizations assume Bernoulli losses and only focus on the determination of the mean throughput. Under similar assumptions and with the same objective in mind, several authors have used fluid models of TCP, rather than discrete Markov chains [14,15]. They obtained similar results.

Several propositions have appeared to handle more general loss processes than Bernoulli. In [16], the authors use a Markov chain to model the evolution of the congestion window of one TCP connection, and a discrete batch Markov arrival process is used to model the rest of the traffic. In [17], the authors concentrate on the congestion window just before a loss and they recover the mean throughput via Palm calculus, using a stationary ergodic loss process leading to another generalized square-root formula. This line of thought is extended in [18,19] to different TCP variants, but with a Bernoulli losses assumption.

The models mentioned above only consider the case of a single TCP connection. A few papers [20,21] concentrate on the case of multiple competing TCP connections with analytical results on fairness and on the impact of synchronization on performance, but their results are again mainly focused on first-order statistics. Finally, in [22], the authors use a fluid model of $N$ TCP sources reproducing the AIMD behavior. Using products of matrices, they derive the steady-state distribution from which they deduce, in addition to the first-order statistics, the autocorrelation function and the covariance matrix.

## 3. Large deviations for Markov chains: Theoretical results

In this section, we expose the fundamental theoretical tools and results supporting the method proposed to describe the TCP throughput's variability. Given the applied nature of the present contribution, simplicity has voluntarily been preferred to rigorous statement of the results developed in a companion mathematical article [2].

Throughout the paper, the congestion-window size in packets and the throughput in packets per RTT are considered equal, and both terms are used interchangeably.

### *3.1. Markov chain model and almost-sure mean*

Consider a discrete Markov chain $(W_i)_{i\in N_+}$ representing the congestion window size (in packets) of a long-lived TCP Reno flow at the $i$-th RTT instant. The state space is the set of possible congestion window values. Due to the maximal congestion window $w_{\max}$ imposed either by the sender or by the receiver, it is finite: $E = \{1, \dots, w_{\max}\}$. The possible transitions of this Markov chain are imposed by the AIMD mechanism, while their probabilities are dictated by the loss process. We introduce the loss probability function $p(w)$ corresponding to the probability that at least one packet of a congestion window is lost given its size $w$. The transition matrix of the Markov chain $P$ is then simply defined for all $w, w' \in E$ by:

$$P_{ww'} = \begin{cases} 1 - p(w) & \text{if } w' = \min(w+1, w_{\max}), \\ p(w) & \text{if } w' = \max(\lfloor w/2 \rfloor, 1), \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

In realistic situations, it is reasonable to assume that this Markov chain is irreducible and aperiodic. We will make this assumption in the rest of the article.

It is known that such a Markov chain possesses a unique steady-state distribution $\pi = (\pi_w)_{w\in E}$ [23]. Moreover, by Birkhoff's ergodic theorem, the sample mean at scale $N$, $\overline{W}^{(N)} = 1/N \cdot \sum_{i=1}^{N} W_i$, converges almost-surely (*i.e.* for almost every realization) toward the mean of this steady-state distribution when the window size $N$ (*i.e.* the time) grows to infinity:

$$\overline{W}^{(N)} = \frac{1}{N}\sum_{i=1}^{N} W_i \xrightarrow[N\to\infty]{\text{a.s.}} \overline{W}^{(\infty)} = \sum_{w\in E} w \cdot \pi_w. \tag{3}$$

Markov chains have been extensively used in the last decade as good models for the TCP congestion window's evolution (see Section 2), but the results have essentially remained limited to the prediction of the mean throughput (in packets per RTT), that is the almost-sure mean congestion window of Eq. (3) (in packets). In this work, we also use a Markov chain model, and we propose a method to describe the throughput's deviations around the a.s. mean.

### *3.2. Large-deviations*

As a consequence of Eq. (3), the probability that the mean throughput $\overline{W}^{(N)}$ at scale $N$ is different from $\overline{W}^{(\infty)}$ tends to zero as $N$ tends to infinity: it is called a *rare event*, though it may happen at a finite scale $N$. The elegant theory of large deviations offers a precise way to estimate the probability of such events. We first briefly expose classical results in the particular case of Markov chains introduced in the previous section to model TCP traffic; then we present the new result that we recently proved [2] which makes possible the present exploitation in the applied context of a single TCP flow's traffic characterization.

#### *3.2.1. Classical results*

For an irreducible aperiodic Markov chain $(W_i)_{i\in N_+}$ on a finite state space $E$, it is known that a large-deviations principle (LDP) holds [24]:

$$\lim_{\epsilon\to 0}\lim_{N\to\infty}\frac{1}{N}\log\mathbb{P}\left(\overline{W}^{(N)}\in[\alpha-\epsilon,\alpha+\epsilon]\right)=f(\alpha), \tag{4}$$

where the function $f(\cdot)$, called the *large-deviations spectrum*, is a concave function satisfying:

$$\begin{cases} f(\alpha)<0 & \text{if } \alpha\neq\overline{W}^{(\infty)},\\ f(\alpha)=0 & \text{if } \alpha=\overline{W}^{(\infty)}. \end{cases} \tag{5}$$

It can be deduced (at least numerically) from the transition matrix of the Markov chain,[1] *i.e.* in our case, the theoretical large-deviations spectrum $f$ can be simply deduced from the loss probability function $p$.

In words, the LDP gives an estimation of the asymptotic probability that the sample mean at scale $N$ is around a value $\alpha$:

$$\mathbb{P}\left(\overline{W}^{(N)}\sim\alpha\right)\underset{N\to\infty}{\sim}\mathrm{e}^{Nf(\alpha)}. \tag{6}$$

If $\alpha\neq\overline{W}^{(\infty)}$ it describes the exponential decay of the probability of rare events, with rate $f(\alpha)$ independent of the scale $N$; if $\alpha=\overline{W}^{(\infty)}$ the probability tends to 1, coherently with Eq. (3). For non-achievable values of the throughput (*e.g.*, here, negative values or values larger than $w_{\max}$), the function $f$ is $-\infty$, which gives a probability zero of achieving such throughputs even at finite scales.

Eq. (6) offers a precious way to characterize the deviations from the almost-sure mean at sufficiently large scales. However, the probability $\mathbb{P}$ involved refers to ensemble quantities (*i.e.* proportions observed on a large number of independent realizations); and it does not directly apply to characterize the variations of the throughput within a single TCP flow.

#### *3.2.2. A recent result on almost-every realization*

To describe the throughput's variations within a single TCP flow, we use an ergodic form of the LDP that we recently demonstrated to hold on almost every realization [2], stated in Theorem 1. In addition to being valid on one realization, this result also allows specifying the observable deviations on a finite-size realization; it is why we state it with a finite realization of size $N$ (although the theoretical validity holds for asymptotic values only).

We consider a finite-size realization $(W_i)_{i\in\{1,\dots,N\}}$ of the finite-state, irreducible, aperiodic Markov chain introduced earlier. At scale $n$, it is divided into $k_n=\lfloor N/n\rfloor$ consecutive intervals of size $n$, and the sample mean over the $j$-th such interval is denoted by $\overline{W}_j^{(n)}=1/n\cdot\sum_{i=(j-1)n+1}^{jn}W_i$. The sequence $(\overline{W}_j^{(n)})_{1\leq j\leq k_n}$ forms the averaged throughput at scale $n$, whose variations we characterize in this work, based on the following result.

