AFRL-IF-RS-TR-2005-422
Final Technical Report
January 2006

MULTISCALE TRAFFIC PROCESSING TECHNIQUES FOR NETWORK INFERENCE AND CONTROL

Rice University

Sponsored by
Defense Advanced Research Projects Agency
DARPA Order No. K187

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2005-422 has been reviewed and is approved for publication

APPROVED: /s/

GREGORY HADYNSKI
Project Engineer

FOR THE DIRECTOR: /s/

WARREN H. DEBANY, JR., Tech Advisor
Information Grid Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	JANUARY 2006	Final Apr 2000 – Jun 2005

4. TITLE AND SUBTITLE
MULTISCALE TRAFFIC PROCESSING TECHNIQUES FOR NETWORK INFERENCE AND CONTROL

5. FUNDING NUMBERS
C - F30602-00-2-0557
PE - 62301E
PR - K187
TA - 18
WU - C1

6. AUTHOR(S)
Richard Baraniuk, Rudolf Riedi, Edward Knightly, Robert Nowak

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Rice University, MS-16
6100 Main Street
Houston Texas 77005

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington Virginia 22203-1714

AFRL/IFGC
525 Brooks Road
Rome New York 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

AFRL-IF-RS-TR-2005-422

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Gregory Hadynski/IFGC/(315) 330-4094/ Gregory.Hadynski@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

This final report overviews Rice University's accomplishments in the project "Multiscale Traffic Processing Techniques for Network Inference and Control" in the Defense Advanced Research Projects Agency's, DARPA, Network Modeling & Simulation program. The overall goal of this project was the development of new traffic models, analysis techniques, and model-based processing algorithms for dealing with the type of bursty traffic that can saturate a link or network. Rice's new models are based on the powerful theory of multifractals and represent a step forward in traffic modeling technology. In addition to providing a better understanding of the traffic mechanisms that cause burstiness, they have integrated model fitting, synthesis, inference and prediction within one rigorous statistical framework. Their new models, designed to match salient traffic characteristics at many levels of abstraction and over a broad range of time scales, offer realism while remaining analytically and empirically tractable, statistically robust, and computationally efficient.

14. SUBJECT TERMS
Network traffic analysis, bandwidth estimation, congestion control

15. NUMBER OF PAGES
35

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents

List of Figures

1.0 Overview

This final report overviews our accomplishments in the project "Multiscale Traffic Processing Techniques for Network Inference and Control" in the Defense Advanced Research Projects Agency's (DARPA) Network Modeling & Simulation (NMS) program. This project focused experts from the fields of networking, statistical signal processing, and applied mathematics towards the goal of analyzing, modeling, and improving network performance relying solely on Commercial-Off-The-Shelf (COTS) network technology.

The overall goal of this project has been the development of new traffic models, analysis techniques, and model-based processing algorithms for dealing with the type of bursty traffic that can saturate a link or network. Our new models are based on the powerful theory of multifractals and represent a revolutionary step forward in traffic modeling technology. In addition to providing a better understanding of the traffic mechanisms that cause burstiness, we have integrated model fitting, synthesis, inference and prediction within one rigorous statistical framework. Our new models, designed to match salient traffic characteristics at many levels of abstraction and over a broad range of time scales, offer unprecedented realism while remaining analytically and empirically tractable, statistically robust, and computationally efficient, allowing control of vast and continually evolving networks, all within the context of a scalable and easily deployable COTS network core.

Parametric traffic models, which distill a complicated workload process into a number of simple parameters, have led to many new insights into network behavior. While traditional homogeneous Markov/Poisson processes cannot realistically capture traffic variability, burstiness, and long-range dependence (LRD), a number of promising new frameworks have been developed recently.

The take-off point for this project was our Multifractal Wavelet Model (MWM) for network traffic traces. Like wavelet-based models for fractional Brownian motion (fBm), a Gaussian LRD process, the MWM is based on the efficient computational tree structure of the wavelet transform. In contrast to fBm models, which analyze and synthesize traffic in terms of an additive superposition of multiscale Gaussian innovations, the MWM employs multiplicative innovations to naturally analyze and synthesize bursty, non-Gaussian, positive traffic. The MWM can be interpreted in terms of iteratively partitioning the total traffic load into finer and finer scale aggregations.

A natural explanation for the close match between the MWM and real traffic has been provided by the rich theory of multifractals. A multifractal is a random process whose (higher-order) moments are self-similar under rescaling. Technically, the multifractal spectrum measures in a signal the occurrence frequency or the density of `atomic bursts' of a certain strength, as measured by their local Holder exponent. The close match of the spectra of real and MWM traffic, especially for the key large bursts, provides yet another quantitative proof of the accuracy and realism of the MWM approach. In one powerful multiscale framework, multifractal analysis encapsulates large deviations theory (for predicting rare burst events), local smoothness estimation (for measuring strength of bursts), and higher-order moments (for characterizing non-Gaussianity and for relating burst

strength with burst frequency) and allows us to study traffic at both macroscopic and microscopic time scales.

Thanks to the simplicity of our multifractal models, we have derived informative schemes for inferring network path dynamics exclusively from end-to-end measurements. End-to-end measurements are of critical importance in large networks, since data collection and analysis at every node is a hopelessly complex task.

In our first effort we modeled the path between two network nodes of interest as dominated by a single "bottleneck" link and seek to estimate the multifractal properties of the competing cross-traffic through the bottleneck (see Figure 1). Of course, this one cross-traffic stream represents the amalgamation of all cross-traffic faced between the two nodes. Modeling the cross-traffic as an MWM process, for example, we used the delay spread between a set of probe packets sent along the path to form Bayesian estimates of both the instantaneous traffic strength and its statistics at different aggregation levels. Both active and passive probing is supported. Inspired by the multiscale nature of the successful MWM model we have designed and improve probing schemes based on the novel idea of variable rates within a train of probes. For verification we instrumented the Internet as well as a fast version of ns2, the widely accepted network simulator.

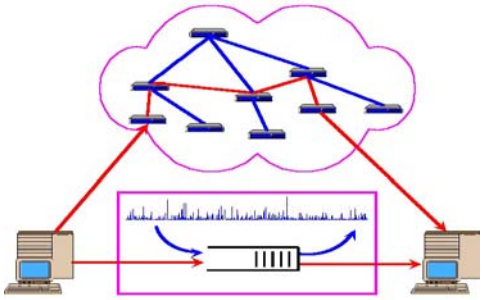


Figure 1: Network abstraction for modeling, bandwidth inference, and connection level analysis

Using path dynamics for congestion control is not a new idea in itself. In the past, however, such control has been based on losses instead of true traffic loads and on the traffic LRD parameter only, not its multifractal structure. Immediate applications of our more accurate approach include improved congestion control algorithms, pricing on a connection basis, and load balancing at web-servers servicing several flows. Little is understood of the origins of network traffic's multifractal behavior. In our second effort we aim at shedding light on this fundamental issue. We are developing and continuously expanding a multifractal analysis toolbox and with it conduct a careful statistical analysis to explore the potential causes, including network variability, multiplexing, protocols, and feedback. As an item of particular interest we have designed a tool to separate connections using the same link into two classes, one formed by only a few aggressive "alpha" connections and the second formed by the more smooth bulk of "beta" connections.

Our analysis of many traces concludes that typically only one alpha connection is active at a time and that alpha connections are essentially large file transfers (elsewhere called "elephants") over fast paths. Since the alpha connections are entirely responsible for the burstiness of traffic, this separation schemes explains the impact of network topology on the complexity of traffic dynamics, an issue that was not recognized up to now. Besides novel approaches to control, this separation scheme has given rise to network-user driven models, allowing for more realistic simulations and ranking of network configurations and topologies.

Launching from the foundation provided our recent leading-edge research on network measurement, multifractal signal analysis, multiscale random fields and signal processing, our effort consisted of four closely integrated research thrusts that directly attack several key networking challenges of the DARPA NMS program:

Thrust 1: Multiscale Traffic Analysis and Modeling Techniques

Thrust 2: Spatio-Temporal Available Bandwidth Estimation

Thrust 3: Connection-level Analysis and Modeling of Network Traffic

Thrust 4: Safe High-Speed Congestion Control

2.0 Multiscale traffic analysis and modeling

We have designed novel models that capture the multiscale variability and burstiness of high-speed network traffic. Leveraging *wavelets* and the powerful theory of *multifractals*, we are integrating model fitting, synthesis, and prediction into one unified statistical framework. Our new models, designed to match salient traffic characteristics at a prescribed level of abstraction, offer unprecedented realism while remaining analytically tractable, statistically robust, and computationally efficient. Using multifractal models, we studied how large traffic flows interact and distribute their burstiness. Furthermore, we are investigating, analyzing, and characterizing the (adverse) modulation TCP/IP places on *application-level* traffic.

Accomplishments: Burstiness in high-speed network traffic increases the queuing at routers and degrades performance. We have developed two powerful traffic models, the ***multifractal wavelet model*** (MWM) and the ***alpha/beta ON-OFF model***, that accurately and efficiently capture the statistical burstiness of high-speed traffic. In the alpha/beta decomposition, we have used connection-level information to identify that a very small percentage of connections cause nearly all of the bursts in traffic traces. The rest of the connections aggregate into a simple fractional Gaussian noise process (see Figure 2). We feel this discovery could have far-reaching implications in a number of areas, including modeling, queuing, scalability, and understanding the effect of network topology on traffic, synthesis, and inference (probing schemes, intrusion or hot spot detection). As of late 2004, the MWM model has been available in an open-source distribution from the spin.rice.edu/DARPA web page. These tools are currently being integrated into the ns-2 network simulator, and a complete open-source analysis toolbox has been released. We discuss the alpha/beta model in greater detail in Thrust 3.

