Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 524960, 12 pages
doi:10.1155/2009/524960

Research Article

A New Scheme to Improve Overall TCP Throughput with Vertical Handover between 3G Cellular Packet Networks and Wireless LANs

Chang Hyeon Lim[1] and Ju Wook Jang[2]

[1] *DMC R&D Center, Samsung Electronics Co., LTD., Dong Suwon P.O. Box 105, Gyeonggi-do 443-742, South Korea*
[2] *Department of Electronic Engineering, Sogang University, Seoul 121-742, South Korea*

Correspondence should be addressed to Ju Wook Jang, jjang@sogang.ac.kr

Received 20 December 2008; Revised 28 May 2009; Accepted 5 August 2009

Recommended by Ingrid Moerman

One major difference between 3G cellular networks and Wireless LANs is in how packet losses are dealt with. 3G cellular networks employ selective repeat ARQ (SR-ARQ) which may result in packet-reordering due to local retransmissions. Wireless LANs employ stop-and-wait ARQ which may drop some packets in case the buffer overflows. While much work has been performed for TCP throughput enhancement with vertical handover between wired networks and wireless networks, little attention is paid to the vertical handover between two different wireless networks. It is expected for majority of new mobile devices to be able to connect to both 3G cellular networks and wireless LANs by being equipped with multiple wireless interfaces. If this is the case, on-going TCP transmission may move between 3G cellular networks and wireless LANs. In this paper, we propose a scheme to improve TCP throughput with vertical handoff between 3G cellular networks and wireless LANs. For this we develop an algorithm which differentiates packet-reordering from packet losses and congestion losses from wireless losses. Simulation of our algorithm using ns-2 shows significant increase in throughput as large as 80% compared to TCP-SACK. The proposed scheme is able to respond to the change of loss pattern and then improve the TCP throughput during handover.

1. Introduction

Vertical handover [1] refers to the process of changing interfaces for a mobile node with multiple network interfaces to move from one network to another network. Recent mobile devices with two or more interfaces are able to connect to different types of networks. Previous researches on vertical handover have addressed the handling of the difference in bandwidth-delay products between two different types of networks (e.g., 3G cellular networks and WLANs) [2–4]. Known TCP enhancement schemes over wireless packet losses [5–8] have paid little attention to vertical handover cases. Differentiation of congestion losses from wireless losses over a network which consists of a wired part and a wireless part is performed in [5, 6], and the detection of packet-reordering due to the retransmission at the link layer on a wireless last hop is dealt with in [8].

In this paper, we consider a new problem that an on-going TCP flow on a mobile node may experience packet losses of different natures over vertical handover depending on link layer retransmission schemes involved. For example, the packet can be lost due to wireless error over WLAN or due to congestion loss over the wired part. Or packets may be regarded as loss due to reordering which happens over 3G cellular networks due to its recovery mechanism. To obtain best TCP throughput over vertical handover, we need to accurately identify these three different causes of packet loss, congestion loss, wireless loss, and packet-reordering.

To address this problem, we propose a new TCP congestion control scheme which adapts to the link layer transmission schemes involved in vertical handover. Specifically, we deal with handover between 3G and WLAN which, we expect, would be most prevailing in the near future mobile environment. The proposed scheme is a modification of loss differentiation (LD) scheme [7] to improve its performance over handover between 3G and WLAN. The proposed scheme deals with wireless/wireless handover problem, while the original LD differentiates wireless loss from congestion loss and thus with steady-state wired/wireless problem. If we employ the LD in its original form for this new problem, the LD does not function as expected. The key parameters of LD have been tuned over time to the network and do not fit new network immediately after the handover. TCP-DCR (Delayed Congestion Response) [8] works well over 3G cellular networks, but not for handover. F-RTO (Forward Retransmission Timeout) [9] can be used over handover. However, the performance of proposed scheme excels the F-RTO. The proposed scheme significantly improves any of the algorithms mentioned above for handover from WLAN to 3G cellular packet network or vice versa. Figures 11, 12, 13, and 14 show the simulation results before handover and after handover for various scenarios. For example, our scheme (proposed scheme II) excels DCR by 80% in throughput averaged for 250 seconds after handover. Furthermore, our scheme outperforms the simple combination of LD and DCR (proposed scheme I) by about 60% in throughput averaged for 250 seconds after handover.

We summarize our contribution as follows. Proposed schemes are not mere combinations of known algorithms (LD, TCP-DCR and/or FRTO). First, we solve a new problem. We try to solve the problem of TCP throughput degradation over vertical handover between 3G cellular packet networks and Wireless LANs, while the original LD [7] is developed for TCP improvement for wired/wireless paths. In other words, our scheme deals with wireless/wireless handover problem, while the original LD deals with steady-state wired/wireless problem.

Second, we do not combine three algorithms in their original forms. For example, if we employ the LD in its original form for this new problem, the LD does not function as expected. On handover, the key parameters of LD (RTTmin, RTTavg, RTTdev, etc.) might not closely reflect the current condition of the new wireless network for some time. The reason is that the parameters have been tuned to the old wireless network before handover and therefore are not supposed to be used for the new wireless network. We modified the LD to deal with the problem of LD over handover. Similarly, TCP-DCR [8] works well over 3G cellular networks, but not for handover. F-RTO [9] can be used over handover. However, our scheme overlaps with F-RTO only in the detection of handover. We may avoid using F-RTO by using one of known cross-layer approaches [9] for detection of handover. We choose to incorporate the handover detection part from the F-RTO since we want to retain layer independency (so that our scheme may work independently from link layer).

Third, our scheme significantly improves any of the algorithms mentioned above for handover from WLAN to 3G cellular packet network or vice versa. Figures 11 through 14 show the simulation results before handover and after handover for various scenarios.

The rest of this paper is organized as follows. We define the problem of TCP throughput enhancement over vertical handover between 3G networks and WLAN in Section 2. Related works are summarized in Section 3. Proposed schemes are presented in Section 4. The performance evaluation of our schemes is provided in Section 5. Section 6 concludes the paper.

2. Problem Definition

The retransmission technology at a wireless link layer can hide the packet loss from the upper layer, for example, TCP. Despite this retransmission, packet losses are still exposed to the TCP layer [10, 11] as shown in Figure 1. The packet losses can reduce the overall TCP throughput no matter where they come from. We consider two types of network: IEEE 802.11-based wireless LAN and 3G packet network.

IEEE 802.11-based wireless LAN adopts the Stop-and-Wait ARQ [12] into their link retransmission technique. In this scheme, when a packet is lost, the sender (access point) retransmits the lost packet several times to a mobile node. The packet loss is exposed to the TCP layer when the number of retransmission trials exceeds the constraint [11]. Then, the TCP sender corresponding to the mobile node halves its congestion window. This leads to reduced TCP throughput.

3G packet network can hide most of packet losses to TCP layer due to selective-repeat ARQ [13, 14]. However, this retransmission technology may generate out-of-order packets. On packet-reordering, the TCP sender retransmits the lost packet and reduces its sending window even though the receiver receives all of packets without losses. Some packets are still lost even with packet loss rate of 1% due to three reasons: queue overflow at link layer, timeout, and the number of transmitting packets more than the size of the retransmission queue [10].

