TCP Performance Improvement with ACK Pacing in Wireless Data Networks*

Seongho Cho[1], and Heekyoung Woo[2]

[1] School of Electrical Engineering and Computer Science,
Seoul National University, Seoul, Republic of Korea
`shcho@popeye.snu.ac.kr`
[2] Division of Information Technology Engineering,
Soonchunhyang University, Asan, Republic of Korea
`woohk@sch.ac.kr`

Abstract. Wireless data services are fundamental in ubiquitous networks. In wireless data networks, the TCP congestion mechanism can mistake a wireless channel error or blackout for congestion on the route. Several wireless TCP protocols have been proposed considering TCP performance degradation in wireless systems. However, these protocols are too complex to deploy in wireless systems. In this paper, we propose the ACK Pacing algorithm for long lasting loss periods like a hand-off or a blackout. By pacing ACK packets sending, we prevent bursty data delivery to the old path and a spurious timeout occurrence. Also, paced ACK packets trigger the route update to the new path faster. The proposed algorithm can be implemented in the wireless-node-side which is easily applicable to wireless systems. From simulation results, ACK Pacing improves TCP performance compared with many TCP variations. The ACK Pacing algorithm shows good performance during a long lasting loss period.

1 Introduction

Wireless mobile communication has become widespread with the pervasive 2nd generation wireless system and wireless LAN. In particular, the upcoming 3rd and 4th generation systems combine wired and wireless network services. These future telecommunication systems eventually evolve into All IP networks, which are IP-based integrated networks. In All IP networks, any mobile terminal can get ubiquitous network services with an IP protocol. In these systems, the mobility of the IP protocol is a key technical issue.

Mobile IP [1] has been proposed for the mobility of the future network. Mobile IP registers the current location of a mobile node to the HA (Home Agent) and relays every packets to the mobile node at the registered location. This protocol can provide the mobility of the terminal in a macro area. In a cellular system, frequent movement from the hand-off requires more efficient mobility

---

* This work was supported by the University IT Research Supporting Program under the Ministry of Information & Communication of Korea.

management. For frequent mobility management, Micro Mobile IP protocols, such as Cellular IP [2], and HAWAII [3], have been proposed. In these systems, the mobility among base stations is reported to the wireless gateway router and the macro area movement is managed by the previous Mobile IP. By hiding the frequent movement into a cellular area, it reduces the overhead of registering the current location to the HA and ensures seamless communication with the correspondent node.

An IP-based mobile communication system takes advantage of the reusability of abundant upper layer protocols used in wired networks. In wired networks, most of the packet loss and abnormal delay occur from router congestion. A TCP congestion control method is designed to adjust the sending rate in a low physical error environment. However, TCP can suffer from performance degradation in a lossy wireless link, because TCP regards the packet loss as the router congestion. Since wireless channel errors can occur randomly, reducing the transmission rate from all packet losses is not necessarily related to congestion in networks. Throughput reduction is often unnecessary for losses from a wireless channel error. Also, the mobile node can experiences long lasting wireless error such as a hand-off or a blackout. These periods cause a TCP timeout and throughput degradation.

To solve these problems, there have been several approaches [4, 5, 6, 7, 8, 9, 10, 11, 12, 13] taken on wireless TCP performance improvement. Some protocols break the end-to-end principle of the transport layer. The flow is divided into a wired part and a wireless part by base stations or supervising hosts. This can cause instability in the TCP connection. Some protocols need the extra information at supervising hosts or base stations. That's because the end system can't distinguish the wireless channel error from wired router queue drop. The aid of supervising hosts or base stations increases the overhead of wireless systems. And some protocols require the wired-node-side modification. However, this wired-node-side modification is not applicable because TCP is already adapted to nodes in wired networks. Some protocols cannot adapt to the fair share of sending rate for the route change. By keeping the old path information such as RTT and congestion window size, TCP can be unstable for a new path.

We propose the TCP performance improvement mechanism with ACK Pacing for a long lasting loss period. Long lasting channel losses can be predicted from the low measured pilot signal channel. The mobile node measures the pilot signal strength and paces ACK packets sending when the measured signal strength is under a certain threshold. We can reduce packet losses by delaying the bursty data transmission on a high loss channel. When a mobile node moves to the new base station, paced ACK packets reduce data transmission to the old station and prevent bursty losses. Also, paced ACK transmission through the new base station can rapidly initiate the route update to the new path. Our method doesn't violate the end-to-end principle without intermediate nodes' operation.