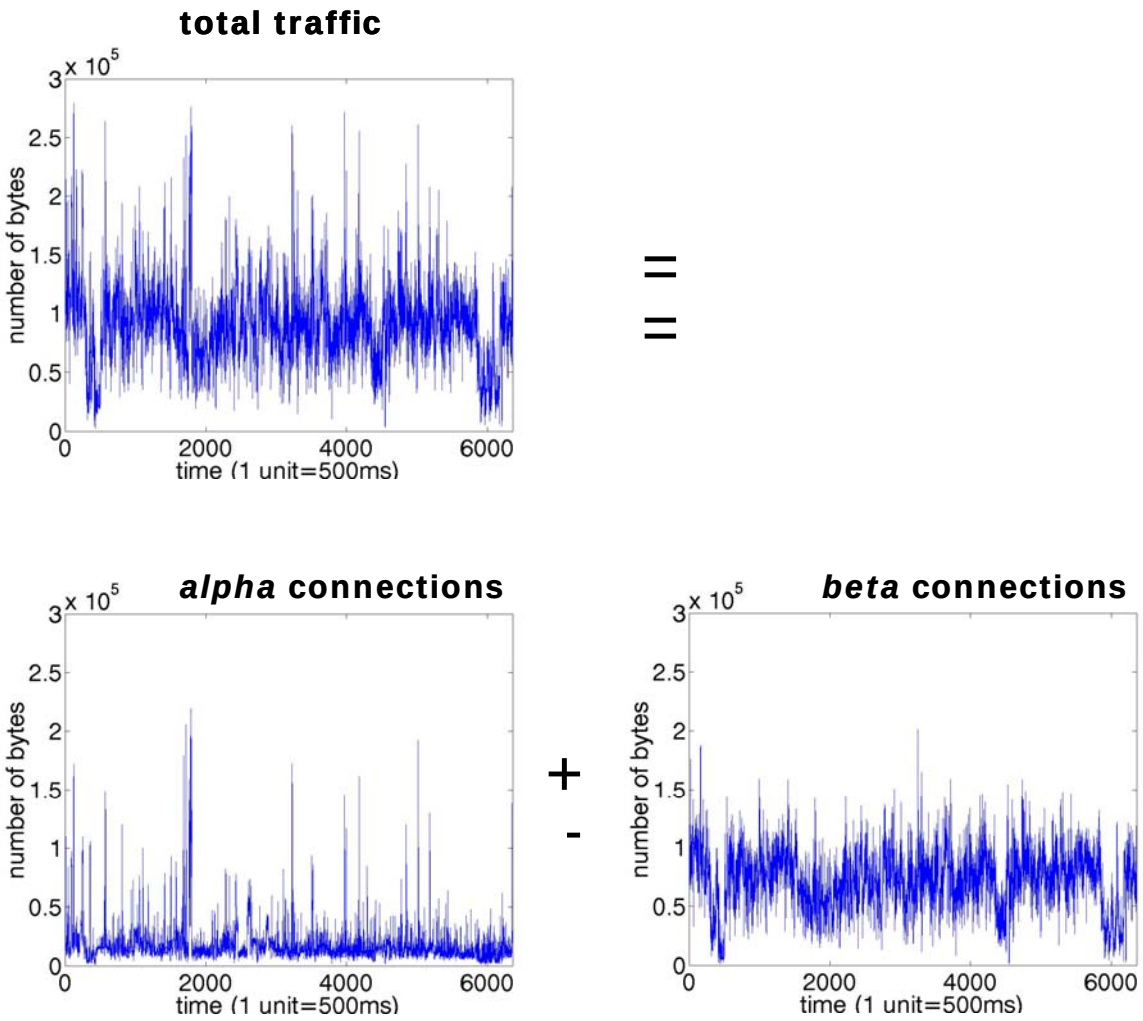


Figure 2: Connection-level traffic decomposition into an aggressive alpha component and residual beta component

2.1 Multiscale Queuing Analysis

We have developed a novel multiscale framework to estimate the tail probability of a queue fed by an arbitrary traffic process. This framework provides strong contributions to both the theory and application of queuing analysis, and creates a strong foundation for understanding the interrelationship between strong correlation or Long-Range-Dependence (LRD), burstiness, traffic marginals and queuing theory.

Using traffic statistics at a small number of time scales, our analysis extends the theoretical concept of the critical time scale and provides three practical approximations for the tail queue probability: the *max*, *product*, and *sum* approximations. These approximations are non-asymptotic; that is they

apply to any finite queue threshold and are thus valuable tools for network design. Simulations with LRD traffic models and real Internet traces demonstrate the accuracy of the queuing approximations (see Figure 3). The max, product, and sum approximation closely track the tail probability of a queue fed by traffic from a multifractal model.

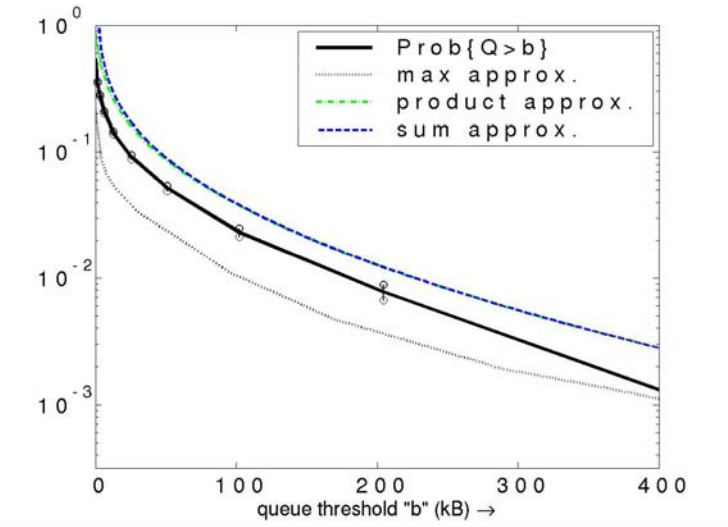


Figure 3: Accuracy of queuing approximations

While our approach applies to any traffic process, it is particularly apt for LRD traffic. For LRD fractional Brownian motion, we prove that a sparse exponential spacing of time scales, such as powers of two, yields optimal performance. This finding has significant implications for practice. Because traffic variability at a wide range of time-scales can impact queuing it is imperative to keep track of statistics over a large range of time-scales. One solution is to collect statistics at a dense set of uniformly spaced time scales covering a wide range. Although this solution would provide accurate queuing approximations, the associated overhead in terms of computation and overhead can be excessive. Our result proves that by using a sparse exponential set of time scales we optimally balance the tradeoff between the accuracy of queuing approximations and overhead.

Simulations reveal that the marginals of traffic at multiple time scales have a strong influence on queuing that is not captured well by its global second-order correlation in non-Gaussian scenarios. Earlier work had focused largely on the impact of the second-order correlation structure of traffic on queuing and established the importance of the short-term as well as the long-term correlation structure on queuing. Our experiments prove that in non-Gaussian scenarios it is important to consider the tails of marginals at different time-scales in addition to the second-order correlation structure.

2.2 Optimal Sampling of Multiscale Stochastic Processes

We have developed efficient algorithms and theory for optimally sampling multiscale processes. Our work applies to problems where we must estimate the global average of some physical process (over space or time) from measured point samples of the process. Examples of such processes are network quantities such as data traffic loads and even environmental quantities such as temperature, pressure and pollution levels, agricultural quantities such as the insect population density.

The multiscale stochastic process we studied consisted of a set of univariate random variables which are organized like the nodes of a tree (see Figure 4). Each node represents a random variable equal to the sum of nodes at the next lower level connected to it by edges. In typical scenarios, different levels on the tree represent different scales in time or space. Nodes at the bottom are called *leaves* and the topmost node is the root. We associate each node with the average of a physical process over some region with nodes at higher scales corresponding to larger regions.

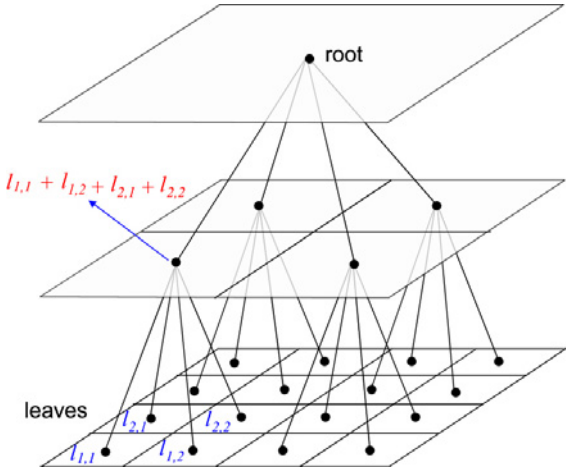


Figure 4: Multiscal stochastic process

Multiscale stochastic processes can vary in their topology, that is the number of nodes at each scale and the connectivity between nodes, as well as in the joint probability distribution of the nodes. We analyze two types of multiscale stochastic processes: *independent innovations trees* and *covariance trees*. In independent innovations trees, nodes at one scale are related to nodes at the next lower scale through independent additive innovations. In covariance trees pairs of leaf nodes with the same distance between them have the same covariance. If the correlation between leaf nodes of a covariance tree increases with distance we say that the tree has negative correlation progression. If the opposite is true then we say the tree has positive correlation progression.

The question we address is: Among all possible sets of leaves of size *n*, which set provides the best Linear Minimum Mean Squared Error (LMMSE) for the root? Equivalently, where must we strategically sample the process in order to estimate its global average optimally?

Accomplishments: We construct an efficient O(*n*) algorithm to determine an optimal leaf set of size *n* for independent innovations trees called *water-filling*. Exploiting key statistical properties of independent innovations trees we establish that water-filling is optimal. Note that the general problem of determining the *n* random variables that provide the best linear estimate of another random variable is NP-hard. In contrast, our setting is one example of this general problem which is solved in polynomial time by the water-filling algorithm.

Using the water-filling algorithm we prove that uniformly spaced leaf nodes are optimal for certain independent innovations trees called scale-invariant trees as well as for covariance trees with positive correlation progression. The optimal solution can, however, change with the correlation structure of the tree. In fact for covariance trees with negative correlation progression we prove that uniformly spaced leaf nodes give the worst possible LMMSE! For covariance trees with positive correlation progression we prove that the set of clustered leaf nodes provides the worst possible LMMSE.

Our findings have implications in a range of different applications. One example is sensor networks. Such networks consist of electronic sensors that are distributed over a large region to measure some physical process. In the context of our work the area is subdivided hierarchically into sub-areas with the nodes of a tree representing the average of the physical process in different sub-areas. Every sensor gives a local sample of the process which we model as a leaf node of the tree. The purpose of the network is to measure the global average of the process which we model as the root of the tree.

For economic reasons it is desirable to design sensor networks that use a small number of sensors, yet cover a large area with sufficiently accurate estimates. An ideal design of such networks simply places sensors at locations corresponding to an optimal leaf set given a budget of number of sensors to be used.

3.0 Spatio-Temporal Available Bandwidth Estimation

Using multifractal traffic models, we have developed new theory and methods for understanding and inferring network-relevant properties of end-to-end paths and connections, especially their dynamics. This is the first step in making applications that are *network-aware*. In contrast to current approaches to path modeling, we explicitly model the causes of queuing latency and loss by introducing a model for competing packets. Our approach to infer the competing cross-traffic load utilizes an innovative exponentially spaced probing sequence of *packet chirps* that balance the trade-off between overwhelming the network with probes and obtaining statistics rich enough for accurate estimates. Passive monitoring theory is providing a novel means to analyze and model the *interaction of large network flows* that multiplex through queues along a common path.

Accomplishments: Our first edge-based probing tool, ***pathChirp***, estimates the available bandwidth on a network path. The available bandwidth of a network link is the average unused bandwidth of the link. For instance a 100Mbps link that is 70% utilized has 30Mbps of available bandwidth. The available bandwidth of a network path is the smallest available bandwidth of all links of that path.

For example the path depicted in Figure 5 has 20Mbps available bandwidth. Knowledge of the available bandwidth plays a crucial role in optimizing different network-aware applications such as grid computing.

