A Fast Handoff Using State Information in Mobile VoIP Environment

Kiyoung, Kim† and Jae-Pyo, Park††,

Dept. of Software, Seoil College, Korea†
Information Media Technology Research Center, Soongsil University, Korea††

Summary
Recent growth of the wireless Internet provides various multimedia services including video, audio, and voice. Especially, the VoIP has been become an important technology supporting transmission of voice data over wireless environment. In this paper, we propose a fast handoff scheme based on state information for guarantee QoS of VoIP services in wireless environment. The proposed handoff reduces delay of the session timer by checking session states and avoids the termination due to the handoff delay in the SIP. The function of the session timer, tracks the states of the current session, is not supported initial version, recognizes the failure of the session and release resource reserved for the failed session, which results in this resource is reused by the other sessions. The proposed scheme performing handoff, mobile node informs the proxy server of its handoff state. The proxy sever doesn't change the period of UPDATE requests adding the handoff state to the exiting states which utilized by set UPDATE request interval. Simulation results show that the proposed scheme reduces the number of state transition and preserves time interval.

Key words:
VoIP, Handoff, Delay, *SIP,* QoS

1. Introduction

The Internet has been becoming a communication media which is transferring data, voice, and video. The Transmission Control Protocol and Internet Protocol (TCP/IP), which was designed initially for non-real-time data transferring, is not appropriate for real-time data transferring. However, the advance of transmission medias, e.g. fiber optic cables, makes up for the defect of the TCP/IP.

The Public Switched Telephone Networks (PSTN), which are a kind of traditional voice communication networks, had already been digitalized internally except for analog lines between telephone subscribers and telephone offices. The popularization of high-speed Internet connections cannot anymore make subscribers connect to a PSTN for voice communications. Many researches for voice communication using the Internet have studied for a long time. Some commercialized Voice over IP (VoIP) technologies for voice transmission using the cable Internet let the number of users increase asymptotically. Representative protocols used for the VoIP are the recommendation H.323 of the International Telecommunications Union (ITU) and the Session Initiation Protocol (SIP) of the Internet Engineering Task Force (IETF). The SIP in current days, which is more lightweight and which can be easily implemented, is widely used unlike the past that the H.323 was used in initial VoIP services.

In recent days, VoIP Services with mobility in the wireless environment are focused owing to the advance and generalization of mobile technologies.

In order to provide VoIP services in the wireless environment dissimilar to the cable environment, the handoff delay time must be considered. A mobile device can be in motion in the wireless environment unlike the cable environment that has no additional delay after a call was connected. Thus, it is need to connect into a new wireless connection point of a reached network, when a node is leaving from current wireless connection point of connected network. This delay time for connecting is called by the handoff delay time or handover delay time.

Low-speed moving or short distances doesn't cause generally the handoff delay time. the handoff delay time, however, should be decreased until users don't recognize delay or extinction of communication in order to replace cellular phone services [1,2].

The handoff delay time is caused by a set of a channel searching, an authentication, and a reconnection. When a mobile node is reached in a new wireless network, it should search a channel supported by an new Access Point (AP), be passed an authentication process, and reconnect to the previous AP. Even if a authentication process also generates the handoff delay time, it is under the influence of security algorithms or security protocols. In this paper, we propose a handoff algorithm supporting VoIP services in the wireless environment with mobility by reducing the channel searching time and the reconnection time without the authentication delay time, and analyze the performance by the number of handoffs and the number of mobile nodes.

Manuscript received December 5, 2008
Manuscript revised December 20, 2008

This paper is divided into five chapters. chapter Ⅱ explains the VoIP and wireless technologies. The chapter Ⅲ proposes a new algorithm. The chapter Ⅳ analyzes the algorithm by performance evaluation. The chapter Ⅴ describes the conclusion and the direction of future works.

