Computers and Mathematics with Applications 64 (2012) 1522–1531

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Efficient hierarchical SIP mobility management for WiMAX networks

Yu-Hsin Cheng [a,c], Fu-Min Chang [b], Shang-Juh Kao [a,*]

[a] Department of Computer Science and Engineering, National Chung Hsing University, Taichung 402, Taiwan

[b] Department of Finance, Chaoyang University of Technology, Taichung 402, Taiwan

[c] Department of Information Networking and System Administration, Ling Tung University, Taichung 402, Taiwan

ARTICLE INFO

Keywords:
Mobility management
WiMAX
Handoff
Session initiation protocol
Hierarchical session initiation protocol

ABSTRACT

By adopting Session Initiation Protocol (SIP) in WiMAX networks, when the mobile node (MN) moves to a foreign network, the MN sends a re-INVITE message to the corresponding node (CN) to re-establish the connection. This re-connection time is the most costly factor for a handoff. To effectively reduce the re-connection latency, a hierarchical SIP (HSIP) mobility management incorporated with MAC layer operations is proposed. As proposed in the HSIP architecture, several Base Stations (BSs) are collectively managed by an HSIP server to form an administration domain. When an MN roams within a domain, which is the most common mobility case, a re-INVITE message is not necessary, hence a significant traffic reduction can result. To demonstrate the applicability of the proposed HSIP mobility mechanism, an evaluation using the NS2 simulator was performed. Handoff delay and signaling overhead are investigated in both single-handoff and multiple-handoff occurrences. When the ratio of intra-domain to inter-domain handoffs is increased from 1 to 14, the proposed HSIP mobility mechanism can improve up to 13% in average handoff delay and 35% in average signaling overhead as compared with traditional SIP mobility management.

1. Introduction

The maturity of IEEE 802.11 technology and its low setup cost has enabled it to be successfully applied to WiFi (Wireless Fidelity) networks. Many cities have now set up wireless network access points to offer users the convenience of always-on Internet [1]. The technology's limited coverage and scalability, however, have limited wireless Internet access to specific areas. The specification of the IEEE 802.16 [2] WiMAX (Worldwide Interoperability for Microwave Access) standard provides for coverage of up to 40 miles and a transfer rate of up to 70 Mbps. The superior coverage and scalability means that it will gradually become the network technology of choice for future IP-based mobile communications networks [3]. Two fundamental standards are supported by IEEE 802.16. One is IEEE 802.16-2004, used for fixed wireless services, and the other is IEEE 802.16e [4], used for mobile communications. Users can either directly connect to the Base Station (BS) or indirectly connect to the BS through a Subscriber Station (SS).

In the All-IP network, support for mobility management is urgently needed. Currently, the Mobile IP [5] and SIP [6] protocols are the two most commonly used mobility mechanisms. Mobile IP may suffer from the triangular routing problem. In Mobile IP, when an MN moves to a foreign network, it is associated with a Care of Address (CoA) by the foreign network agent. The MN must register its CoA with the Home Agent (HA) which will receive all packets sent to the MN and then use tunneling to pass the packets to the MN through its current CoA. Using mobile IP, the involvement of the corresponding node

* Corresponding author.
E-mail addresses: sjkao@cs.nchu.edu.tw, gaulin2010@gmail.com (S.-J. Kao).

0898-1221/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2012.03.107

(CN), the MN, and the HA may suffer from the triangular routing problem. If an MN's movement causes frequent changes in its CoA, the corresponding handoffs could result in a large number of dropped packets. A dedicated software proxy module is required to pass packets sent to the MN's fixed address to its new location. For services that promise consistently high QoS, this proxy forwarding implies a guaranteed delivery requirement, making it unsuitable for real-time applications.

The SIP protocol [6] is used for setting up and controlling voice transmission over the network. The SIP Mobility mechanism is used to meet the mobility requirement for SIP voice telephony. Using the SIP mechanism for mobility support, packet delay, or packet loss may interrupt mobile voice conversation [6]. Reducing handoff delay during mobile voice conversation is therefore essential to maintaining conversation quality. In the original SIP, if the MN is moving frequently, then the number of re-INVITE requests that the MN must send to the CN increases. This may increase system load or even cause handoff to fail.