A mobile node with two or more different types of the link layer interface (e.g., wireless LAN and 3G packet network) needs to be able to handle the packet loss and packet-reordering problem. Many known schemes on the packet losses over wireless links mainly deal with classification of the causes of packet losses [6, 7]. They have improved TCP throughput over wired and wireless network. Packet-reordering due to the advanced link layer retransmission is considered as a loss by TCP [8]. Generally, end-to-end loss differentiation schemes could keep their TCP congestion window against packet-reordering. However, they may redundantly retransmit the packet due to packet-reordering.

Our assumption on the packet loss based on literatures is as follows:

(1) packet loss model over WLAN: Gilbert 2-state Markov chain [15],

(2) packet loss model over 3G: uniformly random loss,

(3) packet-reordering over 3G:

   (a) uniformly random event,

   (b) swapping a packet for the next packet [8, 16].

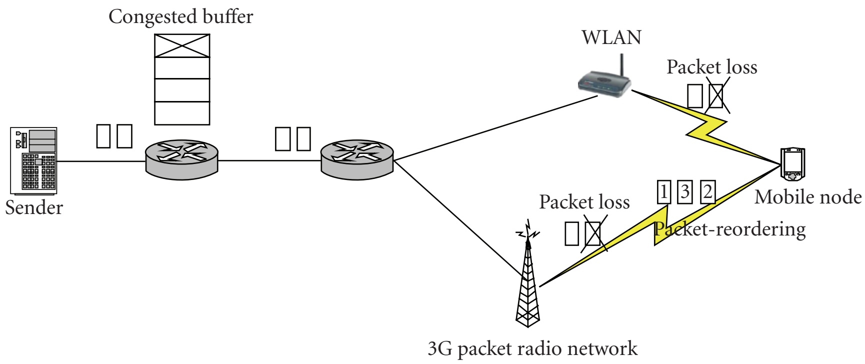


Figure 1: Three packet loss scenarios: (1) congestion losses over the wired network, (2) packet loss over WLAN, and (3) packet loss and packet reordering over 3G packet network.

3. Related Work

The delayed congestion response scheme (TCP-DCR) [8] distinguishes packet-reordering from congestion losses over 3G networks. This scheme waits for the normal ACK to handle packet-reordering to reduce unnecessary retransmission on three duplicate ACKs generated due to reordered packets. However this scheme cannot distinguish wireless losses from congestion losses. Actually, packet losses as well as packet-reordering happen on 3G network.

The loss differentiation schemes differentiate congestion losses from wireless losses [6, 7]. These schemes improve TCP throughput over wireless losses by classifying the cause of packet losses. However, they also classify packet-reordering into packet losses, and therefore retransmit unnecessary packets.

IEEE 802.21 framework [17] deals with vertical handover: a mobile node discovers a new access link, registers itself to the new network, disconnects itself from the old access link, and communicates with corresponding host over the new access link. However, IEEE 802.21 framework does not include the transport layer to support vertical handover. TCP is still exposed to losing packets due to difference among access links and several sources of packet losses (e.g., congestion, wireless errors, and packet-reordering).

Recent research on TCP modifications for vertical handover has focused on compensating for the difference of link delay or bandwidth-delay product between two different wireless links [2–5]. These problems arise when a mobile node moves to the network with a link layer technology different from the current network. If a mobile node moves to the network with a smaller bandwidth-delay product, the mobile node may lose packets since its current congestion window is larger than the new bandwidth-delay product.

In [2], a mobile node estimates the bandwidth-delay products (BDP) of the both paths before and after the handover. After performing the vertical handover procedure, the mobile node then returns the ACK at the rate according to the bandwidth of the new wireless access link. When the BDP becomes smaller, the TCP receiver adjusts congestion window (cwnd) by sending the triple duplicate acknowledgements (ACKs). When the BDP becomes larger, the receiver returns many ACKs to increase the cwnd rapidly. This compensation needs supports of link layer or of the access node of wireless network. In [3], network layer controls the route of ACKs during vertical handover. In particular, for handover from a fast link to a slow link, the ACKs of the first few slow-link packets are routed to the old fast link to reduce the gap between old RTT and new RTT. For handover from a slow link to a fast link, similar procedure is performed. In [4], the TCP receiver sets the *awnd* (acknowledged window) to the minimum value of the receiver window size and the estimation of the maximum BDP over GPRS network at the receiver-side, and reports it to the sender. The maximum BDP is calculated as follows: the TCP receiver counts the amount of TCP data received from the instant; a timestamp is put on an acknowledgement (ACK) to the instant; the TCP packet carrying the echoed timestamp has been received [4]. This scheme follows end-to-end semantics and can be easily implemented at mobile node.

There has been a TCP solution to facilitate smooth handover with the initialization of the TCP sender between two different types of networks [5]. Their TCP solution is for a TCP receiver to trigger the vertical handover event, and for the corresponding TCP sender to initialize the congestion window, slow-start threshold, and retransmission timer on the vertical handover event. This scheme can help a TCP flow adjust its congestion window to the changed BDP.

These previous researches have focused on only the difference of bandwidth-delay product (BDP) between two different packet networks when the vertical handover occurs. The shrinking of BDP increases the number of dropped packets in proportion to the difference between the size of TCP's congestion window and the size of the new BDP. We know that the BDP can be changed when a mobile node connects to a new network with a link layer different from the previous network. This problem should be addressed.

However, the focus of our paper is not on the adaptation of TCP for the change in the BDPs of the links but on the adaptation of TCP for the change in the link-level retransmission schemes on handover. Moreover, we are not alone in assuming the same fixed delay in wired part of both paths in simulation setup for TCP adaptation over vertical handoff [18]. We assume that the change of BDP over the vertical handover does not happen, or that, at least, the difference of the BDPs has been compensated. According to [19], the BDP of 3G cellular network and WLAN could be similar as 15 Kbytes and 14 Kbytes, respectively.

4. Proposed TCP Modification

In this section, two new TCP congestion control schemes to solve TCP throughput degradation over vertical handover are proposed. Proposed scheme I (also called DCR-LD) is basically a combination of our loss differentiation (LD) scheme [7] and the TCP-DCR [8]. Proposed scheme II (also called DCR-FRTO-LD) is a modification of LD which borrows handover detection algorithm from the Forward-Retransmission Timeout (F-RTO) [9] and incorporates the TCP-DCR scheme for improved throughput over 3G.

Proposed scheme I improves TCP throughput over wireless packet losses and packet-reordering. On receiving three duplicate ACKs, the DCR part of the proposed scheme I differentiates real packet losses from packet-reordering. On duplicate ACK, it waits for a new ACK for the duration one estimated RTT. If a TCP sender receives a new ACK, the duplicate ACK is considered due to packet-reordering, and the TCP sender keeps its current window size. Otherwise, the loss differentiation scheme is invoked in an estimated RTT to identify the cause of the loss. If the measured RTT just before the duplicate ACK is lower than the threshold, the loss is classified as wireless loss, and then the congestion window is allowed to retain the current value. Otherwise, the TCP sender regards the loss as congestion loss and halves the congestion window. Consequently, proposed scheme I is able to identify the three types of packet losses, packet-reordering, wireless loss, and congestion loss.

F-RTO [9] is for the TCP sender to keep the congestion window with detection of delay spike due to vertical handover and to reset the measured RTT and its deviation to initial values. When a timeout event happens, F-RTO algorithm sends two probe packets of new sequence numbers. If a normal ACK is received, F-RTO assumes that the vertical handover has occurred, and the TCP timeout retransmission procedure is cancelled. Otherwise, the retransmission procedure is invoked. We may avoid using F-RTO in our proposed scheme II by using one of known cross-layer approaches [18] for detection of handover.

