A study of routing-based distributed mobility management in supporting seamless data transmission in smart cities

Sunguk Lee, Ronnie D. Caytiles, Byungjoo Park
Department of Multimedia Engineering, College of Engineering, Hannam University, Daejeon, South Korea

ABSTRACT
The current trend of people having their multimedia-capable wireless and mobile devices roam within highly urbanized areas (i.e., smart cities) has led to the wide deployment of overlapping heterogeneous wireless network technologies. In this regard, various challenges have emerged, such as the number of mobile devices that connect or disconnect a wireless network domain, the diversity of time of their connections, frequent handovers as mobile devices frequently change their locations, the heterogeneity of wireless network technologies, cooperation challenges between the overlapping heterogeneous wireless network technologies, the increasing volume of multimedia traffic, security and privacy issues, and many more. This paper focuses on the deployment of a routing-based distributed mobility management (DMM) scheme to address the constraints and limitations of centralized wireless architectures for smart cities. The comparative analysis with centralized mobility management solutions shows significant alleviation in performance as to handover latency and packet losses, thus providing seamless handovers to maintain quality of service (QoS) for multimedia services.

Keywords:
Data transmission
Distributed mobility management
Routing-based DMM
Smart city
Wireless networks

1. INTRODUCTION
Wireless technologies have opened a new computing paradigm that has and will continue to change the way people will be dealing with their daily living activities associated with banking, shopping, food delivery services, entertainment, education, government services, and other related services. Nowadays, most transactions and services can be done online, and often through wireless and mobile networks. People just need their multimedia-capable devices with them all the time to access the applications and services that they need, and they can do so even when they are on the move (e.g., while on their car, on the bus, or while riding trains).

In addition, the notion of people moving and living in highly urbanized areas (i.e., smart cities) has brought a significant increase in the number of roaming multimedia-capable wireless and mobile devices [1], [2]. The volume of multimedia traffic for wireless network applications and services also continues to rise. In this regard, multimedia applications and services are challenged with the increasing volume of multimedia traffic, diversity of data transmitted, the number of handovers as mobile devices tend to move in and out of the coverage range of various access networks, scalability, cooperation between wireless network technologies, and many more.

Thus, in order to compensate for these challenges, various heterogeneous wireless technologies have been widely deployed with overlapping coverage areas. These heterogeneous wireless access networks can be
under different domains, administrations, management, and internet service providers (ISPs). Moreover, the core architecture for the deployment of such wireless access networks gradually changes from hierarchical and centralized systems towards distributed systems [3]. The move towards distributed mobility management (DMM) systems has alleviated the challenges and limitations that were inherent with hierarchical and centralized systems, such as scalability constraints, reliability constraints, sub-optimal routing, a higher packet loss rate, higher signal overhead, a lack of granularity in the mobility management service, and a more complex network deployment [4], [5].

In DMM architectures, both client-based and network-based architectures have been largely deployed to compensate for the identified limitations and constraints of centralized mobility management architectures. These approaches can be practical solutions that could provide reliable data transmissions while addressing the inherent limitations identified with centralized mobility management (CMM) solutions. However, as these approaches utilize tunneling methods, further limitations may also be acquired in terms of increased packet sizes and processing time; thus, it may result in longer handover latency that can cause delays, retransmissions, and packet losses.

This paper focuses on the deployment of routing-based distributed mobility management support for multimedia-capable devices roaming over heterogeneous wireless networks in highly urbanized areas, such as smart cities. This approach employs a tunnel-free routing path for delivering data packets to roaming multimedia-capable mobile devices over heterogeneous wireless networks. In this scheme, multiple and distributed access routers (Ars) are deployed at the access network level in order to minimize a network bottleneck and address the single point of failure issue that may occur with the CMM solutions.

The rest of this paper is organized as follows: section 2 outlines the overview of hierarchical and CMM, highlighting their limitations and the need for a robust distributed mobility management scheme; the routing-based DMM support for multimedia-capable wireless and mobile devices is outlined in section 3. The comparative analysis is discussed in section 4. Finally, section 5 outlines the concluding remarks.

2. THE NEED FOR DISTRIBUTED MOBILITY MANAGEMENT SOLUTIONS

The traditional wireless network system architectures are generally based on hierarchical architectures that utilize a central mobility anchor (CMA), which is responsible for the management and control of multimedia-capable mobile devices as well as the routing of multimedia data [5]. The CMA keeps the mapping information between the home address (i.e., permanent identifier) and the care-of address (i.e., the temporary internet protocol (IP) address given when the mobile node (MN) is away from its home network) of the MN. In addition, all packets destined for the MN must traverse the CMA, making the data traffic suboptimal and potentially overloading the CMA with network resources. The overview of the CMM architecture is shown in Figure 1.

