Wireless Pers Commun (2008) 44:383–401
DOI 10.1007/s11277-007-9363-3

A Mobile Bandwidth-Aggregation Reservation Scheme for NEMOs

Jui-Tang Wang · Yuan-Ying Hsu · Chien-Chao Tseng

Abstract A NEtwork that is MObile (NEMO) usually consists of at least one Mobile Router (MR) attached to the infrastructure to manage all external communication for of all nodes inside a NEMO. Because a NEMO moves as a whole, previous mobile ReSource reserVation Protocols have two problems in supporting quality of services (QoS) for NEMOs; that is, mobility unawareness and excessive signal overhead. In this paper, we first address these two problems and then propose a Mobile Bandwidth-Aggregation (MBA) reservation scheme to support QoS guaranteed services for NEMOs. In order to resolve these two problems, MBA makes an MR the proxy of all nodes insides a NEMO and has the MR aggregates and reserve the bandwidth required for all node inside a NEMO. Mathematical analysis and simulation results show that the proposed MBA scheme can significantly reduce the signal overhead for reservation maintenance. Furthermore we also present three hypothetic policies of tunnel reservations for NEMOs, and conduct simulation to evaluate these policies in terms of blocking probabilities and bandwidth utilizations.

Keywords NEMO · Mobile router · QoS · RSVP · MRSVP · HMRSVP

A preliminary version of this work presented at the 2006 IEEE International Conference on Communications, Volume 2, June 2006, Page(s):693–698

J.-T. Wang (✉) · C.-C. Tseng
Department of Computer Science, National Chiao-Tung University, Hsin-Chu 300, Taiwan
e-mail: rtwang@csie.nctu.edu.tw

C.-C. Tseng
e-mail: cctseng@csie.nctu.edu.tw

Y.-Y. Hsu
Telcordia Applied Research Center, Taiwan Co.,Taipei, Taiwan
e-mail: yyhsu@tarc-tw.research.telcordia.com

1 Introduction

A *NEtwork that is MObile (NEMO)* [1] is a moving network segment or subnet. A NEMO normally consists of at least one mobile router (MR) that is responsible for all external communication of all nodes inside a NEMO, henceforth referred to as NEMO nodes. In the NEMO Basic Support [1], an MR uses Mobile IP [2,3] to ensure session continuity, even as the MR changes its point of attachment to the infrastructure, for all NEMO nodes. With Mobile IP, the MR establishes a bi-directional tunnel between itself and its home agent (HA). The HA of MR (HAoMR) will intercept all data traffic destined for NEMO nodes and delivered to NEMO nodes via the bi-directional tunnel. Oppositely, the MR will intercept and forward all data originated from NEMO nodes to the HAoMR through the tunnel. In other words, the MR handles mobility for the entire NEMO and thus all NEMO nodes can access Internet continuously without any modifications. However, all NEMO nodes move together with an MR when the MR changes its point of network attachment. Such integral mobility makes it difficult for the current mobility-aware Quality of Services (QoS) supporting protocols, such as Mobile ReSource reseVation Protocol (MRSVP) [4] and Hierarchical MRSVP [5], to provide QoS for NEMO nodes. In the following paragraphs, we address two main problems of supporting QoS for NEMO nodes.

The first problem of supporting QoS for NEMO nodes is that NEMO nodes are not aware of the integral mobility. Both MRSVP and HMRSVP achieve QoS requirement for mobile nodes by making advance resource reservations at networks that a mobile node may visit in the near future. However, in a NEMO, the MR is the only network entity that is aware of the changes in network attachments. As a consequence, other NEMO nodes are not aware of network changes. Unfortunately, NEMO nodes are responsible for the maintenance of QoS connections. If NEMO nodes can not detect network changes, they will not be able to retain the reservations.

The other problem of supporting QoS for NEMO nodes is the excessive signal overhead introduced by the NEMO nodes if each individual NEMO node maintains its own resource reservation itself. We noted above that all packets destined for or originated from NEMO nodes will transit through the bi-directional tunnel between the MR and its HAoMR. Therefore, the signals a NEMO node uses to maintain a reservation will need to go through the bi-directional tunnel. If each NEMO node has its own reservation, the maintenance of these individual reservations might introduce excessive signal traffic and waste the scarce bandwidth of the MR's wireless link.

In this paper, we propose a Mobile Bandwidth-Aggregation (MBA) reservation scheme to support QoS for NEMO nodes. In MBA, an MR serves as the resource reservation proxy that handles mobility and reserves bandwidth for the bi-directional tunnel on behalf of the NEMO nodes. Therefore MBA can resolve the mobility un-awareness and excessive signal problems in supporting QoS for NEMO nodes. For simplicity, we explain the MBA scheme in terms of mobile RSVP in this paper. However, it could work with other reservation protocols [6–9] as well.