The rest of the paper is organized as follow. In the next section, we classify the previous work on wireless TCP. The loss separation methods from the wireless and wired network and the loss and delay controlling methods using characteristics of wireless networks are explained. In Section 3, we propose TCP ACK pacing method for a long lasting loss period. In Section 4, we simulate and compare our algorithm with TCP Tahoe, Reno, NewReno, and Delayed ACK in Micro Mobile IP networks. In Section 5, we conclude and present some future research issues on the subject.

2 Previous Work

2.1 Loss Separation Methods

In wireless cellular networks, packet losses can occur from wireless channel conditions or from the movement of the terminal. However, the existing TCP congestion control mechanism reduces the sending rate because it regards the wireless loss as the congestion notification. This reduction of the sending rate can deteriorate the TCP throughput. By separating losses from wireless networks with losses from wired networks, TCP performance can be improved. I-TCP [4], M-TCP [5], SNOOP [6], ELN [7], W-TCP [8], and TCP Westwood [9] can be classified in this group.

I-TCP (Indirect-TCP) and M-TCP (Mobile-TCP) get help from the base station. I-TCP separates the connection at the base station. The base station maintains two connections; one for a mobile host and a base station and another for a wired host and a base station. The base station receives data from the sender at a wired network as if it is a mobile node. And the base station transmits data to the mobile node by pretending it is a wired sender. M-TCP keeps the retransmission timeout timer at the base station which is used by the TCP sender for checking a timeout event. And the base station sends ACK packets before the timer is expired.

SNOOP is a link layer assisted TCP improvement protocol. The base station inspects all packets and keeps all the unACKed packets. If there are wireless losses, the base station retransmits the lost packets. This method hides the wireless packet losses from the mobile node and prevents a timeout event occurrence or a duplicate ACK. ELN (Explicit Loss Notification) keeps the wireless loss information at the base station, and the base station sets the ELN bit to the duplicate ACK. For the ELN-set duplicate ACK, the TCP sender regards the packet loss as a wireless channel error and doesn't decrease the window size. I-TCP, M-TCP, SNOOP and ELN have a flaw of keeping information on each TCP connection and having additional memory spaces in the base station for data or ACK packets.

Unlike upper protocols, without the aid of the base station, wireless loss estimation methods exist. From the packet arrival interval, TCP can separate the wireless loss from the wired loss. W-TCP (Wireless-TCP) is a rate-based wireless congestion control protocol. W-TCP regards the bursty losses as router congestion and random losses as the wireless channel error. In each case, the receiver measures the received interval of packets and notifies the sender to control the sending rate by using the estimate.

TCP Westwood is a sender-side modification for wireless TCP. From the received ACK packet, the sender estimates the current sending rate. For the congestion notification, the sender decides the congestion window size from the estimated sending rate. W-TCP and TCP Westwood should estimate exactly and need a fine-grained timer. And in a dynamically changing network, the estimation can be wrong.

2.2 Loss or Delay Controlling Methods

There are several methods to control packet losses or delay in a wireless network environment. In particular, the hand-off increases the packet losses and slows down the route update to the new base station. To solve these problems, Path Prediction [10], ACK Regulator [11], Freeze-TCP [12], and TCP-Probing [13] have been proposed.

Path Prediction provides a method of predicting the new base station and transmitting data to the new base station. When a mobile node moves to the new base station, packets transmitted to the old base station are lost until a route update occurs. To prevent misrouted losses, the wireless supporting nodes predict the hand-off pattern of the mobile node and transmit data packets to the predetermined path. The performance of the Path Prediction method is affected by the accuracy of predicting the new base station. Duplicate data transmission via multiple base stations wastes the precious wired and wireless resources.

ACK Regulator focuses on the fact that triple or more losses affect the performance of TCP. ACK Regulator at the supervising host controls the sender by regulating the ACK transmission rate. The ACK transmission rate is determined by the current data buffer size, and it prevents the sender from experiencing the bursty data losses due to an insufficient data buffer size. This method requires that the supervising hosts maintain the data and ACK buffer for all connections. And there is also the control overhead for the ACK regulating rate calculation.