CMAs such as the home agent (HA) in mobile IPv6 [6], [7] and local mobility anchor (LMA) in proxy mobile IPv6 [8], [9] provide centralized anchoring for both the control plane and data plane in hierarchical network architectures such as the foreign agent (FA) and mobile access gateway (MAG). The CMM approaches can be manageably efficient as all of the network controls are anchored into a single entity (i.e., the CMA, such as HA or LMA), but some considerations can impede the efficiency of handovers and mobility of MNs in the heterogeneous wireless network. These limitations may include low scalability (e.g., complexity for expansion as modifications on the settings in the CMA may affect the performance on the entirety of the wireless network system), the challenge for reliability (e.g., CMAs can be susceptible to single point of failure), suboptimal routing (e.g., data traffic must traverse the CMAs before reaching its intended destinations), higher packet loss rate (e.g., congestion is expected on CMAs, thus, retransmissions can be frequent), signaling overhead (e.g., a number of mobility-related signaling must be shared by the MNs and the wireless network entities during handover procedures), complex network deployment (e.g., it becomes more complex as controls of data traffic or routing and MN’s mobility management are managed by a single entity which is a CMA), security (e.g., the CMAs can be the focal point for attacks).

DMM refers to the deployment of geographically distributed mobile anchors placed closer to the users [10], [11], as shown in Figure 2. The control and data infrastructures were distributed among these mobile anchors, allowing the MN to move between them. The distributed mobility management systems can be deployed in client-based, network-based, and routing-based architectures.

In a client-based DMM architecture, a number of HAs are widely deployed at the edge of access networks, making them closer to multimedia-capable MNs [12], [13]. The MN utilizes multiple addresses comprised of a permanent address (i.e., home address, or HoA), which is associated with its central HA, and additional addresses (i.e., temporary care-of addresses, or CoAs), associated with and configured at each visited wireless access network. Each active CoA has a binding with the HoA to make sure that the MN continues to be reachable whenever it moves across heterogeneous wireless networks. Whenever the MN moves,
a bidirectional tunnel is then created between the previous HA it was attached to and the HA that the MN is newly attached to, thus guaranteeing the continued delivery of data packets from its ongoing and previous sessions.

As shown in Figure 3, MN initially connects with HA1, setting up HoA as its permanent address and being able to communicate with the first correspondent node (CN1). When the MN moves to the domain of HA2, it configures CoA1 as its new locally anchored and temporary IP address, which will be used for new sessions and communications with CN2. The newly acquired CoA1 must have a binding with the MN’s HoA in order for the MN to have continuous reachability. A bi-directional tunnel is configured between HA1 and HA2 to provide session continuity for the communication between the MN and CN1.

The client-based DMM architecture addressed the high handover latency and packet delay issues for the standard MIPv6; however, the MNs are highly involved in the mobility management configurations (i.e., this is a requirement for client- or host-based mobility management protocols). Thus, extensions and additional intelligence on the MN are required. The network-based DMM architecture is based on the proxy mobile IPv6 (PMIPv6) [7], where the mobility anchors (MAs) are moved to the edge of the wireless access network; that is, it employs ARs closer to the users [14]. It is primarily designed to address the limitations and issues that come with CMM architectures [4], [15], which include limited scalability, locality, suboptimal load balancing and routing, a higher packet loss rate, high signal overhead, deployment complexity, and fault tolerance. These architectures can be implemented in either fully distributed or partially distributed models. In a fully distributed model, both control and data infrastructures are managed by the MAs (i.e., tightly coupled). Thus, each AR implements both LMA (i.e., responsible for anchoring and routing the local data traffic for each MN) and MAG (i.e., responsible for receiving the data traffic of the MNs through tunneling operations) functionalities.
On the other hand, in a partially-distributed model, the data plane is split with the control plane. The model deploys a centralized control plane that is managed by the mobility management entity (MME), which is responsible for managing the mobility sessions of the MNs [16]. The data plane, on the other hand, is distributed and managed by the corresponding gateways. The ARs at the edge of the wireless network are handling the data forwarding [8]. The partially-distributed DMM model is depicted in Figure 4.
The partially-distributed DMM model includes two major entities: (i) the MA and AR (MAAR), which is deployed at a single-hop distance from the MN. Its main purpose is to forward data to and from the wireless networks. In addition, MAAR implements the functionalities of an AR, MAG, and LMA for the data plane in PMIPv6. The MAAR also assigns and anchors an IP prefix to each of the connected MNs. (ii) The centralized mobility database (CMD), which serves as the central entity that stores and manages the MN’s mobility sessions as they roam within the bounds of the DMM domain. In addition, the data packets will no longer go through the CMD as they are transmitted to their intended destinations.