2. Related Works

2.1 VoIP Overview

Standardizations for VoIP Services are led by the IETF and the ITU. In the early days, the H.323 recommended by the ITU was used commonly for developing. However, in current days, the SIP, that has advantages of simplicity and scalability, suggested by the IETF is generally exploited to implement the VoIP services. The SIP is a protocol that is used to initiate, modify, and terminate sessions between users. The SIP performs the call establishment based upon the user location service, the call participant management, and the feature invocation. Standards for the VoIP are comprised of the H.323 announced by the ITU, the SIP in standardization progress by the IETF, the MGCP suggested by the IETF-MEGACO. Although the H.323 has a market share of more than 90 percent in the past, the SIP has been replacing in today, and is expected as a representative protocol for the VoIP [1].

Table 1 : VoIP devices and their functions

Device	Signaling gateway	Media gateway	Call Agent
Functions	Signal converting for transmitting SS7 into packet networks Supporting of various signal formats SSu, ISUP, TCAP Providing network access services	Transferring multimedia data packets Supporting various media types Control and processing media streams on heterogeneous networks	Media Gateway Controller S/W Call Processing Address Mapping Choosing a Media Gateway

The VoIP consists of protocols, a device which binds PSTNs and the Internet, and a Call Agent which handles nodes connected to the Internet. Fig. 1 illustrates the relationship between PSTNs and the Internet.
Elementary technologies of a VoIP system are comprised of a Signaling Gateway, a Media Gateway, and a Call Agent. The Signal Gateway is a device that interconnects PSTNs and the Internet by converting signals used in the PSTN such as Signaling System 7 (SS7) to appropriate protocols, but is not at Internet networks. The Call Agent is a device that provides the H.323, the SIP, the Session Description Protocol (SDP), and so on., and interconnects calls of two nodes in the Internet. The Call Agent can be operated with Signal Gateway when calls requested by devices in the Internet. Table 1 summarizes VoIP devices and their functions.

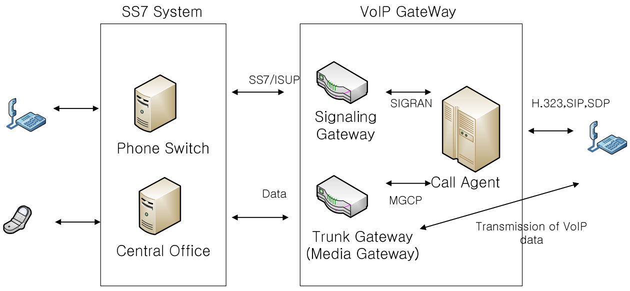

Fig. 1 Components of the VoIP

In order to adapt the VoIP which is optimized for the cable environment into the wireless environment, connections with Access Points (AP) that are connection points to cable environment and reconnections that are caused by changing locations of mobile nodes are required. In following section, we describe wireless technologies for interconnecting with cable networks.

2.2 Wireless LAN Technologies

Recently, standards for local area communications have been expanded from cable environment to wireless environment. Many of customers requests the wireless environment with mobility in addition to choose own communication media because of the increasement of easy to use and the enhancement of communication quality, which is similar to the cable environment, of the wireless environment.
The IEEE 802 LAN/MAN, also known as 8902.x, is a representative standard for wireless network suggested by the IEEE 802 Standards Committee. The IEEE 802 Standards Committee, who enactments standards related functions and processes in the layer 2 of the Open Systems Interconnection (OSI) reference model, is involved lately in standardizations for the Personal Area Network (PAN) and the Bluetooth. Current working groups are comprised of the IEEE 802.1 Working Group up to the IEEE 802.7 Working Group. Each working group conducts own cases and standards, and submits results with a release number based on the number of own working group. The IEEE 802.1 standard and the IEEE 802.2 standard for cable and

wireless networks define security specifications and the Logical Link Control (LLC) respectively. The IEEE 802.11 Working Group, who deals with the Wireless Local Area Network (WLAN) standardization, and the IEEE 802.15, who treats the PAN standardization, have a plan to be departmentalized into some small-scale Task Groups (TG), and to supply improve standards and technical supports. The IEEE 802.11 standard had been released in 1999, and provides 2 Mbps speed using 2.4 GHz band through the Frequency Hopping Spread Spectrum (FHSS) and the Direct Sequence Spread Spectrum (DSSS) [3].