In proposed scheme II, the measured RTT, its deviation, and the minimum RTT are reset to new measured values to prepare LD into the new wireless environment. The current congestion window is retained. Furthermore, LD is not invoked for several rounds until the parameters needed for LD are tuned to the new wireless link.

For completeness of presentation, we present our loss differentiation scheme [7] and its proposed modification for handover in Section 4.1. The modified scheme is used in proposed schemes I and II.

4.1. Loss Differentiation Scheme and Its Modification for Handover. In Section 4.1.1, we briefly review our loss differentiation scheme [7] which is originally proposed to infer the cause of each packet loss as congestion loss over wired part or wireless loss for wired/wireless networks. Then the modification of LD for vertical handover is presented in Section 4.1.2.

4.1.1. Loss Differentiation Scheme. Let T_{cur} denote the RTT measured immediately before the current packet loss. It is used in (1) as an indicator. The RTT consists of the propagation delay and the queueing delay [10]. Let T_{avg} and T_{dev} denote an exponentially weighted moving average of RTT and the deviation, respectively. The initial values of T_{avg} and T_{dev} are set to 0 and 0.5 s, respectively. They are updated by $T_{\text{avg}} = (7/8)T_{\text{avg}} + (1/8)T_{\text{cur}}$ and $T_{\text{dev}} = (3/4)T_{\text{dev}} + (1/4)|T_{\text{avg}} - T_{\text{cur}}|$. Let $T_{\min}$ denote the propagation delay which is reset by the expiration of the retransmission timer. The current packet loss is determined to be a congestion loss if (1) is satisfied:

$$\frac{T_{\text{cur}} - T_{\text{avg}}}{T_{\text{dev}}} > 2\left(\frac{T_{\min}}{T_{\text{cur}}}\right)^{k} - 1. \quad (1)$$

We assume that the connection between the source and the destination is on a fixed path throughout the connection's lifetime.

The LD [7] accurately classifies congestion and random losses over varying network parameters, that is, buffer size, wireless random loss rate, and so forth. This scheme can reliably classifies the congestion loss with an accuracy close to 100% even with a small buffer of size the 20% of the bandwidth-delay product and has less dependency on network parameters compared with other schemes in literature.

4.1.2. Proposed Modification of Loss Differentiation Scheme for Handover. In the original LD [7], we have assumed that the connection between the source and the destination is on a fixed path throughout the connection's lifetime. However, on vertical handover, a mobile node may change its last wireless hop to a different link layer. This implies that the characteristics of the path, that is, propagation delay and the cause of packet losses, may change. We modify the LD in such a way that, on vertical handover, the parameters for the loss differentiation are reset for LD to accurately infer the cause of packet losses on the changed wireless link.

We show via simulations the efficiency of the modified loss differentiation scheme when the vertical handover occurs. The parameters setup and the environment are the same as Table 1 in Section 5. Since the loss differentiation depends on the measured RTT and its statistical values, it is important how soon statistical values are to become stable after handover. Equation (2) is derived from (1):

$$T_{\text{cur}} > T_{\text{avg}} + T_{\text{dev}}\left(2\left(\frac{T_{\min}}{T_{\text{cur}}}\right)^{k} - 1\right). \quad (2)$$

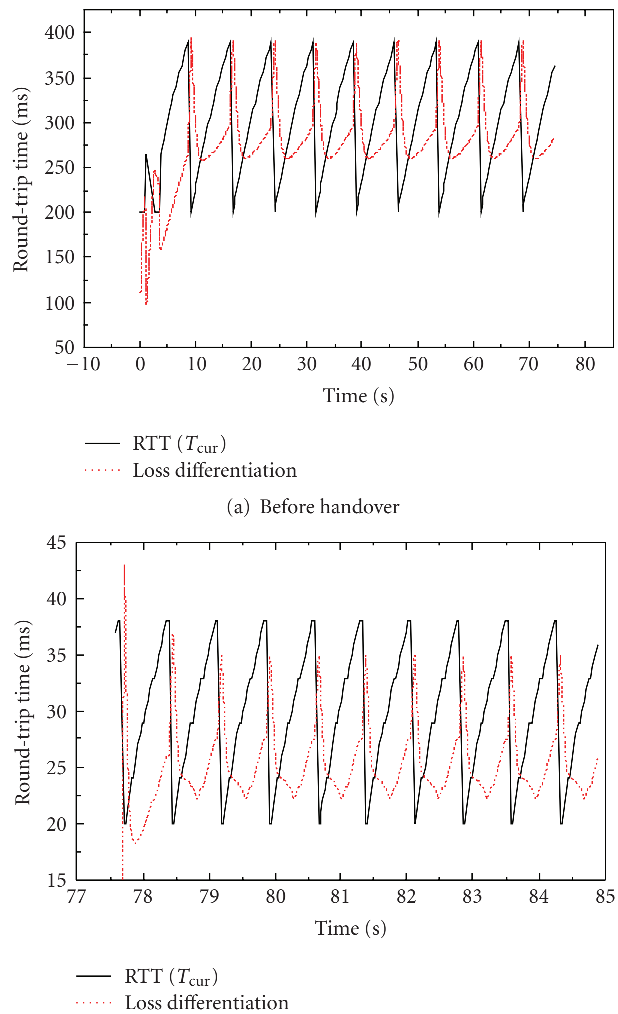
The left-hand side of (2) is the measured RTT. The right-hand side of (2) is the threshold of loss differentiation. Each term in (2) is updated on receipt of an acknowledgement packet. A packet loss is regarded as a congestion loss if the measured RTT(T_{cur}) is greater than the threshold, otherwise, a wireless loss.

We modifies the retransmission timeout procedure in TCP to set statistical values, for example, T_{avg} and T_{dev}, to the initial values on vertical handover, and to set the propagation delay T_{min} to a large value. We monitor the measured RTT and the threshold. The purpose of this monitoring is to determine if the modified loss differentiation scheme accurately identifies congestion loss even when the statistical values are yet to be stabilized after handover.

In our simulation scenario, it is assumed that the spurious timeout during vertical handover is set to be equal to 3 seconds and that the bandwidth-delay product of 3G network is the same as WLAN (the reasoning behind this assumption is provided by the last paragraph of Section 3). The timeout event of handover is invoked by pausing the simulation for several seconds. To simulate 3G links, bandwidth and delay are set to 1 Mbps and 200 milliseconds, respectively. To simulate WLAN, bandwidth and delay are set to 10 Mbps and 20 milliseconds. There is no cross traffic and no wireless packet loss.

Figure 2 shows the vertical handover from 3G network to WLAN. The loss differentiation scheme correctly indicates noncongestion when T_{cur} is close to T_{min} and congestion when T_{cur} is over the threshold in (2). Actually, this case is to simulate that the mobile moves from a slow link to a fast link. Thus, T_{min} adapts into a lower value over handover to the faster link.

The case of the vertical handover from WLAN to 3G network needs requires that the minimum RTT T_{min} is set to a great initial value on handover. The original loss differentiation scheme [7] may wrongly classify the cause of packet losses on handover due to a large difference of minimum RTT between a fast link and a slow link. As shown in Figure 3, the modified loss differentiation scheme works well in the case of the vertical handover from WLAN to 3G.



