Impact of Layer 2 Behavior on TCP Performance in WLAN

Kazuya Tsukamoto†, Ryouichi Ijima†, Shigeru Kashihara††, and Yuji Oie†††
†Graduate School of Computer Science and Electronics,†††Department of Computer Science and Electronics,
Kyushu Institute of Technology, Kawazu 680-4, Iizuka, 820-8502 Japan, TEL:+81-948-29-7687, FAX:+81-948-29-7652
††Graduate School of Information Science, Nara Institute of Science and Technology,
8916-5, Takayama-cho, Ikoma, 630-0192 Japan, TEL:+81-743-72-5213, FAX:+81-743-72-5219
Email: †{kazuya, r-ijima}@infonet.cse.kyutech.ac.jp, ††shigeru@is.naist.jp, †††oie@cse.kyutech.ac.jp

***Abstract*— Ubiquitous networks enabling access by Mobile Nodes (MNs) at any time and at any location will consist of several different wireless LANs (WLANs), so it is likely that an MN will traverse several WLANs during communication. In such a case, avoiding performance degradation during the handoff becomes a critical issue. The majority of existing studies deal with handoff decisions based on upper layer information such as packet loss, but none of the studies extensively examine the criteria for a handoff decision and investigate how the criteria can affect the performance degradation just before the handoff is executed. In this paper, we first extensively investigate the TCP performance under the criteria for the handoff decision employed by existing studies. Then, we deal with Layer 2 behavior and suggest that a cross-layer approach, which enables an MN to perceive the deterioration of a wireless link quickly and exactly, is necessary to avoid the performance degradation during a handoff. Finally, we identify how the movement of the MN affects the overall TCP performance of its WLAN and its own TCP performance. Through simulations, we show that degradation during a handoff can effectively be avoided by exploiting Layer 2 information.**

I. Introduction

Wireless LANs (WLANs) based on IEEE 802.11 [1] are spreading widely due to their low cost, simplicity of installation and broadband connectivity. WLAN services are becoming available at many public spaces, called hotspots, such as coffee shops and waiting areas for public transport, and these WLAN hotspots are managed by different organization or companies. In the near future, these hotspots will continue to spread until they form a complete overlay providing wide-area coverage, and they will be the underlying basis of ubiquitous networks.

In a ubiquitous network, Mobile Nodes (MNs) can access the Internet from any location via these WLAN hotspots. There are two factors essential to achieving transparent mobility over the Internet at any time and at any location. First, for the many existing applications employing TCP as the transport protocol, an MN must be able to execute a handoff between different WLANs and not cause performance degradation during the handoff. Second, the movement of an MN must also not cause the performance degradation of a WLAN.

A number of technologies such as Mobile IP [2], mobile Stream Control Transmission Protocol (mSCTP) [3] and their enhanced protocols, including our previous study [4], have been proposed for allowing MNs to traverse different IP networks while maintaining uninterrupted communications, and most of these technologies make the handoff decision based on upper layer information such as packet loss [2], which can be obtained from Layer 3 or 4. Although the existing studies focus on the performance **during** and **after** the handoff, none of them extensively examine the criteria for deciding how the handoff is to be executed, and none investigate how these criteria can affect TCP performance degradation just before the handoff is executed.

Recently, some new enhanced protocols of Mobile IP and mSCTP have been proposed for executing an optimal handoff using different handoff decisions, which are based upon the change in either the jitter (the variation of the packet interarrival time) [5], Smoothed Round Trip Time (SRTT) [6], or signal strength [7]. However, because these parameters change dynamically due to various factors such as congestion in a wired network and any intervening objects located in a wireless network, they will not be appropriate as handoff decision criteria.

In this paper, we first extensively investigate the TCP performance discussed in existing studies, which make the handoff decision based on upper layer information such as packet loss. With this information, other information can indicate whether the transmission performance on Layer 2 is getting better or worse. Therefore, we deal with Layer 2 behavior, and suggest that a cross-layer approach, which enables the transport layer to perceive the deterioration of a wireless link quickly and exactly, is necessary to avoid the performance degradation during a handoff. Finally, we identify how the movement of the MN affects the overall TCP performance of its WLAN and its own TCP performance. Through simulations, we show that these degradations during a handoff can effectively be avoided by exploiting Layer 2 information.