Freeze-TCP applies the Zero Window Size (ZWS) mechanism, one of the TCP flow control mechanisms, at the hand-off time. When a mobile node moves to the new base station, TCP advertises the receiver windows size as zero. The ZWS-received sender stalls the data transmission and keeps the previous congestion window size and RTT (Round Trip Time) until the hand-off ends. This mechanism has a problem with transmission pausing during a hand-off period. Also the stored TCP connection information can induce congestion in the new path.

TCP-Probing checks the network state using probing packets, and decides the congestion control from the probing results. When a sender receives the congestion notification, the sender stops data transmission and exchanges an extra control packet with the receiver. If the control packer is not returned, the congestion control mechanism is initiated. On the other hand, if the control packet is returned in an estimated RTT, it resumes data transmission with the previous TCP state. This protocol needs both the sender and receiver modification and has probing packet overhead.

3 ACK Pacing during a Long Lasting Loss Period

3.1 Protocol Design Criteria

As we considered before, previous work had several deployment problems, such as the end-to-end principle violation, the overhead to base stations or supervising hosts, and the adaptability to the current system. Therefore, we designed a new protocol considering the following design criteria.

First, the end-to-end principle should be kept without any aid of base stations or supervising hosts. If the flow is divided into a wired and wireless part, it is possible the TCP flow can be unstable. Second, base stations or supervising hosts should not keep flow information. If the base station keeps the TCP information or data/ACK buffer, the overhead of base stations is increased, which is not suitable for the current trend of low cost and small size for base stations. Third, only the information at the mobile node should be used. An extra control message increases the complexity and overhead. In the cellular system, the base station transmits the pilot signal for the mobile node to measure the channel state and the distance from the base station. This kind of lower layer information can be helpful for performance improvement. Fourth, only the wireless-node-side modification should be required. Several wired-node-side or both-side modifications, such as TCP Westwood, TCP-Probing, and SACK [14], have been proposed to improve TCP performance. However, it will take a long time to deploy a new mechanism in too many servers using previous TCP implementation. Finally, a wireless TCP protocol should adapt naturally the sending rate to the fair share of the new path. After the hand-off period, the path to the new base station can be congested. Therefore, a new transmission rate should be selected using a congestion control mechanism. However, previous wireless TCP congestion controls, such as Freeze-TCP, preserve the previous state before the hand-off and transmit data with the previous transmission rate.

3.2 ACK Pacing Algorithm

We propose an ACK Pacing mechanism for a long lasting loss period. The following two parameters are considered. First, we decide the period of the ACK pacing. The mobile node measures the pilot signal from the base station. If the signal strength is below a certain level, the terminal moves to the new base station. We define this hand-off signal level as L_Handoff. And we select the signal strength level of the pacing period as L_Pacing as (1- α)L_Handoff, where $0 < \alpha < 1$. The ACK pacing begins, when the pilot signal strength is detected below L_Pacing. And the ACK pacing is finished when the pilot signal strength of the new base station is above L_Pacing. This ACK Pacing period is displayed as the gray period as shown in Figure 1.

Second, the ACK pacing time is considered. The route update period for the mobile node can affect the performance of ACK Pacing. Considering the route update period of the micro mobility system and the MAC characteristics, we select the ACK pacing time as 100 ms.

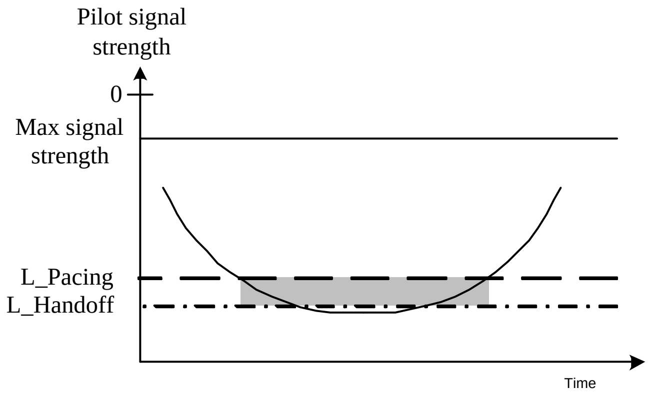


Fig. 1. The ACK Pacing Period Decision with the Pilot Signal Strength

When the mobile node is a sender, data can be paced during a low signal strength period for the same effect. And this can also improve TCP performance.