(a) Before handover

(b) After handover

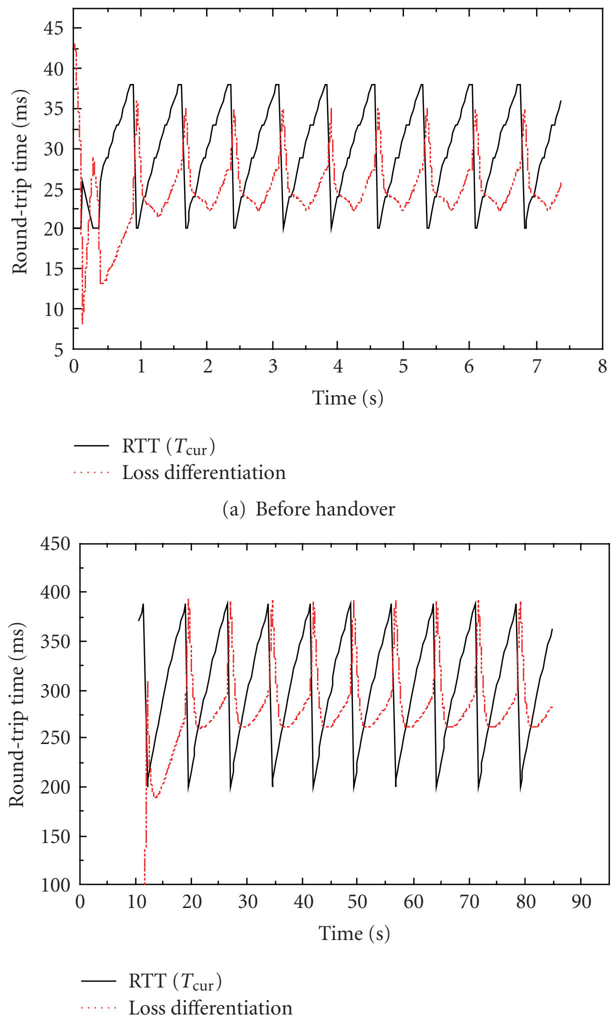
Figure 2: Loss differentiation on vertical handover from 3G network to WLAN.

4.2. Proposed Schemes. The proposed scheme I is a combination of LD [7] and TCP-DCR [8]. The proposed scheme II is a modified loss differentiation scheme (introduced in Section 4.1.2) which borrows handover detection scheme from F-RTO [9] and incorporates TCP-DCR [8] for performance improvement for 3G.

Figure 4 shows the state transition diagram of the proposed Scheme II. The diagram includes states for reset, slow-start, fast recovery, congestion avoidance, F-RTO Transmit Two Packets (handover detector borrowed from F-RTO), DCR (distinguishes between packet-reordering and packet losses via delayed retransmission), and the modified loss differentiation. The congestion window starts from the reset state and then grows in slow-start state.

The event of three-duplicate ACKs leads to fast recovery of halving the congestion window, and then the state moves to congestion avoidance state. Timeout leads to F-RTO state which is to avoid delay spike due to vertical handover. F-RTO [9] retransmits the packet that triggered timeout on entering retransmission timeout and waits for the ACK. The sender's receipt of two new ACKs aloows the sender to transmit new packets continuously in congestion avoidance state. However, on receiving one or two duplicate ACKs, the state goes to reset. Whenever the state goes to F-RTO, T_{min} is set to a large value and is updated per every RTT. T_{avg} is set to a new measured RTT. T_{dev} is set to a half of the new measured RTT. This reset operation allows the modified loss differentiation scheme to adapting itself to the change in propagation delay due to vertical handover to new wireless network. Furthermore, the modified loss differentiation is not invoked by three-duplicate ACKs for n rounds. This pause prevents the congestion window from reacting to congestion losses or the number of dropped packets with a set of parameters which have become invalid on vertical handover. The parameters include T_{min}, T_{avg}, and T_{dev}. In our simulation, we set n to be equal to 1. In our simulation, the values of the parameters measured for the first time after the handover are used to correctly infer the cause of packet losses.

From slow-start state, the state changes into congestion avoidance which grows the congestion window by 1 when the



(b) After handover

Figure 3: Loss differentiation on vertical handover from WLAN to 3G network.

size of congestion window goes higher than the threshold. Three-duplicate ACKs involve the modified LD and DCR. When the third duplicate ACK is received, the LD determines the cause of a packet loss, and the sender with DCR waits for a normal ACK for the duration of one RTT. The waiting duration can be adjusted according to network condition, for example, wireless loss of retransmitted packets and we use the duration used in TCP-DCR [8]. If the sender receives a normal ACK, the state stays in congestion avoidance. Otherwise, the state moves to the fast recovery if the modified LD identifies congestion loss or stays in the congestion avoidance otherwise (wireless loss). In congestion avoidance, the timeout involves F-RTO as described above.

In summary, this proposed scheme II uses the handover detection mechanism from F-RTO to reset the parameters to be used by the modified LD and incorporates the TCP-DCR scheme to improve throughput over 3G network. The LD scheme uses the parameters such as T_{cur}, T_{avg}, and T_{dev} which already exist in the original TCP. T_{min} is the minimum value of T_{cur}. The DCR uses only T_{avg} to wait for recovery of lost packet at link layer. F-RTO scheme does not use any information. The proposed scheme II needs no extra information other than available in a TCP sender. Thus the proposed scheme II can be implemented easily at the sender-side without involving any change in intermediate routers or receivers.

4.3. Behavior on Packet Loss. The behaviors of proposed scheme II are shown in Figures 5, 6, and 7. On receipt of the third duplicate ACK, the scheme determines whether the network is congested or not. A flag indicating congestion is set to true if inequality (1) is satisfied, or set to false, otherwise. We recall the equation as follows:

$$\frac{T_{cur} - T_{avg}}{T_{dev}} > 2\left(\frac{T_{min}}{T_{cur}}\right)^k - 1. \tag{3}$$

The scheme then sets the timer to one RTT to wait for link level retransmission of the packets requested in the 3rd duplicate ACK and continues transmitting new packets without reducing the congestion window. When a new ACK is received within one RTT, the sender cancels the timer as shown in Figure 5. This part overlaps with the scheme in TCP-DCR [8].

However, the retransmission procedure is involved when duplicate ACKs are still received after one RTT, and then the decision of retaining or halving the current window is made according to the congestion flag. If it is false, the TCP sender retains the current congestion window as shown in Figure 6. If it is true, the TCP sender halves the congestion window as shown in Figure 7.

The DCR part can handle certain patterns of packet-reordering, but not all possible patterns. For example, packet-reordering of 1-3-2 can be handled since it is a swapping of a packet with the next packet. DCR waits for one RTT after receiving duplicate ACKs and thus can handle such packet-reordering. However, DCR scheme cannot handle the packet-reordering of 1-5-4-2-3 because waiting for one RTT might not be enough for this case.

4.4. Behaviors on Timeout Event. On the timeout event, detailed behaviors are shown in Figures 8 and 9 for a vertical handover and a normal timeout, respectively.

F-RTO scheme is involved to send two probe packets and to wait for ACKs. If a new ACK arrives, the sender regards the timeout event as due to vertical handover (see Figure 8) and resets RTT values (T_{min}, T_{avg}, and T_{dev}) to the initial values. However, if duplicate ACKs arrive, the sender retransmits the lost packets immediately with setting congestion window to 1 (see Figure 9).