II. Handoff decision criteria of existing studies

As stated in the introduction, a number of existing technologies have been proposed to allow MNs to traverse different IP networks without interrupting communication. One example is Mobile IP, which is being proposed as a standard by the Internet Engineering Task Force (IETF). In Mobile IP networks, each MN detects its own movement by utilizing

Router Advertisement packets, which are broadcast from an Access Point (AP) infrequently (typically one per second). Major movement detection mechanisms such as Lazy Cell Switching (LCS) [2] and Eager Cell Switching (ECS) [2] have been proposed in Mobile IP: these mechanisms can detect the movement of an MN based on the loss of the Router Advertisement packets. However, its infrequency results in an increase of the handoff decision latency. Furthermore, most enhanced protocols of Mobile IP, such as Fast Handover Mobile IP (FMIP) [8] and Hierarchical Mobile IP (HMIP) [9], have been proposed to improve the performance **during** and **after** the handoff; in other words, they have not extensively examined the criteria for the handoff decision, i.e., the condition triggering the handoff decision.

Another proposed study is mSCTP, which is the mobile extension of the Stream Control Transmission Protocol (SCTP) and is newly equipped with functions to dynamically add or delete the IP addresses of MNs, thereby supporting mobility during the handoff. In mSCTP, the issues of handoff decision are not discussed in detail [3].

Some new enhanced protocols deal with the handoff decision and propose a new movement detection mechanism to quickly perceive the change in the transmission condition. The methods proposed by Cunningham [5] and Kelly [6] make the handoff decision based upon the change in either jitter or SRTT of a stream of packets. However, the jitter and SRTT may change dynamically due to congestion in the wired Internet; that is, these methods cannot exactly perceive the change in the transmission condition of a wireless link. To solve this issue, reference [7] bases the decision for handoff on the signal strength obtained from Layer 1, which can obtain information from a wireless link directly. Although signal strength is one type of information that indicates the condition of a wireless link, properly estimating the occurrence of packet loss from signal strength in advance is very difficult at the MN, because the signal strength fluctuates abruptly due to the distance from the AP and any interfering objects located between the MN and the AP.

On the other hand, in [10], we have already proposed to employ the number of retries experienced by a frame on Layer 2 as the handoff decision criterion for Voice over IP (VoIP) applications on a WLAN. However, our paper did not investigate other Layer 2 information except for the number of frame retries, and did not consider TCP applications. Therefore, in this paper, we first extensively investigate TCP performance under the criteria for handoff decisions employed in existing studies. Then, we focus on various Layer 2 behaviors and compare these behaviors with the TCP performace, thereby finding an effective way to prevent performance degradation during a handoff.

III. Layer 2 information

We consider that the handoff decision should be based upon the information obtained from Layer 2 because Layer 2 information indicates the condition of a wireless link precisely. However, the handoff is decided at an upper layer than Layer

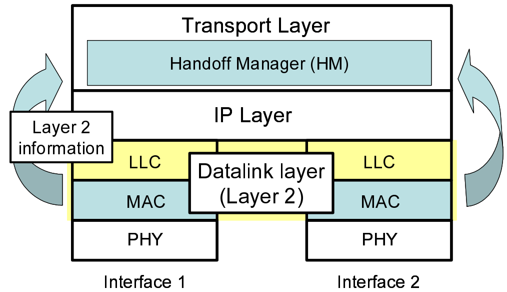

Fig. 1. Cross-layer architecture.

3 in order to maintain the independence from a wide range of datalink technologies. Therefore, notification between these layers is essential to optimally execute a handoff; however, the information held in each layer cannot be accessed from different layers due to the concept of the traditional layered architecture. In this paper, we suppose that the benefit of introducing a cross-layer approach [11] is greater than the cost paid for its benefits [12], and thus we employ the cross-layer approach to achieve the interaction between these layers. In our concept, a Handoff Manager (HM) on the transport layer decides the handoff based on Layer 2 information, as shown in Fig. 1. Note that our cross-layer concept can be applied between both end-to-end hosts only, and so it cannot be applied to an intermediate router such as the AP.