**Theorem 1.** *For a given value $\alpha\in\mathbb{R}$, if the number of intervals at scale n is large enough in the sense:*

$$k_n\geq\mathrm{e}^{n(-f(\alpha)+\delta)},\quad \textit{for some } \delta>0, \tag{7}$$

*then the following LDP on the proportion of intervals where the throughput achieves a value around $\alpha$ holds almost-surely (i.e. on almost-every realization):*

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\frac{\#\left\{j\in\{1,\dots,k_n\}:\overline{W}_j^{(n)}\in[\alpha-\epsilon,\alpha+\epsilon]\right\}}{k_n}=f(\alpha), \tag{8}$$

*where the large-deviations spectrum f is the same as in the classical LDP of Eq. (4).*

[1] The large-devations spectrum $f$ is the Legendre–Fenchel transform of the logarithmic moment generating function $\Lambda$ : $f(\alpha)=\inf_{q\in\mathbb{R}}(\Lambda(q)-\alpha q)$; where $\Lambda(q)$ is obtained for any real $q$ as the spectral radius' logarithm of the matrix $A(q)$ with elements $(A(q))_{ij}=\exp(qj)\cdot P_{ij}$ ($P$ being the transition matrix of the Markov chain), see [24].

*Moreover, if the number of intervals $k_n$ is too small in the sense that $k_n \leq \mathrm{e}^{n(-f(\alpha)-\delta)}$ for some $\delta > 0$, then the set $\{j \in \{1, \ldots, k_n\} : \overline{W}_j^{(n)} \in [\alpha - \epsilon, \alpha + \epsilon]\}$ is almost-surely empty (for $\epsilon$ small enough and n large enough).*

Theorem 1 was originally stated under much wider conditions than those of a finite-state irreducible and aperiodic Markov chain: it is valid for any stationary process possessing sufficient mixing properties[2] and for which a classical LDP holds (see [2]). We present it here in a simplified form adapted to our purpose. Instead of discussing the proof, which is outside the scope of this paper and can be found in [2], we elaborate on its interpretation and application in the context of network traffic.

### 3.2.3. Interpretation and discussion

Theorem 1 is an ergodic interpretation of the classical LDP: it gives an estimation of the proportion of intervals of size $n \gg 1$, within a single realization of size $N \gg n$, where the throughput achieves a value around $\alpha$:

$$\frac{\#\,\{j \in \{1, \ldots, k_n\} : \overline{W}_j^{(n)} \sim \alpha\}}{k_n} \underset{1 \ll n \ll N}{\sim} \mathrm{e}^{nf(\alpha)}, \tag{9}$$

where the large-deviations spectrum $f$ is the same as the classical LDP and then satisfies conditions (5). The large-deviations spectrum reflects the throughput's distribution within a flow, at a "sufficiently large" scale $n$ (see next section for concrete values in practical situations). For instance, the spectrum's width reflects the traffic's variability: a large width corresponds to many significant deviations from the a.s. mean; while a negligible width corresponds to an almost deterministic evolution where the throughput at scale $n$ stays close to the a.s. mean.

Strikingly, the function $f$, which conveys all the information related to the throughput's distribution at any given scale, is precisely independent of this scale. This scale-invariance property is reminiscent of a scaling law derived from the LDP. Practically, to verify that this scale-invariance property holds on a particular traffic trace, we will compute an estimation of the large-deviations spectrum of Eq. (8) from the trace at different scales $n$ and observe whether the estimations at different scales superimpose.

The exponential decay of rare events given by the classical LDP (Eq. (4)) is a standard result. Yet, for lack of an ergodic form as the one provided by Theorem 1 (Eq. (8)), its use for TCP traffic characterization has never been fully exploited. Indeed, Theorem 1 justifies the possibility to observe large deviations around a value $\alpha$ on a single TCP connection's throughput, but also clarifies the "price to pay" to do so: the necessary number of intervals $k_n$ must grow exponentially fast with scale $n$ and with exponential rate $-f(\alpha)$ (condition (7)).[3]

On a finite-size realization of size $N$, the number of intervals at scale $n$ is constrained ($k_n = \lfloor N/n \rfloor$), and condition (7) is satisfied only if the rate $-f(\alpha)$ is small enough, *i.e.* if $\alpha$ lies between the bounds:

$$\begin{cases} \alpha_{\min}(n) = \min\{\alpha : \text{condition (7) is satisfied at scale } n\} = \min\{\alpha : \lfloor N/n \rfloor \geq \mathrm{e}^{-nf(\alpha)}\}, \\ \alpha_{\max}(n) = \max\{\alpha : \text{condition (7) is satisfied at scale } n\} = \max\{\alpha : \lfloor N/n \rfloor \geq \mathrm{e}^{-nf(\alpha)}\}. \end{cases} \tag{10}$$

This defines an interval $[\alpha_{\min}(n), \alpha_{\max}(n)]$ around the almost-sure value $\overline{W}^{(\infty)}$, of observable throughputs at scale $n$. The second part of Theorem 1 shows that throughputs outside this interval are almost-surely not observed.

The condition of Eq. (7) is natural and can be intuitively retrieved with a simple reasoning: to have a chance to be observed at least once, a $\mathrm{e}^{nf(\alpha)}$-rare event must be trialed at least $\mathrm{e}^{-nf(\alpha)}$ times. Then, for a fixed realization size $N$, the interval $[\alpha_{\min}(n), \alpha_{\max}(n)]$ that follows narrows down around $\overline{W}^{(\infty)}$ as $n$ grows and it can be used in two ways. If the scale $n$ is fixed, it bounds the observable throughputs on a realization of size $N$. On the other side, if the range of acceptable throughput values is fixed, it allows determining the minimal scale beyond which the averaged throughput is guaranteed to remain in this range.

In the next sections, we show the experimental evidence of the practical validity of these results on TCP traffic traces, both in terms of large-deviations spectrum and the bounds $\alpha_{\min}$ and $\alpha_{\max}$.

## 4. One TCP Reno connection in controlled situations

The numerical and experimental results of this section aim at illustrating the potential and the versatility of our method in a large set of situations.

### 4.1. Bernoulli losses and comparison with Padhye's model: Simulation approach

We start with the widely-used ideal case of Bernoulli losses that we treat with MATLAB simulations. We simulate a realization of a Markov chain of length $N = 2^{18}$, representing the congestion window sampled at an arbitrary unspecified RTT. The transition matrix is as in (2) with Bernoulli loss probability function $p(w) = 1 - (1 - p_{\text{pkt}})^w$. The packet loss rate

[2] Finite-state irreducible and aperiodic Markov chains are strongly mixing with exponentially decaying coefficients [25], which is sufficient for the result of [2] to hold.

[3] Since $\delta$ can be arbitrarily small in Eq. (7), we take in practice $\delta = 0$.