3.3 Comparison with Previous Work

We compare ACK Pacing with the other wireless TCP protocols. The characteristics of each protocol are classified in Table 1. The simplicity of BS means whether the base station keeps the information for the TCP connection. The wired-side and wireless-side modification means that which side of protocol change is needed. The end-to-end principle means that a protocol uses only packets from each end node. The path adaptation means that a protocol adapts the state of the new path without just keeping the old path information. And the control overhead means that a protocol uses an extra control packet at the link layer or the transport layer. As shown in Table 1, ACK Pacing is the simplest algorithm compared with previous protocols.

Table 1. Characteristics of Wireless TCP Protocols

Protocol	Simplicity of BS	Wired-side	Wireless-side	End-to-end Principle	Path Adaptation	Control Overhead
		Modification				
I-TCP	x	x	x	x	x	x
M-TCP	x	x	x	x	x	x
SNOOP	x	x	x	o	x	o
ELN	x	x	o	x	x	o
W-TCP	o	o	o	o	o	x
TCP Westwood	o	o	x	o	o	x
Path Prediction	x	x	x	o	x	o
ACK Regulator	x	x	x	x	x	x
Freeze-TCP	o	x	o	o	x	o
TCP-Probing	o	x	o	o	o	o
ACK Pacing	o	x	o	o	o	x

4 Simulation Results

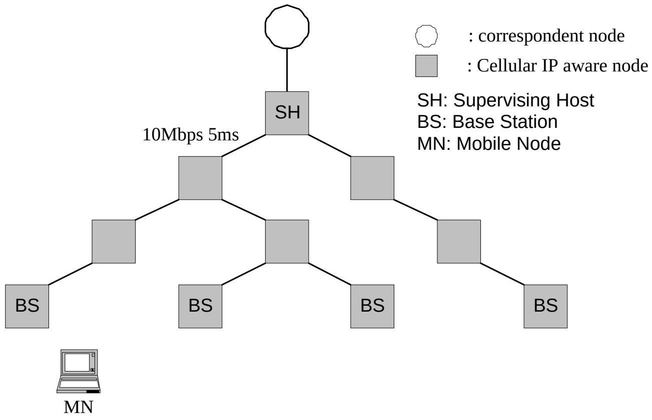
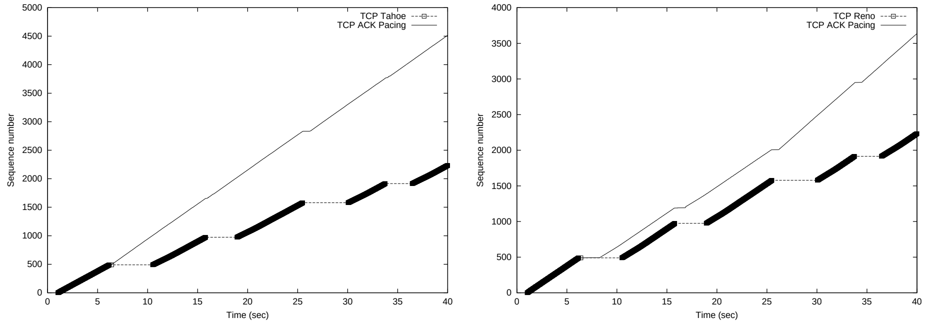