The rest of this paper is organized as follows. In Sect. 2, we briefly describe the concepts of RSVP [10], RSVP tunnel [11], MRSVP [4], and HMRSVP [5]. We then introduce the problem in using previous mobile RSVP protocols to support QoS for NEMOs. After the problem descriptions, we then propose the MBA scheme in Sect. 3. Section 4 presents our mathematical analysis model and simulation results. Finally, we conclude the paper in Sect. 5.

2 Related Work

In this section, we first present the concepts of RSVP [10], RSVP tunnel [11], and the two representative mobile RSVP protocols, namely MRSVP [4], and HMRSVP [5]. Following the description of MRSVP and HMRSVP, we then introduce the problem in supporting QoS for NEMO Nodes.

2.1 RSVP

RSVP [10] is a signaling protocol for setting up resource reservations in the Internet. RSVP uses two types of messages, Path and Resv, to setup resource reservation states on the nodes along the path between a sender and a recipient. Initially, the sender learns the IP address of the recipient using some out-of-band mechanism and sends a Path message to the recipient to find a path all the way from the sender to the recipient for a particular flow. When a router receives the Path message, it records the upstream router from which it receives the message and forwards the message to a downstream router, relying on the underlying routing protocol. The Path message then traverses downward from one router to another, and finally arrives at the recipient. The recipient responds with a Resv message to make a resource reservation for the specific flow. The Resv message traverses in reverse along the same path as the Path message originally traverses. Upon receiving a Resv message, each router on the path reserves resources for the specific flow if sufficient resources are available.

However, host mobility introduces two impacts on the original RSVP signaling protocol. First, RSVP is not aware of mobility. According to the original RSVP signaling protocol, a mobile node (MN) cannot adjust the resource reservation path dynamically when the MN changes its point of network attachment. In other words, when an MN enters a new network, the resources the MN reserved previously are no longer available and it needs to request resources in the new network. However the request may take too long or even fail to fulfill the QoS required for the MN. Second, IP-in-IP encapsulation makes RSVP messages invisible. Mobile IP uses an IP-in-IP encapsulation [12] to route IP packets to an MN that is away from its home network. This encapsulation will conceal the original RSVP messages, Path and Resv, and make them unrecognizable to the routers along the bi-directional tunnel between the MN and its home agent, Therefore the routers on the path will not reserve resources for the MN.

2.2 RSVP Tunnel

Terzis et al. [11] propose RSVP Tunnel to resolve the RSVP message invisibility problem. The underlying principle of RSVP Tunnel is to establish a nested RSVP session between the tunnel end-points, namely entry and exit points, for each original RSVP session. That is, the original Path and Resv messages (referred to as end-to-end RSVP messages in Sect. 3) will trigger the tunnel entry and exit points to exchange an extra pair of Path or Resv messages, called tunnel Path and tunnel Resv messages. Tunnel Path and Resv messages are carried in normal IP, instead of IP-in-IP datagram, and thus the routers on the tunnel can recognize these messages and reserve resources for the MN accordingly.

Although RSVP Tunnel can resolve the RSVP signal invisibility problem, it still does not resolve the mobility unawareness problem of RSVP. When an MN moves to a new network, the MN needs to request resources in the new network. As a consequence, the MN's service may be disrupted or terminated because the MN may spend substantial time or even fail in making a new resource reservation.

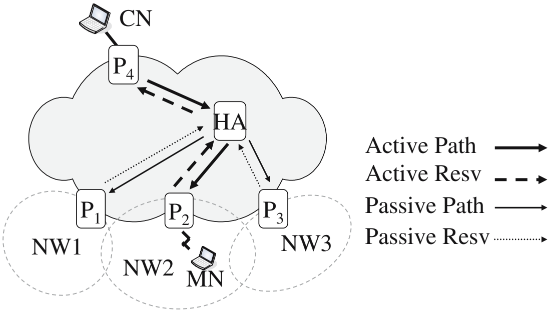


Fig. 1 MRSVP active and passive reservations

2.3 Mobile RSVP

Mobile ReSource reserVation Protocol (MRSVP) is a resource reservation protocol proposed by Talukdar et al. [4] to support multi-media delay-sensitive streaming applications in wireless Integrated Services Internet [13]. The resource reservations made by MRSVP can be classified as either active or passive. An active reservation is on the path to the MN's current location and has an actual data flow, whereas a passive reservation is on the path to a neighbor location of the MN and does not have an actual data flow.

MRSVP attempts to make passive reservations in advance at the neighbor locations where an MN may visit next. If MRSVP reserves the resources for the MN successfully in a neighbor location, the MN can use the pre-reserved resources directly without issuing further resource reservation requests upon entering the location.

MRSVP introduces proxy agents to make resource reservation for the MNs. A proxy agent at an MN's current location is called the *local proxy agent* of the MN; the proxy agents at the MN's neighbor locations are called *remote proxy agents*. The local proxy agent of an MN is responsibility for setting up an active resource reservation whereas the remote proxy agents of an MN are responsibility for making passive reservations. As shown in Fig. 1, a receiver MN, its HA and a corresponding sender node CN exchange Path and Resv messages to establish an active reservation and several passive reservations. In this case, we assume Mobile-IP reversed tunneling [14] is in use and HA serves as the anchor node of MN. Therefore, HA aggregates the MN's active and all passive reservations, and makes only a single reservation with CN.