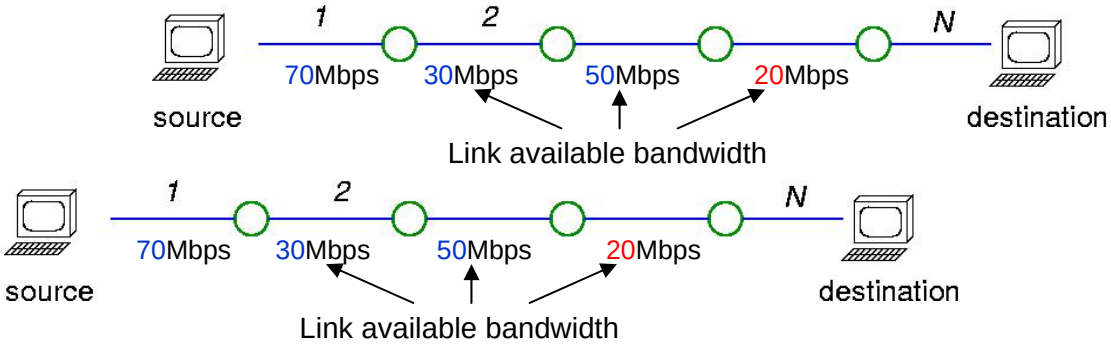
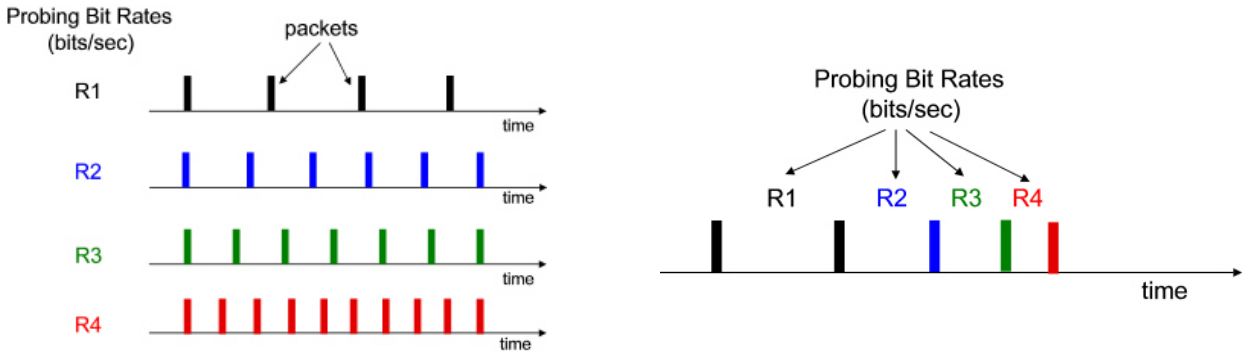


Figure 5: Available bandwidth of a network path

PathChirp employs an effective way to estimate available bandwidth termed the principle of *self induced congestion.* This principle relies on the fact that routers buffer incoming packets in queues before transmitting them on output links. If the incoming packet bit rate exceeds the transmission rate of the outgoing link then packets fill up the corresponding queue and face queuing delays. According to the principle of self-induced congestion if we inject probe packets into a path at a bit rate faster than the available bandwidth then the path's queues will congest leading to increasing delays. The path's queues will not congest, however, if the probing bit rate is less than the available bandwidth. One can thus infer the available bandwidth by injecting trains of probe packets into the network, each at a different bit rate as illustrated in Figure 6(a), and find the minimum bit rate at which we start to see increasing delays of probe packets. Such a scheme would however introduce a large number of probe packets into the network and consume significant bandwidth resources. The simplistic probing scheme in (a) uses one packet train for each probing bit-rate.



(a) Simplistic probing scheme

(b) Packet chirp

Figure 6: Packet Chirp - efficient exponential flight pattern of packets

In contrast, pathChirp is highly efficient because it uses special packet trains called chirps (see Figure 6(b)). The packet chirp in (b) combines several probing bit-rates into a single train. In a chirp the inter-spacing between packets decreases exponentially. Because the probing bit rate is inversely proportional to the time interval between probe packets, the probing bit rate within a chirp increases exponentially fast. This property allows a chirp to probe the path at a wide range of probing rates using few packets, allowing accurate estimates of available bandwidth based on the self-induced congestion principle while introducing only a light load on the network.

PathChirp's estimates the available bandwidth from a careful analysis of the profiles of the end-to-end queuing delay of probe packets, called *signatures*. Figure 7 depicts a typical signature. Notice several regions of increasing and decreasing delay within a signature. Making apt use of the principle of self-induced congestion, pathChirp assigns instantaneous estimates of available bandwidth to the different regions and averages these to get its final estimate of available bandwidth.

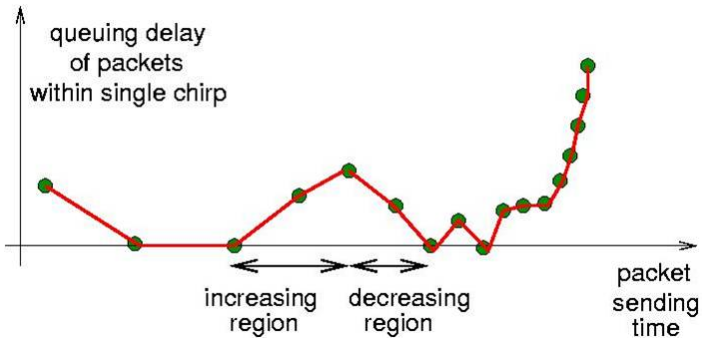


Figure 7: A chirp queuing delay signature

Our test results with pathChirp in simulation as well as on the real Internet have been encouraging; in particular we have been able to accurately estimate the available bandwidth on a high-speed link using about 10x fewer packets than the state-of-the-art tool *pathload* and many times fewer than *iperf*. Figure 8 illustrates pathChirp's ability to track changes in available bandwidth on the CAIDA Gigabit testbed; observe how the pathChirp estimate rises and falls in proportion to the introduced cross-traffic. Our paper on pathChirp received the *Best Student Paper Award* at the *Passive and Active Measurement Workshop* in April 2003. At the time of this report, pathChirp is under test at a number of labs and is being integrated into the ns-2 network simulator. Open-source software is available at www.spin.rice.edu/Software/pathChirp. TNO Information and Communication Technology Division, The Netherlands, is applying for standardization of a variant of pathChirp through the ITU Telecommunications Standardization Sector (COM 12 – C draft2 – E, July, 2005).

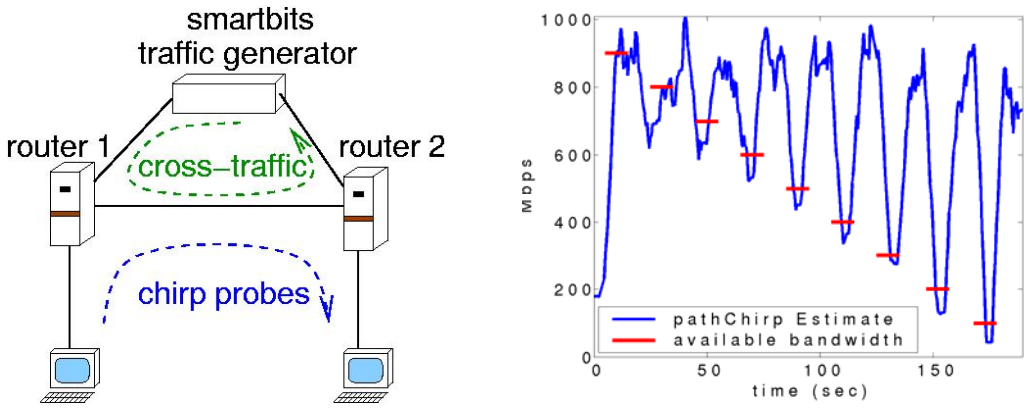


Figure 8: pathChirp provides accurate estimate of bandwidth on CAIDA Gigabit testbed

Our second edge-based probing tool, ***STAB*** (for Spatio-Temporal Available Bandwidth estimator), locates the tight link (link with lowest available bandwidth) both in space and tracks changes in its location over time. Open-source software is available at www.spin.rice.edu/Software/STAB.

Building on our successful bandwidth estimation tool pathChirp, STAB provides an essential step towards a more precise location of hot spots in the network, an improved detection of drastic changes in network usage and related anomalies. Its superior capabilities, being able to locate tight links both in space and time, allows us to probe an entire network over a reduced set of paths.

This advanced capability of our algorithm for estimating available bandwidth along a path in a packet network provides a critical step towards our far end goal, to build a chirp web capable of localizing hot spots in a network through a conservative amount of probing traffic.

Tight link localization is an algorithm that identifies which link along a network path provides the smallest bandwidth and limits thus the entire path. It leverages two technologies: packet-tailgating for probing of internal nodes of a network that is combined with a low-weight self-induced congestion for bandwidth estimation.

While tailgating has been previously used in end-to-end tools towards capacity estimation and topology identification our team is the first to employ the technique for bandwidth estimation. In a nutshell, tailgating means to send a large "probe packet" followed immediately back-to-back by a small "tailgating" packet. While both packets are sent to the same destination, the probe packet is limited purposely to travel only a certain number, say *m*, hops along the path towards the destination. This is done by setting the Time-To-Live (TTL) field in the probe packet equal to *m*. No such restriction is imposed on the tailgating packets that travel all the way to the receiver.

Such pairs of probe and tailgating packets are now sent according to a bandwidth estimation scheme of choice. While in principle the technique would work in combination with any bandwidth estimator, it is imperative to instrument one that is low weight, meaning that it imposes a minimal amount of probing traffic on the network. This point cannot be stressed enough, since the

localization of tight links requires probing the path repeatedly, going through an entire estimation sequence for each hop along the path. Our choice of bandwidth estimator is naturally the low-weight pathChirp algorithm, which was developed in this team and which has won much interest as well as a best paper award in the network measurement and modeling community.

The tight link localization technique exploits now the following fact: a given number of pairs of probes with their tailgating packet are sent along the path towards the receiver in exponentially increasing rate, thus forming what we call a tailgating chirp (see Figure 9). The large probe packet will be queued as congestion builds up caused by the probe packets ahead of it, according to the available bandwidth at a link (principle of self-induced congestion). The queuing delay adds to the propagation delay and can be used to infer the available bandwidth. However, the large packets will be dropped at router number *m* along the path due to the TTL. Yet, the tailgating packets following the probes back-to-back experience the same delay and continue their journey towards the destination. Being very small in size, they induce no further congestion and arrive at the destination with delays that are affected only by some additional noise down from the router that dropped the large probe packets. Since the probes only induce congestion at the first *m* links, their delays help us compute the available bandwidth of the segment of the path from the source to link *m*, called the sub-path available bandwidth up to link *m*.

