HANDOFF MANAGEMENT: A Critical Function in Mobility Management for Fourth Generation (4G) Wireless Networks

Nwachi-Ikpor

Received: 16 December 2013 Accepted: 5 January 2014 Published: 15 January 2014

Abstract

Efficient mobility management techniques are critical to the success of next-generation wireless systems. Handoff management, which is one of the two basic functions of mobility management, has become more critical in fourth generation wireless networks which support multimedia services. The paper treats basic issues involved in handoff management aspect of general mobility management in wireless communication systems. The relevance of mobility management, handoff management, and general mobility management protocols are explained. The taxonomy of handoff mechanisms, causes of delays in handoffs, and security in handoff procedures are elicited. The paper concludes highlighting some open areas of research in providing seamless services.

Index terms—

One of the research challenges for nextgeneration wireless systems is the design of intelligent mobility management techniques that take advantages of IP-based technologies to achieve global roaming among various wireless networks. Mobility management enables mobile wireless networks to locate roaming terminals for call delivery and to maintain connections as the terminal is moving into a new service area. Thus, mobility management supports mobile terminals (MTs), allowing users to roam while simultaneously offering them incoming calls and supporting calls in progress (Akyildiz & Ho, 1996).

Mobility management contains two components: location management and handoff management. Location management enables the system to track the attachment points of MTs between consecutive communications. Handoff (or handover) management enables the network to maintain a user's connection as the MT continues to move and change its access point to the network. Moreover, when a user is in the coverage area of multiple wireless networks, for example, in heterogeneous wireless environments, handoff management provides always best connectivity (Gustafsson, 2003) to the user by connecting the user to the best available network (Zhang, 2003). In nextgeneration wireless systems, there are two types of mobility for MTs: intra-system (intra-domain) and intersystem (inter-domain) mobility. Intra-system mobility refers to mobility between different cells of the same system. Intra-system mobility management techniques are based on similar network interfaces and protocols. Inter-system mobility refers to mobility between different backbones, protocols, technologies, or service providers. Based on intra-and inter-system mobility, the corresponding location management and handoff management can be further classified into intra-and inter-system location management and handoff management.

Efficient mobility management techniques are critical to the success of next-generation wireless systems. Efficient location management design implies minimized signaling overhead for location update and paging as well as minimized update and paging delay. particular, handoff latency is critical for real-time applications such as voice, real-time video, and streaming services and packet loss during handoff is important for both real-time and non real-time applications. Hence, handoff management has become more critical in fourth generation (4G) wireless networks which support multi-media services. For instance, services such as FTP require zero packet loss during handoff. Similarly, Internet-based gaming services require very low handoff latency. Therefore, efficient handoff management design implies minimized handoff failure rate, packet dropping rate, and handoff latency. In addition, Quality-of-Service (QoS) requirements, scalability, and robustness are also important.

With the increasing demands for new data and real-time services, wireless networks should support calls with different traffic characteristics and different Quality of Service (QoS) guarantees. In addition, various wireless technologies and networks exist currently that can satisfy different needs and requirements of mobile users. Since these different wireless networks act as complementary to each other in terms of their capabilities and suitability for different applications, integration of these networks will enable the mobile users to be always connected to the best available access network depending on their requirements.

This integration of heterogeneous networks will, however, lead to heterogeneities in access technologies and network protocols. To meet the requirements of mobile users under this heterogeneous environment, a common infrastructure to interconnect multiple access networks will be needed. Although IP has been recognized to be the de facto protocol for nextgeneration integrated wireless, for inter-operation between different communication protocols, an adaptive protocol stack is also required to be developed that will adapt itself to the different characteristics and properties of the networks (Akyildiz et al., 2004a). Finally, adaptive and intelligent terminal devices and smart base stations (BSs) with multiple air interfaces will enable users to seamlessly switch between different access technologies.