With passive reservations, an MN can use resources immediately upon entering a new network. However, MRSVP may make unnecessary resource reservations if it reserves bandwidth in advance at all neighbors.

2.4 Hierarchical Mobile RSVP

Hierarchical Mobile RSVP [5] (HMRSVP) adopts the hierarchical concept of Mobile-IP regional registration [15] and makes resource reservation in advance for an MN only when the MN resides in the overlapped area of the boundary cells between two regions. A region could be an enterprise or a campus network that consists of a set of routers or subnets. In the base Mobile-IP protocols [2,3], an MN registers with the MN's HA each time when it changes its point of network attachment. In the cases where the HA is far away, the registration process may become too expensive. In order to support low latency and smooth handoff, Mobile-IP regional registration localizes the registration process within a region when an MN makes an intra-region movement, by organizing Mobility Agents (MAs) in a region hierarchically in accordance with the routing topology of the region. The MA situated on the region gateway is called the Gateway Mobility Agent (GMA) of the region. Due to the hierarchical nature and network-based routing properties of Internet, GMA and regional MAs together can process registrations locally and hide the mobility of a visiting MN from the MN's HA and CNs when the MN moves within the region. Because the reservation setup time within a region is likely to be short, HMRSVP does not make advance resource reservations when MNs move within the region. In stead, HMRSVP makes passive reservations in advance only when an MN may make an inter-region movement when the MN resides in the overlapped area of the boundary cells of two regions.

2.5 Problems in Supporting QoS for NEMO Nodes

As mentioned previously, mobility unawareness and excessive signal overhead are the two problems in supporting QoS for NEMO nodes. A NEMO moves as a whole and the MR is the only network entity of the NEMO that is aware of the net-work changes. Other NEMO nodes are not aware of such integral mobility. The un-awareness of integral mobility of NEMO nodes makes MRSVP or HMRSVP inapplicable to supporting QoS for NEMO nodes because both MRSVP and HMRSVP rely on the MN's ability to detect the network changes.

Even if a NEMO node could detect the network changes, integral mobility will still cause bursty signal problem if each individual NEMO node establishes its own resource reservations. When a NEMO moves from one network to another all NEMO nodes that require QoS will need to make new resource reservations in the new network. The new reservations may be not necessary if the NEMO adopts MRSVP or HMRSVP and make advance reservations beforehand.

However, even with MRSVP or HMRSVP, the maintenance of individual reservation will still cause excessive signal overhead. RSVP requires the sender and the receiver to issue, respectively, a Path and a Resv message periodically to refresh the reservation states. Such periodical refreshing will generate excessive amount of signals if each NEMO node maintains the resource reservations itself.

The mobility unawareness and excessive signals problems make previous mobile RSVP protocols [4,5] inapplicable to supporting QoS for NEMO nodes. Therefore, in the following section, we propose a Mobile Bandwidth-Aggregation (MBA) reservation scheme that can eliminate these two problems and support QoS for NEMO nodes.

3 Mobile Bandwidth-Aggregation Reservation Scheme

As we mentioned previously, all data traffic of a NEMO goes through the bi-directional tunnel between the MR and HAoMR. Therefore, MBA chooses the MR and HAoMR as the resource reservation proxies that handle mobility and reserve bandwidth for the bi-directional tunnel on behalf of the NEMO nodes. Therefore NEMO nodes need not be aware of the integral mobility. Furthermore, the MR and HAoMR can aggregate the Path and Resv messages to eliminate the excessive signal problem.

In the following subsection, we first present the basic concept and the reservation procedure of MBA in Subsects. 3.1 and 3.2. Then we describe MBA Bandwidth Adjustment, Reservation Maintenance and Bandwidth Reservation Policies in Subsects. 3.2–3.4.

3.1 Overview of MBA Reservation Scheme

In this section, we use the example shown in Fig. 2 to explain the MBA reservation scheme. As shown in the Fig. 2, MN_1 and MN_2 are two NEMO nodes that attach to the MR of a