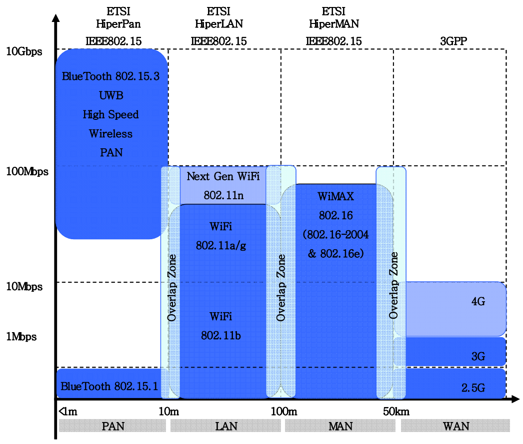

Fig. 2 Wireless technologies

Principal wireless technologies for the local area, in current days, are implementation technologies of IEEE 802.11 standards and Bluetooth standards. Although they have some interference for each other, their applications are naturally separated because of the difference of aspects of their services. Fig. 2 classifies current wireless technologies by speed and bandwidth.

The typical wireless technologies used by the VoIP are implementations of IEEE 802.11 standards. Mobile nodes can utilize cellular networks and Wi-Fi networks by the use of duplex-modes. These technologies allow mobile nodes that were approached a Wi-Fi network to connect automatically into it.

They transfer reorganized IP formats from cellular signals into the Internet through the tunnelling approach for non-connectionless states and compatibilities between cellular networks and Unlicensed Mobile Access (UMA) [4]. Duplex modes enable customers to supply voice communications with low-cost than cellular networks. The Wi-Fi can be used for multimedia services containing large volume because the Wi-Fi supports 54 Mbps speed. The IEEE 802.11b provides data rates up to 11 Mbps. The IEEE 802.11g provides data rates up to 54 Mbps. The IEEE 802.11n provides date rates 103 Mbps up to 320 Mbps. The IEEE 802.11 supply high and stable bandwidths and connectionless services in addition to replace traditional cable networks. The IEEE 802.11, however, doesn't consider the mobile environment. Hence, if a node migrates to a new network from a current network, it has to re-establish a new connection with the new network. The reconnection with the new network makes customers recognize the delayed or disconnected communication. A series of processes for the reconnection is the handoff. The delay for the reconnection is the handoff delay [4,5]. A communication is disconnected during the handoff delay time. Long handoff delay times obstruct stable communications. Many of ideas and techniques for the reduction of the handoff delay time have been suggested because it is possible to guarantee good connection quality by reducing the handoff delay time. The reduction of the handoff delay time can be realized by decreasing the AP searching time, the authentication time, and the connection reorganization time.

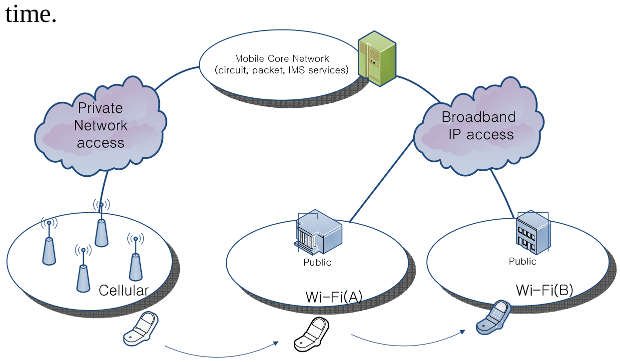

Fig. 3 UMA Network

The Unlicensed Mobile Access (UMA) protocol, which is developed by the 3rd Generation Partnership Project (3GPP), provides circuit-based voice transmissions over IP access networks.