In the IEEE 802.11 specifications, the MAC layer is responsible for the access control of a wireless channel. Therefore, we are able to obtain various Layer 2 information from the MAC layer directly. Next, we describe each item of Layer 2 information, which has the potential to become an optimal criterion for handoff decisions.

A. Queue length of the buffer

Each WLAN interface has a buffer to transmit/receive the packets to/from the PHY layer or IP layer. The sender can detect successful transmission by receiving an ACK frame in response to a transmitted data frame, that is, by the stop-and-wait mechanism. Therefore, when a data frame or ACK frame is lost in a WLAN, its subsequent frames, which are received from the IP layer, are stored in the buffer until the transmission of the lost frame is successful. Thus, in the case that the condition of a wireless link is getting worse, the queue length of the buffer may increase.

B. Transmission time

The MAC layer can get both the time of sending a data frame (*sending time*) and the time of receiving an ACK frame (*receiving time*), because the MAC is responsible for the access control of the wireless channel. In this paper, this interval time between the *sending time* and *receiving time* on Layer 2 is defined as the *transmission time*. When data frames or ACK frames are lost, the interval time increases; that is, this interval time includes the retry delay on Layer 2. Note that, if the Request to Send/Clear to Send (RTS/CTS) is employed, the transmission time also includes the delay of the RTS/CTS frame exchanges.

C. Number of frame retries

When data frames or ACK frames are lost, the sender retransmits the same data frames until the number of retries reaches a predetermined limit. If the RTS/CTS is applied, the retry limit is set to 4; otherwise, it is 7. Therefore, the data frame can be retransmitted a maximum of 4 or 7 times (the initial transmission and 3/6 retries), if necessary. If the sender does not receive an ACK frame within the retry limit, the data frame is treated as a lost packet and the TCP sender retransmits the same packet by its retransmission control of TCP. That is, we can see that frame retries occur before packet loss, irrespective of the RTS/CTS.

Frame retries occur for either of the following two reasons: (1) degradation of signal strength and (2) collision with other frames. Note that, with RTS/CTS, collisions between data frames, namely, a hidden terminal problem, never occur due to the exchange of the RTS/CTS frames. TCP applications with a large packet size commonly employ the RTS/CTS mechanism, because the collision probability of a large packet is higher than that of a small one.

IV. Simulation Model

We now describe the simulation model employed here. We use Network Simulator version 2 (ver 2.27). In our simulation, we assume the open space environment without reflecting walls, and employ a realistic model to evaluate the effect of the movement of an MN, as shown in Fig. 2. Each of the five MNs establishes a TCP (NewReno) connection with a Corresponding Node (CN) over a wireless link for file transfer communication, with a packet size of 1500 bytes. In our simulation, we employ the RTS/CTS mechanism due to the large packet size, and treat two realistic cases: an MN moves away from the AP (Case 1), and three MNs move away from the AP simultaneously (Case 2). In both cases, the other MNs (except for the moving MN) are located close to the AP, and packets sent by these MNs never experience data frame retransmission due to degradation of the signal strength.

Through these two simulations, we evaluate the MN movement effects from two points of view, (1) a moving MN and (2) a WLAN, where the distance between the AP and an MN varies from 0 to 22 m. First, we extensively study how the TCP goodput performance is related to features such as packet loss and RTT, observed on Layer 4, to identify how the movement of the MN affects the performance under the criteria employed in the existing studies. Next, we examine features associated with Layer 2 and how they affect the TCP goodput performance to find an optimal handoff decision criterion enabling the prevention of performance degradation during a handoff. Finally, we evaluate if the optimal handoff decision criterion is effective even when multiple MNs move simultaneously.

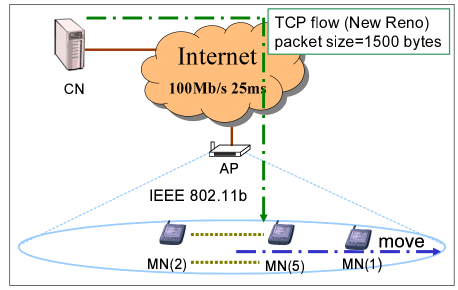

Fig. 2. Simulation Model.