For efficient delivery of services to the mobile users, the next-generation wireless networks require new mechanisms of mobility management where the location of every user is proactively determined before the service is delivered. Moreover, for designing an adaptive communication protocol, various existing mobility management schemes are to be seamlessly integrated. Each of these schemes utilizes IP-based technologies to enable efficient roaming in heterogeneous network (Chiussi et al., 2002). Therefore, efficient handoff mechanisms are essential for ensuring seamless connectivity and uninterrupted service delivery.

1 II. Issues in Wireless Network Handoff Management a) Importance of Mobility Management

Mobility in wireless networks can take different forms (Akyildiz et al, 1996) such as:

? Terminal Mobility: the ability for a user terminal to continue to access the network when the terminal moves;

? User Mobility: the ability for a user to continue to access network services from different terminals under the same user identity when the user moves;

? Service Mobility: the ability for a user to access the same services regardless of where the user is

In addition, a terminal or a user may be considered by a network to have "moved" even if the terminal or the user has not changed its physical location. This may occur when the terminal switched its connection from one type of wireless network to another, e.g., from a wireless local area network to a cellular network.

Mobility management is the fundamental technology to enable the seamless access to nextgeneration wireless networks and mobile services. Future IP-based wireless networks support all types of multimedia services including real-time services such as voice and video streaming as well as non-real-time services such as email, web-browsing, and FTP. Basic requirements of mobility management in next-generation wireless networks should include: first, the support of all forms of mobility; second, the support of mobility for both real-time and non-real-time applications; third, the support of users seamlessly moving across heterogeneous wireless networks in the same or different administrative domains; fourth, the support of an on-going user application session to continue without significant interruptions as the user moves. This session continuity should be maintained when a user changes its network attachment points or moves from one type of wireless network to another; and last, the support of global roaming, i.e. the ability for a user to move into and use different operators' networks. Finally, location management in next-generation wireless networks is critical to provide location based services.

In order to satisfy the above requirements, nextgeneration wireless systems with mobility management should have two basic functional capabilities:

? Location Management: This is a process that enables the system to determine a mobile device's current location, i.e. the current network attachment point where the mobile device can receive traffic from the system.

? Handoff Management: This is a process that Similarly, efficient handoff management support implies minimum latency and packet loss during handoff. In change involves the roaming into another network with a different operator, then network access control is also involved in the handoff process. Network access control includes authentication (verify the identity of a user), authorization (determine whether a user should use the network service), and accounting (collecting information on the resources used by a user).

2 b) Handoff Management

Handoff management is the process by which a mobile node keeps its connection active when it moves from one access point to another. There are three stages in a handoff process.

First, the initiation of handoff is triggered by either the mobile device, or a network agent, or the changing network conditions. The second stage is for a new connection generation, where the network must find new resources for the handoff connection and perform any additional routing operations. Finally, data-flow control needs to maintain the delivery of the data from the old connection path to the new connection path according to the agreed upon QoS guarantees. Depending on the movement of the mobile device, it may undergo various types of handoff. In a broad sense, handoffs may be of two types: (i) intra-system handoff (horizontal handoff) and (ii) inter-system handoff (vertical handoff). Handoffs in homogeneous networks are referred to as intra-system handoffs. This type of handoff occurs when the signal strength of the serving BS goes below a certain threshold value. An intersystem handoff between heterogeneous networks may arise in the following scenarios (Mohanty, 2006) -(i) when a user moves out of the serving network and enters an overlying network, (ii) when a user connected to a network chooses to handoff to an underlying or overlaid network for his/her service requirements, (iii) when the overall load on the network is required to be distributed among different systems.

The design of handoff management techniques in all-IP based next-generation wireless networks must address the following issues: (i) signaling overhead and power requirement for processing handoff messages should be minimized, (ii) QoS guarantees must be made, (iii) network resources should be efficiently used, and (iv) the handoff mechanism should be scalable, reliable and robust.