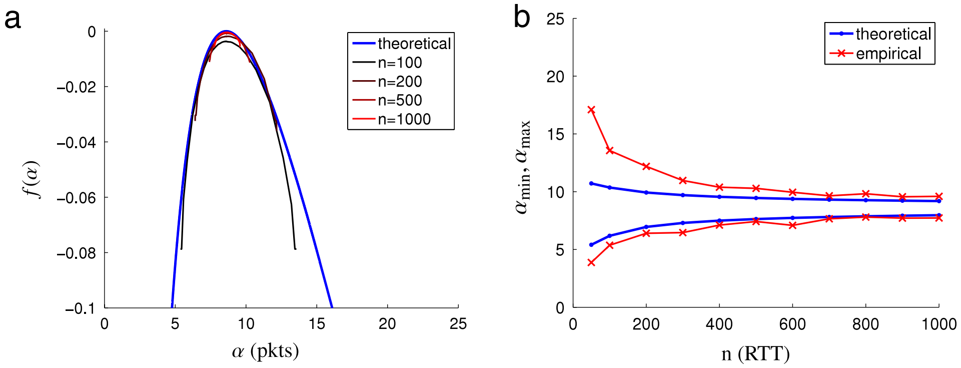


**Fig. 1.** One TCP Reno connection under Bernoulli losses with packet loss rate $p_{\mathrm{pkt}} = 0.02$ (obtained via MATLAB simulations): (a) large-deviations spectra, (b) minimal and maximal deviations.

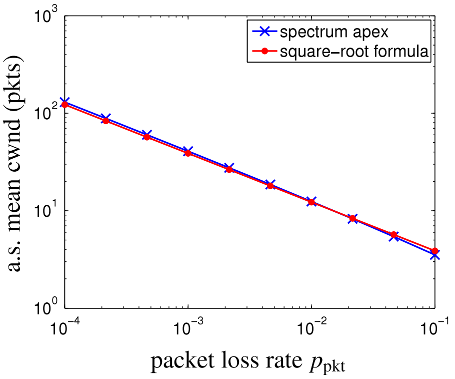


**Fig. 2.** Almost-sure mean congestion window of one TCP connection with Bernoulli losses: comparison of the spectrum apex with the square-root formula of [1] without timeout (Eq. (1)).

is fixed to $p_{\mathrm{pkt}} = 0.02$ and we impose a maximal congestion window $w_{\max} = 1000$ packets.[4] The theoretical spectrum is directly calculated from the Markov transition matrix, whereas the empirical spectra at different scales are estimated from the time series evolution of the congestion window.

Results are displayed in Fig. 1. First of all, let us notice the good match between the theoretical prediction and the estimations obtained for scales $n$ larger than 100 RTTs. Beyond technical validation of our procedure, these superimpositions are a clear evidence of the fact that the limit in Eq. (8) does converge toward a steady regime attained, in current practice, above the scale $n \sim 100$ RTTs. This minimal scale should actually depend on the dynamic speed of the system (*i.e.* the mean interval between two consecutive losses), but, as we shall see in the sequel, the scale $n = 100$ RTTs seems to be a typical threshold in many situations.

As expected, the apex of the curves in Fig. 1(a), corresponding to 8.6 packets (per RTT) in our simulations, is very close to the value predicted by the square-root formula and equal in that case to $\sqrt{3/2p_{\mathrm{pkt}}} = 8.66$. To fully confirm this observation, the plots of Fig. 2 represent the evolution with the packet loss rate, both of the almost-sure mean congestion window obtained by the square-root formula and of the apex of the theoretical spectrum. The almost perfect superimposition of the two lines clearly demonstrates the coherence of the elaborated method we propose with simpler models focusing on first-order statistics only.

Returning to the spectra of Fig. 1, the non-negligible widths of their support show that the almost-sure mean value is insufficient to fully describe the throughput's evolution. Indeed, at scale $n = 100$ for instance, the spectrum indicates that the probability to find an interval whose mean throughput is around 11 packets per RTT equals $e^{-0.01\times 100} = 0.37$, and cannot be reasonably neglected. Moreover, from the plots of Fig. 1(b), we conclude that the minimal and maximal observable deviations on a finite-size realization are fairly well estimated, although a significant prediction bias on $\alpha_{\max}$ exists at the smaller scales. This deviation notwithstanding, our results support relevance of the definitions of Eq. (10), and more fundamentally of Eq. (7) in practical situations.

Together with the spectra, the good prediction of the bounds $\alpha_{\min}$ and $\alpha_{\max}$ at each scale gives a fine description of the throughput variations that are observable on a finite-size trace, and of their corresponding probability. In the following, we

[4] The maximal congestion window is chosen sufficiently large to play no role at such packet loss rate.

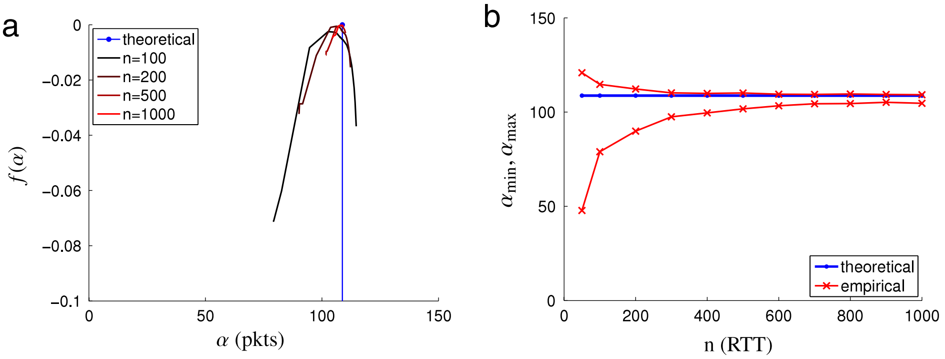


**Fig. 3.** One TCP Reno connection with constant UDP cross-traffic of rate 833 Mbps (obtained via controlled experiments with RTT = 12 ms): (a) large-deviations spectra, (b) minimal and maximal deviations.

experimentally demonstrate that the proposed method generalizes beyond the simple case of Bernoulli loss assumption, and yields as accurate descriptions of the throughput's variability in more realistic situations.

### 4.2. Congestion losses: Experimental approach

In this section, we use the large scale experimental facility Grid'5000 [26] to demonstrate the appropriateness of our model on real TCP traffic and the accuracy of the throughput's characterization. Grid'5000 is a fully controllable testbed constituted of 5000 CPUs spread over 9 geographical sites in France and interconnected with dedicated 10 Gbps links.

In all our experiments, the traffic is sent from the site of Lyon to the site of Rennes (RTT = 12 ms), and goes through a single bottleneck of 1 Gbps. The size of the limiting buffer is 96 packets, and corresponds to a 1.2 ms maximal queueing delay which does not significantly modify the RTT value.

One Long-lived TCP connection shares the bottleneck with two possible types of UDP cross-traffic:

- constant cross-traffic, generated by a single UDP source at 833 Mbps;
- random cross-traffic, generated by 32 ON/OFF UDP sources, with mean idle OFF times and mean active ON periods of 0.1 s and 0.12 s respectively. Each source's rate is limited at 40 Mbps, leading to a mean aggregate traffic of 580 Mbps.