In Figure 4, the MN initially connects with the serving MAAR1, which registers the MN at the CMD using an exchange of a proxy binding update (PBU) and a proxy binding acknowledgment (PBA) message. CN1 then communicates directly with the MN through the configured IP address (e.g., HoA for initial attachment). When the MN moves to MAAR2, a PBU message is sent to the CMD in order to register the MN’s new connection. Then, the CMD transmits directives on setting up the proper routing configuration to both MAAR2 and MAAR1. A bidirectional tunnel between MAAR1 and MAAR2 is then established to provide session continuity between the MN and CN1. Data packets from new sessions (e.g., communications with CN2) are received through the newly acquired IP address (e.g., CoA1) during attachment in MAAR2.

Both client-based and network-based DMM architectures utilize tunneling methods to deliver the intended packets to the MNs. This approach can be a sensible solution for efficient data transmission and address the limitations implicated by the CMM architectures; however, the employed tunneling processes may also incur additional limitations that affect the quality of service (QoS) of data transmissions. The tunneling method requires encapsulation and decapsulation processes that could increase packet sizes as well as the processing time, that is, the time required for establishing tunnels between network entities. Hence, it may result in longer handover latency that can cause delays, retransmissions, and packet losses.

3. THE ROUTING-BASED DMM SUPPORT

This paper focuses on the deployment of a routing-based DMM architecture to support seamless handovers and robust data transmissions for multimedia-capable devices roaming within urbanized areas such as smart cities. The architecture removes the functionality of MAs and allows the standard routing process to configure a new routing table whenever the MN moves [17], [18]. Initially, the MN obtains an IP address (HoA) whenever it attaches to an AR, which will be used in routing data packets all throughout its movement within the DMM domain. The HoA allocated with the MN from the AR it is attached to is through the standard dynamic host configuration protocol (DHCP). A domain name system (DNS) update is then done by the MN in order to bind the acquired HoA to its hostname. On the other hand, the serving AR updates the reverse pointer in ip6.arpa space to map the MN’s hostname to its IP address. This is done in order for the MN and its serving AR to control both the forward and reverse mappings. The obtained IP address is then advertised to its peer ARs within the DMM domain using an intra-domain protocol such as the border gateway protocol (BGP). BGP enables the wireless network to exchange routing and reachability information between autonomous systems.

The operations of the routing-based DMM architecture are summarized in Figure 5. When the MN moves, the new AR it attaches to looks up an IP address using the MN’s hostname that is obtained during the authentication. Then, a reverse lookup is performed by the new AR in order to confirm that a previous AR has already assigned the IP address to the MN’s hostname. A routing update is then performed once it is confirmed. Then, the new AR creates a BGP update message that contains the MN’s IP address and sends it to its peer ARs, announcing the new route for the data packets. Figure 5(a) depicts the data traffic flows while the MN moves across two different ARs within the DMM domain, while Figure 5(b) depicts the handover process when the MN moves its attachment from AR1 to AR2. Since the routing-based DMM makes use of a tunnel-free routing protocol as compared with the client-based and network-based DMM, an optimal route (i.e., through routing updates) can be used between the MNs and their CNs. In addition, session interruptions during MN’s handover can be minimized as the previous AR forwards the data packets to the currently serving AR.

In routing-based DMM architecture, multiple ARs are deployed and distributed one-hop distance with the MNs (i.e., at the edge of the network and closer to the users). The ARs extend the capabilities of HAs and MAARs in client-based and network-based DMMs, respectively, that perform both mobility and routing management for the MNs. The ARs also allocate the IP address (i.e., network prefix) to be used by the MN while on the DMM domain. The MN attaches to another serving AR whenever it is within its range and is almost leaving the range of the previous AR. Session continuity is guaranteed as data packets are continuously delivered using the same routing path (i.e., through the previous AR) as long as the MN is not yet totally out of the previous AR’s coverage area (see data flow number 2 in Figure 5(a)).