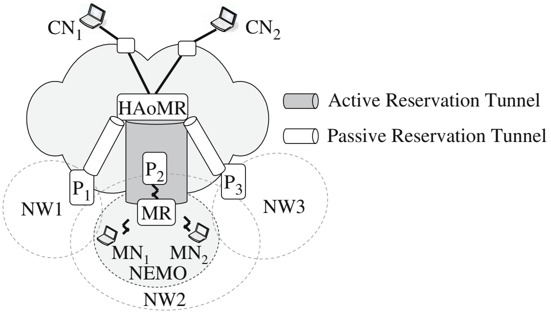


Fig. 2 An example of MBA scheme

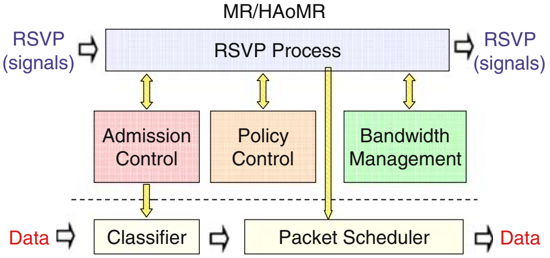


Fig. 3 MBA modules of MR and HAoMR

NEMO that is visiting a foreign network NW$_2$. Assume that MN$_1$ communicates with a corresponding node CN$_1$ first, and then MN$_2$ communicates with CN$_2$. As mentioned previously, all data traffic of MNs goes through the bi-directional tunnel between the MR and HAoMR. In order to provide QoS for MN$_1$ and MN$_2$, the MR and HAoMR will establish an active tunnel reservation in NW2 and a passive tunnel reservation in each of the neighbor networks, NW1 and NW3.

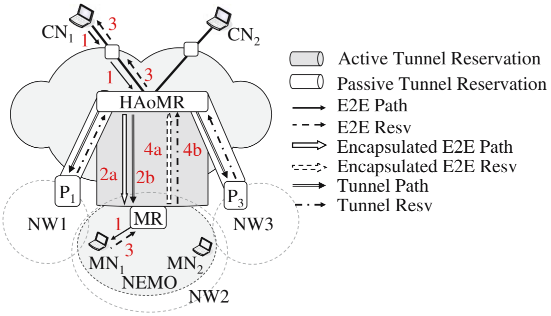
The management of the active/passive tunnel reservations in MBA is exactly the same as the one of the active/passive reservations in MRSVP. The passive tunnel reservation of an MR on the new network will change from passive to active upon the MR moves from a network to another. At the same time, the active tunnel reservation on the old network will become a passive tunnel reservation. If the MR moves away further, the passive tunnel reservation will be obsolete and the resources reserved previously for the passive tunnel reservation will be reclaimed by the old network.

However unlike the active/passive reservations in MRSVP and HMRSVP, the tunnel reservations in the MBA reservation scheme are shared by all NEMO nodes inside a NEMO; the MR and HAoMR of the NEMO are responsible for requesting and managing the resources of the tunnel reservations and forwarding data packets on behalf of the NEMO nodes. Fig. 3 shows the MBA modules on an MR or HAoMR. The MBA modules consist of RSVP Message Process, Admission and Policy Control, Tunnel Bandwidth Management, and Packet Classifier and Scheduler.

When receives an RSVP Path/Resv message, RSVP Message Process performs RSVP Tunneling protocol to configure a tunnel reservation and a nested end-to-end reservation simultaneously. In addition, it is also responsible for periodical refreshing of the tunnel reservations and the end-to-end reservations on behalf of the NEMO nodes.

Tunnel Bandwidth Management estimates the bandwidth required to fulfill the request of NEMO nodes, and informs RSVP Message Process to adjust the tunnel bandwidth dynamically. In addition to the aggregated bandwidth of all active sessions of NEMO nodes, Tunnel

Fig. 4 MBA tunnel reservations (Receiver is a NEMO node)



Bandwidth Management will also request some extra bandwidth, depending on the reservation policy in use, to serve the NEMO nodes that may visit the MR in the future.

Admission Control performs QoS negotiation, and decides whether to reject or grant a new session with a specific QoS class. Policy Control determines how to drop the data packets to fulfill the requested QoS if the bandwidth reserved for the tunnel is not sufficient to deliver data packets in time. Packet Classifier performs packet classification in accordance with the QoS classes of the data packets. Scheduler determines the packet delivery schedule based on the packet classes.

In the following paragraphs, we use Fig. 4 as an example to illustrate the reservation procedures of MBA. Assume that CN$_1$ issues an E2E Path message to MN$_1$ in the NEMO. Following the NEMO Base Support Protocol [1], the message will be intercepted by HAoMR. Further assume that, MRSVP is used to reserve the resources on the path between HAoMR and CN$_1$. Therefore, without loss of generality, we describe only the messages flows and procedures MBA uses to reserve resources for the tunnel between HAoMR and MR, henceforth referred to as tunnel HAoMR-MR.

On receiving the E2E Path message, HAoMR will forward the message to MR via IP-in-IP encapsulation. In addition, HAoMR will also send an Active Tunnel Path message to MR to setup an active tunnel reservation, and a Passive Tunnel Path to each of the proxies in neighbor networks, NW1 and NW3, to setup a passive tunnel reservation for MR.