The packet size of the TCP connection is fixed to 1500 Bytes. Duration of the experiment is set to 1 h, during which TCP parameters (congestion window and detected losses) are collected every 3 ms via web100 reports.[5] To avoid transient behaviors, we discard the first 60 s, and we retrieve from the remaining stationary trace the congestion-window size's time series resampled at the RTT rate. Results presented in the following stem from the analysis of these processes.

All the results of this section correspond to the Reno variant of TCP.

*Deterministic losses.* When the cross traffic is constant, the TCP throughput hits the available bandwidth (147 packets per RTT here) periodically, and the loss process is fully deterministic. The associated Markov chain itself is periodic and does not strictly enter the framework of Theorem 1. However, as the congestion window oscillates between 73 and 147 packets, its mean calculated over sufficiently large time windows should be constant and equal to $(147 + 72)/2 = 110$ packets. The corresponding theoretical spectrum is said to be degenerate, meaning that it reduces to a single point as depicted in Fig. 3. Effectively then, the empirical spectrum clearly converges toward the theoretical spectrum. The non-zero width of the empirical spectrum is due to a non-integer number of periods of the congestion-window evolution within a time window of size $n$, and it naturally decreases as the number of such periods increases with $n$.

*Non-Bernoulli random losses.* To produce random loss conditions that do not meet the Bernoulli assumption, we generate a random ON/OFF cross-traffic as described above. We first verify that the cross traffic aggregated at the buffer scale (1.2 ms) is normal (see Fig. 4(a)), due to a sufficient number of sources; and reasonably uncorrelated at time lags beyond 0.1 s, the mean ON time (see Fig. 4(b)). Note that we chose exponentially distributed ON and OFF periods to guarantee short-range dependent cross-traffic. Thus, as the loss probability does not depend on previous losses, the conditions of a Markov-chain-based model are satisfied. The impact of a long-range dependent cross-traffic will be discussed in the next section.

As a basic component of our model, we now specify the loss probability $p(\cdot)$, as a function of the congestion window. To do so, we can directly use the empirical observation displayed in Fig. 4(c). However, to still permit numerical studies when this empirical observation is not available, we also propose an alternative approach based on the following simple model. Assume that TCP packets are regularly spaced within each RTT period, leading to a locally constant TCP connection's bandwidth equal to $\mu_{\text{TCP}}(w) = \frac{w \cdot \text{pktsize}}{\text{RTT}}$ in bps; where $w$ is the current congestion window's size in packets. Within an RTT period, the total traffic simply equals the UDP cross-traffic shifted by $\mu_{\text{TCP}}(w)$. It inherits in particular Gaussian characteristics and variance from this cross-traffic. We then estimate the loss probability given the current congestion window's size $w$ as the probability

[5] See http://www.web100.org/.

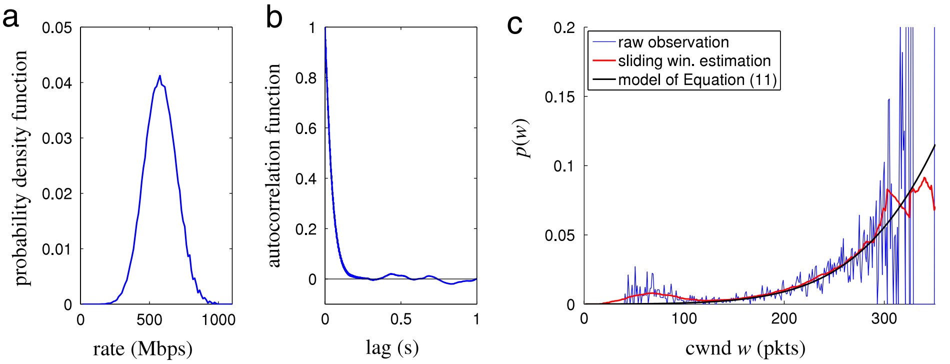


**Fig. 4.** Statistical characteristics of the ON/OFF UDP cross-traffic aggregated at the buffer scale (1.2 ms): (a) Probability density function estimation, (b) Autocorrelation function estimation. Loss probability function for one TCP Reno connection competing with this cross-traffic: (c) For the model of Eq. (11), we took the estimated value $\sigma = 108$ Mbps.

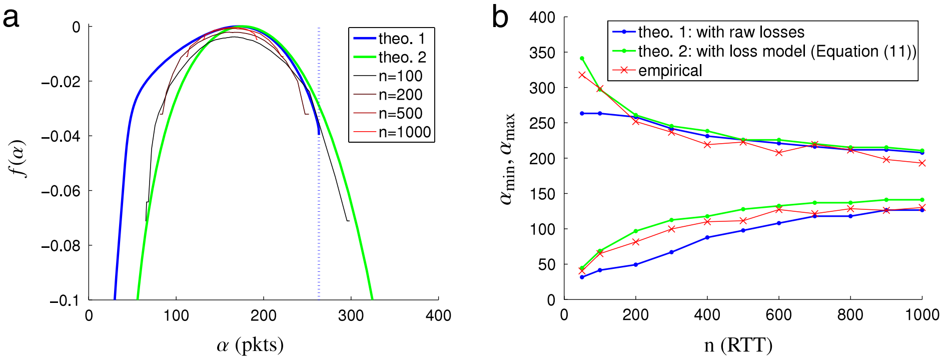


**Fig. 5.** One TCP Reno connection with random ON/OFF UDP cross-traffic of mean rate 580 Mbps (obtained via controlled experiments with RTT = 12 ms): (a) large-deviations spectra, (b) minimal and maximal deviations. The two theoretical curves are deduced from the Markov chain transition matrix using either the raw observed loss probability function of Fig. 4 (theo. 1, blue); or the model of Eq. (11), also displayed on Fig. 4 (theo. 2, green). If the raw observed loss probability is used, the loss probability 1 at a window of size 350 pkts theoretically limits the congestion window averaged in large time windows to a maximal value of $350 \times 3/4 = 262.5$ pkts due to the AIMD mechanism. This limit is materialized on the large-deviations spectra plots (a) by the dashed vertical bar (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for this total traffic aggregated at buffer scale to exceed the link capacity $C$, that is:

$$\begin{aligned} p(w) &= \mathbb{P}(\text{total traffic at buffer scale } > C), \\ &= \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{C - (\mu_{\text{TCP}}(w) + \mu_{\text{UDP}})}{\sigma\sqrt{2}}\right)\right); \end{aligned} \tag{11}$$

where $\sigma$ is the standard deviation of the UDP cross-traffic at buffer scale and erf denotes the error function arising from the cumulative Gaussian distribution.[6] Note that this reasoning can easily adapt to any distribution of the cross-traffic, and is not limited to the Gaussian case.