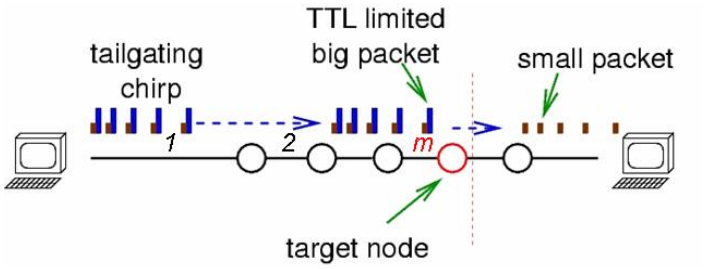


Figure 9: Tailgating chirp

STAB locates the tight link using a plot of sub-path available bandwidth against *m*. Such a plot typically decreases only when *m* equals the location of a *thin* link. A link qualifies as a thin link if it has less available bandwidth than all preceding links. The thin link farthest away from the source is the tight link of the entire path. Thus the last link at which we see a noticeable decrease in the sub-path available bandwidth is the tight link.

We successfully demonstrated the technique in ns-2 simulation as well as on the Internet both over space and over time. The simulation topology depicted in Figure 10**,** called the *double web-farm* topology, is designed to ensure that the tight link location changes over time. Each web-farm consists of 420 clients downloading data from 40 servers.

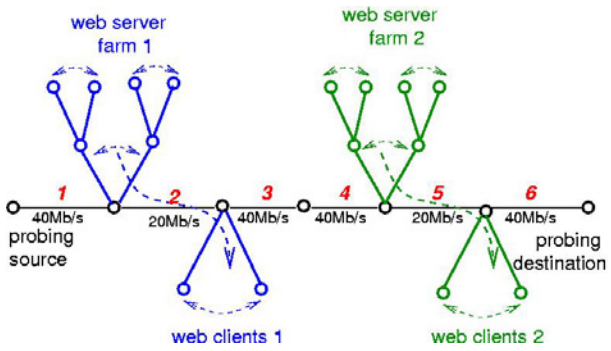
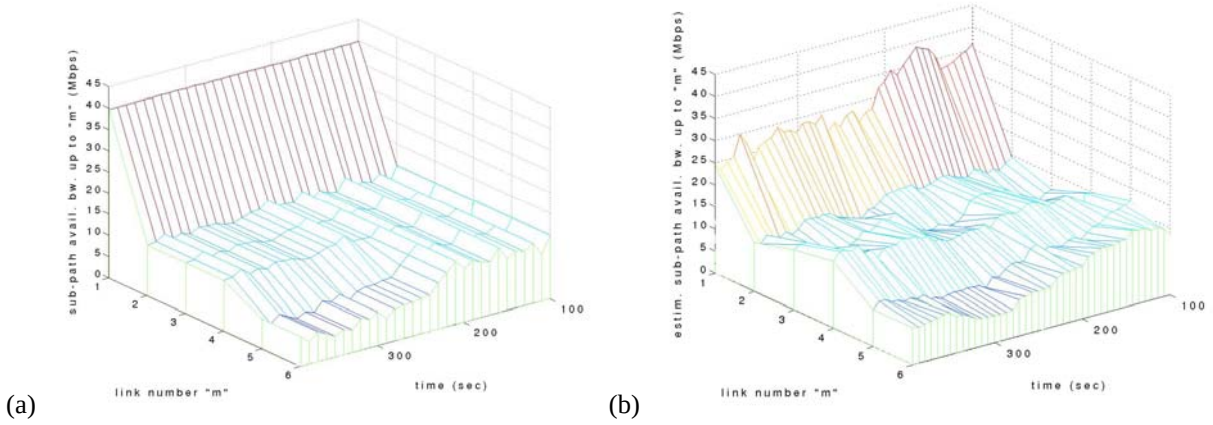


Figure 10: Double web-farm topology

Figure 11(a) depicts the actual sub-path available bandwidth up to link m for different intermediate links m and their variation over time. In the first half of the simulation, that is up to time t=200s, only the first web-farm generates traffic. As a result link 2 is the tight link of the path and consequently the available bandwidth plot flattens out after link 2 at any time prior to t=200s. In the second half of the simulation both web-farms generate traffic. Because the second web-farm generates more traffic than the first, link 5 now becomes the tight link.



(a) (b)

Figure 11: (a) Actual and (b) STAB estimates of sub-path available bandwidth during simulation

Observe from Figure 11(a) that the available bandwidth plot dips at link 5 after time t=200s. From Figure 11(b) we see that STAB estimates the sub-path available bandwidth well. Observe that prior to time t=200s the estimates flatten after link 2 while after t=200 the estimates dip at link 5 due to the traffic from the second web-farm.

Using the estimates of sub-path available bandwidth we compute the probability that different links on the path are thin links. Recall that a link m qualifies as a thin link if it has less available bandwidth than all preceding links, and that the thin link farthest away from the source is the tight link of the entire path. Figure 12(a) plots the probability of different links being thin links at time instant t=180s, which belongs to the first half of the simulation. We see that link 2 is almost certainly a thin link while the other links have low probabilities of being thin links. This strongly

suggests that link 2 is the last thin link of the path and hence the tight link. Figure 12(b) plots the probability of different links being thin links at time instant t=360s, which belongs to the second half of the simulation. Now both links 2 and 5 are almost certainly thin links unlike other links. Clearly at this time instant link 5 is most likely the path's tight link.

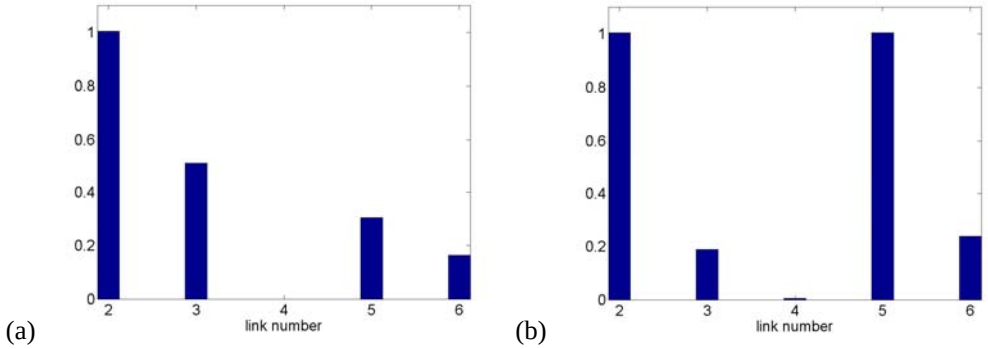


Figure 12: Probability of different links being thin links at time instants (a)t=180s and (b)t=360s

In one Internet experiment STAB simultaneously located the tight link on two paths, one from the University of Illinois to Rice and the other from the University of Wisconsin to Rice (Figure 13). STAB's results for both paths were consistent; it located the same link as the tight link of both paths, that is link 14 of the UIUC→Rice path and link 13 of the UWisc→Rice path. MRTG data obtained from the different routers on the paths confirmed STAB's estimates.

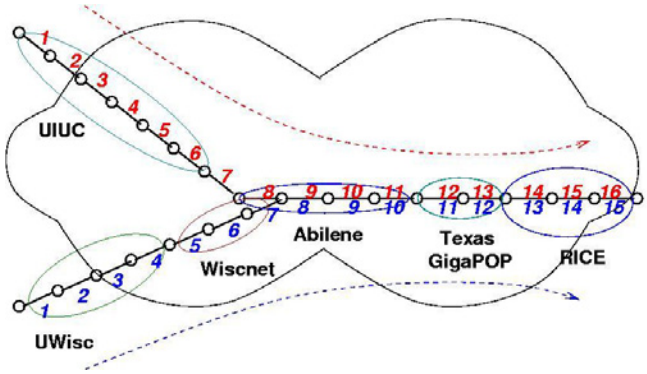


Figure 13: Locating the tight link on two Internet paths from U of Illinois and U of Wisconsin to Rice

4.0 Connection-level Analysis and Modeling of Network Traffic

Our key idea is a decomposition of traffic into an aggressive "alpha" component and a passive "beta" component. In a nutshell, we have used connection-level information to identify that a very small percentage of connections cause nearly all of the bursts in traffic traces. The rest of the connections aggregate into simple a fractional Gaussian noise process. We feel this discovery has

some far-reaching implications in a number of areas, including modeling, queuing, scalability, effect of network topology on traffic, synthesis, and inference (probing schemes, intrusion or hot spot detection).

More detailed analysis of many traces from various places of the Internet has revealed that the aggressiveness of alpha connections can be traced back mainly to one crucial parameter which determines the bandwidth of TCP connections, the Round Trip Tip (RTT) or distance of sender and receiver. This major discovery not only allows for a simple first-order translation of network topology to the burstiness observed in traffic dynamics. It also suggests an appealing and simple model for alpha traffic: while beta traffic is obtained as the celebrated limit of ON-OFF sources in the limit of "very many sources", i.e. fractional Gaussian noise, alpha traffic can be understood as a limit of the same ON-OFF model, however in the limit of "fast time", i.e., stable Levy motion. This modeling framework allows for simple queuing analysis in the framework of self-similar traffic, however, with the added difficulty of having to consider stable, non-Gaussian processes.

Our advances in the theoretical aspects of the alpha-beta decomposition have led to a better understanding of impact of traffic bursts, i.e., bursts of alpha traffic on the queuing and the performance of the network. Our results indicate that a rate limiter (in the form of flow control at either end of a path or in the form of an appropriate AQM scheme in dedicated routers) can be effective. Thereby it is sufficient to constrain the arrival rate of a high-bandwidth flow to the available bandwidth.

4.1 The alpha-beta ON-OFF model:

Based on our understanding of alpha and beta traffic analysis, we developed a physically motivated traffic model using observations at the connection level. By connection-level, we mean all the packets that have the same source and destination IP addresses, port numbers and protocol. Connection level measurements contain much more information than the simple aggregates; we study the statistics of session parameters such as duration, file size and rate, and their impact on aggregate traffic. We call our traffic model the *alpha-beta ON-OFF* model. This traffic model is the result of a fruitful marriage between alpha-beta traffic analysis and the classical ON-OFF traffic model.

Before we describe the alpha-beta ON-OFF model, we briefly review the celebrated ON-OFF model for network traffic. The ON-OFF model is based on connection-level information in the traffic. The model successfully captures the second-order correlations of traffic, in particular their Long-Range Dependence (LRD). The traffic is obtained as a superposition of a large number of ON-OFF sources, with heavy-tailed ON and/or OFF periods. A conclusion commonly drawn based on the ON-OFF model is that at large timescales (beyond several RTTs), transport protocol mechanisms like TCP do not have a significant impact on the network traffic. Unfortunately, however, the classical ON-OFF model fails to capture several key aspects of the network traffic, including its spikiness at small timescales (by small scale spikiness, we refer to the deviation of the distribution from Gaussianity).