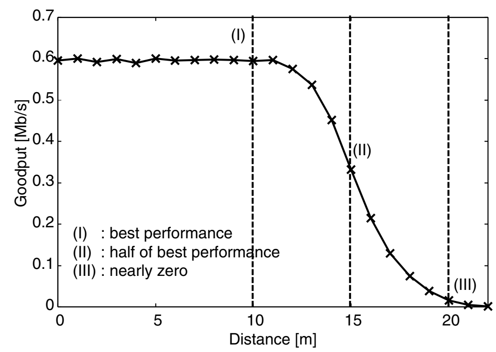

Fig. 3. Goodput.

V. Simulation Results

A. Movement of one MN: Layer 4 information

First, we extensively study how the TCP goodput performance is related to features such as the packet loss and RTT observed on Layer 4.

1) Goodput: Figure 3 shows how TCP goodput performance changes as one MN moves away from the AP. As shown in Fig. 3, the TCP goodput begins to decrease beyond 10 m (I), and at around 15 m, the goodput rapidly drops to 0.3 Mb/s (half of the best performance) (II). Beyond that, at around 20 m, the goodput decreases to nearly zero, which means that the MN can hardly communicate to the CN (III).

2) Packet loss: Figure 4 shows how often packet losses due to duplicate ACKs and expiration of the timeout timer (timeout) are detected by TCP. As shown in Fig. 4, duplicate ACKs are detected at a fixed rate, even when the MN is located close to the AP (0 m). However, because this is a common phenomenon in TCP congestion control, an MN should not make a handoff decision based on the packet losses caused by duplicate ACKs. Packet loss caused by timeout begins to occur around 15 m, and after that, its rate increases drastically.

Comparing Fig. 4 with Fig. 3, we can see that the TCP goodput rapidly drops to 0.3 Mb/s (half of the best performance), even though packet loss rarely occurs at around 15 m. These results suggest that existing studies, which decide a handoff based upon packet loss, could not maintain TCP performance during the handoff.

3) RTT: Figure 5 shows the change in the 90 % confidence interval of RTT. As shown in Fig. 5, the RTT begins to increase

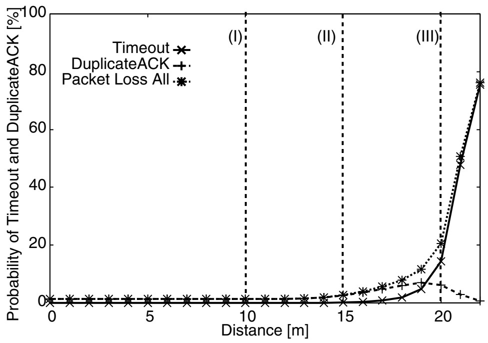

Fig. 4. Packet loss.

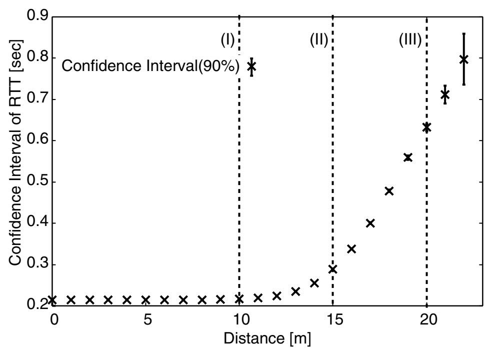

Fig. 5. 90% confidence interval of RTT.

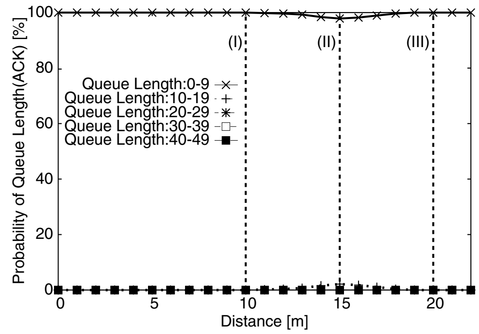

Fig. 6. Queue length of the buffer.