In Fig. 4(c), the model of Eq. (11) is compared to the empirically observed loss probability function. We observe a very good agreement for large congestion windows (roughly beyond 100 packets), supporting in particular the assumption that packets are spanned over the RTT. This spanning is due to the randomness of the packet sending mechanism, directly inherited from the natural randomness of real systems (CPU usage, queueing delays, etc.). The slight discrepancy between the model and the data, for shorter windows, is due to the non-negligible correlation of the cross-traffic at small time lags discussed earlier. This short-term persistence effect favors losses in consecutive RTT periods and thus accentuates the probability of TCP loss at rather small congestion windows.

Returning to the time series defining the evolution of the congestion window size, Fig. 5(a) markedly illustrates the scale-invariance property of the throughput distribution, through a clear superimposition of the empirical spectra obtained at different time scales ranging from 100 to 1000 RTTs. Moreover, the good agreement with the theoretical spectrum expression deduced from the Markov-chain model bears out the adequacy of this latter at accounting for the AIMD algorithm,

[6] Recall that a Gaussian random variable $X$ of mean $\mu$ and variance $\sigma^2$ has a cumulative distribution function: $\mathbb{P}(X \le C) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{C-\mu}{\sigma\sqrt{2}}\right)\right)$.

the principal factor responsible for the throughput's variations. Let us stress that our simplified Markov model remains satisfactory, even though more complex mechanisms, such as the fast recovery, exist in real TCP connections.

The loss probability overshoot experienced at small window sizes explains the difference existing for small $\alpha$'s, between theoretical predictions and empirical estimations of the large-deviations spectra. Because the loss model of Eq. (11) does not take into account this overshoot, it leads to an underestimated number of intervals with small mean throughputs. On the other hand, direct use of the empirical loss function in our Markov-chain model leads to overestimating the chances of repeatedly experiencing losses at small congestion windows over long periods whereas in reality, consecutive-losses periods responsible for the overshoot are followed by periods without loss increasing the congestion window (recall the persistence effect of the cross-traffic). The good global match between empirical and theoretical spectra confirms that this short-term persistence, that could easily be accounted for using an $m$-dependent Markov chain, has actually very slight consequences in our experiment.

Remarkably though, this actual loss probability overshoot does not seem to significantly bias the estimation of the observable throughput bounds $\alpha_{\min}$ and $\alpha_{\max}$ of Eq. (10), as shown in Fig. 5(b).

As it was the case for Bernoulli losses, Fig. 5 shows a very wide spectrum, characteristic of a highly varying throughput between 125 and 190 packets, averaged over 1000 RTTs. Unquestionably, and despite its asymptotic nature, the result of Eq. (8) is conclusive and effective in practice for scales larger than 100 RTTs.

## 5. Further validation and extension

In this section, we demonstrate the accuracy of TCP Reno throughput prediction provided by our method in more complex environments, including long-range dependent cross-traffic and real Internet traffic. We also propose an extension to different TCP variants.

### *5.1. One TCP Reno connection with long-range dependent cross-traffic*

We start with controlled experiments on the Grid'5000 testbed, similar to the ones presented in the previous section, where we use 45 ON/OFF TCP sources instead of introducing artificial UDP cross-traffic. The OFF times are still exponential (of mean 0.48 s), but in accordance with the situation commonly observed in real Internet traces [27], the ON periods are drawn from a heavy-tailed distribution of tail index 1.1 (and of mean 0.48 s). It is known that such heavy-tailed flows yield strongly long-range dependent aggregate traffic [28], contrasting with the supposed uncorrelated traffic of the previous section. Duration of the experiment is set to 2 h.

Then, focusing on one long TCP flow (more than one hour), the loss probability function (as deduced from the web100 report) does not match a Bernoulli assumption (with the measured packet loss rate: 0.6%), nor it has a clearly identifiable shape (Fig. 6(a)). Moreover, the long-term correlations of the cross-traffic it experiences theoretically invalidate the Markov property of the congestion window's evolution, which implicitly relies on the assumption that the loss probability depends only on the current congestion window's size.

Yet, Fig. 6(b) and (c) show that our Markov-chain-based model (using the empirically observed loss function), leads to very accurate predictions both in terms of large-deviations spectrum and minimal and maximal deviation bounds. This confirms the intuition that, despite the slow decrease of the cross-traffic autocorrelation function, its value at large enough time lags is small enough to consider that two consecutive loss events are independent. As a result, long-range dependent cross-traffic, very common in the Internet, does not hamper in practice the use of the our Markov model.

### *5.2. Real web traffic*

As the final stage of assessment of our model's adequacy, we now turn to real traces observed in Internet traffic, and confront effective TCP performance with our prediction. To this end, we analyze the traffic corresponding to the download of a large file from the xmission mirror[7] in the USA, to the INRIA Paris-Rocquencourt research center in France, on March 19, 2010 between 8 am and 10:30 am (UTC time). About 20 routers separate these two locations, the mean RTT is 175 ms and the mean rate during the transfer is 4 Mbps.

During the transfer, the packet-level trace is recorded using *tcpdump*, from which we recover the number of packets in consecutive time windows of size RTT that we assimilate to the congestion window time series. The TCP variant used on the xmission server is Reno, and identifying an AIMD structure of the congestion window's evolution yields the estimation of the loss probability function displayed on Fig. 7(a). Then, Fig. 7(b) and (c) show that the resulting theoretical predictions almost perfectly match the empirical estimations of the spectra and observable-deviations bounds $\alpha_{\min}$ and $\alpha_{\max}$ at different scales. The robustness with respect to complex and non-controlled conditions of the Internet (non-negligible and variable queueing delays, possibly long-range dependent traffic, etc.), combined with the simplicity of the Markov-chain setting, is a strong asset of the proposed model that warrants its applicability and effectiveness in many real situations of interest. Finally, Fig. 7(b) and (c) show that, again, the asymptotic regime of Eq. (8) is fairly well achieved at scales larger than $n \sim 100$ RTTs, which is of practical relevance in a host of different applications.

[7] http://mirror.xmission.com.