T-Mobile USA, Inc. in Washington prepares for Cellular/Wi-FI services. Cingular Wireless, Inc. merged with AT&T Wireless in 2004 and Sprint Nextel Inc. in Kansas also have plans for similar services. Many service providers in the Europe including Telecom Italia, Inc. and TeliaSonera, Inc. also have various researches in progress for the duplex mode. VoIP services over the Wi-Fi is expected to the highly-use of services because they can supply more high-speed data rates than 4 Mbps speed of the 3G CDMA cellular technology.

2.3 Wireless VoIP Handoff

The UMA environment supports the cooperation with cellular networks and Wi-Fi networks. It, however, is optimized for communication devices with long stays within connected networks, but not with continued movements. These movements cause generally handoff delay times. A handoff delay time can be divided into three times that are the AP searching time, the authentication time, and the connection reorganization time. Fig. 4 illustrates a typical flow of a handoff delay time.

Searching approaches of APs are comprised of the active searching approach, that lets a moving node entered into a network broadcast ProbeRequest messages for searching APs, and the passive searching approach, that lets a moving node be recognized by receiving Beacon messages from APs in networks [6].

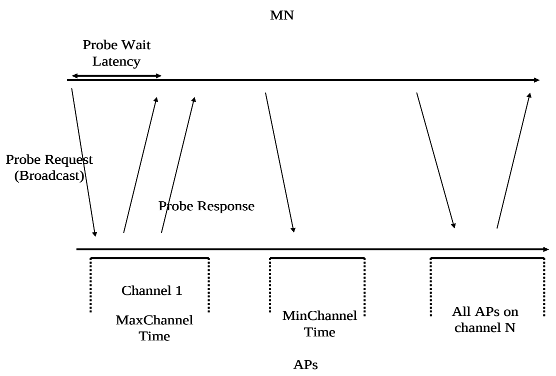

Fig. 4 Channel searching delay time

A moving node decides an AP, which is one of the found APs by the above AP searching approaches, to be connected through the Received Signal Strength Indication (RSSI). The moving node tries to establish another AP after a connection process is finished. The moving node sends ReassociationRequest messages to the new AP when the authentication between the moving node and the new AP is successfully completed. The AP received ReassociationRequest messages performs message exchanges, defined in the Inter-Access Point Protocol (IAPP) [6], with other APs.

Fig. 5 illustrates a network structure supporting the mobility. A node tries to search available APs, and to connect to an AP with strong signal and good quality in the arrived network which contains three APs. Researches for the handoff delay reduction can be divided into two layers that are the layer 2 and layer 3. In actually, The IEEE 802.11f defines the IAPP, which consists of a series of the AP searching, the authentication, and the connection reorganization, for supporting the handoff.

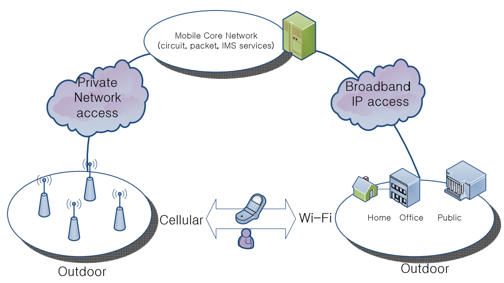

Fig. 5 VoIP devices and Handoffs

An AP searching, or a channel searching, can be performed by trying to search the next channel after the minimum channel searching time or the maximum channel searching time is expired. A moving node, for example, broadcasts channel detection messages to find available channels, and waits for a while because it doesn't know how many APs are in the arrived network. The waiting time can be defined as maximum channel searching time or minimum channel searching time. If the moving node cannot catch the respond for channel detection messages in the waiting time, it tries to scan the next channel.