In this paper, a hierarchical SIP (HSIP) architecture is proposed. In HSIP, a WiMAX network is partitioned into several domains, each contains an HSIP server and multiple BSs. And, each MN retains two addresses. One is the Local Address (LAddr) and the other is the External Address (EAddr). The LAddr of the MN is registered to the SIP Local Registrar and the EAddr of the MN is registered to the SIP Home Registrar. Through the address mapping mechanism, packets can be forwarded from the MN's EAddr to its LAddr. Only if MN moves to a different domain, does it need to change its EAddr and send a re-INVITE to re-establish its connection with the CN. As the re-connection time is the most costly factor for a handoff, applying the proposed design, the number of re-INVITE requests could be reduced. Especially when the MN is moving within a domain, which is the most common mobility case, no re-INVITE message is necessary. Consequently, the handoff delay can be shortened.

The rest of this paper is organized as follows. Section 2 introduces related work in mobility management. Our proposed system architecture and mobility management are presented in Section 3. Simulation results and performance analysis are included in Section 4. Finally, conclusions are given.

2. Related work

IEEE 802.16 technology also well known as WiMAX is poised to support the next step in the wireless evolution [7]. IEEE 802.16e technology is further used for mobile communications [8]. IEEE 802.16e proposes several mechanisms to facilitate handoff between MN and BS. For example, BS could notify an MN of nearby BSs. The MN can then measure the signal strength of potential BSs and communicate with them in advance. When the handoff is in progress, the current serving BS uses a backbone network to pass information about the MN to the target BS. The above mechanism can reduce the time spent on signal measurements and handshaking between the target BS and the MN. In general, a handoff can be initiated either from the MN using an MOB_MSHO_REQ or from the serving BS using an MOB_BSHO_REQ. Either way triggers a similar series of message exchanges. The message flows of the serving BS in initiated handoff are shown in Fig. 1.

In Fig. 1, the serving BS first regularly broadcasts an MOB_NBR_ADV containing information about nearby BSs, such as channel frequency, a channel's low-level parameters, and network services. This provides the MN with a list of potential BSs it can connect to, and the MN uses its own criteria such as Carrier to Interference plus Noise Ratio (CINR) or the Received Signal Strength Indication (RSSI) to select the target BS. If a handoff is confirmed, the MN sends an MOB_MSHO_REQ to the serving BS and requests handoff. After the MN receives an MOB_MSHO_RSP from the serving BS, it sends the MOB_HO_IND to the serving BS to advise that it is switching to another BS.

When MN moves from its serving BS to the target BS, the MAC Layer switch from the original channel to the current channel is first performed. Once MN completes its MAC layer handoff, if both serving BS and target BS are on different IP domains, extra messages must be exchanged over the Network Layer to update the MN's current IP address. Only then can a CN continue to communicate with the MN.

Mobile IP offers a solution to handoff in the Network Layer and supports mobility without changing the existing TCP protocol. It has therefore become the de facto standard for solving the mobility problem in the 802.16e environment [9–12]. Although Mobile IP successfully supports non-real-time services such as FTP and HTTP, it is not well suited to real-time services such as VoIP and video conferencing.

Many real-time application services now apply the SIP standard for establishing multimedia connections [13–17]. SIP offers a comprehensive mobility management mechanism that supports Terminal Mobility, Session Mobility, Personal Mobility, and Service Mobility [18]. Its main advantages are that no changes are required to the operating system and no fixed IP address is needed to support mobile communications. The SIP registration mechanism is similar to that of Mobile IP, but instead of a fixed IP address, SIP combines the user's SIP Uniform Resource Identifier (URI) with a temporary IP address for identification. Once the MN starts roaming, it re-registers this identification with the SIP Server. This allows the SIP Server to track the MN's current location. Once the MN has acquired a new IP address from the new network, it can send a re-INVITE to the CN to re-establish the connection. The message flows for this process are shown in Fig. 2.

Using SIP, whenever the MN moves to a different network domain, it acquires a new IP address and the connection must be re-established. When handoff occurs, delays are introduced in different layers [6]: a link layer delay, a moving detection delay, a new IP assignment delay, an auto configuration delay, an SIP re-INVITE delay, and a Real-time Transport Protocol (RTP) packet transmission delay. This is the main weakness of SIP mobility support. The new IP assignment delay in particular will exceed 1 s, whereas the SIP re-INVITE delay depends on the distance between two endpoints. If the MN is moving frequently, it may result in additional re-INVITE requests, which causes message overhead and unacceptable delay.