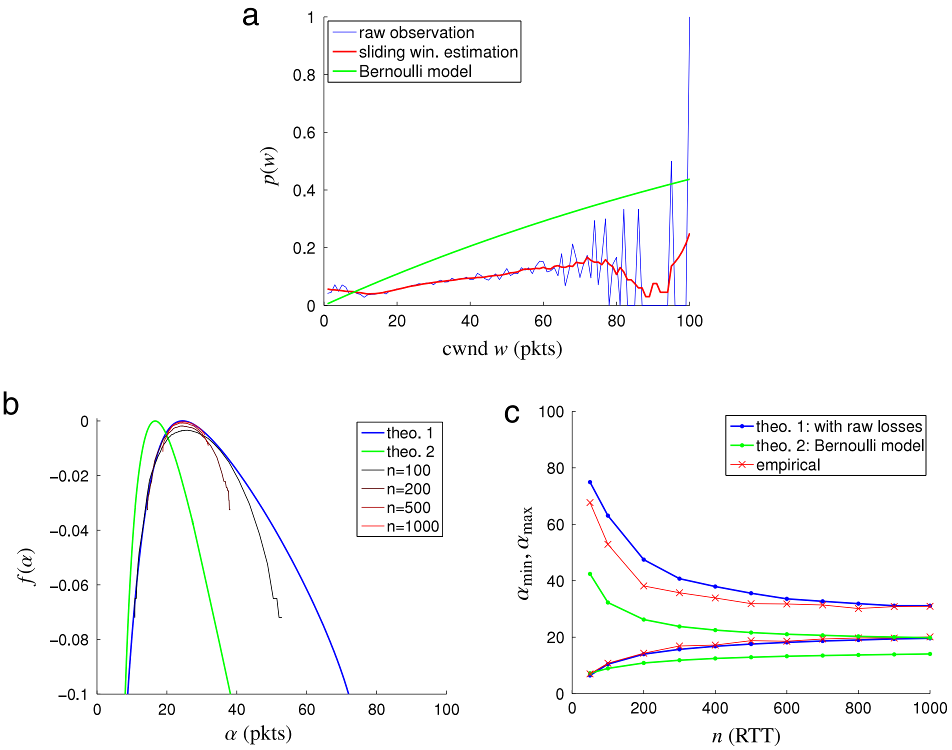


**Fig. 6.** One long TCP flow among 45 ON/OFF TCP sources (obtained via controlled experiments with RTT = 12 ms): (a) loss probability function (the measured packet loss rate 0.6% is used for the Bernoulli model), (b) large-deviations spectra, (c) minimal and maximal deviations.

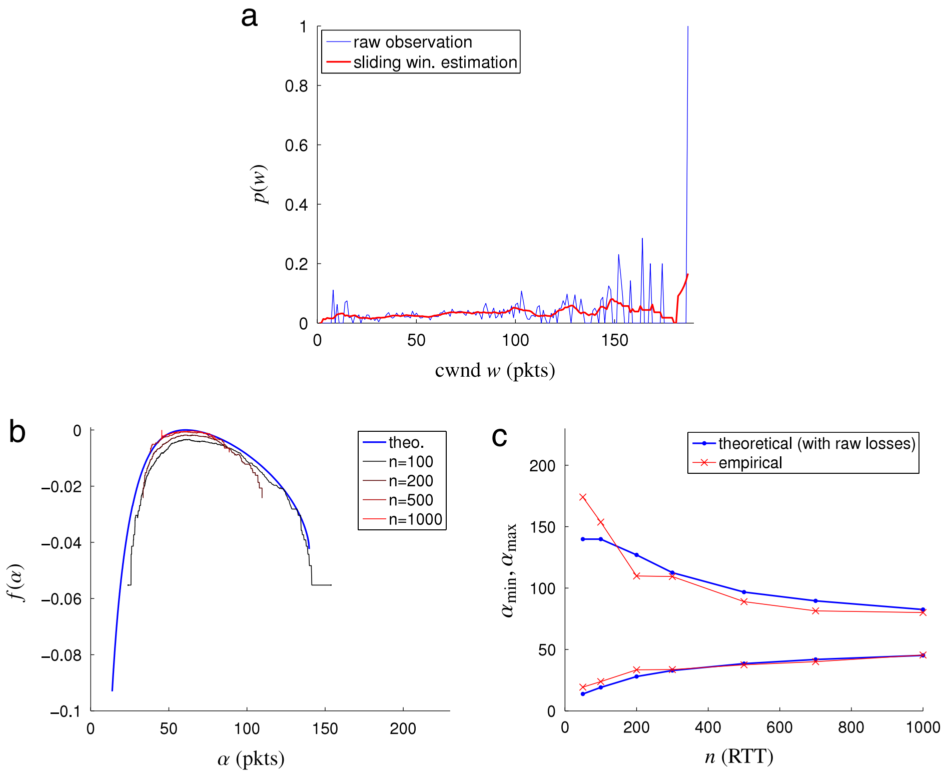
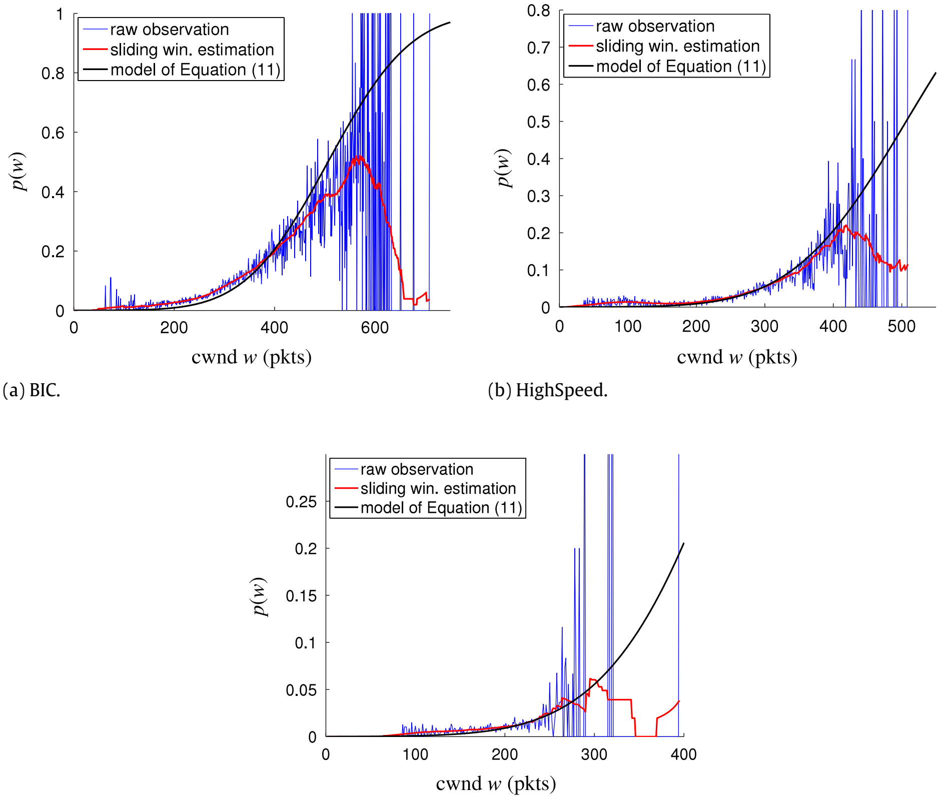


**Fig. 7.** One real-world TCP flow from the Internet: (a) loss probability function, (b) large-deviations spectra, (c) minimal and maximal deviations.



(a) BIC.

(b) HighSpeed.

(c) Cubic.

**Fig. 8.** One TCP connection in ON/OFF UDP cross traffic: loss probability as a function of the congestion window for (a) BIC, (b) HighSpeed, (c) cubic (*cf.* also Fig. 4(c) for the Reno case). For the model of Eq. (11), we took the estimated value of the standard deviation $\sigma = 108$ Mbps.

### *5.3. Other TCP variants than Reno*

We finally present in this section experimental evidence that the method developed above extends to the characterization of TCP traffic using several non-Reno variants: BIC, HighSpeed and Cubic.