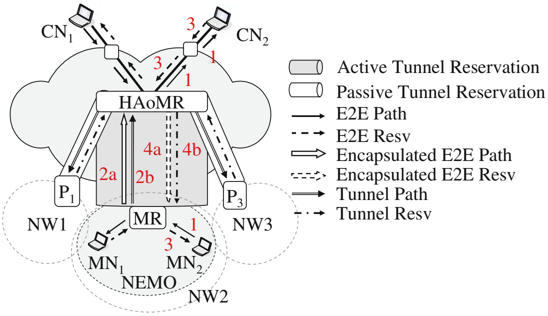
Each router on the path of the tunnel forwards the encapsulated E2E Path message, as a normal IP packet, downstream to MR. On the contrary, the routers on the tunnel will recognize the tunnel Path message as an RSVP message and will perform the path finding function as described in the original RSVP protocol [10].

When MR receives the encapsulated E2E Path message, it decapsulates the message and forwards the original E2E Path message to MN$_1$. However, when MR receives the tunnel Path message, it will not forward the message to MN$_1$, but creates a soft Path state, instead, for HAoMR-MR tunnel.

In response to the original E2E Path message, MN$_1$ replies an E2E Resv message to CN via MR. In a similar way, when MR receives the E2E Resv message, it will encapsulate the message and forward the encapsulated E2E Resv message to HAoMR. In addition, MR will also issue a tunnel Resv message to HAoMR. Thus, all routers on the tunnel path and HAoMR, on receiving the tunnel Resv message, can reserve the desired resources for the tunnel HAoMR-MR if the resources requested by MR are available.

Moreover, when HAoMR receives the encapsulated E2E Resv message, it decapsulates the message and forwards the original E2E REsv message to CN. The E2E Resv message will inform the routers on the path from HAoMR to CN$_1$ to setup a resource reservation for the CN$_1$-MN$_1$ session. As a consequence, an E2E reservation exists between CN$_1$ and MN$_1$

Fig. 5 MBA tunnel reservation (Sender is a NEMO node)



and is nested by a tunnel reservation between HAoMR and MR. Using the above nested RSVP sessions, MR can reserve resources for the HAoMR-MR tunnel and fulfill the QoS requirements of the NEMO nodes (only MN$_1$ now).

Furthermore, HAoMR will also establish two passive reservations with the proxy agents at neighbor networks of NW2 that the NEMO currently stays. Once the NEMO moves, the E2E reservations do not need to be re-established. Instead, MR and HAoMR need only cooperate to switch the passive tunnel reservation that MR visits to active and the original active tunnel reservation to passive. If MR moves further, a passive reservation may become obsolete and the routers and proxies on the path of an obsolete reservation will reclaim the resources after a predefined period.

3.2 MBA Bandwidth Adjustment

Further assume that CN$_2$ initiates a QoS session with MN$_2$ by exchanging a pair of E2E Path/Resv messages through tunnel HAoMR- MR as shown in Fig. 5. If MR and HAoMR have reserved bandwidth that is sufficient to fulfill the request, they simply encapsulate the E2E messages and forward the encapsulated message through the tunnel. In this case, they need not exchange tunnel Path/Resv messages to make a new tunnel reservation or modify the original tunnel reservation.

However if MR and HAoMR have not reserved enough bandwidth, they need to exchange a pair of tunnel Path/Resv messages to request more bandwidth for the tunnel reservation. If there is not enough bandwidth existing on the wireless links of the MR or along the path of the tunnel HAoMR-MR, the Admssion Control module of MR or HAoMR will deny the request. Therefore, the policies of tunnel bandwidth reservation will affect both the signal overhead and blocking probability of the QoS sessions requested by NEMO nodes. In Sect. 3.4, we present three hypothetic reservation policies and discuss the performance of MBA under these policies in Sect. 4.

3.3 MBA Reservation Maintenance

If MBA uses RSVP as the underlying reservation protocol, then the tunnel reservation in MBA are soft reservations; that is, an MR and its HAoMR need to send Tunnel Path/Resv messages periodically to refresh the soft states maintained by the routers along the path of the tunnel. Because a tunnel reservation is shared by all NEMO nodes served by an MR, MBA adopts a batch refreshing scheme to reduce the signal overhead for the maintenance of tunnel reservations.

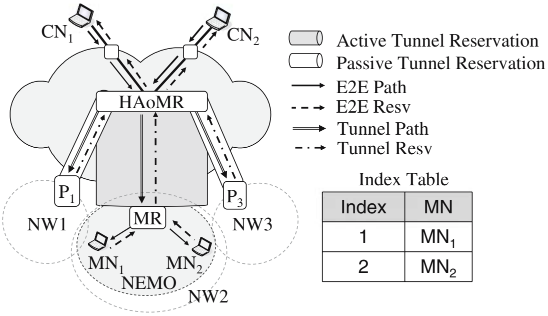


Fig. 6 MBA reservation maintenance

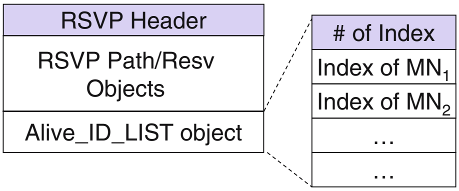


Fig. 7 Alive_ID_List in a tunnel Path/Resv message

Figure 6 shows the MBA reservation maintenance operations. In MBA, MRs, or HAoMRs serve as proxies that refresh the E2E reservations on behalf of the NEMO nodes and send Tunnel Path/Resv messages periodically to keep tunnel reservations alive. Therefore, in addition to the normal RSVP fields, a tunnel Path/Resv message will also carry an Alive_ID_List object as shown in Fig. 7. An Alive_ID_List is an index list of NEMO nodes or an index list of CNs that have an ongoing E2E session inside the tunnel HAoMR-MR.