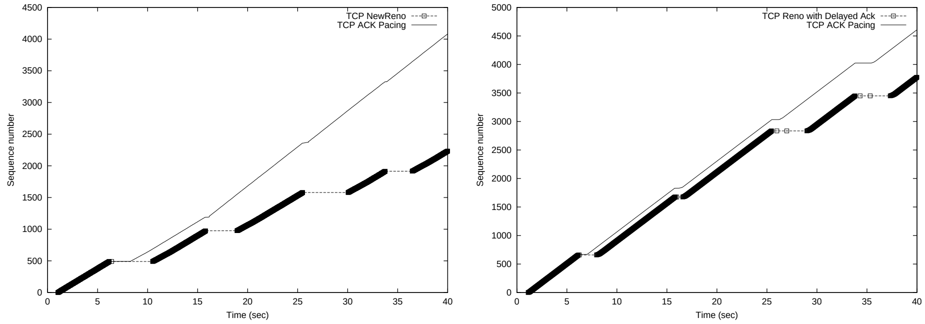
Fig. 2. Simulation Topology

Our simulation is based on Cellular IP [3, 15], one of the Micro Mobile IP protocols. Cellular IP is a routing protocol developed by Erricsson and Columbia University. We use the ns-2 simulator [16] and the simulation topology is as shown in Figure 2. The wireless supporting network consists of supervising nodes, intermediate nodes, and base stations. There exists a correspondent node at the wired network. The simulation topology is designed asymmetrically to identify the effect of the rerouting time on TCP performance. Each wired link has a 10 Mbps bandwidth and 5 ms delay, and the wireless link has a 5% packet loss. The TCP connection is set up between the correspondent node at the wired network and the mobile node at the wireless network. The mobile node moves from the left base station to the right base station repeatedly at 70 km/h. The other assumption follows that of Cellular IP.

We compare the throughput of TCP Tahoe, Reno, NewReno, and Reno with the Delayed ACK option with ACK Pacing. Figure 3 shows the sequence number increase pattern of each TCP version and ACK Pacing. As shown in Figure 3 (a), TCP Tahoe is highly improved by the ACK Pacing mechanism. TCP Tahoe has a fast retransmission mechanism for the triple duplicate ACKs, and reduces the congestion window size to zero. While TCP Reno and NewReno have a fast recovery mechanism, they reduce the congestion window size to half of the maximum window size before the losses. This mechanism is designed considering triple duplicate ACKs as a light congestion notification, and allows the transmission of a packet per each RTT. Because the hand-off period is longer than the normal packet loss, ACK Pacing can be occurred multiple times in a blackout period. After ACK Pacing, the fast retransmission mechanism prevents data



(a) TCP Tahoe (b) TCP Reno



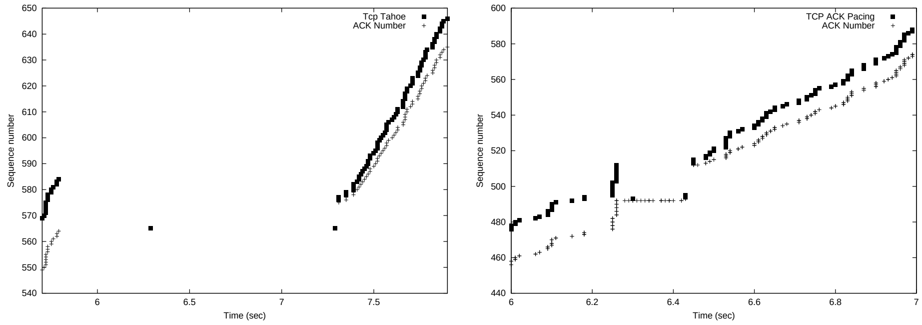
(c) TCP NewReno (d) TCP Reno with Delayed ACK option

Fig. 3. The Throughput of TCP variations and ACK Pacing

packets bursty. Therefore, TCP Tahoe with ACK Pacing has better performance than TCP Reno with ACK Pacing.

We also simulate the Delayed ACK option with TCP Reno. We select the TCP Reno because TCP Reno is currently adapted by most of wired nodes. The delayed ACK option also has the effect of delaying the ACK transmission like ACK Pacing. This effect reduces the timeout occurrence during the hand-off period compared with TCP Reno. Even though the Delayed ACK option has an effect on timeout prevention, dynamic control by ACK pacing at a hand-off time more improves the TCP performance, as shown in Figure 3 (d).

To find the effect of ACK Pacing more clearly, we select the highly improved case of TCP Tahoe. We illustrate the sequence number of data and ACK packets of TCP Tahoe and TCP Tahoe with ACK Pacing. The mobile node moves to the new station after 6.17 seconds. In Figure 4 (a), TCP Tahoe transmits data in bursts to the old base station before the hand-off and it causes 2 timeouts. This phenomenon is known as the spurious timeout at hand-off time [17]. However, TCP Tahoe with ACK Pacing doesn't induce bursty data losses and a timeout as shown in Figure 4 (b). Before the hand-off, ACK transmission is paced and it prevents data losses in bursts.



(a) TCP Tahoe

(b) TCP Tahoe with ACK Pacing

Fig. 4. The sequence number of data packets and ACK packets at hand-off time

ACK Pacing prevents a timeout occurrence and reduces the new route update time with the paced ACK delivery. ACK Pacing delays data transmission to the old path. Especially, as the congestion window size of the sender is larger, paced ACK can lead to less bulk transmission to the old path at one time and reduce bursty packet losses. The paced ACKs are delivered to the new path and they trigger the route update more quickly. In the cellular network, the periodic route update message or the data from the receiver trigger the route update. When a mobile node transmits the data to the base station, the route information is updated by the data transmission. If there is no data transmission, the route information is not updated until the periodic routing update message is exchanged. During this period, there exists a stale period of route information.