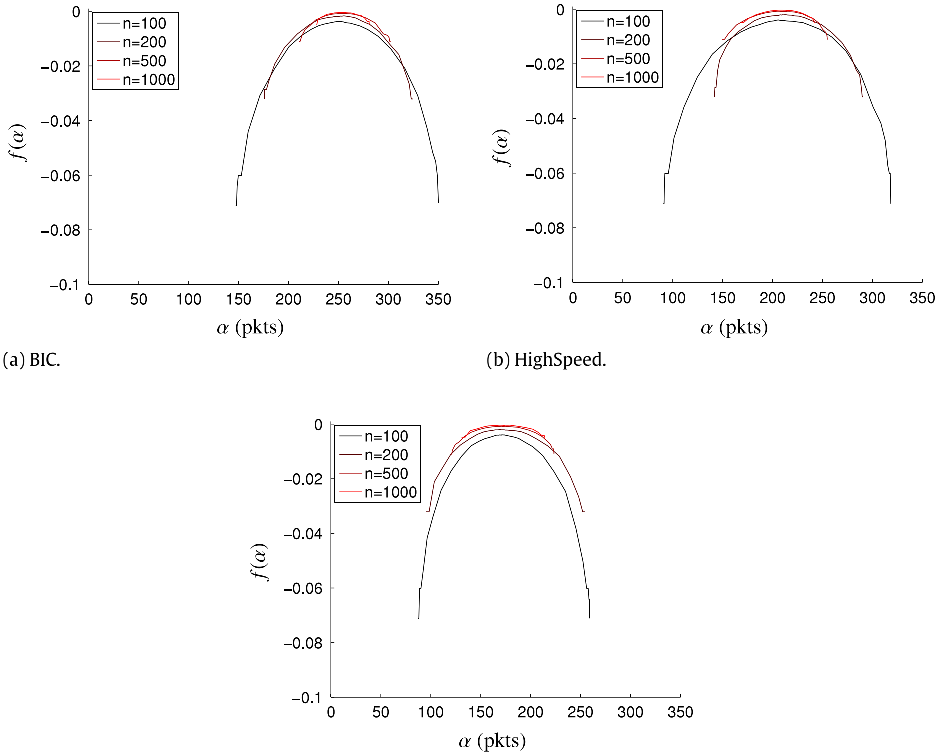
Under the same experimental conditions as in Section 4.2 with ON/OFF UDP cross-traffic, Fig. 8 shows that Eq. (11) still accurately models the loss probability function for the tested variants. The deviation observed for large congestion window sizes is simply due the length of the trace that is too short to significantly sample these rare occurrences.

Finally, Fig. 9 shows, for each TCP variant, a clear superimposition of the empirical spectra estimated at different scales, which attests for the presence of a scale-invariance property, induced by a LDP as the one expressed in Eq. (8). To permit a theoretical prediction of the corresponding large-deviations spectrum, an appropriate Markov-chain model of the congestion-window evolution would be required. It can be obtained by including additional information into the model, such as the size of the congestion window at the latest loss event in the BIC case. Then, our experimental results show that the method proposed in this work, and extensively tested in the Reno case, will extend to other variants and permit fine characterization of the corresponding TCP variant's performance.

## 6. Conclusion

In this work, we proposed a method to predict TCP throughput's variations. This method relies on a recent ergodic large-deviations result that we theoretically derived in [2], and that we applied here to a simple Markov-chain model of the evolution of TCP Reno's congestion window. The striking accuracy of the model to characterize real traffic in terms of these elaborated large-deviations properties reinforces the adequacy between Markov models and TCP data that had been observed in prior works focusing on simpler properties such as the mean throughput (see *e.g.* [1]).

More precisely, we showed how to compute, from the loss probabilities imposed by the network conditions, a theoretical prediction of the so-called *large-deviations spectrum*, a scale-invariant function which contains precise information on TCP performance. Then, it is possible to deduce, at any scale, two essential properties of a single TCP connection's throughput: its distribution and the extremal values observable on a finite-size trace. Such information significantly extends the existing predictions of the mean throughput and consequently is expected to leverage on systems' design and prediction. In particular, it is of special interest in congestion or starvation problems where anticipation of the flows' dynamic is essential.



(a) BIC.

(b) HighSpeed.

(c) Cubic.

**Fig. 9.** Large-deviations spectra of one TCP connection with ON/OFF UDP cross traffic, for different TCP variants: (a) BIC, (b) HighSpeed, (c) cubic (*cf.* also Fig. 5(a) for the Reno case).

We presented extensive experimental evidence of the accuracy of our performance prediction in common idealized situations such as that of Bernoulli losses. We also demonstrated that it remains perfectly adequate in more complex environments (*e.g.* long-range dependent cross-traffic) and, most importantly, on real Internet traffic. Finally, we presented experimental results showing that the scale-invariance property at the root of the proposed method's validity is also present and observable in non-Reno TCP traces.

A first direction to extend this work will be to build the appropriate Markov model to allow theoretical predictions of the large-deviations spectrum for these different TCP variants. Then, an extension to higher-dimension Markov chains will be considered to handle the case of multiple competing TCP connections. In this case, our method will apply to finely characterize several performance metrics, including the important notion of *fairness* at different scales. It will be useful, for example, to study *TCP friendliness* of new variants developed for the future Internet.

## Acknowledgements

Part of the work presented in this paper was done while the first author was with Université de Lyon/École Normale Supérieure de Lyon, Lyon, France. The authors are very grateful to the ALADDIN-G5K initiative whose developers and engineers provided us with full support to perform our experiments on the Grid'5000 infrastructure. They also would like to warmly thank Pr. Richard Baraniuk of the DSP department at RICE university (TX, US) for allowing us to use their computer resource to perform experiments. They finally thank anonymous reviewers for their useful comments that helped improve the paper.

## References

[1] J. Padhye, V. Firoiu, D. Towsley, J. Kurose, Modeling TCP throughput: a simple model and its empirical validation, in: Proceedings of ACM SIGCOMM'98, 1998, pp. 303–314.
[2] J. Barral, P. Loiseau, Large deviations for the local fluctuations of random walks and new insights into the "randomness" of Pi, Preprint. arXiv:1004.3713 (April 2010).
[3] P. Loiseau, P. Gonçalves, G. Dewaele, P. Borgnat, P. Abry, P. Vicat-Blanc Primet, Investigating self-similarity and heavy-tailed distributions on a large scale experimental facility, IEEE/ACM Transactions on Networking 18 (4) (2010) 1261–1274.
[4] C. Casetti, M. Meo, A new approach to model the stationary behavior of TCP connections, in: Proceedings of IEEE INFOCOM'00, 2000, pp. 367–375.
[5] A. Wierman, T. Osogami, J. Olsén, A unified framework for modeling TCP-Vegas, TCP-SACK, and TCP-Reno, in: Proceedings of IEEE MASCOTS'03, 2003, pp. 269–278.

[6] F. Baccelli, D.R. McDonald, A stochastic model for the throughput of non-persistent TCP flows, in: ACM Valuetools'06: Proceedings of the 1st International Conference on Performance Evaluation Methodolgies and Tools, 2006.

[7] F. Baccelli, K.B. Kim, D.R. Mcdonald, Equilibria of a class of transport equations arising in congestion control, Queueing Systems: Theory and Applications 55 (1) (2007) 1–8.