Each index in an Alive_ID_List object will map to a soft state that contains information for an MR or HAoMR to generate an E2E Path/Resv message on behalf of an NEMO node or a CN. An MR uses the index list to inform its HAoMR the soft states of which NEMO nodes the HAoMR needs to refresh for. Similarly, the HAoMR includes an Alive_ID_List in a Tunnel Path/Resv message to inform the MR the soft states of which CNs the MR needs to refresh for.

The indices are created and refreshed periodically by an MR (or HAoMR) when the MR (or HAoMR) receives E2E Path/Resv messages. The MR (or HAoMR) reclaims the indices if it does not receive matched refreshing messages before the expiration time. In the following paragraphs, we use the above example shown in Figs. 4 and 5 to explain the creation and usages of an Alive_ID_List. Initially, when MR and HAoMR setup a tunnel reservation and a nested E2E reservation for the session CN$_1$-to-MN$_1$, MR and HAoMR will create an index for MN$_1$ and CN$_1$, respectively. Similarly, later when MR and HAoMR request more resources for the tunnel reservation and establish a nested E2E reservation for the session MN$_2$-to-CN$_2$, MR and HAoMR, respectively, will also create an index for MN$_2$ and CN$_2$.

Following the RSVP protocols, CN$_1$, CN$_2$, MN$_1$, and MN$_2$ will periodically sends E2E Path/Resv messages to refresh the E2E reservations. When MR (HAoMR) receives an E2E Path/Resv message of an ongoing session, it will not tunnel the E2E messages to HAoMR (MR). Instead, it will simply refresh the corresponding index. The index of an active MN (CN) will appear in the Alive_ID_List object, when MR (HAoMR) sends a tunnel Path/Resv message to HAoMR (MR). Because MR and HAoMR intercept and do not forward the E2E Path/Resv messages, which are sent periodically by MNs or CNs, MBA can reduce the signal overhead significantly.

3.4 MBA Bandwidth Reservation Policies

As mentioned earlier, all NEMO nodes served by an MR share the bandwidth of the tunnel HAoMR-MR. When an MR receives a request to establish an E2E QoS session to/from an NEMO node, it needs to acquire more bandwidth to fulfill the request if it has not reserved enough bandwidth in the tunnel HAoMR-MR. Therefore, the policies of tunnel bandwidth reservations will affect both the signal overhead and blocking probability of the QoS sessions requested by NEMO nodes.

In this subsection, we present three bandwidth reservation policies for the tunnel HAoMR-MR. These policies are Static Reservation, Dynamic Reservation and Hybrid Reservation policies.

3.4.1 Static Reservation Policy (SRP)

In this policy, an MR reserves a static amount of bandwidth for NEMO nodes. With SRP, an MR reserves a constant amount of bandwidth and does not change the reserved bandwidth no matter a session begins or terminates. If the reserved bandwidth is not enough to fulfill the request of a new session, the MR will simply block the new session. This policy is simple. However, it is not flexible. It may have poor bandwidth utilization if the MR has reserved too much bandwidth or a high blocking rate otherwise. This policy is suitable for a NEMO with frequent joining and leaving of NEMO nodes, such as a bus for example.

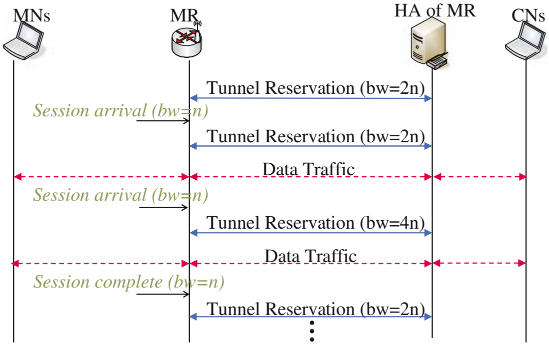
3.4.2 Dynamic Reservation Policy (DRP)

This policy allows an MR to adjust the reserved bandwidth dynamically when necessary. Initially, an MR does not reserve any bandwidth for the tunnel HAoMR-MR if no sessions request QoS services. When a session that requests n units of bandwidth arrives, the MR would reserve n units for the tunnel to fulfill the requests. On the other hand, the MR releases the reserved bandwidth when a session terminates. If more than one session arrives or terminates at the same time, the MR needs to make only one-time bandwidth adjustment. However, frequent adjustments may still introduce huge signal overhead, and thus the policy is suitable for NEMOs with low variation in bandwidth requests, such as the NEMOs inside trains or airplanes for example. DRP has the best bandwidth utilization, but it may suffer from high blocking rates if the MR can not acquire bandwidth in time.