Using the alpha-beta traffic analysis technique we develop a physically motivated traffic model that accurately matches the both large and small time scale behavior of network traffic but retains the simplicity of the ON-OFF model. The resulting model is the alpha-beta ON-OFF model. The analysis involves decomposing the network traffic into two components according to the connection bandwidth: the large bandwidth alpha component contributes all the small scale spikes and is highly non-Gaussian, while the small bandwidth beta component contributes most traffic and is Gaussian. This decomposition explains the LRD of network traffic at very large time scales as well as the highly non-Gaussian marginals and multi-fractal scaling on small time scales. The alpha-beta analysis parsimoniously accounts for user behavior, network topology, and the heterogeneous distribution of network bandwidths. Our analysis suggests that the bursts in traffic are not due to a ``conspiracy'' of connections. This is a significant finding, because the classical ON-OFF model produces bursts only by having a large number of ON-sources active at the same time. The alpha-beta ON-OFF model can be viewed as an extension of the ON-OFF model to accommodate small-scale spikes.

We now motivate the need for a new traffic model. It is well known that there exists a strong correlation between the rate and the file size for the overall traffic. This indicates that the users choose the file sizes to download based on the speed of the connection. However, the correlation between file size and the rate is greatly reduced when we look only at the alpha sessions. Hence we conclude that for the alpha sessions, the download time, which is small, is not a significant factor that limits a user's behavior. Thus, we may assume that the rate and the file size are independent for the purpose of simplified modeling and synthesis. Our observations also reveal that for beta connections the file size depends on the bottleneck bandwidth (rate) and the duration of the download that the user is willing to tolerate. We need a model that captures the independence of rate and duration for the beta sessions; and independence of file size and rate for the alpha connections. Therefore for a more realistic traffic model, the key parameters of the ON-OFF model need to be modeled differently for the alpha and beta components. The alpha-beta ON-OFF model uses this intuition to set the file size, rate and duration of the alpha and beta connections.

Alpha-beta ON-OFF model consists of the sum of two ON-OFF models: one for the alpha component and one for the beta component. Interestingly, the three controlling parameters duration = file size / rate are used in different combinations for each of the two models. We model the alpha sessions as having ON-periods with independent size and rate. Random sizes and rates are chosen independently from their respective empirical marginal distributions in the alpha connections of the real trace. The duration of the ON-period is determined by duration = file size / rate. We model the beta sessions as having ON-periods with independent duration and rate. Random durations and rates are chosen independently from their respective empirical marginal distributions in the beta connections of the real trace. The file size of the on-period is determined by file size = rate $\times$ duration.

We choose the number of alpha and beta connections to match that of the real trace. The starting times of the alpha and beta ON periods are chosen uniformly over the duration of the trace with wrap-around if the on-period exceeds the total time of the trace. The synthetic alpha trace is the

superposition of the alpha ON-periods and the synthetic beta trace is the superposition of the beta ON-periods. The total trace is the sum of these two superpositions.

We showed that the synthetic traces obtained from the alpha-beta ON-OFF model match well with the real traces. The metrics we used were closeness of energy plots (also called variance-time plots) and closeness of queue length distributions when we feed the traffic to an infinitely long queue. Lastly, the synthetic traffic resembles the real traffic visually; the synthetic traffic contains the small scale burstiness that is associated with real internet traffic.

4.2 Implications of alpha-beta ON-OFF model:

The implications of the alpha-beta ON-OFF model to congestion control and what-if-scenarios are relevant and even surprising.

The network-driven user: Let us start by drawing a telling relationship between network and the user layer. Our evidence suggests separating user behavior into two regimes. We base the discussion on the three basic parameters of file size, sending rate, and duration. The beta regime, on the one hand, is characterized in the network layer by rate-limitation, which results in the user layer as *patience* being the apparent parameter ruling behavior. For the network-user-driven ON-OFF model this translates into approximating the joint distribution of file, size and rate as being characterized by independent rate and duration. The alpha regime, on the other hand, corresponds in the network layer to high available bandwidth, which translates in the user layer into a free choice of file transfers according to *interest*. For the model this means that file size and rate are independent.

What-if scenarios: It has been shown that most of the alpha sessions come from short RTT connections. By looking at the histogram of the connection RTT, a network operator can immediately comment on the relative strengths of the alpha and beta components. The relationship between the connection RTTs and the strength of the alpha component has some important implications when we consider Content Distribution Networks (CDNs). CDNs change the RTT distributions of sessions, and we expect more flows to have shorter RTTs. It is well known that alpha connections tend to build up large queues. This calls for a mechanism at routers (AQM) to detect and control alpha connections so that other connections are not unduly affected. These issues are particularly relevant as we expect more alpha flows in the future, due to higher available bandwidth to users, thanks to digital subscriber line (DSL), cable modem, and fiber-to-the-home technologies.

Another interesting aspect is the potential of *alpha-DOS* attacks where hosts close to a server leverage small RTTs to drain the network resources (bandwidth), depriving other clients of service. One way to detect such an attack would be to identify transfers in the alpha regime that persist beyond the length of time a typical human user would tolerate.

Effectiveness of congestion control: Among the two components, only the beta component appears to be effectively controlled by TCP. Alpha connections profit greatly from small RTTs to obtain large bandwidth. Operating as it does with small TCP windows, loss-based congestion control does not restrict alpha connection since they recover much more quickly to drops than the average beta user. This calls for alternative methods of congestion control than those provided by TCP.

Predicting traffic patterns in a high-speed TCP environment: An interesting question is to predict a priori the expected volumes of alpha and beta components of traffic in a high speed TCP environment. Many high-speed TCP schemes have been proposed recently to adapt to high bandwidth-delay product environments such as High-Speed TCP (HSTCP), Scalable TCP (STCP), eXplicit Control Protocol (XCP), FAST-TCP and Binary-Increase TCP (BI-TCP). For TCP in a high bandwidth-delay environment, we predict that the strengths of the alpha and beta components will depend on TCP's bias toward short RTT connections. In the current TCP-Reno, alpha connections cannot arise from large RTT connections because they lose out to short RTT flows due to TCP bias. Based on the analysis of TCP's RTT bias, we expect that the alpha-beta differentiation between flows will be more pronounced for STCP and HSTCP (which have large RTT bias) than for BI-TCP (which has RTT bias comparable to TCP-Reno).

We have completed a C++ version of an "alpha-beta trace analyzer" which allows for fast off-line analysis of traffic traces on the connection level, providing thus a simple tool for the networking researchers and providers to obtain a decomposition of network traffic into alpha- and beta-components, which provides hands-on information of the sources of burstiness in the traffic flow. This is useful for congestion control and more efficient use of resources. Several peers from the NMS program have shown interest in verifying the alpha-beta paradigm with their data. We have also developed an interface which will allow make the tool accessible and easy to handle. Having our peers test the tool as well as the modeling assumption in various scenarios has been instrumental towards identifying the physical causes of the disruptive bursty fluctuations in traffic flow.

We have completed a C+ version of an "alpha-beta trace analyzer" which allows for fast off-line analysis of traffic traces on the connection level, providing thus a simple tool for the networking researchers and providers to obtain a decomposition of network traffic into alpha- and beta-components, which provides hands-on information of the sources of burstiness in the traffic flow. This is useful for congestion control and more efficient use of resources. Several peers from the NMS program have shown interest in verifying the alpha-beta paradigm with their data. We have also developed an interface which will allow make the tool accessible and easy to handle. Having our peers test the tool as well as the modeling assumption in various scenarios has been instrumental towards identifying the physical causes of the disruptive bursty fluctuations in traffic flow.

4.3 Queuing of alpha-beta traffic

First, let us revisit the celebrated ON-OFF model for network traffic, a process introduced very early by B. Mandelbrot. In this framework, a traffic source is modeled as being sending traffic at a

constant rate (the ON state), or as being silent (the OFF state); here, the durations of the ON and OFF states are both assumed heavy tailed; variations are studied in the literature. It is well known that in the limit of an infinite number of sources the aggregate (sum) of such sources becomes a Gaussian process with LRD, converging to fractional Brownian motion at large time scales. This explains the success of the latter in network traffic modeling.

On the other hand, the aggregate of a fixed number of ON-OFF sources converges to Levy stable motion in the limit of infinitely large time scales, i.e., as clocks tick infinitely fast. With RTT being the ``clock" of the transport protocol, Levy motion seems thus the appropriate approximation of the alpha component of the traffic. The large crowd of beta connection with roughly equal sending rates and ``clocks" aggregate to the beta-component which is well represented by fractional Brownian motion.

Thus, we are lead to the *self-similar burst model* where we pose network traffic as the superposition of two self-similar processes, a fractional Brownian motion and a Levy stable motion. Assuming their scaling parameter **H** is equal, this sum is again self-similar and allows to apply scaling techniques from queuing theory as well as the concept of critical time scales. Thus it is possible to approximate the tail probability of **Q**, the stationary queue length of a link with constant capacity. We find that probability queue size (**Q**) become larger than a threshold (**b**) behaves roughly like a Weibull law for moderate threshold **b** and like a Pareto law for large threshold.

Second, in the *ON-OFF burst model* we pose the contributions of alpha connections not as a Levy stable motion but simply as one ON-OFF source with particularly large rate. Again, we set the beta component to be fractional Brownian motion. A natural approach to queuing is then to consider the alpha source as reducing the link capacity. The queue length of the ON-OFF burst model is, thus, equal to the queue length of fractional Brownian motion entering a link with variable service rate. The somewhat involved analysis can be summarized as follows.

Denote by $\mathbf{R_\alpha}$ the sending rate of the alpha source when in the ON state, and by **C** the link capacity. If the average arrival rate $\mathbf{R_\beta}$ of the beta component is smaller than $\mathbf{C - R_\alpha}$ then the alpha source has no crucial influence on the queuing behavior of the ON-OFF burst model. Intuitively, we may argue that the link is in this case still stable even in the presence of the alpha source since the overall arrival rate is then smaller than the capacity.

If the given condition is not met and $\mathbf{R_\beta > C - R_\alpha}$ then the queue becomes unstable during the ON periods of the alpha source. In particular, if the ON periods of the alpha source are Pareto distributed, then the queue tail probability will be Pareto as well.

We have shown using simulations and models (See Figure 14) that when mixtures of alpha and beta connections exist at a router, the beta connections receive only a fraction of their fair share (sometimes as low as 20%) of bandwidth. Our results suggest a situation best described as "the rich (broadband/ethernet) get richer, the poor (modem) get poorer."