[8] M. Mathis, J. Semke, J. Mahdavi, T. Ott, The macroscopic behavior of the TCP congestion avoidance algorithm, ACM SIGCOMM—Computer Communication Review 27 (3) (1997) 67–82.

[9] T.V. Lakshman, U. Madhow, The performance of TCP/IP for networks with high bandwidth-delay products and random loss, IEEE/ACM Transactions on Networking 5 (3) (1997) 336–350.

[10] S. Fortin-Parisi, B. Sericola, A Markov model of TCP throughput, goodput and slow start, Performance Evaluation 58 (2–3) (2004) 89–108.

[11] I. Kaj, J. Olsén, Throughput modeling and simulation for single connection TCP-Tahoe, in: Proceedings of the 17th International Teletraffic Congress, ITC 17, 2001.

[12] B. Sikdar, S. Kalyanaraman, K.S. Vastola, Analytic models for the latency and steady-state throughput of TCP Tahoe, Reno, and SACK, IEEE/ACM Transactions on Networking 11 (6) (2003) 959–971.

[13] A. Kumar, Comparative performance analysis of versions of TCP in a local network with a lossy link, IEEE/ACM Transactions on Networking 6 (4) (1998) 485–498.

[14] V. Misra, W.-B. Gong, D. Towsley, Stochastic differential equation modeling and analysis of TCP-windowsize behavior, in: Proceedings of Performance'99, 1999.

[15] A. Budhiraja, F. Hernández-Campos, V.G. Kulkarni, F.D. Smith, Stochastic differential equation for TCP window size: analysis and experimental validation, Probability in the Engineering and Informational Sciences 18 (1) (2004) 111–140.

[16] A.E. Kamal, Discrete-time modeling of TCP Reno under background traffic interference with extension to RED-based routers, Performance Evaluation 58 (2–3) (2004) 109–142.

[17] E. Altman, K. Avrachenkov, C. Barakat, A stochastic model of TCP/IP with stationary random losses, IEEE/ACM Transactions on Networking 13 (2) (2005) 356–369.

[18] A. Blanc, K. Avrachenkov, D. Collange, G. Neglia, Compound TCP with random losses, in: Proceedings of Networking'09, 2009, pp. 482–494.

[19] A. Blanc, K. Avrachenkov, D. Collange, Comparing some high speed TCP versions under Bernoulli losses, in: Proceedings of PFLDnet'09, 2009.

[20] O. Ait-hellal, E. Altman, D. Elouadghiri, M. Erramdani, N. Mikou, Performance of TCP/IP: the case of two controlled sources, in: Proceedings of ICCC'97, 1997, pp. 469–477.

[21] P. Hurley, J.-Y. Le Boudec, P. Thiran, A note on the fairness of additive increase and multiplicative decrease, in: Proceedings of the 16th International Teletraffic Congress, ITC 16, 1999, pp. 467–478.

[22] F. Baccelli, D. Hong, AIMD, fairness and fractal scaling of TCP traffic, in: Proceedings of IEEE INFOCOM'02, 2002, pp. 229–238.

[23] L. Breiman, Probability, SIAM, 1992.

[24] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Springer, 1998.

[25] R.C. Bradley, Basic properties of strong mixing conditions. A survey and some open questions, Probability Surveys 2 (2005) 107–144.

[26] R. Bolze, F. Cappello, E. Caron, M. Daydé, F. Desprez, E. Jeannot, Y. Jégou, S. Lanteri, J. Leduc, N. Melab, G. Mornet, R. Namyst, P. Vicat-Blanc Primet, B. Quetier, O. Richard, E.-G. Talbi, I. Touche, Grid'5000: a large scale and highly reconfigurable experimental grid testbed, International Journal of High Performance Computing Applications 20 (4) (2006) 481–494.

[27] M.E. Crovella, A. Bestavros, Self-similarity in World Wide Web traffic: evidence and possible causes (extended version), IEEE/ACM Transactions on Networking 5 (6) (1997) 835–846.

[28] W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson, Self-similarity through high-variability: statistical analysis of ethernet LAN traffic at the source level, in: Proceedings of ACM SIGCOMM'95, 1995, pp. 100–113.

**Patrick Loiseau** received a degree of Professeur-Agrégé de Sciences-Physiques (2005), a M.S. degree of physics (2006), and a Ph.D. degree of computer science (2009) from École Normale Supérieure de Lyon. He also received a M.S. degree of mathematics (2010) from Université Pierre et Marie Curie (Paris 6) and École Polytechnique. He is currently working as a post-doctoral fellow at INRIA Paris-Rocquencourt.

His main research interests are in probability and in statistical analysis and modeling of network traffic and performance. This includes wavelet-based analysis of scaling phenomena, long-range dependent and multifractal models, large deviations for Markov processes, and queuing theory with correlated input processes. He is also interested in statistical estimation methods and their application to network measurement problems; and in the analysis and modeling of the heart-rate variability using control theory.

**Paulo Gonçalves** graduated from the Signal Processing Department of CPE Lyon, France in 1993. He received the M.S. and Ph.D. degrees in signal processing from INPG, France, in 1990 and 1993 respectively. While working toward his Ph.D. degree, he was with ÉNS Lyon. In 1994–96, he was a Postdoctoral Fellow at Rice Univ., US. Since 1996, he is an associate researcher at INRIA, first with Fractales (1996–99), then with is2 (2000–2003) and now with team RESO at the Parallel Computing Lab. (LIP), ÉNS Lyon. From 2003 to 2005, he was on leave at IST Lisbon, Portugal.

His research interests are in multiscale analysis and in wavelet-based statistical inference. His principal application is in metrology and deals with grid-traffic statistical characterization and modeling for protocol quality assessment and control.

**Julien Barral** received the Professeur-agrégé degree in Mathematics in 1995, and the M.S., Ph.D. and Habilitation à Diriger des recherches degrees from Paris-Sud University respectively in 1995, 1997 and 2005.

From 1999 to 2001, he was a Maître de conférences with Montpellier-II University and did his reserach in the Convex Analysis team. Then he was Associate Researcher with INRIA, from 2001 to 2005 in Fractales team, and from 2005 to 2009 in Sisyphe team. He is currently a Professeur in Mathematics with Paris-Nord University, where he is a member of Ergodic Theory and Dynamical Systems team.

**Pascale Vicat-Blanc Primet** is senior researcher at the National Institute of Research in Computer Science (INRIA) since 2005. Since 2002, she has been leading the INRIA RESO team (22 researchers and engineers) within the LIP laboratory of École Normale Superieure de Lyon. Since the beginning of 2008, she is also leading the ‘‘Semantic Networking’’ research team of the INRIA-Bell Labs common laboratory.

Her research interests include High-Speed and High-Performance Networks, Internet protocols’ design and architecture, Quality of Service, network and traffic measurement, Network programmability and virtualization, and Grid networking. She is a member of the scientifical committee of Grid5000’s/ALADDIN—French Computer Science Grid initiative. She has published more than 80 papers in International Journal and Conferences in Networking and Grid computing. She obtained her Habilitation à Diriger les Recherches from Université de Lyon in 2002, her Ph.D. (88) in Computer Science, MsC (84) and Engineer diploma (84) in CS from INSA de Lyon.