Statistical Science
2006, Vol. 21, No. 4, 485–486
DOI: 10.1214/088342306000000592
Main article DOI: 10.1214/088342306000000583

Rejoinder: Monitoring Networked Applications With Incremental Quantile Estimation

John M. Chambers, David A. James, Diane Lambert and Scott Vander Wiel

1. DIVERSITY OF MONITORING GOALS AND CONSTRAINTS

There are many kinds of networks, each with many types of variables and monitoring goals. Our paper addressed only one of the countless possible combinations of network and monitoring goals. We are grateful to the discussants for expanding our paper by providing insights into other network monitoring problems that present different challenges to statisticians.

Denby, Landwehr and Meloche (DLM) describe three network monitoring problems, each with different requirements for detection speed, communication constraints and scalability. The Voice over Internet protocol (VoIP) application, for example, requires good scalability, low overhead and quick responses to problems that manifest in a variety of quality-of-service (QoS) metrics. Monitoring service-level agreements, on the other hand, needs a prompt signal when path transit times become too long—a more focused goal than the VoIP problem. Our monitoring problem is most similar to DLM's third example, monitoring call centers through flexible reporting of historical reliability and performance. These problems typically have a wide variety of analytic goals, some of which are not determined until an analyst begins to drill through high-level summaries into data slices that show unusual behavior.

Whereas DLM concentrate on full-path QoS for VoIP, Lawrence, Michailidis and Nair (LMN) describe a QoS problem in which path measurements are used to estimate link-level characteristics, presumably for the purpose of managing the network, perhaps by modifying routing tables, adding key links or upgrading hardware at nodes.

To the list of monitoring problems that we and the discussants have described, we would add detection of worm outbreaks (Bu, Chen, Vander Wiel and Woo, 2006), dynamic thresholding of error counts (Lambert and Liu, 2006), fraud detection (Cahill, Lambert, Pinheiro and Sun, 2002) and call blocking events (Becker, Clark and Lambert, 1998). And there are certainly others that we are overlooking.

The variety of applications raised by the reviewers and our own experience demonstrate that there is no canonical statistical problem in the domain of monitoring networks for performance and reliability. In our application, the software architects imposed a hard constraint that the summary records had to have a fixed length and would be transmitted at regular intervals. Also, the requirement for a very small footprint stemmed from the need for the agent software to run on personal computers that may be old and slow and may be connected to the network by a low bandwidth link. While the quantile estimates must be reasonably accurate, the growth plan for the business placed much more emphasis on ease of implementation for new features and upgraded architecture to improve scalability. Therefore, improvements to quantile accuracy had to be made with relatively low development (software coding) cost. The simplicity of Incremental Quantiles (IQ) was obviously attractive.

2. DATA COMPRESSION

DLM, LMN and Yu all discuss connections that the IQ algorithm has to methods for compressing and sketching data streams. Although compression was not likely to be used in our application, it is critical for sensor networks, for example, where data transmission is much more costly. We hope that Yu and others will pursue statistical compression methods that allow updating summaries without decompression.

3. SMOOTHING AND DETECTION PERFORMANCE

LMN advocate that, for monitoring purposes, "the procedure should be devised to estimate the current scenario" and then outline how exponentially weighted moving averages (EWMAs) could be formed using either quantiles or cumulative distribution functions (CDFs).