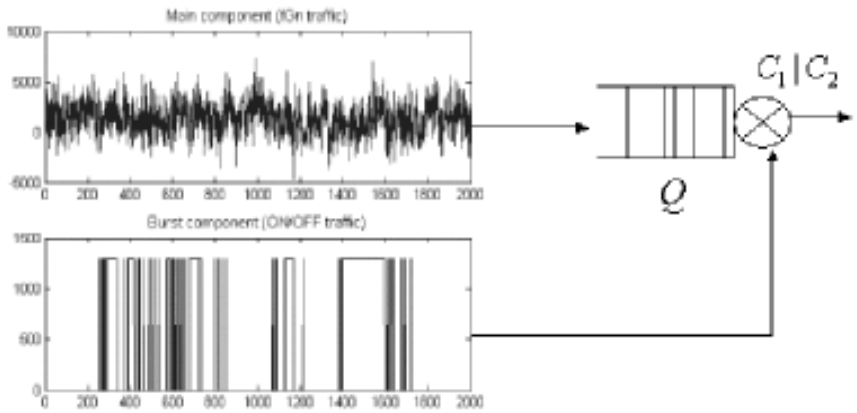


Figure 14: Modeling alpha-connections as reduction of queue capacity assesses effect of heavy alpha on performance

5.0 Safe High Speed Congestion Control

Another area we have explored is the problem of congestion control on high speed long distance networks. TCP, the default protocol for congestion control on the Internet, has been in existence since the 1980s, and its basic algorithm is largely unchanged from then. The Internet, on the other hand, has grown by several orders of magnitude. In 1989, an Internet core link might have run at 56Kbps, and stretched possibly across California. In the modern Internet, there are links that run as fast as 10Gbps, and can cross the Pacific Ocean, connecting continents. Furthermore, there are many new exciting applications, such as grid computing, high energy physics experiments, and large biological simulations, that are demanding extraordinarily high bandwidth over very large distances. It is a testament to the foresight of the original TCP designers that the algorithm has been able to scale to as far as it has, but the limits of the basic TCP algorithm are starting to constrain the possibilities for these new applications on the Internet. The modern links where TCP is starting to show its age are those that have both a high bandwidth, and a long delay, so called *high bandwidth-delay-product* links.

One of the challenges in designing a new TCP algorithm, however, has been the requirement that any new congestion control algorithm must co-exist with the old algorithm. Since the main way in which TCP needs to be adjusted is to make it more aggressive in these high *bandwidth-delay-product* links, maintaining a sense of fairness with the old algorithm has been quite a challenge. Many schemes, such as HSTCP, have been proposed that try to adjust the scalability of TCP by simply making it more aggressive on large networks. . For standard TCP, the throughput decreases by half on a congestion event, and it can take an extended period for the protocol to recover. HSTCP quickly recovers from congestion events, but also causes them with the same rapidity. While these schemes do improve throughput, they have been prone to fairness problems when competing with the old TCP.

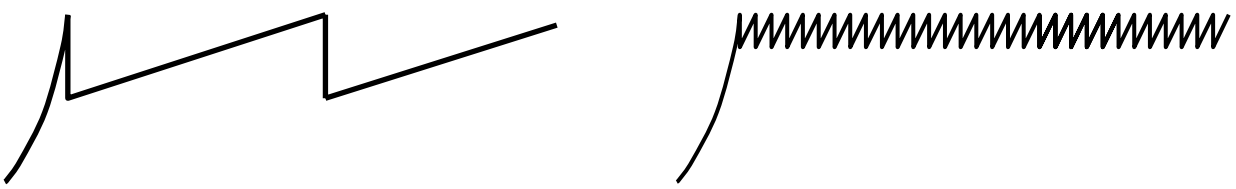


Figure 15: Congestion window behavior for standard TCP on left, and for HSTCP on the right

To address these concerns, we have developed a new delay sensitive congestion avoidance rule for TCP that promises excellent utilization, efficiency, and acquisition of available bandwidth, with significantly improved safety features compared to other high speed TCP algorithm. We find that since standard TCP uses only packet loss information to control its sending rate, we find that by incorporating packet delay information, we are able to design and intelligent algorithm for specifying sending rules. This algorithm, TCP AFRICA is a hybrid algorithm that uses an aggressive, scalable window increase rule to allow quick utilization of available bandwidth, but also uses packet round trip time measurements to predict eminent congestion events. Once the link is sensed to be highly utilized, the protocol reacts by resorting to the conservative congestion avoidance techniques of the old TCP. This hybrid approach to congestion control allows TCP AFRICA to scale to networks with high delay-bandwidth-products, while maintaining a high degree of fairness and safety towards the existing TCP protocol.

By having two congestion avoidance modes, TCP AFRICA is able to quickly utilize available bandwidth when the link is underutilized, and it is also able to delay the onset of a self-induced congestion event, allowing TCP AFRICA to enjoy the high throughput for a longer period than would be possible with HSTCP. Also, since the number of packet losses incurred in a congestion event is proportional to the increase step the previous RTT, the congestion events induced by TCP AFRICA cause fewer packet losses. This effect, in addition to easing the demand on router buffers, also helps prevent the synchronized loss events that can prevent standard TCP from achieving acceptable utilization against super-aggressive loss based protocols, such as HSTCP, STCP, and the scaleable mode of TCP AFRICA.

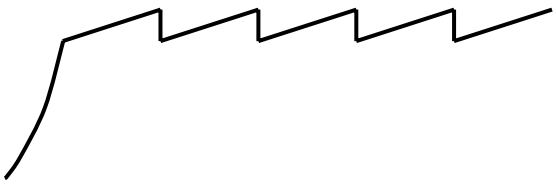


Figure 16: Congestion window behavior for TCP AFRICA

Another major design consideration for TCP AFRICA is robustness to the misbehavior of delay information. While the direct use of packet delay information in congestion control has been discussed since the mid 1990s, there has been no extensive deployment of such a protocol. There

are many possible concerns with using delay information, just as there are many possible benefits. Unusual network phenomenon such as reverse path cross traffic, ACK compression, the use of the TCP delayed-acknowledgement option, route changes, and other phenomenon can seriously compromise the accuracy of the estimated packet round trip time. The hybrid nature of TCP AFRICA helps mitigate the risks of using delay information for congestion control. Since TCP AFRICA only uses delay information to switch between an increase rule based on standard TCP and an increase rule based on High-Speed TCP, it can never be less scaleable and efficient than standard TCP, or more aggressive than HSTCP, which has been proposed as a standard in its own right. Thus, even with noisy or possibly corrupted delay information, TCP AFRICA never performs too badly, or acts too dangerously.

To verify and study the properties of TCP AFRICA, we have implemented extensive ns-2 simulations examining the interactions between TCP AFRICA and the standard SACK TCP found on the Internet. Our experiments have examined properties such as efficiency of utilization, fairness with regular TCP at different network speeds and loss rates, and throughput bias between flows of different round trip time. Our experiments indicate that while TCP AFRICA does not divide bandwidth perfectly with standard TCP, it gives standard TCP a bandwidth share much closer to the amount of bandwidth standard TCP would achieve if not in the presence of a high-speed protocol.

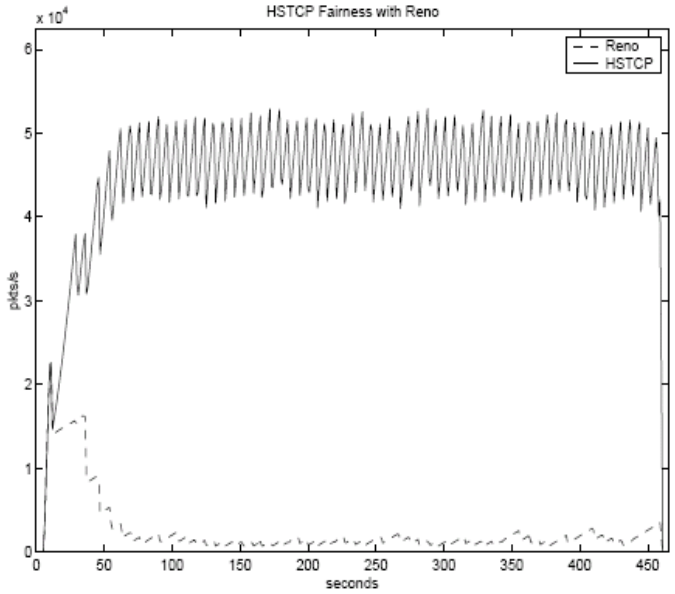


Figure 17: ns-2 experiment between HSTCP and TCP-Reno on a 622 Mbps link

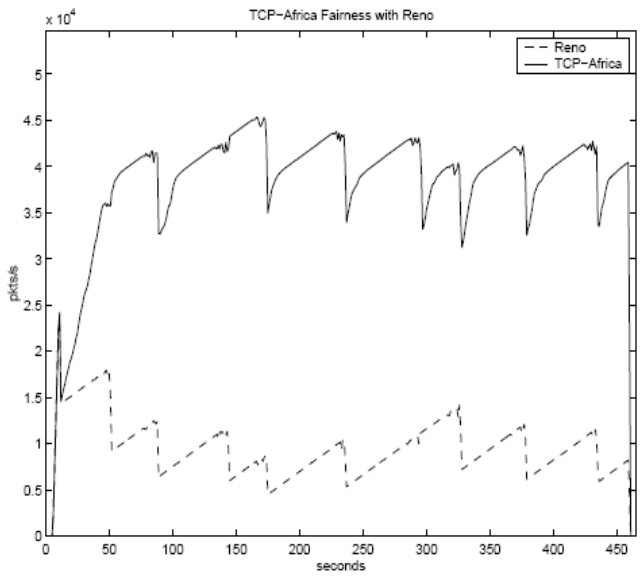


Figure 18: ns-2 experiment between TCP-AFRICA and TCP-Reno on a 622 Mbps link

To facilitate conducting live Internet experiments with TCP AFRICA, we have developed a user space TCP over UDP emulation tool. This software was designed to allow us to rapidly prototype new congestion control algorithms without requiring the extensive kernel programming and administrative permissions necessary to design and test a new TCP/IP stack. The emulator sends UDP packets containing random data between a sender and receiver according to its congestion control rules, in this case TCP AFRICA. The lightweight and efficient receiver daemon has been designed to be easily launched at remote sites, while all congestion control is handled by the sender, just as in normal TCP.

Using our emulator software, we have conducted limited live Internet experiments of TCP AFRICA on 100Mbps networks. These experiments have indicated that TCP AFRICA does not take an undue share from traditional Reno flows at these speeds. TCP AFRICA also demonstrated the ability to rapidly utilize available bandwidth.

6.0 Deliverables

This project has produced four classes of deliverables: (1) new theory and technologies, (2) software simulations of these technologies, (3) experimental testing and reference implementations, and (4) publications in the form of talks, technical reports, and papers.

Multiscale traffic analysis toolbox

We have completed development and release an open-source toolbox for analyzing high-speed network and Grid traffic traces using wavelets, multifractals, and the alpha/beta traffic decomposition.

End-to-end path modeling and inference toolbox

We have completed validation, testing, and software release of *pathChirp* and *STAB*.

Alpha-beta decomposition

We have developed and tested a C+ version of an "alpha-beta trace analyzer" that enables fast off-line analysis of traffic traces on the connection level.

TCP AFRICA

We have developed ns-2 simulation code and real-network emulation code for TCP AFRICA.

