

Internet Traffic Data

William S. Cleveland

<http://cm.bell-labs.com/stat/wsc>

Bell Labs, 600 Mountain Av., Murray Hill, NJ 07974

wsc@research.bell-labs.com

Internet traffic data are exciting because they measure an intricate, fast-growing network connecting up the world and transforming society.

And an understanding of the statistical properties of Internet traffic are vital for network performance, network design, traffic engineering, security, and quality of service. Two ingredients are required for this understanding: frameworks for traffic measurement that produce data bearing on the Internet issues, and statistical models for the data. This talk will review the standard research framework for measurement, packet header capture, and will review the major themes that have been pursued in analyzing packet header data, largely time-series issues revolving around long-range dependence. Then, recent discoveries about the fundamental nature of Internet traffic will be presented in detail.

For Internet engineering, the most important aspect of the discoveries is that as the number of active connections multiplexing (superimposing) their packets on an Internet link increases, packet inter-arrivals and sizes go from long-range independent to dependent.

This has important implications for router design and for determining the amount of traffic that can be sent over a link for a fixed level of packet loss. But much more broadly, studies of Internet engineering technologies that depend on the statistical properties of packet traffic need to consider how performance changes as the number of multiplexed connections changes.