5. Performance Evaluation

5.1. Simulation Setup. In this section, we evaluate the TCP throughput performance of TCP-Sack, DCR, DCR-FRTO, proposed scheme I and proposed scheme II through the ns-2 simulations (Version 2.27) [20]. The proposed scheme I is

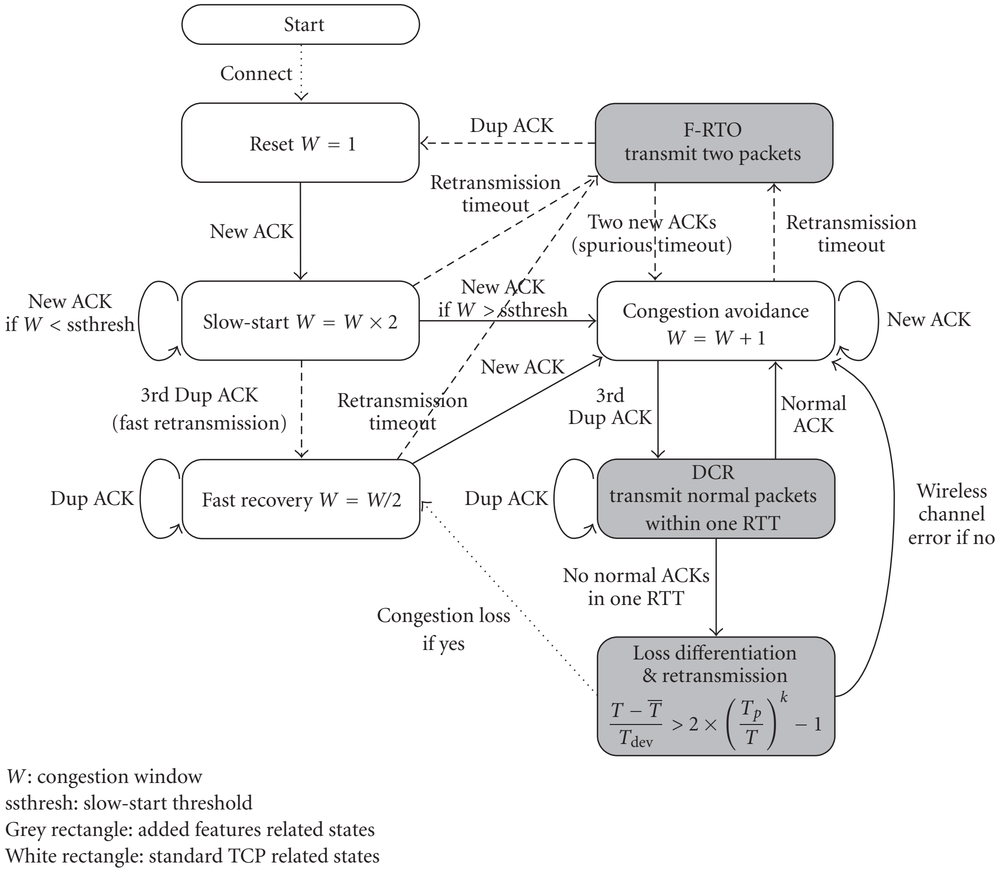


Figure 4: Proposed TCP modification in response of vertical handover. After returning congestion avoidance from F-RTO, loss differentiation should not be involved by 3rd Dup ACK for n rounds.

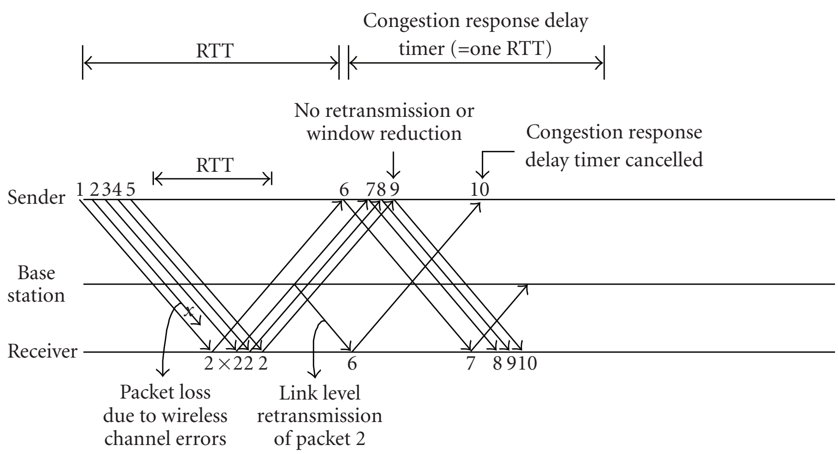


Figure 5: Behavior of the proposed scheme II: packet-reordering due to retransmission at wireless link layer.

a simple combination of TCP-DCR and LD. The proposed scheme II is introduced in Section 4.2.

The network model is shown in Figure 10. Two scenarios of vertical handover are considered. The first scenario is 3G-to-WLAN handover which denotes handover from 3G to WLAN, and the second one is (W-to-3G) which denotes handover from WLAN to 3G. For each scenario, we also consider two different situations which are noncongestion and congestion. In noncongestion scenarios, the proposed schemes are able to improve TCP throughput on both wireless losses and packet-reordering compared to TCP-SACK and TCP-DCR. Congestion scenarios are more realistic cases with forward cross traffic causing congestion losses over wired paths while wireless packet losses and

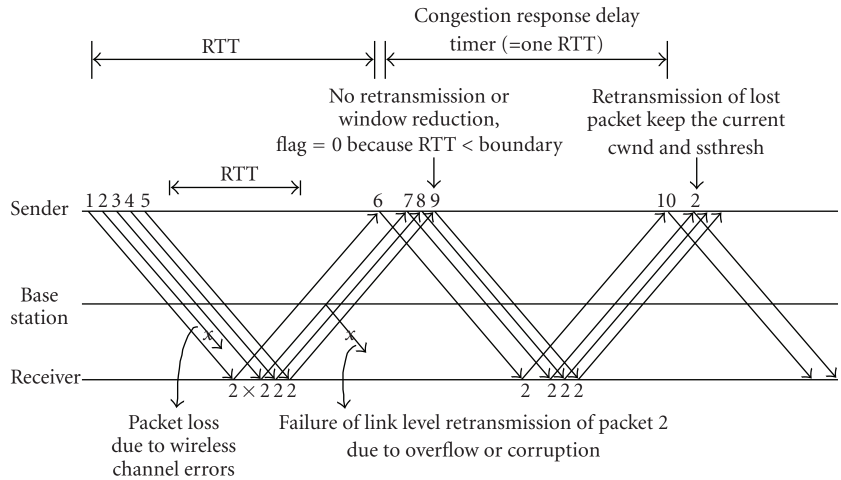


Figure 6: Behavior of the proposed scheme II: a wireless loss.