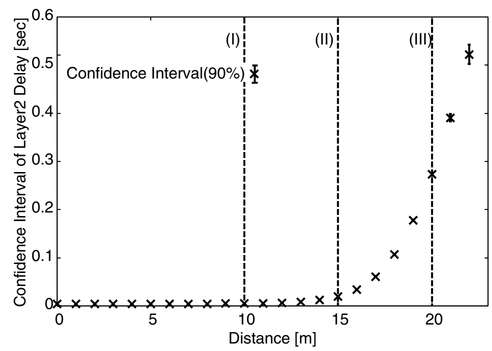

Fig. 7. 90% confidence interval of Transmission time.

just after 10 m, where the goodput begins to decrease. However, the change is small until the distance between the AP and the MN reaches 15 m, which is where the packet loss begins to occur (Fig. 4). Beyond that, the RTT increases drastically due to the occurrence of packet losses. Furthermore, the increase of RTT may also frequently occur due to the congestion of the wired Internet (minus the wireless condition). Therefore, the method which makes a handoff decision based on the change in RTT may cause a ping-pong effect. This leads to a severe performance degradation during the handoff. These results suggest that some existing studies [6], which decide a handoff based on the change in RTT, also could not maintain TCP performance during the handoff.

B. Movement of one MN: Layer 2 information

Next, we examine some features associated with Layer 2 described in Sec. III, and further investigate how they affect the TCP goodput performance.

1) Queue length of the buffer: Figure 6 shows how the queue length of the buffer on Layer 2 changes when an MN moves away from the AP. The queue length increases slightly around 14 m, just before the packet loss occurs (Fig. 3). However, because the increase is quite small, the degradation of the wireless condition could not be perceived exactly by the change in the queue length. In our simulation, the MN rarely receives the frames transmitted from the AP due to congestion at the AP, and thus the number of TCP ACK packets transmitted from the MN decreases. Therefore, the number of frames waiting for transmission in the buffer decreases, thereby leading to the small queue length. Furthermore, because the TCP goodput decreases to 75 % of the best performance at around 14 m, where the queue length begins to increase, we can note that the TCP goodput performance degradation could not be avoided by exploiting the queue length of the buffer.

2) Transmission time: Figure 7 shows the change in the 90 % confidence interval of the transmission time within a WLAN network. The transmission time does not increase until the packet loss begins to occur at around 15 m, and then it increases drastically, like the RTT described in Sec. V-A.3. From these results, we can see that the TCP goodpout may decrease to half of the best performance, even when the MN makes the handoff decision based on the transmission time. These results suggest that the transmission time is not appropriate as a criterion for handoff decisions.

3) Number of frame retries: Figure 8 shows how often one or multiple retries are required. Note that the packet loss occurs when the fourth retry fails (with RTS/CTS). The frame retry begins to occur at around 8 m, and the TCP goodput also begins to decrease soon after the occurrence of the frame retry. These results show that degradation of the TCP goodput performance begins even when a frame retry occurs at least once. Therefore, we suggest that degradation of the TCP goodput performance could effectively be avoided

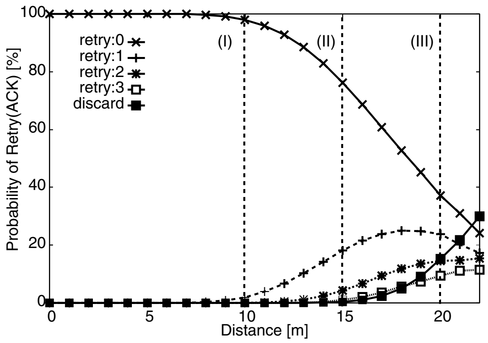

Fig. 8. Frame retry ratio.

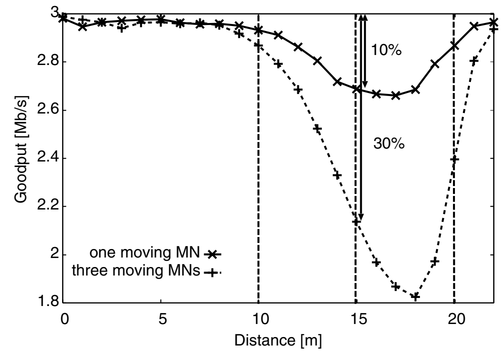

Fig. 9. Total goodput of WLAN.

by exploiting the frame retry when an MN executes a handoff, and the frame retry has the potential to become an optimal criterion for handoff decisions.

C. Movement of multiple MNs

Figure 9 shows how the total goodput of a WLAN is affected by the movement of one or three MNs. In a future mobile environment, many MNs will often move into or through a WLAN, resulting in the occurrence of simultaneous and frequent handoffs among multiple MNs. Under that environment, preventing the performance degradation of a WLAN during the handoff will be a critical issue for effective communication.

