from the guest editors

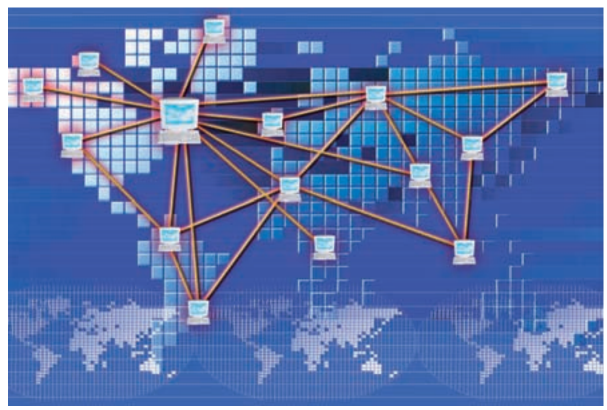
Athina Petropulu, *Drexel University*, and
Robert Nowak, *Rice University*

Signal Processing for Networking

The Internet—the Information Superhighway—is not only radically changing the way we communicate, educate, and conduct research and business, it is transforming the very fabric of our society. The Internet offers both tremendous opportunities and challenging dilemmas. In February 2000, a distributed denial of service (DDOS) attack brought Yahoo to its knees. Major Internet news centers were severely overloaded, causing major slowdowns in Internet traffic in the hours following the attacks on the World Trade Center and Pentagon last September. On a more positive note, at the Super Computing Bandwidth Challenge held last November in Denver, an unprecedented 3.3 gigabits per second transmission rate was sustained for several minutes over a high-speed network. Careful analysis, engineering, and optimization were the key to this success and will be the key to improving and enhancing Internet performance and security in the years to come.

This issue of *IEEE Signal Processing Magazine* focuses on an exciting new area of signal processing devoted to the study, modeling, and analysis of networks and network traffic. The reader will find many familiar SP concepts and algorithms playing key roles in this new application area. Moreover, we believe that the SP research community has an enormous wealth of experience and know-how to contribute to future generations of networks and the Internet, making them faster, more reliable, and more secure.

Network maintenance, design, provisioning, quality of service, and security all benefit from careful modeling and data analysis. This presents two major SP challenges: network traffic analysis and network performance mapping. The first challenge, network traffic modeling and analysis, is intriguing for the following reasons. For a long time, before the Internet, the dominant type of network traffic had been circuit-switched voice traffic. The characteristics of such traffic are well understood, and several analytical models have been proposed for studying its implications on network infrastructure. This type of traffic is here referred to as *traditional*. With the appearance of Internet and high-speed networks, however, traffic now originates from multimedia sources, with characteristics distinctly different from traditional teletraffic. We will refer to this new kind of traffic as *data traffic*. Data traffic exhibits large and frequent deviations about its average level, a feature that gives rise to unique and interesting statistical properties, most notably self-similar behavior or long-range dependence (i.e., a much slower decay of correlations than in traditional, Markovian situations), and impulsiveness (i.e., much slower decay of complementary distribution



than in the Gaussian case). Both of these unusual features, which have been observed almost invariably for data traffic in a variety of different settings, can have significant consequences in network design and management. Moreover, these issues cannot be effectively dealt with by traditional teletraffic theory. For example, they can lead to dramatically different queuing performance and buffer overflow probabilities. Along with the realization that we are now faced with a new type of traffic came the need for analytical characterization and modeling. Good analytical models can provide insight into certain traffic behavior and can be used to predict traffic consequences (e.g., buffer overflow and packet delays). Thus, they are extremely useful in designing and maintaining a network.

An analysis of the massive, heterogeneous, and decentralized system

1053-5888/02/$17.00©2002IEEE

that is the Internet poses a second formidable challenge: mapping network connectivity and performance as functions of space and time. To date, a wide variety of Internet maps have been produced using existing networking tools. Most of these standard mapping techniques, however, provide only a partial picture of the Internet. The decentralized nature of the Internet makes quantitative assessment of network performance very difficult. Due to many factors, including the expense of special-purpose measurement equipment, the burden of transmitting measurements from internal network nodes to a central processing center, and confidentiality and proprietary issues, mapping methodologies cannot depend on the cooperation of individual computers and routers to freely transmit vital network performance statistics such as traffic rates, delays, and dropped packet rates. This leads to the following intriguing question: Can accurate maps of network performance be derived from limited numbers of measurements at a limited set of measurement sites?

The first and second articles in this issue provide the theoretical framework for studying the aforementioned high-speed traffic characteristics in a way that is familiar to the SP audience. The first article, "Long-Range Dependence and Heavy-Tail Modeling for Teletraffic Data," focuses on the basic notions of long-range (or long-term) dependence and heavy-tailed marginal distributions (or impulsiveness). Several models that have been proposed in the network literature to capture and/or explain these features are discussed. A significant part of the contribution is devoted to illustrating the inherent difficulty in quantifying precisely these effects from empirical data.

The second article, "Multiscale Nature of Network Traffic," deals primarily with scaling issues related to the observed self-similar nature of network traffic. Although different notions can be attached to it, the idea of "scaling" basically refers to the fact that no single time scale completely captures the rich behavior of network traffic, which displays complex characteristics at scales of milliseconds to several tens of minutes. This results in nonstandard statistical behaviors that include long-range dependence or long memory, which are discussed in the first article, as well as multifractal properties related to a high variability of traffic traces at small scales. Since scaling processes involve many scales simultaneously, multiresolution (wavelet-based) techniques offer a natural framework for their statistical analysis, processing, and modeling. The article provides the reader with a survey of most of the scaling concepts relevant to network traffic, and it focuses on the capabilities of wavelet-based techniques for efficiently exploring the richness of such data.

The third article, "Internet Tomography," examines the problem of deriving spatial maps of network performance and connectivity from traffic measurements made at a relatively small number measurement sites. The general methodology used in such mapping problems is referred to as network tomography because of its analogy to tomographic medical imaging techniques. Network tomography falls in the realm of statistical inverse problems, an area that has long been of interest to signal and image processing researchers.

Signal processing expertise, from areas such as image reconstruction, system identification, and sensor array signal processing, can provide tremendous insight into networking inverse problems. The articles in this issue survey the state-of-the-art in network tomography and demonstrates how SP theory and methods can be brought to bear on this important new problem.