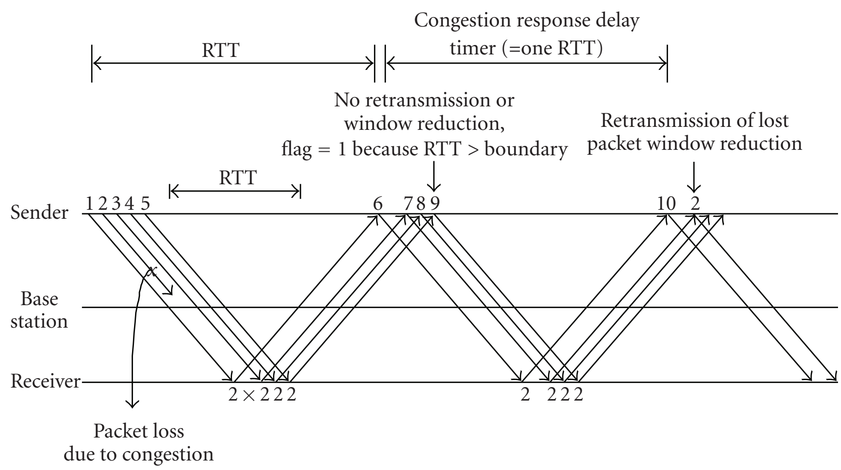


Figure 7: Behavior of the proposed scheme II: a congestion loss.

packet-reordering still happen on wireless paths. The reason for considering congestion scenarios is to test the ability of the proposed schemes to differentiate the congestion loss on handover where parameters experience rapid change. In these congestion scenarios, we mainly focus on the throughput improvement by the proposed scheme II on congestion losses compared to the proposed scheme I as well as known schemes in literature.

In the simulation, we employ a pause event for 30 seconds as a vertical handover from a network to another network. The change in network parameters follows the pause event. The simulation time is equal to 500 seconds. Simulations run 30 times. 95% confidence interval is shown along with the average TCP throughput.

In the simulations, we use the Gilbert two-state Markov chain to model the packet loss process over WLAN [15]. In particular, we assume that $p_{w,\mathrm{WLAN}}$ is set equal to 0.1% when the channel is in the Good state and is set equal to 10% when the channel is in the Bad state, deterministically. The duration in the Good state is assumed to be equal to 1 second whereas the duration in the Bad state is assumed to be equal to 100 milliseconds. After the duration, the state transits to the Good state or the Bad state with a probability $P = .5$. The correlated loss model is generally used for modeling the packet loss model of IEEE 802.11 wireless LAN [15].

For the packet loss model over 3G link, we set the packet reordering rate to be as high as 5% and the packet loss rate to be as low as 0.5% with uniform distribution. The packet-reordering can be simply produced by swapping a packet for the next packet. We follow the simulation scenario in TCP-DCR [8] assuming that the underlying mechanism is a simple link-level retransmission scheme, possibly NACK-based, that does not attempt in-order delivery [8, 16].

For the congestion scenario, we inject the forward cross traffic consisting of several TCP connections to cause 1.5% congestion losses rate. This cross traffic also increases the end-to-end delay over the path due to queuing delay. We summarize the configurations in Table 1.

The TCP throughputs are represented as percentages of the maximum bandwidth of the corresponding networks. If a mobile node performs handover from 3G cellular network of 1 Mbps to WLAN of 10 Mbps as in Table 1, 50% TCP

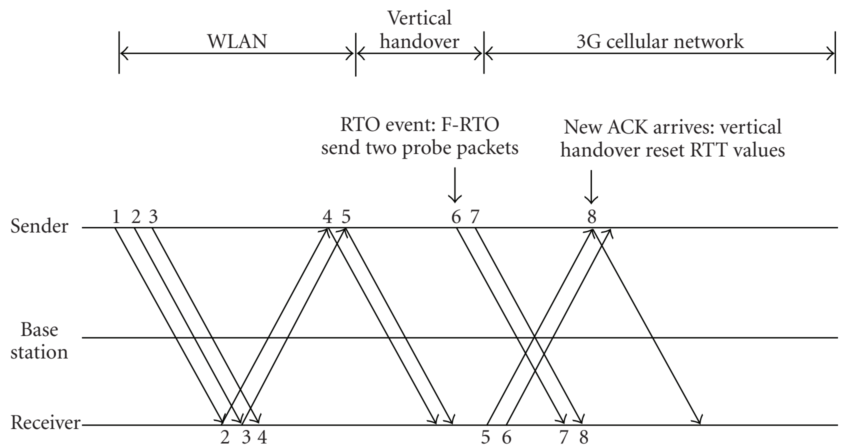


Figure 8: Behavior on vertical handover.

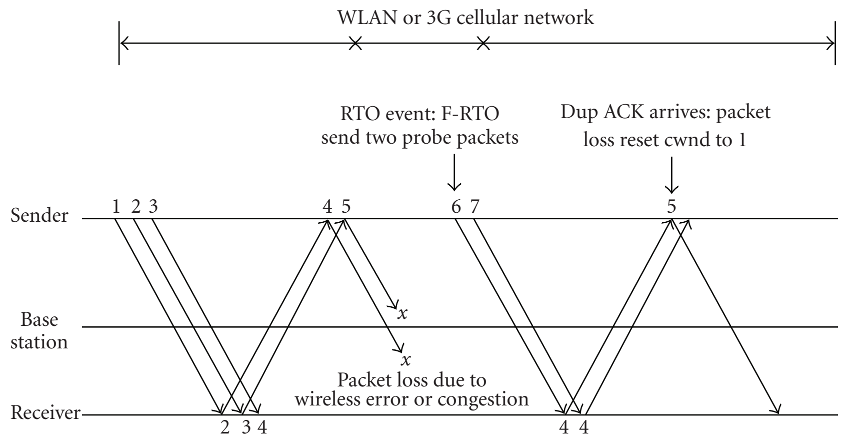


Figure 9: Behavior on normal timeout.

throughput before handover is translated to 0.5 Mbps while 40% TCP throughput after handover is translated to 4 Mbps. Please refer to Figures 11–14.

5.2. Throughput Improvement with Vertical Handover under Noncongestion. We evaluate the throughput of a single TCP connection during lifetime including one single handover event between WLAN and cellular network. The throughput of the TCP connection can be affected by wireless packet loss over WLAN and cellular network, packet-reordering over cellular network, and delay spike due to handover process.

Figure 11 shows TCP throughputs for proposed schemes (I and II) against known schemes when a mobile node performs the handover from cellular network to WLAN. Before the handover, the proposed schemes I and II improve the throughput by 26% compared with DCR. This result indicates that the modified loss differentiator improves TCP throughput of the DCR scheme. After the handover to WLAN, the packet-reordering does not happen any more. The only possible loss pattern is wireless loss (noncongestion). The proposed schemes improve the throughput by 60% compared with other schemes (Sack, DCR, and DCR-FRTO).

Figure 12 shows TCP throughputs for proposed schemes (I and II) against known schemes when a mobile node performs the handover from WLAN to cellular network. Before the handover, proposed schemes I and II improve the throughput over WLAN by about 85% compared with Sack, DCR, and DCR-FRTO. The difference between proposed schemes I and II is negligible. Since there are only wireless losses over WLAN (noncongestion), throughputs for the other TCP schemes without loss differentiator are limited. After the handover, the loss pattern changes to coexistence of packet losses and/or packet-reordering. Proposed schemes I and II can achieve the throughput as high as up to 60% of the link bandwidth, and they improve the throughput by about 87% compared with DCR and DCR-FRTO. DCR scheme improves the throughput by 100% compared with Sack, but DCR scheme cannot help halving its congestion window on packet losses in 3G communication system [10].