The channel searching, in other words, is a set of processes that are searching channels with some waiting times. If a node tries to search APs in a unused channel, it causes the delay time. Therefore, the reduction of this delay time provides decreased handoff delay times and fast handoffs.

Ⅲ. VoIP for Wireless Environment

The SIP which is used for VoIP services is a signal processing protocol, functions of which are to create, update, and finalize multimedia sessions and calls , for the application layer. The SIP defines User Agent (UA) and Network Server (NS). The UA includes User Agent Client (UAC) and covers SIP requests. User Agent Server (UAS) receives and responds SIP requests and it is designed to operate under SIP terminals. The SIP contains a proxy server, which delivers SIP requests to it destinations, and charges authentications, and a redirect server register, which responds to requesters by their destination addresses. The UA records the state of sessions into the redirect server register to support mobility of customers.

The original SIP doesn't have any function that checks the state of established sessions. The SIP extension, in order to solve this problem, includes a SIP session timer to check the state of SIP sessions. The SIP extension checks the state of sessions by the use of UPDATE requests, or sends UPDATE requests to the proxy server to maintain active sessions.

UPDATE requests for the session maintenance are sent by every predefined period time. If the period time is too short, it causes the extravagance of the network resource because of frequent UPDATE requests. If the period time is too long, it causes the excessive possession for the resource of the proxy server. It is possible to manage the SIP session timer with a dynamic period to check efficiently the state of sessions, however, in case of the mobile environment, UPDATE requests may be received in the period time because of the handoff delay. This makes the proxy server finalize pertinent sessions. Thus, it is required to notify the proxy server of handoff progresses of moving nodes to prevent to finalize sessions. In addition, it is need to re-adjust the period time considering the handoff delay for UPDATE requests to avoid misconceiving handoffs for network collisions during the handoff process.

Although the both a mobile device and a new AP can detect the occurrence of the handoff, it is more efficient to regulate the period time for UPDATE requests by a mobile device storing the current session.

In this paper, we define four states, which are Good (G), Bad (B), Down (D), and Handoff (H), for a mobile device by extending three states [7] presented by the earlier studies. The state G is the call connection state with the below of a specific critical value. The state B is the call connection state with the above of a specific critical value. The state D is the disconnected call state. The critical value is defined as the delay time. The call connection state is depended upon states of network and the mobile device.

The state G, for example, can be changed to the state B when the call connection state is not good or to the state D when the call connection state is disconnected. The state D doesn't have state changes anymore, and finishes. The state B can be changed to the state G. The state H causes when a mobile device is in the handoff state. All of states excepts for the state D can be changed to the state H. The state H can be changed to the others.

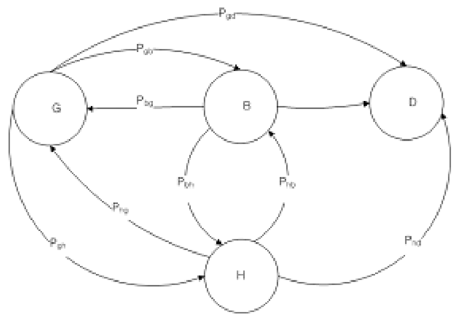

Fig. 6 The Status diagram with four states

The sum of probabilities to be transited from each state is 1. The probability is in the following formula.

$$
\begin{aligned}
&P_{gb} + P_{gh} + p_{gd} = 1 \\
&P_{bg} + P_{bh} + P_{bd} = 1 \\
&P_{hg} + P_{hb} + P_{hd} = 1
\end{aligned}
\qquad (1)
$$

A mobile node transmits the handoff state through the UPDATE request to a proxy server to be recognized the handoff state. The proxy server adjusts the period for UPDATE requests to the handoff delay time. The proxy server, after the completion of the handoff, re-adjusts the period for UPDATE requests by the call connection state. The suggestion, which defines four states, can solve the possibility that the state G may be misunderstood the state B because of the handoff delay time.