3.4.3 Hybrid Reservation Policy (HRP)

Hybrid Reservation Policy (HRP) is a combination of SRP and DRP. In HRP, an MR requests a static increment of bandwidth dynamically whenever it does not have spare bandwidth to accommodate new requests. Fig. 8 illustrates an example of how HRP works with a static increment of 2n bandwidth. Initially, the MR reserves a static amount of 2n bandwidth, which is enough to fulfill the requests of the first two sessions. After the second requests, MR requests to increase the bandwidth to 4n in total. Similarly, the departure of two sessions will trigger MR to release the reserved bandwidth. HRP can provide higher flexibility by altering the value of static increment, and could be proper for both high and low variation NEMOs.

Fig. 8 Hybrid reservation policy



4 Performance Evaluation

As mentioned previously, previous mobile RSVP protocols possess a mobility unawareness problem and are not applicable to NEMOs. Therefore, we proposed the MBA scheme that can resolve the problem. Furthermore, the MBA scheme can also reduce the signal overhead in reservation maintenance. Therefore, we would like to study the effects of MBA on reducing the signal overhead. Besides, we also conduct a preliminary evaluation of the blocking probabilities and resource utilization of the three reservation policies. Furthermore, the MR in MBA is similar to the MN in MRSVP [4] and Hierarchical MRSVP [5]. As a consequence, the forced termination probabilities, during handoff, will be the same as the ones in MRSVP or Hierarchical RSVP, depending on the underlying mobility scheme. Therefore, for the readers who are interested at the forced termination probabilities for handoff, please refer to MRSVP [4] and Hierarchical MRSVP [5] for the performance results.

4.1 Mathematics Analysis

In this subsection, we present an analysis model to evaluate the signal overhead of MBA and use simulation to verify the analytical results. In MBA, an MR and its HAoMR aggregate the bandwidth required for the NENO nodes and reserve bandwidth for the HAoMR-MR tunnel. Furthermore the HAoMR-MR tunnel is transparent to NEMO nodes, and thus NEMO nodes in MBA could behave exactly the same as the nodes with the ordinary RSVP. Therefore, in the following analysis, we focus on the performance analysis of the HAoMR-MR tunnels only. We assume that the inter-arrival time of QoS session requests and the reservation holding time (the connection service time) both follow an exponential distribution with a mean of $1/\lambda$ and $1/\mu$, respectively. We present the analytical models for the sessions with requests of identical bandwidth in Sect. 4.1.1 and with requests of multiple classes, in which each QoS session can request up to a predefined amount of bandwidth, in Sect. 4.1.2. From the mathematic models, we can derive the signal overhead for reservation maintenance in a NEMO and the blocking probabilities that an MR may reject a reservation request, respectively, in Sects. 4.2 and 4.3.

4.1.1 Sessions with Identical Requests

In this case, we assume all sessions have the same QoS requirement and each session requests a single unit of bandwidth. Furthermore, because the reservation policies make no difference

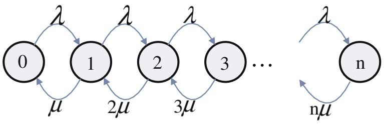


Fig. 9 State transition diagram for identical requests

in signal overhead for reservation maintenance, we also assume the MR uses SRP and reserve n units of bandwidth beforehand.

Figure 9 shows the state transition diagram for the case of identical requests. In the figure, each state represents the number of QoS sessions the MR has already offered. Because the MR has reserved only n units of bandwidth for the tunnel, it will block new requests when there exists n active sessions.

We could derive the probability of each state as follows.

$$p_i = \begin{cases} \frac{1}{i!}\left(\frac{\lambda}{\mu}\right)^i p_0 & i \le n \\ 0 & i > n \end{cases} \tag{1}$$

where p_0 is the probability of the initial state and the equation of p_0 is

$$p_0 = \frac{1}{\sum_{i=0}^{n}\left(\frac{\lambda}{\mu}\right)^i \frac{1}{i!}}. \tag{2}$$

Next, we can calculate the average number of active sessions, m, in a NEMO as follows.

$$m = \sum_{i=0}^{n} i \times p_i \tag{3}$$

4.1.2 Sessions with Multi-class Requests

We further analyze the signal overhead for the case of multi-class requests where active sessions could request different number of bandwidth units. Multi-class requests are more realistic because a session might request a larger bandwidth for video services or a smaller bandwidth for voice or text services. However, this case is more complicate since the number of active sessions is no longer equal to the bandwidth units reserved by an MR. Therefore we can use a two-tuple state transition diagram, as shown in Fig. 10, to model the reservation behavior of active sessions. In the diagram, a state (i, j) means that the MR has accepted i sessions and has reserved j units of bandwidths in total for all sessions. For example, the state (2, 3) represents two sessions are currently active and have jointly requested 3 units of bandwidth.