7.0 Publications Supported by this Grant

2005

S. Sarvotham, R. Riedi, and R. Baraniuk, "Network and User-Driven On-Off Source Model for Network Traffic," *Computer Networks*, Special Issue on 'Modeling Network Long Range Dependent Traffic: Characterization, Visualization, and Tools,' *Computer Networks* 48 (2005), p 335-350.

V. Ribeiro, R. Riedi, and R. Baraniuk, "Optimal Sampling Strategies for Multiscale Stochastic Processes," submitted to *Proceedings of First Erich Lehman Symposium*, Institute of Mathematical Statistics, 2005.

V. Ribeiro, R. Riedi, and R. Baraniuk, "Multiscale Queuing Analysis," submitted to *IEEE Transactions on Networking*, under review.

R. King, R. Riedi, and R. Baraniuk, "TCP-Africa: An Adaptive and Fair Rapid Increase Rule for Scalable TCP," *IEEE INFOCOM*, March 2005.

R. King, R. Riedi, and R. Baraniuk, "Evaluating and Improving TCP-Africa: An Adaptive and Fair Rapid Increase Rule for Scalable TCP," *International Workshop on Protocols for Fast Long-Distance Networks – PFLDNET'05*, Lyon, France, February 2005.

P. Chainais, R. Riedi and P. Abry, "Warped infinitely divisible cascades: Beyond power laws," *Traitement du Signal* **22**(1), 2005

P. Chainais, R. Riedi and P. Abry, "On Non-Scale-Invariant Infinitely Divisible Cascades," *IEEE Trans IT*, **51** (3), pp 1063--1083, (March 2005)

Y. Tsang and R. Nowak, “Optimal Network Tomography,” *IEEE INFOCOM 2005 Student Workshop*, March 2005.

Paulo Goncalves and Rudolf H. Riedi, “Diverging moments and parameter estimation,” *Journal American Statistical Association JASA* (accepted for publication Feb 2005)

2004

R. Castro, M. Coates and R. Nowak, "Likelihood Based Hierarchical Clustering," *IEEE Transactions in Signal Processing*, August 2004.

R. Castro, M. Coates, G. Liang, R. Nowak and B. Yu, "Network Tomography: Recent Developments," *Statistical Science*, 2004.

A. Kuzmanovic, E. Knightly and R. L. Cottrell, “HSTCP-LP: A Protocol for Low-Priority Bulk Data Transfer in High-Speed High-RTT Networks,” *Second International Workshop on Protocols for Fast Long-Distance Networks (PFLDnet'04)*, February 2004.

A. Kuzmanovic and E. Knightly and R. L. Cottrell. “Bulk Data Transfer in High- Speed High-RTT Networks,” *Proceedings of the 2nd International Workshop on Protocols for Fast Long-Distance Networks*, February 2004.

A. Kuzmanovic and E. Knightly, “A Performance vs. Trust Perspective in the Design of End-Point Congestion Control Protocols,” *IEEE ICNP 2004*, Berlin, Germany, October 2004.

M. Rabbat, R. Nowak, and M. Coates, “Multiple Source, Multiple Destination Network Tomography,” *IEEE Infocom*, March, 2004.

Rudolf H. Riedi, A. Keshavarz-Haddad, S. Sarvotham and Richard G. Baraniuk, “Fractals in Networking: Modeling and Inference,” *Proceedings of Fractal 2004, Conference on Fractals and Complexity in Nature*, Vancouver, Canada, April 2004.

V. J. Ribeiro, R. H. Riedi, and R. G. Baraniuk, “Locating Available Bandwidth Bottlenecks,” *IEEE Internet Computing Magazine*, **8** (5), pp 34--41, September 2004.

V. J. Ribeiro, R. Riedi, and R. G. Baraniuk, “Spatio-Temporal Available Bandwidth Estimation with STAB,” *ACM Sigmetrics/Performance*, June 2004.

V. Ribeiro, R. H. Riedi, and R. G. Baraniuk, “Multiscale Queuing Analysis of Long-Range-Dependent Network Traffic,” *TR2004-07*, Rice University, Dept. of Statistics, August 2004, submitted for publication Aug 2004.

Y. Tsang, M. Yildiz, R. Nowak and P. Barford, “Network Radar: Tomography from Round Trip Time Measurement,” *ACM Internet Measurement Conference*, October, 2004, Taormina, Sicily, Italy, pp.175-180

2003

R. Castro and R. Nowak, "Likelihood Based Hierarchical Clustering and Network Topology Identification," *Fourth International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR 2003)*, Lisboa, Portugal, July 2003. (Available in Lecture Notes in Computer Science 2683, Springer 2003)
M. Coates, M. Rabbat, and R. Nowak, "Merging Logical Topologies Using End-to-end Measurements," in Proceedings of ACM SIGCOMM Internet Measurement Conference, Miami Beach, Florida, October 2003.

A. Keshavarz-Haddad, “Effects of Traffic Bursts on the Network Queue,” *MS Thesis*, Department of Electrical and Computer Engineering, Rice University, 2003.

A. Kuzmanovic and E. Knightly, “Low-Rate TCP-Targeted Denial of Service Attacks (The Shrew vs. the Mice and Elephants),” *Proc. of ACM Sigcomm '03)*, August 2003.

A. Kuzmanovic and E. Knightly, “TCP-LP: A Distributed Algorithm for Low Priority Data Transfer,” *Joint Conference of the IEEE Computer and Communications Societies (IEEE Infocom'03)*, April 2003.

M. Rabbat, R. Nowak, and M. Coates, “Multiple Source Network Tomography: A Hypothesis-Testing Approach,” *IEEE Workshop on Statistical Signal Processing*, September 2003.

M. Rabbat, “Multiple Sender Network Tomography,” *MS Thesis*, Department of Electrical and Computer Engineering, Rice University, 2003.

V. Ribeiro, R. Riedi, and R. G. Baraniuk, “Spatio-Temporal Available Bandwidth Estimation for High-Speed Networks,” *ISMA 2003 Bandwidth Estimation Workshop (Best)*, CAIDA, La Jolla, CA, December 2003.

V. Ribeiro, R. Riedi, J. Navratil, R. L. Cottrell, and R. Baraniuk, “pathChirp: A Light-Weight Available Bandwidth Estimation Tool for Network-Aware Applications,” Proceedings LACSI Symposium, Santa Fe, NM, October 2003.

V. Ribeiro, R. Riedi, R. Baraniuk, “Optimal Sampling Strategies for Multiscale Models with Application to Network Traffic Estimation,” *IEEE Statistical Signal Processing Workshop*, September 2003.

V. Ribeiro, R. Riedi, R. Baraniuk, J. Navratil, R. L. Cottrell, “pathChirp: Efficient Available Bandwidth Estimation for End-to-End Paths,” *Passive and Active Measurement Workshop (PAM)*, March 2003. (*Winner of best student paper award*).

Y. Tsang, M. Coates and R. Nowak, "Network Delay Tomography," *IEEE Transaction of Signal Processing in Networking*, Aug. 2003, Volume 51, Issue 8, pp. 2125-2136

2002

P. Abry, R. G. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, “The Multiscale Nature of Network Traffic: Discovery, Analysis, and Modeling,” *IEEE Signal Processing Magazine*, May 2002.

W. Willinger, V. Paxson, R. H. Riedi and M. S. Taqqu, “Long-Range Dependence and Data Network Traffic”, in *Long range dependence: theory and applications*, eds. Doukhan, Oppenheim, Taqqu (Birkhauser 2002) ISBN: 0817641688

R. H. Riedi, “Multifractal Processes,“ in *Long range dependence : theory and applications*, eds. Doukhan, Oppenheim, Taqqu (Birkhauser 2002) ISBN: 0817641688

R. Castro, M. Coates and R. Nowak, "Maximum Likelihood Identification of Network Topology from End-to-end Measurements," *DIMACS Workshop on Internet Measurement, Mapping and Modeling*, DIMACS Center, Rutgers University, Piscataway, New Jersey, February 2002.

R. Castro, M. Coates and R. Nowak, "Maximum Likelihood Identification of Network Topology from End-to-End Measurements," *Technical Report TREE0109*, Department of Electrical and Computer Engineering, Rice University, May 2002.

M. Coates and R. Nowak, "Sequential Monte Carlo inference of internal delays in nonstationary data networks," *IEEE Transactions in Signal Processing*, Volume: 50 Issue: 2, Feb. 2002 Page(s): 366 -376.

M. Coates, R. Castro, R. Nowak and Y. Tsang, "Maximum Likelihood Network Topology Identification from Edge-Based Unicast Measurements," *Proc. ACM Sigmetrics*, Marina del Rey, California, June 2002.

M. Coates, A. Hero, R. Nowak, and B. Yu, “Internet Tomography,” *IEEE Signal Processing Magazine*, May 2002.

M. Coates, M. Raat and R. Nowak, "Discovering General Logical Network Topologies," *Technical Report TREE0202*, Department of Electrical and Computer Engineering, Rice University, November 2002.

M. Coates, R. Castro, M. Gadhiok, R. King, R. Nowak, E. Rombokas and Y. Tsang, "Maximum Likelihood Topology Identification from Edge-Based Unicast Measurements," *Technical Report TREE0107*, Department of Electrical and Computer Engineering, Rice University, Nov. 2001.

M. Coates, R. Nowak and Y. Tsang, "Nonparametric Estimation of Internal Delay Densities from Unicast End-to-end Measurement", *Technical Report TREE0106*, Department of Electrical and Computer Engineering, Rice University, October 2002.

M. Rabbat, R. Nowak, and M. Coates, “Network Tomography and the Identification of Shared Infrastructure,” *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2002.

Y. Tsang, M. Coates and R. Nowak, “Nonparametric Internet Tomography,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, Florida, May 2002, Volume 2, pp. 2045-2048.

S. Sarvotham, R. H. Riedi and R. G. Baraniuk, “Multiscale Connection-level Analysis of Network Traffic,” Proceedings of the 36th Conference on Signals, Systems and Computers, Asilomar, CA, November 2002.

Xin Wang, S. Sarvotham, R. H. Riedi and R. G. Baraniuk, “Network Traffic Modeling using Connection-Level Information,” *Proceedings SPIE ITCom*, Boston, MA, August 2002

P. Chainais, R. Riedi and P. Abry, “Compound Poisson cascades,” *Proc. Colloque "Autosimilarite et Applications"*, Clermont-Ferrant, France, May 2002

S. Sarvotham, X. Wang, R. Riedi, and R. Baraniuk, "Additive and Multiplicative Mixture trees for Network Traffic Modeling", *Proceedings ICASSP Orlando*, FL, (May 2002)

X. Wang, S. Sarvotham, R. Riedi, and R. Baraniuk, “Connection-Level Modeling of Network Traffic,” *Proceedings DIMACS Workshop on Internet and WWW Measurement, Mapping and Modeling*, Rutgers, NJ, (February 2002).