3 c) General Mobility Management Protocols

Mobile IP is the most widely used protocol for macro-mobility management. In addition to Mobile IP, three macro-mobility architectures are discussed in the section. These protocols are: Session Initiation Protocol (SIP)-based mobility management, multi-tier hybrid SIP and Mobile IP protocol, and network inter-working agent-based mobility protocol.

i. Mobile IP Mobile IP **??**Perkins, 2008) is the most well-known macro mobility scheme that solves the problem of node mobility by redirecting the packets for the MN to its current location. It introduces seven elements: 1. Mobile node (MN) -a device or a router that can change its point of attachment to the Internet. 2. Correspondent node (CN) -the partner with which MN communicates. 3. Home network (HN) -the subnet to which MN belongs. 4. Foreign network (FN) -the current subnet in which the MN is visiting. 5. Foreign agent (FA) -provides services to the MN while it visits in the FN. 6. Care-of-address (CoA) -defines the current location of the MN; all packets sent to the MN are delivered to the CoA. 7. Mobile IP protocol has three steps: (i) agent discovery, (ii) registration, and (iii) routing and tunneling.

Over the past several years a number of IP micro-mobility protocols have been proposed, designed and implemented that complement the base Mobile IP (Campbell & Gomez, 2001) by providing fast, seamless and local handoff control. IP micro-mobility protocols are designed for environments where MHs changes their point of attachment to the network so frequently that the base Mobile IP mechanism introduces significant network overhead in terms of increased delay, packet loss and signaling. For example, many real-time wireless applications, e.g. VOIP, would experience noticeable degradation of service with frequent handoff. Establishment of new tunnels can introduce additional delays in the handoff process, causing packet loss and delayed delivery of data to applications. This delay is inherent in the round-trip incurred by the Mobile IP as the registration request is sent to the HA and the response sent back to the FA. Route optimization (Perkins & Johnson, 2001) can improve service quality but it cannot eliminate poor performance when an MH moves while communicating with a distant CH. Micro-mobility protocols aim to handle local movement (e.g., within a domain) of MHs without interaction with the Mobile IP-enabled Internet. This reduces delay and packet loss during handoff and eliminates registration between MHs and possibly distant HAs when MHs remain inside their local coverage areas. Eliminating registration in this manner also reduces the signaling load experienced by the network. The micro-mobility management schemes can be broadly divided into two groups: 1. tunnel-based schemes and 2. routing-based schemes. In tunnel-based approaches, the location database is maintained in a distributed form by a set of FAs in the access network. (Misra et al., 2002) are tunnel-based micromobility protocol. Routing-based approaches forward packets to an MH's point of attachment using mobilespecific routes. These schemes introduce implicit (snooping data) or explicit signaling to update mobilespecific routes. In the case of Cellular IP, MHs attached to an access network use the IP address of the gateway as their Mobile IP CoA. The gateway decapsulates packets and forwards them to a BS. Inside the access network, MHs are identified by their home address and data packets are routed using mobile-specific routing without tunneling. Cellular IP (CIP) and handoff-aware wireless access Internet infrastructure (HAWAII) are routing-based micro-mobility protocols.

4 Global Journal of Computer Science and Technology

5 d) Handoff Management Protocols

Handoff or handover is a process by which an MN moves from one point of network attachment to another. Handovers can be classified as either homogeneous or heterogeneous. A heterogeneous handover occurs when an MN either moves between networks with different access technologies, or between different domains. As the diversity of available networks increases, it is important that mobility technologies become agnostic to link layer technologies, and can operate in an optimized and secure fashion without incurring unreasonable delay and complexity. Supporting handovers across heterogeneous access networks, such as IEEE 802.11 (Wi-Fi), global system for mobile communications (GSM), code-division multiple access (CDMA), and worldwide interoperability for microwave access (WiMAX) is a challenge, as each has different quality of service (QoS), security, and bandwidth characteristics. Similarly, movement between different administrative domains poses a challenge since MNs need to perform access authentication and authorization in the new domain. Thus, it is desirable to devise a mobility optimization technique that can reduce these delays and is not tightly coupled to a specific mobility protocol. In this section, we describe different types of handovers and investigate the components that contribute to a handover delay. Some inter-technology and media-independent handover frameworks are then described.