F-RTO facilitates the TCP sender's detection of the spurious timeout inferred from delay spike without packet losses and prevents the TCP connection from entering the timeout process. When the mobile node moves from WLAN to 3G cellular network, DCR-FRTO and proposed scheme

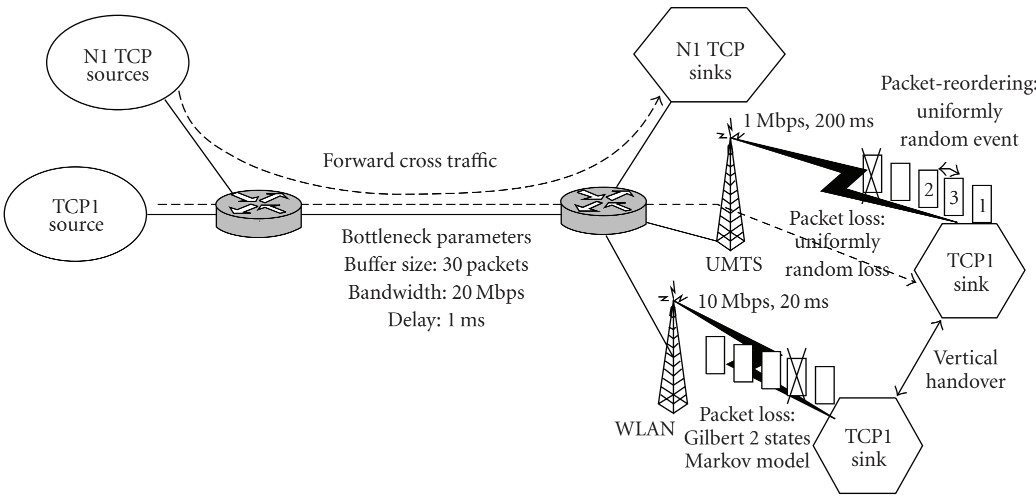


Figure 10: Network model for vertical handover.

Table 1: Simulation Configurations.

Term	3G cellular	WLAN
Packet loss model	Uniform random loss model with rate 0.5%	Gilbert Markov 2-state model with good state 0.1%, bad state 10%, and transition rate 0.5
Bandwidth	1 Mbps	10 Mbps
Delay	200 ms	20 ms
Packet-reordering model	Uniform and random packet-reordering probability of 5%. Swap a packet for the next packet. (5% is chosen out of the channel error rate range from 0% to 8% in [8, Figure 5, page 523])	—
Wired part		
Congestion loss model		Forward cross traffic (10 connections) with congestion loss rate 1.5%
Bandwidth		20 Mbps
Delay		1 ms
Router buffer		30 packets
Vertical Handover		
Vertical handover Event		At 250 s, simulation is paused during 30 seconds

II slightly improves the throughput by 1.2% and 1.8% compared with DCR and proposed scheme I, respectively. However, when the mobile node moves from cellular network to WLAN, the improvement can be negligible. Without F-RTO, the congestion window is shrunk to 1 due to RTO on a vertical handover. But, the high bandwidth of the new fast network can help the congestion window to be recovered to the previous value in a short time. Thus, F-RTO can affect the improvement in throughput only for the case of vertical handover from WLAN to cellular network.

5.3. Throughput Improvement with Vertical Handover under Congestion. In this scenario, we assume that the wired path is congested with congestion packet loss rate of 1.5% due to the forward cross traffic. The proposed schemes improve the throughput by about 44% compared with DCR over cellular network before handover. However, after handover to WLAN, both of proposed schemes I and II improve the throughput by 480% and 457% compared with DCR and DCR-FRTO, respectively. The reason for this great difference is that unlike proposed schemes, the throughputs of previous schemes are suppressed by congestion after handover from a slow link to a fast link.

Figure 13 shows TCP throughputs for proposed schemes (I and II) against known schemes when a mobile node performs the handover from cellular network to WLAN with congestion in the wired part while Figure 14 shows TCP throughputs when a mobile node performs the handover from WLAN to cellular network. After handover to WLAN, throughput is expected to increase due to larger bandwidth over WLAN. Proposed scheme II can increase the

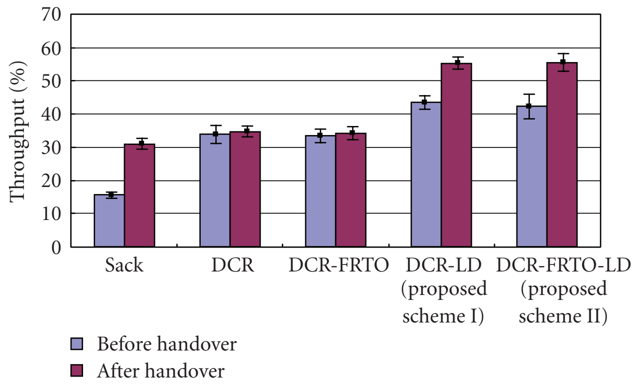


Figure 11: Handover from cellular network to WLAN.

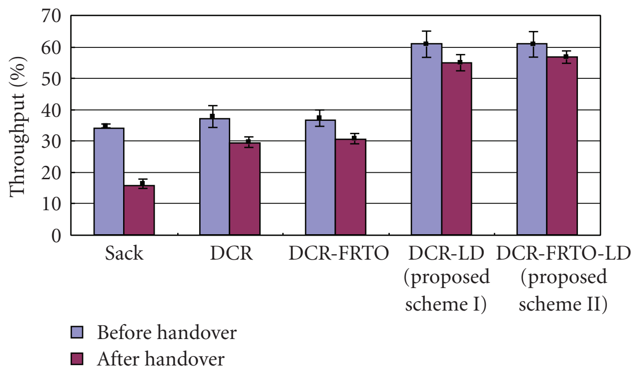


Figure 12: Handover from WLAN to cellular network.

throughput from 229 kbps to 734 kbps. Proposed scheme I increases the throughput from 238 kbps to 459 kbps. In these two cases, the modified LD has made the TCP throughput to be higher with differentiation of wireless losses from congestion losses.

In the proposed scheme I, LD regards most of packet losses as wireless losses just after the handover. The threshold of LD is larger than the newly measured RTT because the mean RTT and the deviation remain high immediately after the handover. This makes TCP keep current congestion window even on a congestion loss, which may eventually result in an unnecessary RTO. The RTO holds transmission of normal packets during timeout and reduces its congestion window to 1 after the timeout. Only when the mean RTT and the deviation are adapted to the WLAN, the LD can correctly detect congestion losses.

This is remedied in proposed scheme II. The mean RTT and the deviation are updated on handover to new network. This allows proposed scheme II to adapt to new network faster than the proposed scheme I. Furthermore, the congestion window just after handover is kept to be as same as a value before handover by F-RTO. This implies that the LD in proposed scheme II can detect congestion losses just after handover. The TCP can avoid RTO. This explains why proposed scheme II improves the TCP throughput by 59.9% compared to proposed scheme I over WLAN

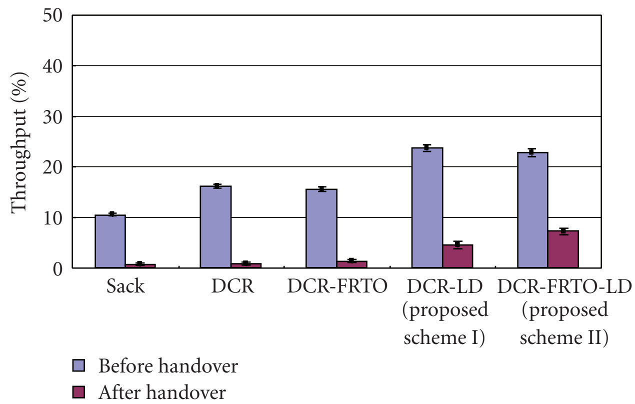


Figure 13: Handover from cellular network to WLAN under congestion.

after the handover. Throughputs for other schemes are reduced by about 50% compared to throughputs before handover.