When one MN moves away from the AP, the total goodput performance decreases to 2.7 Mb/s, which is about 90 % of the best performance, at around 15m, where packet loss just begins to occur. In addition, with the movement of three MNs, the goodput decreases to 2.1 Mb/s, which is about 70 % of the best performance. These results suggest that a decision criterion allowing MNs to perceive the change in the condition of the wireless link plays an important role in executing a handoff.

When the multiple MNs move simultaneously, the total goodput begins to decrease at around 9 m, which is slightly nearer than the case of one MN (around 10 m). Furthermore, the frame retry begins to occur at around 8 m. Thus, we can see that the number of frame retries could be an optimal criterion for handoff decisions, even in the case of the movement of multiple MNs.

VI. Conclusion and Future Work

In this paper, we first investigated TCP goodput performance under the criteria for handoff decisions employed in existing studies. Then, we focused on Layer 2 behavior and identified how the movement of an MN affects overall TCP performance as well as its own TCP performance. Simulation results showed that the existing technologies, which make the decision to execute a handoff based on packet loss and the change in RTT, could cause performance degradation of both one moving MN and a WLAN shared by multiple MNs. In contrast, we found that the number of frame retries has the potential to serve as a criterion for handoff decisions to effectively avoid TCP performance degradation, irrespective of the number of moving MNs, because the goodput performance decreases soon after the frame retries occur.

In our paper, the handoff management architecture cannot be treated in detail. Therefore, several candidate topics for future work are the proposal of various handoff decision algorithms based on the number of frame retries, and the investigation of the performance of our proposed scheme.

Acknowledgment

This work was supported in part by the Japan Society for the Promotion of Science, Grant-in-Aid for JSPS Fellows (17-6551) and Scientific Research (A)(15200005) and in part by the Ministry of Public Management, Home Affairs, Posts and Telecommunications, Japan.

References

[1] IEEE 802.11, 1999 Edition, Available at http://standards.ieee.org/getieee802/download/802.11-1999.pdf
[2] C. Perkins, "IP Mobility Support for IPv4," IETF RFC3344, Aug. 2002.
[3] S. J. Koh, et al., "Mobile SCTP for Transport Layer Mobility," IETF Internet Draft, *draft-reigel-sjkoh-sctp-mobility-04.txt*, Jun. 2004.
[4] K. Tsukamoto, Y. Hori, and Y. Oie, "Mobility Management of Transport Protocol Supporting Multiple Connections," *Proc. of ACM MobiWac2004*, pp. 83-87, Oct. 2004.
[5] G. Cunnigham, et al., "Soft, Vertical Handover of Streamed Video," *IEEE 3G Mobile Communication Technologies (3G 2004)*, Oct. 2004.
[6] A. Kelly, et al., "Delay-Centric Handover in SCTP over WLAN," *Transactions on Automatic Control and Computer Science*, Vol. 49 (63), May 2004.
[7] M. Chang, et al., "Transport Layer Mobility Support Utilizing Link Signal Strength Information," *IEICE Trans. on Communication*, Vol. E87-B, No. 9, pp. 2548-2556, Sep. 2004.
[8] R. Koodli, "Fast Handovers for Mobile IPv6," IETF Internet Draft, *draft-ietf-mobileip-fast-mipv6-08.txt*, Oct. 2003.
[9] H. Soliman, et al., "Hiearchical Mobile IPv6 mobility management (HMIPv6)," IETF Internet Draft, *draft-ietf-mipshop-hmipv6-02.txt*, Jun. 2004.
[10] S. Kashihara and Y. Oie, "Handover Management based on the Number of Retries for VoIP on WLANs," *Proc. of IEEE VTC2005 spring*, CD-ROM, May 2005.
[11] S. Shakkottai et al., "Cross-Layer Design for Wireless Networks," *IEEE Communications Magazine*, Vol. 41, No. 10, pp. 74-80, Oct. 2003.
[12] Vikas Kawadia, et al., "A Cautionary Perspective on Cross-Layer Design," *IEEE Wireless Communications*, Vol. 12, No. 1, pp. 3-11, Feb. 2005.