T. Karagiannis, M. Faloutsos and R. H. Riedi, “Long-Range Dependence: Now you see it now you don't!”, *Global Internet 2002*, November 2002

2001

M. Coates and R. Nowak, “Network Tomography for Internal Delay Estimation,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, Utah, May 2001.

S. Sarvotham, R. Riedi and R. Baraniuk, “Connection-level analysis and modeling of network traffic,” *Proc. IEEE/ACM Internet Measurement Workshop*, San Francisco, CA, November 2001.

Y. Tsang, M. Coates and R. Nowak, “Passive Unicast Network Tomography using EM Algorithms,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, Utah, May 2001, Volume 3, pp. 1469-1472.

V. Ribeiro, R. Riedi and R. Baraniuk, “Wavelets and Multifractals for network traffic modeling and inference,” *Proceedings, ICASSP01*, Salt Lake City, Utah, May 7-11, 2001.

2000

Rudolf Riedi, "On the multiplicative structure of network traffic," *Proceedings of the IMA Conference on Mathematics in Signal Processing*, Warwick (UK), December 2000.

V. Ribeiro, M. Coates, R. Riedi, S. Sarvotham, B. Hendricks and R. Baraniuk, "Multifractal Cross-Traffic Estimation," *Proceedings Specialist Seminar on IP Traffic Measurement, Modeling and Management*, September 2000, Monterey, CA.

8.0 Software Supported by this Grant

pathChirp: Available bandwidth estimation tool, http://www.spin.rice.edu/Software/pathChirp

STAB: Spatio-temporal available bandwidth estimation, http://www.spin.rice.edu/Software/STAB

MWM: Multifractal wavelet model for generating realistic bursty network traffic traces, http://www.dsp.rice.edu/software/mwm.shtml

poisson_gen: Poisson traffic generation tool, http://www.spin.rice.edu/Software/poisson_gen

9.0 Professional Personnel

Principal Investigators
Richard Baraniuk (PI), Rudolf Riedi (PoC), Edward Knightly, Robert Nowak

Postdocs (part of grant duration)
Mark Coates, Maarten Jansen, Xin Wang

Graduate Students (part of grant duration)
Mohammed Ahamed, Rui Castro, Alireza Keshavarz-Haddad, Ryan King, Aline Martin, Michael Rabbat, Vinay Ribeiro, Shriram Sarvotham, Yolanda Tsang

Academic Support (part of grant duration)
Myrl Carlson, Liz Hickman

10.0 Papers presented at conferences and meetings

Listed in chronological order:

"On the multiplicative scaling of internet traffic," Ericsson Workshop on High Speed Networking, Balatonfured, Hungary, May 2000

"Multifractal processes," Physics Dept., Eotvos University, Budapest, Hungary, May 2000

"Multiscale Modeling and Queuing Analysis of Network Traffic, "Budapest University of Technology and Economics, Budapest Hungary, May 2000

"Multifractal Inference of network performance characteristics along a single path," Conference on Stochastic Networks, University of Wisconsin-Madison, June 2000

"Multifractal Cross-Traffic Estimation," Proceedings Specialist Seminar on IP Traffic Measurement, Modeling and Management, September 2000, Monterey, CA

Multiscale Traffic Processing Techniques for Network Inference and Control," DARPA NMS Workshop, September 2000, Albuquerque, NM

"On the multiplicative structure of network traffic," IMA Conference on Mathematics in Signal Processing, Warwick Univ., UK, (Invited Speaker).

"Multiplicative Network modeling and Multifractal estimators," DIMACS workshop on "Wavelet Analysis, Digital Processing, and 3 Dimensional Multiresolution Analysis, Rutgers, NJ, April 2001 (Invited Speaker).

"Impact and Accuracy of Probing for Cross-Traffic Inference," DARPA NMS Workshop, San Diego, CA, April 2001.

"Multifractal Inference of Network Traffic," NISS Workshop on "Modeling and Analysis of Network Data," National Institute for Statistical Sciences, Research Triangle, NC, March 2001. (Invited Speaker).

"Wavelets and Multifractals for network traffic modeling and inference," ICASSP01, Salt Lake City, Utah, May, 2001.

"Multifractal structure and Inference of Network Traffic," EMUlab, Dept. Electronic and Electrical Engineering, University of Melbourne, Australia

"Connection-level Analysis and Modeling of Network Traffic," 16th IEEE Annual Workshop on Computer Communications, Charlottesville, VA, October 2001.

"Connection-level Analysis and Modeling of Network Traffic," ACM SIGCOMM Internet Measurement Workshop, San Francisco, CA, November 2001.

"Connection-Level Modeling of Network Traffic," X. Wang, S. Sarvotham, R. Riedi, and R. Baraniuk, Proceedings DIMACS Workshop on Internet and WWW Measurement, Mapping and Modeling, Rutgers, NJ, (February 2002).

"InterNet Control and Inference Tools at the Edge," Virtual Worlds and Simulation conference, SCS, San Antonio, TX, January 2002.

"Compound Poisson cascades," P. Chainais, R. Riedi and P. Abry, Proc. Colloque "Autosimilarite et Applications," Clermont-Ferrant, France, May 2002

"Connection-Level Modeling of Network Traffic," and "Additive and Multiplicative Mixture trees for Network Traffic Modeling," S. Sarvotham, X. Wang, R. Riedi, and R. Baraniuk, Proceedings ICASSP Orlando, FL, May 2002.

"Network Traffic Modeling using Connection-Level Information," Xin Wang, Shriram Sarvotham, Rudolf H. Riedi, and Richard G. Baraniuk, Proceedings SPIE ITCom, Boston, MA, August 2002.

"Multiscale Connection-level Analysis of Network Traffic," Shriram Sarvotham, Rudolf H. Riedi, and Richard G. Baraniuk, Proceedings of the 36th Conference on "Signals, Systems and Computers", Asilomar, CA, Nov. 2002.

"Long-Range Dependence: Now you see it now you don't!" T. Karagiannis, M. Faloutsos and R. H. Riedi, Global Internet 2002, November 2002.

"Optimal Sampling Strategies for Multiscale Models with Application to Network Traffic Estimation," Vinay J. Ribeiro, Rudolf H. Riedi and Richard G. Baraniuk, Workshop on Statistical Signal Processing SSP03, St. Louis, MO, Sept 2003.

"Bandwidth estimation: challenges in future networks (poster)," SAMSI workshop on heavy traffic and congestion control, Raleigh, NC, Oct-Nov 2003

"Multiscale connection-level analysis of network traffic," Affiliates Day, Depts. of ECE and CS, Rice University, Houston, TX, Oct 2003 (poster)

"Fractals in Networking: Modeling and Inference," Rudolf H. Riedi, A. Keshavarz-Haddad, S. Sarvotham and Richard G. Baraniuk, Fractal 2004, Conference on Fractals and Complexity in Nature, Vancouver, Canada, April 2004.

"Diverging moments and parameter estimation," Paulo Goncalves and Rudolf Riedi, CoTS (Conference of Texas Statisticians), Nacogdoches, TX, March 2004.

"On the origin and impact of scaling in network traffic," CIRM Workshop on Powerlaws in probability and statistics, Luminy-Marseille, France, March 2004.

"Spatio-Temporal Available Bandwidth Estimation with STAB," Summary and Poster at SIGMETRICS Performance '04, Vinay J. Ribeiro, Rudolf H. Riedi and Richard G. Baraniuk.

"Optimal Sampling Strategies for Multiscale Models with Application to Network Traffic Estimation," Vinay J. Ribeiro, Rudolf H. Riedi and Richard G. Baraniuk, Proceedings of the Lehmann Symposium, Houston, TX, April 2004.

"From Mandelbrot's multiplicative cascades to Infinitely Divisible Cascades," Invited Lecture at Summer School on Wavelet and Multifractal Analysis, Corsica, France, July 2004, Rudolf Riedi.

"Optimal Sampling Strategies for Multiscale Models with Application to Network Traffic Estimation," Proceedings of the Lehmann Symposium, Houston, TX, April 2004, Vinay J. Ribeiro, Rudolf H. Riedi and Richard G. Baraniuk.

"Evaluating and Improving TCP-Africa: an Adaptive and Fair Rapid Increase Rule for Scalable TCP," R. King, R. Riedi, and R. Baraniuk, International Workshop on Protocols for Fast Long-Distance Networks - PFLDNET'05, Lyon, France, February 2005.

"Spatio-Temporal Available Bandwidth Estimation with STAB," Summary and Poster at SIGMETRICS Performance 04, Vinay J. Ribeiro, Rudolf H. Riedi and Richard G. Baraniuk.

11.0 Consultative and Advisory Functions

Institution: SLAC, Palo Alto, CA
Personnel: Les Cottrell, Matthew Warrens, Jiri Navratil
Subject: Expand PingEr to traffic inference using Delphi
Meeting: Albuquerque, September 29, 2000, L. Cottrell
Visits: Stanford, Nov. 6, 2001 (L. Cottrell, M. Warrens); Rice, Feb 8-22 (J. Navratil)
Subject: Use Chirps to infer network conditions
Visits: Stanford, March 3-24, May 15-22; Rice student at SLAC

Joint experiments on the Internet: summer 2002
Testing pathChirp over the Internet using PingER: spring 2003

Institution: Telcordia
Personnel: Allan McIntosh, K. Krishnan
Subject: verification of pathChirp and alpha-beta decomposition
Sent code and consulted: Spring 2003
Closer collaboration in planning.

Institution: SAIC
Personnel: Gary Warren
Subject: MWM-code for traffic synthesis.
Date: Rice University, Houston, March 2001; per email.

Institution: LANL, Los Alamos, NM
Personnel: Wu Feng
Subject: Traffic measurements on high speed networks using
Visits: Los Alamos, March, 2001; Rice, January 7, 2002
Meetings: Reston VA, January, 2002; Napa CA, March, 2003

Institution: Sprint Labs, Burlingame, CA
Personnel: Sue Moon
Subject: traffic inference at core routers
Visits: Sprint Labs, November 5, 2001; Rice University, Houston, November 28-29, 2001
Visit: Sprint Labs, 2001-2002 internship of Rice student at Sprint Labs

Institution: Cetacean Networks, Inc., Portsmouth, NH
Personnel: Sean Moore, Principal Scientist
Subject: Multiscale Queuing Analysis for verification.
Date: Rice University, Houston, May 2002; per email.

Institution: GaTech (NMS peer)
Personnel: Thom McLean, G. Riley
Subject: Integration of pathChirp
Date: Rice University, Houston, May-Oct 2003; per email and phone.

Institution: J9
Personnel: Jason Boyer
Subject: Integration of pathChirp
Date: Rice University, Houston, July-Sept 2003; per email and phone.

Institution: Computer Sciences Corporation. Chantilly, VA
Personnel: Stephen C. Tsang, Sr. Systems Engineer
Subject: Proposal for collaboration using MWM in development of VoIP capabilities
Date: Rice University, Houston, February-July 2004; per email and by phone.