Wireless Networks 8, 301–316, 2002

WTCP: A Reliable Transport Protocol for Wireless Wide-Area Networks

PRASUN SINHA
Bell Laboratories, Lucent Technologies, New Jersey, USA

THYAGARAJAN NANDAGOPAL
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, USA

NARAYANAN VENKITARAMAN
Motorola Labs, Schaumburg, IL, USA

RAGHUPATHY SIVAKUMAR
Electrical and Computer Engineering, Georgia Institute of Technology, USA

VADUVUR BHARGHAVAN
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, USA

Abstract. Wireless wide-area networks (WWANs) are characterized by very low and variable bandwidths, very high and variable delays, significant non-congestion related losses, asymmetric uplink and downlink channels, and occasional blackouts. Additionally, the majority of the latency in a WWAN connection is incurred over the wireless link. Under such operating conditions, most contemporary wireless TCP algorithms do not perform very well. In this paper, we present WTCP, a reliable transport protocol that addresses rate control and reliability over commercial WWAN networks such as CDPD. WTCP is rate-based, uses only end-to-end mechanisms, performs rate control at the receiver, and uses inter-packet delays as the primary metric for rate control. We have implemented and evaluated WTCP over the CDPD network, and also simulated it in the *ns-2* simulator. Our results indicate that WTCP can improve on the performance of comparable algorithms such as TCP-NewReno, TCP-Vegas, and Snoop-TCP by between 20% to 200% for typical operating conditions.

Keywords: Wireless transport, WTCP, reliable wireless transmission

1. Introduction

Recent years have witnessed an explosive growth in the use of wireless wide-area networks (WWANs) such as CDPD [23], RAM [23], Ardis [23] and HDR [22], with industry projections topping $2.5 billion by the year 2002. Many large corporations are equipping their mobile work-force with laptops that have WWAN connectivity, thus enabling users to be connected anytime, anywhere. In the typical WWAN deployment scenario, a mobile user connects over the WWAN network to a dedicated proxy in the corporate backbone, which then acts as the service point for all the requests from the mobile user[1]. Providing efficient and reliable connectivity between the proxy and the mobile host over commercial wide-area wireless networks is, thus, becoming a critical issue.

[1] Most of the current deployment uses the proxy model for three reasons: (a) the connection quality of the WWAN network is too poor to sustain typical client-server applications, (b) the amount of data transferred to the mobile host must be filtered because of the orders of magnitude difference in bandwidth between the wired and WWAN connections, and (c) many portable computing devices have display and processing limitations that must be addressed by the proxy before sending the processed/filtered response to the mobile user.

Despite the enormous commercial interest, a typical WWAN network exhibits some or all of the following characteristics: very low and highly variable throughput (between O(100 bps)–O(10 Kbps)), very high and highly variable latency (between O(400 ms)–O(4 s)), bursty and random packet losses unrelated to congestion (1–10%), and occasional blackouts, some of which can exceed 10 s. Under such operating conditions, applications require a generic transport protocol that is amenable to different types of wireless networks which exhibit similar characteristics. However, standard transport protocols such as TCP perform very poorly in such conditions because of two main reasons: first, TCP assigns all packet losses to congestion and throttles down the transmission rate upon detecting a packet loss; and second, TCP sets its retransmission timeout (RTO) based on observed round trip times ($\text{RTO} = (\overline{rtt}) + 4 \times \Delta(rtt)$). For the WWAN networks under consideration, neither mechanism is appropriate. Packet losses may happen either due to adverse channel conditions or congestion, and throttling down the transmission rate in the former case is the wrong thing to do. Furthermore, due to the very low bandwidth, transmission delays contribute significantly to the observed round trip time. Thus,