The handover mechanism for routing-based DMM support is focused on user mobility, where mobile devices can intelligently route latency-insensitive packets. It avoids disruptions and delays on ongoing user sessions and data packet transfers while disregarding the existence of network infrastructures. In addition, the...
MN’s connection can be established by overlapping regions of two ARs. Thus, the MN can receive data packets via two different routing paths as it communicates with its CNs (see data flow numbers 2 and 3 in Figure 5(a)). The availability of more than one active routing path leads to a more seamless and reliable handover performance.

Figure 5. Routing-based DMM operations (a) traffic flows and (b) handover process
4. DISCUSSIONS

This section outlines the advantages of routing-based DMM for supporting seamless and robust data transmission over heterogeneous wireless networks in smart cities. The primary objectives of this scheme are to address the following drawbacks that generally limit the QoS performance of data transmissions:

i) Suboptimal routing: because the IP address (HoA) being used by the MN is anchored at the CMA, data packets will always traverse this central anchor, which may lead to longer routes as compared with a direct path between the MN and its CNs. With distributed deployment, multiple ARs manage the MN’s mobility, and data packets can be routed optimally without traversing the CMA [19], [20].

ii) Scalability problems: scaling centralized routes and maintaining the mobility context for each MN can become more complex, especially with a large number of MNs. The CMA requires enough processing and routing capabilities to deal with the increasing volume of data traffic for all users simultaneously. With distributed deployment, route maintenance for data packets as well as mobility context maintenance among heterogeneous wireless networks can become more scalable [21].

iii) Reliability: centralized anchoring solutions share the problem of being more susceptible to single-point failure and vulnerable to attacks. With distributed deployment, the network can still be functional even if some of its network entities may already be compromised.

iv) Signaling overhead: with centralized solutions, network resources can be wasted with a lot of exchange of mobility-related signaling messages during handovers. With the routing-based DMM approach [22], the exchange of mobility-related signaling messages is optimized. A single network prefix (i.e., IP address) can be used by the MN while roaming a particular DMM domain.

v) Packet loss rate: data packet congestion on CMAs is always expected, as all data packets must traverse these centralized anchors before they can be delivered to their respective MNs [23]. With the routing-based DMM approach, routes are selected based on the best-optimized data paths.

vi) Complex network deployment: network design for centralized architectures becomes more complex as control of data traffic or routing and MN’s mobility management are managed by a single entity (i.e., CMA) [24]. With the routing-based DMM approach, design can become flattened and simpler, and network expansion can be done by simply adding more ARs capable of handling both mobility management for MNs and routing their data packets.

vii) Security: the centralized anchors can be the focal point for attacks. With the routing-based DMM approach, aside from the use of standard IP security (IPSec) [25], [26], mutual authentication and authorization processes are required between the MNs, and the serving ARs are required to prevent potential attacks.

viii) Wasted resources: tunneling methods in CMM and in both client-based and network-based DMM approaches require network resources (e.g., tunnel maintenance, increased packet sizes for encapsulation, increased processing time for tunnel setup, increased processing time for encapsulation and decapsulation processes). With the routing-based DMM approach, the tunneling process is removed; thus, uninterrupted and continuous sessions can be guaranteed between the MNs and their CNs.

The scheme allows inter-domain operations and is able to work between trusted administrative domains (i.e., the routing-based DMM approach exercises compatibility). It is able to support and co-exist with the existing wireless network deployment. In addition, it can co-exist with other mobility management protocols.

5. CONCLUSION

This paper presents a study of the routing-based DMM approach to supporting seamless data transmissions over heterogeneous wireless networks for smart cities. The model allows the deployment of distributed ARs that are capable of managing both the mobility context of MNs as well as their respective data packets. It allows uninterrupted and continuous ongoing sessions for the users by providing tunnel-free and seamless data transmissions. The scheme can address the drawbacks of existing wireless network infrastructures, specifically those of CMM models, such as suboptimal routing, scalability problems, reliability problems, higher signaling overhead, a higher packet loss rate, complex network deployment, security problems, and wasted resources.

ACKNOWLEDGEMENTS

This work was supported by 2023 Hannam University Research Fund.

A study of routing-based distributed mobility management in supporting seamless … (Sunguk Lee)
REFERENCES


A study of routing-based distributed mobility management in supporting seamless ... (Sunguk Lee)