The state H can be changed to the state G, B, or D after the handoff progress is completed. The state D is become if the handoff progress is fall, and a mobile device cannot reconnect. If a mobile device doesn't transmit the UPDATE request in the period time after the handoff progress, the proxy server can recognize the state D. After the handoff progress is normally completed, the mobile device transmits UPDATE requests to the proxy server to ascertain the call connection state, and to re-adjust the period for UPDATE requests.

The proxy server re-adjusts the period for UPDATE requests by the following procedure. The initial period time default time interval is assigned greater value than the handoff delay time after a call connection is established. The period time, thenceforth, is assigned by the call connection state by the UPDATE requests.

Scheme in proxy server

interval = default time interval;
state = Get(UPDATE REQUEST);
Switch(state) {
case G
increases UPDATE request interval time;
case B
decreases UPDATE request interval time;
case D
terminates current session & releases resource of proxy server;
case H
set default UPDATE request interval time
}

Ⅳ. Performance Evaluation

we denote the network state as G, which is the good state, B, which is the bad state, H, which is the handoff state, and D, which is the down state that it is need to return the resource of the proxy server. Each state is followed by the exponential distribution. The probability Ps that a packet is transferred without error is as in the following formula.

$$P_s = \frac{(1-P_{lh})\lambda_h + (1-P_{\lg})\lambda_b + (1-P_{lb})\lambda_g}{\lambda_h + \lambda_b + \lambda_g} \quad (2)$$

We have examined the number of state changes by the handoff probability including the number of UPDATE request messages and the number of handoffs. Variables for the simulation are as in the following.

Table 2 : Variables for the simulation

Variable	Description
MN_n	The number of mobile node
T_h	Handoff delay time
P_{lh}	Packet-loss rates for handoff
P_{lg}	Packet-loss rates for call connection with the good state
P_{lb}	Packet-loss rates for call connection with the bad state
Handoff delay	searching delay time authentication delay time reconnection delay time

Fig. 7 shows the number of message transmissions by the handoff probability.

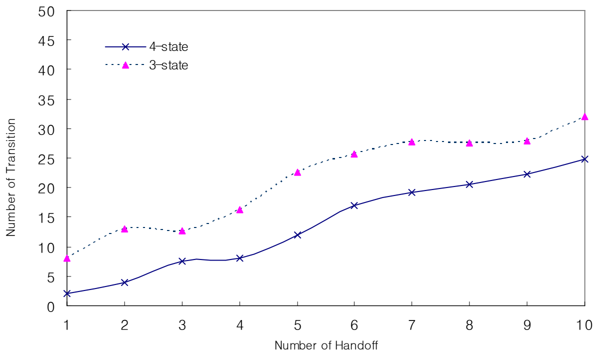

Fig. 7 The number of state changes by handoffs

The suggested approach and the typical approach have the approximate number of message transmissions in case of the low handoff probability that few mobile devices are in motion, or are not in motion. However, the suggested approach has relatively more less the number of message transmissions in case of the high handoff probability that many mobile devices are constantly in motion. The reason of this result can be interpreted that the handoff delay is not determined as the delay caused by the call connection with the bad state.
Fig. 8 shows the number of the period changes for message transmissions by the number of handoffs. The axis x is the number of movements of mobile devices, and can be ranged from 1 to 10. The axis y is the number of the period changes for message transmissions. Because the suggested approach doesn't alter the state based on the handoff state, the number of state changes doesn't be rapidly increased in case of the high number of handoffs. Therefore, the suggested approach is an appropriate technique for the environment with many mobile devices in motion.