In the scenario of the handover from WLAN to cellular link over the congested path, the throughput of proposed scheme II in Figure 14 is still higher than other schemes. Especially, the throughput of proposed scheme II is 4.5% higher than proposed scheme I over the cellular network after the handover. This improvement is due to our modification as above mentioned.

The difference between proposed schemes I and II is negligible in noncongestion scenario (Figures 11 and 12). The large difference in WLAN throughput in Figures 13 and 14 can be explained as follows. In our simulation, the forward cross traffic is added on the wired path which has much shorter latency than the wireless paths (3G and WLAN). The added traffic dominates the TCP traffic on WLAN or 3G since it flows only over wired path of high bandwidth (20 Mbps) and very short delay (1 millisecond). Still, the throughput of TCP traffic over WLAN is much larger than the throughput of TCP traffic over the 3G path since the WLAN has much larger bandwidth (10 Mbps) and much shorter delay (20 milliseconds) than 3G path of 1 Mbps bandwidth and 200 milliseconds delay. Thus, before the handover, the suppressed TCP throughput over 3G is much smaller than that over WLAN. Immediately after the handover to WLAN, the TCP throughput is slow to recover in previous schemes (Sack, DCR, DCR-FRTO) and is faster to recover in proposed schemes, but all still remain small due to the limited TCP throughput over 3G before handover as shown in Figure 13. However, the situation is different when handover is performed from WLAN to 3G as shown in Figure 14. The TCP throughput over WLAN before handover is large compared with that over 3G before handover. Thus the TCP throughput over 3G after handover is large whether the schemes are slow to recover (in previous schemes) or fast (proposed schemes). Thus, the big difference in WLAN throughput can be explained by the difference in TCP throughputs over 3G and WLAN under congested condition.

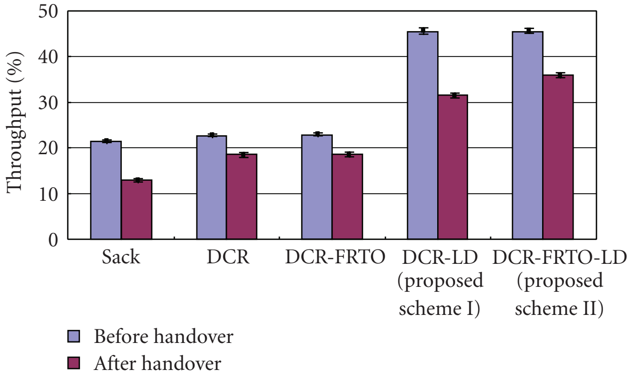


Figure 14: Handover from WLAN to cellular network under congestion.

6. Conclusion

We modified the TCP congestion control scheme to improve throughput over the vertical handover between two most popular wireless networks, that is, 3G cellular networks and wireless LANs. The proposed scheme is an end-to-end scheme which is easy to implement because it uses information readily available to TCP sender. The modified loss differentiation which employs handover detector scheme from F-RTO and incorporates TCP-DCR is shown to greatly improve the TCP throughput against the coexistence of congestion losses, wireless packet losses, and packet-reordering. When the vertical handover occurs, the proposed scheme is robust to the change in loss pattern over wireless link. Our modification can avoid unnecessary retransmission and reduction of congestion window invoked by spurious timeouts during handover. We find that our modification excels other known schemes (Sack, DCR, FRTO) even when the wired part is congested.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0077261).

References

[1] M. Stemm and R. H. Katz, "Vertical handoffs in wireless overlay networks," *Mobile Networks and Applications*, vol. 3, no. 4, pp. 335–350, 1998.

[2] M. Yosuke, M. Takahiro, and Y. Miki, "TCP congestion control with ACK-pacing for vertical handover," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '05)*, vol. 3, pp. 1497–1502, March 2005.

[3] H. Huang and J. Cai, "Improving TCP performance during soft vertical handoff," in *Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA '05)*, vol. 2, pp. 329–332, March 2005.

[4] Y. Gou, D. A. J. Pearce, and P. D. Mitchell, "A receiver-based vertical handover mechanism for TCP congestion control," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2824–2833, 2006.

[5] S.-E. Kim and J. A. Copeland, "Interworking between WLANs and 3G networks: TCP challenges," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '04)*, vol. 2, pp. 1252–1257, 2004.

[6] B. Sardar and D. Saha, "A survey of TCP enhancements for last-hop wireless networks," *IEEE Communications Surveys & Tutorials* , vol. 8, no. 3, pp. 20–34, 2006.

[7] C.-H. Lim and J.-W. Jang, "Robust end-to-end loss differentiation scheme for transport control protocol over wired/wireless networks," *IET Communications*, vol. 2, no. 2, pp. 284–291, 2008.

[8] S. Bhandarkar, N. E. Sadry, A. L. Narasimha-Reddy, and N. H. Vaidya, "TCP-DCR: a novel protocol for tolerating wireless channel errors," *IEEE Transactions on Mobile Computing*, vol. 4, no. 5, pp. 517–529, 2005.

[9] P. Sarolahti, M. Kojo, and K. Raatikainen, "F-RTO: an enhanced recovery algorithm for TCP retransmission timeouts," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 2, pp. 51–63, 2003.

[10] J. Chen and V. C. M. Leung, "Improving end-to-end quality of services in 3G wireless networks by wireless early regulation of realtime flows," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '03)*, Beijing, China, September 2003.

[11] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris, "Link-level measurements from an 802.11b mesh network," in *Proceedings of the International Conference of the Special Interest Group on Data Communication (SIGCOMM '04)*, Portland, Ore, USA, August 2004.

[12] IEEE 802.11 Part 11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specification: higher-speed physical layer extension in the 2.4 GHz band," IEEE, 1999.

[13] R. Fantacci, "Queuing analysis of the selective repeat automatic repeat request protocol wireless packet networks," *IEEE Transactions on Vehicular Technology*, vol. 45, no. 2, pp. 258–264, 1996.

[14] M. Rossi and M. Zorzi, "Analysis and heuristics for the characterization of selective repeat ARQ delay statistics over wireless channels," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 5, pp. 1365–1377, 2003.

[15] A. Gurtov and S. Floyd, "Modeling wireless links for transport protocols," *ACM SIGCOMM Computer Communications Review*, vol. 34, no. 2, pp. 85–96, 2004.

[16] TCP-DCR source codes for ns-2 simulations, http://dropzone.tamu.edu/~sumitha/research.html.

[17] IEEE Std P802.21/D02.00, IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handoff Services, September 2006.

[18] Y. S. Choi, W. Lee, and Y. Z. Cho, "TCP with explicit handoff notification for a seamless vertical handover," in *Proceedings of International Conference on Information Networking (ICOIN '08)*, vol. 5200 of *Lecture Notes in Computer Science*, pp. 831–840, 2008.

[19] A. Gurtov and J. Korhonen, "Effect of vertical handovers on performance of TCP-friendly rate control," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 8, no. 3, pp. 73–87, 2004.

[20] NS-2 network simulator (ver 2). LBL, http://www-mash.cs.berkeley.edu/ns.