Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computer Networks 47 (2005) 219–237

COMPUTER NETWORKS

www.elsevier.com/locate/comnet

Improving TCP performance in integrated wireless communications networks

Kai Xu, Ye Tian, Nirwan Ansari *

Advanced Networking Laboratory, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA

Available online 6 August 2004

Abstract

Many analytical and simulation-based studies of TCP performance in wireless environments assume an error-free and congestion-free reverse channel that has the same capacity as the forward channel. Such an assumption does not hold in many real-world scenarios, particularly in the hybrid networks consisting of various wireless LAN (WLAN) and cellular technologies. In this paper, we first study, through extensive simulations, the performance characteristics of four representative TCP schemes, namely TCP New Reno, SACK, Veno, and Westwood, under the network conditions of asymmetric end-to-end link capacities, correlated wireless errors, and link congestion in both forward and reverse directions. We then propose a new TCP scheme, called TCP New Jersey, which is capable of distinguishing wireless packet losses from congestion packet losses, and reacting accordingly. TCP New Jersey consists of two key components, the timestamp-based available bandwidth estimation (TABE) algorithm and the congestion warning (CW) router configuration. TABE is a TCP-sender-side algorithm that continuously estimates the bandwidth available to the connection and guides the sender to adjust its transmission rate when the network becomes congested. TABE is immune to the ACK drops as well as ACK compression. CW is a configuration of network routers such that routers alert end stations by marking all packets when there is a sign of an incipient congestion. The marking of packets by the CW-configured routers helps the sender of the TCP connection to effectively differentiate packet losses caused by network congestion from those caused by wireless link errors. Our simulation results show that TCP New Jersey is able to accurately estimate the available bandwidth of the bottleneck link of an end-to-end path; and the TABE estimator is immune to link asymmetry, bi-directional congestion, and the relative position of the bottleneck link in the multi-hop end-to-end path. The proactive congestion avoidance control mechanism proposed in our scheme minimizes the network congestion, reduces the network volatility, and stabilizes the queue lengths while achieving more throughput than other TCP schemes.

* Corresponding author. Tel.: +1 973 596 3670; fax: +1 973 596 5680.
E-mail addresses: kx2@njit.edu (K. Xu), yt2@njit.edu (Y. Tian), nirwan.ansari@njit.edu (N. Ansari).

1389-1286/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2004.07.006

Keywords: Wireless TCP; Bandwidth estimation; Explicit congestion notification; Congestion control; Loss differentiation

1. Introduction

Communication networks have evolved greatly in the last decade. Packet switching technologies have eventually merged the traditional voice networks and data networks together into a converged and integrated multimedia network. The horizon of the converged integrated network is extending further to incorporate wired, wireless, and cellular technologies. The all-IP wired and wireless hybrid network is becoming a reality and the wireless network is getting more involved in our daily communications. The revolutionary cellular technologies and the network integration of WAN, LAN, WLAN and cellular networks stretch the Internet beyond the limits of geography and terrain. In integrated WAN + LAN + 3G cellular systems, illustrated in Fig. 1, more data and multimedia communications are carried end-to-end over the existing Internet protocol infrastructure. TCP/IP is the dominant communication protocol suite in today's multimedia applications. Nowadays, most of the Internet traffic is carried by TCP, including traffic generated by Web accesses, e-mails, bulk data transfers, remote terminals, etc. Spurred by the demand of wireless Internet, TCP/IP needs to depart from its original wired network oriented design and evolve to meet the challenges introduced by the wireless portion of the network. Internet Protocol (IP) [1] is a connection-less, best-effort based variable length packet delivery network layer protocol that does not guarantee the reliable, timely and in-order delivery of packets between end stations. TCP, the transmission control protocol [2], is a layer-4 transport protocol that uses the basic IP services to provide applications with an end-to-end connection-oriented packet transport mechanism that ensures the reliable and ordered delivery of data.

Improving TCP's performance in wireless IP communications has been an active research area. The performance degradation of TCP in wireless and wired-wireless hybrid networks, as reported in much research [3,4], is mainly due to its lack of the ability to differentiate the packet losses caused by network congestion from the losses caused by wireless link errors. TCP was originally designed primarily for wired networks. In wired networks, the random bit error rate (BER) is negligible, and congestion is the main cause of packet losses. The TCP sender behavior of adjusting the sending rate of data packets is triggered by the self-clocking acknowledgement (ACK) sent by the corresponding receiver after successfully receiving the data packet. When packet loss occurs at a congested link due to buffer overflow at the intermediate router, either the sender receives duplicate ACKs (DUPACK) or the sender's retransmission timeout (RTO) timer expires. These events activate the sender's congestion control mechanism by which the sender reduces the size of its transmission window, or congestion window (*cwnd*) in TCP terminology, resulting in a lower transmission rate to relieve the link congestion. This is a reactive congestion control scheme, in which the action is triggered by the sender's self-induced congestion. Such TCP sender behavior works fairly well in the wired networks where packet losses are almost always caused by link congestion; and packet losses due to bit errors are usually negligible or, if any, not exceeding one packet loss per *cwnd*. However, in wired/wireless heterogeneous networks, high BER, fading and blackout become non-negligible factors for packet losses. Standard TCP's congestion control and congestion avoidance mechanisms based on the assumption that all packet losses are due to congestion become incapable of handling the mixed packet losses. TCP without modification exhibits throughput degradations when used in wired/wireless heterogeneous networks.

Many analytical and simulation-based studies of TCP performance in wireless environments assume an error-free and congestion-free reverse channel that has the same capacity as the forward channel. Such an assumption does not hold in many real-world scenarios, particularly in the hybrid networks constituted by various wireless