6 e) Taxonomy of Handoff Mechanisms

Different types of handovers may be classified based on three parameters as follows: (i) subnets, (ii) administrative domains, and (iii) access technologies. Inter-technology: this type of handover is possible with an MN that is equipped with multiple interfaces supporting different technologies. An inter-technology handover occurs when the two points of attachment use different access technologies. During the handoff, the MN may move out of the range of one network (e.g., Wi-Fi) into that of a different one (e.g., CDMA). This is also known as vertical handover.

7 i. Intra-technology

This type of handoff occurs when an MN moves between points of attachments supporting the same access technology, such as between two Wi-Fi access points. An intra-technology handover may happen due to intra-subnet or inter-subnet movement and thus may involve the layer 3 trigger.

ii. Inter-domain When the points of attachment of an MN belong to different domains, this type of handoff takes place. A domain is defined as a set of network resources managed by a single administrative entity that authenticates and authorizes access for the MNs. A administrative entity may be a service provider or an enterprise. An inter-domain handover possibly involves an inter-subnet handover also.

iii. Intra-domain Handovers of this type occurs when the movement of an MN is confined within an administrative domain. Intra-domain movement may also involve intrasubnet, inter-subnet, intra-technology, and/or intertechnology handovers as well. iv. Inter-subnet An inter-subnet handover occurs when the two points of attachment belong to different subnets. The MN acquires a new IP address and possibly undergoes a new security procedure. A handover of this type may occur along with either an inter-or an intra-domain handover and also with either an inter-or an intratechnology handover.

v. Intra-subnet An intra-subnet handover occurs when the two points of attachment belong to the same subnet. This is typically a link layer handover between two access points in a WLAN networks, or between different cell sectors in cellular networks. It is administered by the radio network and requires no additional authentication and security procedures. Year 2014

8 E

Each FA reads the incoming packet's original destination address and searches its visitor list for a corresponding entry. If an entry exists, it is the address of next lower level FA. The sequence of visitor list entries corresponding to a particular MH constitutes the MH's location information and determines the route taken by downlink packets. Mobile IP regional registration (MIP-RR) **??**Fogelstroem et al., 2006), hierarchical Mobile IP (HMIP), and intra-domain mobility management protocol (IDMP).

i. Link layer delay Depending on the access technology, an MN may go through several steps with each step adding its contribution to the overall delay before a new link is established. For example, a Wi-Fi link goes through the process of scanning, authentication, and association before being attached to a new access point. For intrasubnet handovers, where network layer configurations are necessary, link layer contributes the maximum to the overall delay.

ii. Network layer delay After completion of the link layer procedures, it may be necessary to initiate a network layer transition. A network layer transition may involve steps such as: acquiring a new IP address, detecting a duplicate address, address resolution protocol (ARP) update, and subnet-level authentication.

iii. Application layer delay The delay of this type is due to reestablishment and modification of the application layer properties such as IP address while using session initiation protocol (SIP). The authentication and authorization procedure such as extensible authentication protocol (EAP) includes several round-trip messages between the MN and the authentication authorization and accounting (AAA) server causing delay in handoff.