5 Conclusion

As wireless technology develops, Internet services for the mobile node will become more widely deployed. The All IP network will be the next generation global system for wired and wireless networks and this will lead to IP-based mobile communication. In the All IP network, the performance of transport layer protocols over IP should be improved for wireless services. For the future system, several TCP improvements for the wireless network have been proposed. Because of the complexity, the base station prevents the wireless TCP from being deployed to wireless systems. Therefore, the new wireless TCP should be simple and easily adaptable to wireless systems. The proposed protocol, ACK Pacing, is a wireless-node-only modification algorithm using measured pilot signal strength without any extra control overhead. In a high loss state such as a hand-off and a blackout, ACK Pacing controls the bursty loss by delaying data transmission. Also, when a mobile node moves to the new base station, the paced ACKs delay data transmission of the stale route information, and trigger the route update to the new path more quickly. Simulation results show the TCP performance improvement of the ACK Pacing algorithm.

For future work, we plan on developing a dynamic mechanism to control the pacing time based on the receiver-estimated RTT to reflect the RTT effect on TCP throughput, and we will simulate under a dynamic environment.

References

1. C. Perkins, “IP mobility support,” *IETF RFC 2002*, 1996.
2. A. G. Valko, “Cellular IP : A New Approach to Internet Host Mobility,” *ACM Computer Communication Review*, Jan. 1999.
3. T. La Porta, et al., “HAWAII: a domain-based approach for supporting mobility in wide-area wireless networks,” *IEEE/ACM Transactions on Networking*, Jun. 2002.
4. A. Bakre and B. R. Badrinath, “I-TCP: Indirect TCP for Mobile Hosts,” *In Proceedings of 15th International Conf. on Distributed Computing Systems (ICDCS)*, May 1995.
5. P. Sinha et al., “WTCP: A Reliable Transport Protocol for Wireless Wide-Area Networks,” *In Proceedings of ACM MOBICOM 1999*, Aug. 1999.
6. H. Balakrishnan, S. Seshan, E. Amir and R. H. katz., “Improving TCP/IP Performance Over Wireless Networks,” *In Proceedings of ACM MOBICOM ’95*, 1995.
7. H. Balakrishnan and R. H. Katz, “Explicit Loss Notification and Wireless Web Performance,” *In Proceedings of IEEE GLOBECOM ’98*, 1998.
8. K. Brown and S. Singh, “M-TCP: TCP for Mobile Cellular Networks,” *In Proceedings of INFOCOM ’96*, 1996.
9. S. Mascolo and C. Casetti, “TCP Westwood: Bandwidth Estimation for Enhanced Transport over Wireless Links,” *In Proceedings of MOBICOM ’2001*, 2001.
10. S. Hadjiefthymiades, S. Papayiannis, and L. Merakos, “Using Path Prediction to Improve TCP Performance in Wireless/Mobile Communications,” *IEEE Communication Magazine*, Aug. 2002.
11. M. C. Chan and R. Ramjee, “TCP/IP Performance over 3G Wireless Links with Rate and Delay variation,” *In Proceedings of MOBICOM ’2002*, 2002.
12. T. Goff, J. Moronski, and D. Phatak, “Freeze-TCP: A True End-to-End Enhancement Mechanism for Mobile Environments,” *In Proceedings of INFOCOM ’2000*, 2000.
13. V. Tsaoussidis and H. Badr, “TCP-Probing: Towards an Error Control Schema with Energy and Throughput Performance Gains,” *In Proceedings of the 8th IEEE ICNP ’2000*, 2000.
14. M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, “TCP Selective Acknowledgment Options,” *IETF RFC 2018*, Oct. 1996.
15. Cellular IP, *http://www.comet.columbia.edu/micromobility*, 2001.
16. ns-2 Network Simulator, *http://www.isi.edu/nsnam/ns*, 2003.
17. A. Gurtov and R. Ludwig, “Responding to Spurious timeouts in TCP,” *In Proceedings of INFOCOM ’2003*, 2003.