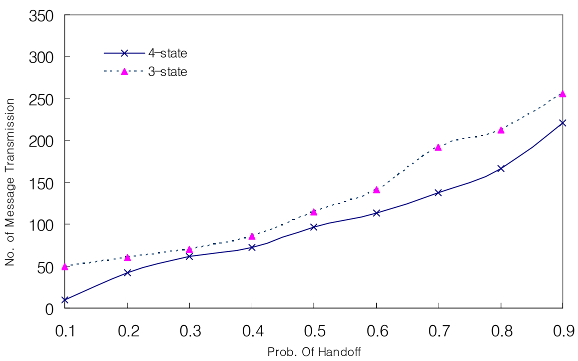

Fig. 8 The number of state changes by handoffs

Ⅴ. Conclusion and Future Works

It is need to consider the handoff which caused by movements of mobile devices dissimilar to the cable environment in order to support VoIP services in the wireless environment In this paper, we propose a handoff technique using the session timer function of the VoIP. The technique has the handoff state in addition to the three call connection states to avoid misunderstanding the call connection state by the handoff delay.

Mobile devices unlike fixed devices have generally the handoff delay during the locomotion. If the delay is treated as the latency caused by the unstable call connection state, although the call connection state is good, the number of message transmissions is increased because the short period of UPDATE request messages for the SIP. As the result, it causes high traffic conditions. In addition, frequent UPDATE message transmissions to unravel un-received packets because of the handoff delay make the call connection state unstable.

We examine the number of call connection state changes by the number of message transmissions and by the number of handoffs with the suggested technique and the previous research. In consequence, the difference of the number of message transmissions is low with the low number of handoffs. the difference, however, is high with the high number of handoffs. Therefore, the suggested technique is more efficient for the environment with many mobile devices in motion. These comparisons, nevertheless, are just based on the handoff probability. It is need to proceed additional researches considering the delay time for substantive call connection states with the Quality of Service (QoS).

References

[1] Lin, Y.-B., et al, "All-IP Approach for Third Generation Mobile Networks. IEEE Network," 2002.

[2] D. Johnson, C. E. Perkins, and J. Arkko, "Mobility Support in IPv6," IETF RFC, 3775, 2004.

[3] "Understanding Wi-Fi," HP White Paper, 2002.

[4] C. Perkins, D. Johnson, "Route Optimization in Mobile IP," internet Draft, Internet Engineering Task Force, Feb. 2000.

[5] Understanding Wi-Fi and WiMAX as Metro-Access Solutions," Intel White Paper. 2004.

[6] IEEE "Recommended Practice for Multi-Vendor-Access Point Interoperability via an Inter-Access Across Distribution Systems Supporting IEEE 802.1 Operation," IEEE Standard 802.11, 2003.

[7] Ai-Chun Pang, et al. "A study on SIP Session Timer for Wireless VoIP," Wireless Communications and Networking Conference, 2005 IEEE, 2005.

[8] C. Perkins, D. Johnson, "Route Optimization in Mobile IP," internet Draft, Internet Engineering Task Force, 2000.

[9] Editor:Lee Garber, "New Mobile Phones Will Work with Cellular and Wi-Fi," IEEE computer Society, pp. 20, 2006.

[10] Understanding Wi-Fi and WiMAX as Metro-Access Solutions," Intel White Paper. 2004.

[11] IEEE 802.15 Working Group for WPAN statndard. http://grouper.ieee.org/groups802/15

Kiyoung, Kim received the B.S. degree in Computer Science from Sangji University in 1996 and M.S and Ph.D. degrees in Computer Science from Soongsil University in 1999 and 2003, respectively. During 1995-1996, he stayed in TGICC Research Laboratory. Currently He is a Assistant Professor in Seoil College, Seoul, Korea.

Jaepyo Park received his BS, MS and Ph.D degrees in computer science from Soongsil University, Seoul, Korea, in 1996, 1998 and 2004, respectively. Currently he is a Researcher of Information Media Technology Research Center in Soongsil University, Seoul, Korea. His research interests include multimedia security, PKI, network security, cyber forensic, and mobile network.