Chapter 2
Load Characterization and Measurement

Abstract  Timely detection of changes in traffic is critical for initiating appropriate traffic engineering mechanisms. Accurate measurement of traffic is an essential step towards change detection and traffic engineering. However, precise traffic measurement involves inspecting every packet traversing a link, resulting in significant overhead, particularly on routers with high speed links. Sampling techniques for traffic load estimation are proposed as a way to limit the measurement overhead. Since the efficacy of change detection depends on the accuracy of traffic load estimation, it is necessary to control error in estimation due to sampling. In this paper, we address the problem of bounding sampling error within a pre-specified tolerance level. We derive a relationship between the number of packet samples, the accuracy of load estimation and the squared coefficient of variation of packet size distribution. Based on this relationship, we propose an adaptive random sampling technique that determines the minimum sampling probability adaptively according to traffic dynamics. Using real network traffic traces, we show that the proposed adaptive random sampling technique indeed produces the desired accuracy, while also yielding significant reduction in the amount of traffic samples, yet simple to implement. We also investigate the impact of sampling errors on the performance of load change detection.

Key words:  adaptive random sampling, traffic measurement, packet trace, estimation accuracy, traffic load, change detection

2.1 Introduction

With the rapid growth of the Internet, traffic engineering has become an important mechanism to reduce network congestion and meet various user demands. Measurement of network traffic load is crucial for configuring, managing, pricing, policing, and engineering the network. Network traffic may fluctuate frequently and often unexpectedly for various reasons such as transitions in user behavior, deployment of
new applications, changes in routing policies, or failure of network elements. It is a daunting task for network administrators to manually tune the network configuration to accommodate the traffic dynamics. Thus, there is a need for automatic tools that enable intelligent control and management of high speed networks. Clearly, accurate measurement of traffic load is a prerequisite to on-line traffic analysis tools.

On-line traffic analysis tools are only implemented on a separate measurement-box. It is because traffic analysis is not a part of router functionality in the current routers, and it is easier to develop and augment analysis functions separately from a router’s built-in software.

The Simple Network Management Protocol (SNMP) [21] provides router status information such as interfaces, routing tables, and protocol states defined by Management Information Base (MIB). Hosts or routers, as SNMP agents, can make requests to pull out MIB variables. In backbone routers, however, SNMP polling interval cannot be set short, since the router’s performance should not be degraded replying SNMP requests, and bandwidth should not be wasted too much for SNMP messages. Also, SNMP is processed at a low priority in a router. Thus, often SNMP is not flexible, and the data is not readily available for timely analysis on-the-fly.

For out-of-the-box (or off-board) measurement systems, inspecting every single packet that traverses a link, however, is extremely costly. It may not keep up with today’s high-speed of links. With off-board measurement devices, huge volumes of data are generated that can quickly exhaust storage space.

Sampling techniques may provide timely information economically, in particular if on-line analysis of the data is needed. However, sampling inevitably introduces errors in the traffic load estimation. Such errors may adversely affect the measurement applications.

In this chapter, we develop an adaptive random sampling technique for traffic load measurement. Our adaptive random sampling technique differs from existing sampling techniques for traffic measurement in that it yields bounded sampling errors within a pre-specified error tolerance level. Such error bounds are important in reducing the “noise” in the sampled traffic measurements. Furthermore, the pre-specified error tolerance level allows us to control the performance of on-line traffic analysis algorithms as well as the amount of packets sampled. This chapter is devoted to the analysis and verification of the proposed adaptive random sampling technique for traffic load measurement. We also study the impact of sampling errors on the performance of on-line traffic analysis - load change detection. Our contributions are summarized as follows.

We observe that sampling errors in estimating traffic load arises from the dynamics of packet sizes and counts, and these traffic parameters vary over time. Consequently, static sampling (i.e., with a fixed sampling rate) cannot guarantee errors within a given error tolerance level. From our analysis, we find that the minimum required number of samples to bound a sampling error within a given tolerance level is proportional to the squared coefficient of variation (SCV) of packet size distribution. Using this relationship, we propose an adaptive random sampling technique that determines the (minimum) sampling probability adaptively based on the SCV of packet size distribution and the packet count. More specifically, time is divided