9 g) Security in Handoff Procedures

Whenever an MN connects to a point of network access, it establishes a security context with the service provider. During the handover process, some or all the network entities involved in the security mechanism may change. Thus the current security context changes as well. The MN and the network have to ensure that they still communicate with each other and they agree upon the keys to protect their communication. However, during handovers in networks like GSM/GPRS and UMTS no authentication is used. This makes the handover procedures vulnerable to a hijacking attack. An attacker can masquerade as an authentic mobile station (MS) just by sending message at the right frequency and time slot during handover. As long as the attacker does not know the encryption and/or integrity keys currently being used, he cannot insert valid traffic into the channel. However, if an attacker can gain access to the key(s) (e.g. because of a missing protection on the backbone network), he can impersonate the MS. In fact, in GSM/GPRS, UMTS and WLAN networks, no standard protection mechanism in the backbone network has been specified. Many GSM operators do not protect the radio link between their fixed networks and the BSs. In UMTS, during a handover, the keys used to protect the traffic between the MS and the previous BS are reused in communication with the next BS. While the keys are being transmitted, they can be intercepted by an adversary, if the wireless link is not protected.

Usually an authentication process happens before location updates and call setups. The same mechanisms cannot however, be applied in establishing connection during a handover process because of the stringent time constraint. In GSM, for example, the time between the handover command and the handover complete or handover failure message is restricted to 0.5-1.5 s. The generation of an authentication response, however, takes about 0.5 s at the MS side. Thus an authentication overhead will cause connection disruption. Mobility and Handoff Management in Wireless Networks 481. As we have seen earlier in this chapter, efficient cell prediction mechanisms can reduce the signaling overhead between the MS and the old BS. The free time slots may be used to forward authentication traffic between the MS, the old BS and the new BS. The MS can precompute an authentication challenge and the encryption and integrity protection keys before the actual change of channel. When the MS and the new BS establish connection, the MS sends the pre-computed authentication response for the new BS to check. If the checking yields positive results, a handover complete message is sent and the old BS releases its resources. Otherwise, a handover failure happens and the MS falls back to the old channel.

10 III.

11 Conclusion

This paper has discussed some essential issues on handoff management in the general context of mobility management in next-generation mobile wireless networks. The mobile IP has been seen as the most widely used protocol for macro-mobility scheme that solves the problem of node mobility.

Future wireless network will be based on all-IP framework and heterogeneous access technologies. Design of efficient handoff management mechanisms will be playing ever important role in providing seamless services. Some open areas of research that will play dominant role include QoS issues, user terminals, handoff management in wireless overlay networks, and cross-layer optimization. [1]

Figure 1:

[Mohanty (2006)] 'A new architecture for 3G and WLAN integration and inter-system handover management. Wireless Networks'. S Mohanty . *IEEE Personal Communications* 2006. November, 2006. 12 (6) p. .

[Akyildiz (2004)] 'A Survey of mobility management in next-generation all-IP-based wireless systems'. F Akyildiz . *IEEE Wireless Communications* 2004. August, 2004.

[Akyildiz (2004)] 'Adapt net: An adaptive protocol suite for the next-generation on wireless internet'. I F Akyildiz . *IEEE Wireless Communications Magazine* 2004a. March, 2004. 42 (3) .

[Gustafsson and Johnson ()] 'Always best connected'. E Gustafsson , A Johnson . *IEEE Wireless Communications* 2003. 2003. 10 (1) p. .

[Zhang ()] *Design and evaluation of a handover "decision strategy for 4G mobile networks*, W Zhang . 2003.

[Misra ()] 'IDMP-based fast handoffs and paging in IP-based 4G mobile networks'. A Misra . *IEEE Communications Magazine* 2002. 4 (3) .

[Camphbell and Gomez ()] 'IP micromobility protocols'. A Camphbell , J Gomez . *ACM SIGMOBILE Mobile Computing and Communication Reviews* 2001. 4 (4) p. .

[Chiussi ()] 'Mobility management in 3G all-IP networks'. F Chiussi . *IEEE Communications magazine* 2002. September, 2000. 40 (9) p. .

[Akyildiz and Ho (1996)] 'On location management for personal communications networks'. F Akyildiz , J S Ho . *IEEE Communications Magazine* 1996. September, 1996.

[Perkins and Johnson ()] *Route optimization in mobile-IP internet draft*, C Perkins , D Johnson . 2001.