Assume that n is the bandwidth an MR has reserved, and k is the maximal number of bandwidth units a session can request. Figure 10 shows the state transition diagram when $n = 7$ and $k = 2$. The initial state of the system is (0, 0). Let λ and μ represent the arrival and departure rates, respectively, of a new session. Suppose each session can request one or two units of bandwidth with the same probability. A state (i, j) may transit to states $(i+1, j+1)$ or $(i+1, j+2)$ with an equal transition rate of $\lambda/2$, except when $(j+1)$ is equal to or greater than 7. For example, if the current state is (2, 3), the new state will be either (3, 4) or (3, 5) when the MR accepts a new session. However, if the current state is (5, 6) the MR will accept only the session with one unit request. Furthermore if the current state is (5, 7), the MR will not accept any new session request since it has already used up all bandwidth.

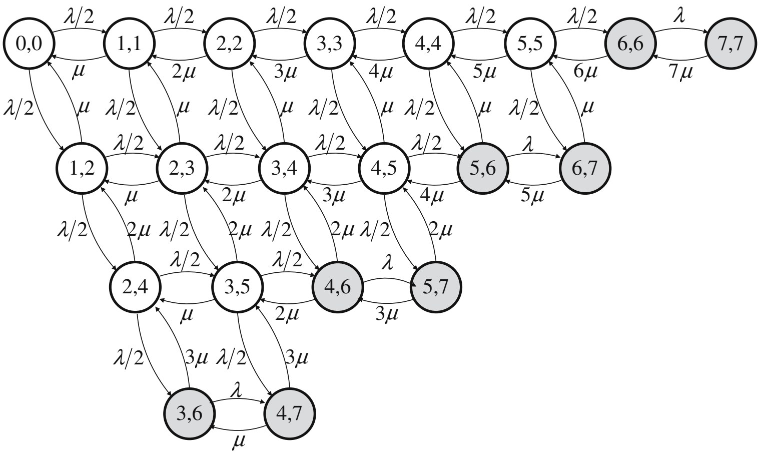


Fig. 10 State transition diagram ($n = 7, k = 2$)

Table 1 One-step transition matrix

	(0,0)	(1,1)	(1,2)	(2,2)	...	(6,6)	(6,7)	(7,7)
(0,0)	0	$\lambda/2$	$\lambda/2$	0	...	0	0	0
(1,1)	μ	0	0	$\lambda/2$	...	0	0	0
(1,2)	μ	0	0	0	...	0	0	0
(2,2)	0	2μ	0	0	...	0	0	0
...	...	...	...	...	...	...	...	...
(6,6)	0	0	0	0	...	0	0	λ
(6,7)	0	0	0	0	...	0	0	0
(7,7)	0	0	0	0	...	7μ	0	0

Table 2 Normalized one-step transition matrix

	(0,0)	(1,1)	(1,2)	(2,2)	...	(6,6)	(6,7)	(7,7)
(0,0)	$1 - \rho/7$	$\rho/14$	$\rho/14$	0	...	0	0	0
(1,1)	1/7	$6/7 - \rho/7$	0	$\rho/14$	...	0	0	0
(1,2)	1/7	0	$6/7 - \rho/7$	0	...	0	0	0
(2,2)	0	2/7	0	$5/7 - \rho/7$	...	0	0	0
...	...	...	...	...	...	...	...	...
(6,6)	0	0	0	0	...	$1/7 - \rho/7$	$\rho/7$	0
(6,7)	0	0	0	0	...	0	1/7	0
(7,7)	0	0	0	0	...	1	0	0

Similarly, when an active session terminates, it will release one or two units of bandwidth with an equal probability of 1/2.

From the above state transition diagram, we can obtain a corresponding one-step transition matrix as shown in Table 1. The one-step transition matrix can be further normalized as Table 2 by applying by Jensen algorithm [16].

Let $P^{(1)}$ denote the normalized one-step transition matrix. We can apply $P^{(1)}$ n times to obtain the n-step transition matrix, $P^{(n)}$, as

$$P^{(n)} = \overbrace{P^{(1)} \times P^{(1)} \times \cdots \times P^{(1)}}^{n}. \tag{4}$$

Furthermore, let $\pi = (1, 0, \ldots, 0)$ be the vector that represents the state probabilities before any transition, and $\pi^{(n)}$be the state probabilities after n-step transition. Then, we have $\pi^{(n)} = \pi^{(0)} \times P^{(n)}$. By taking the limit of n, we can obtain the steady state probability vector, π, as follows.

$$\pi = \lim_{n=\infty} \pi^{(n)} = \pi^{(0)} \times \lim_{n=\infty} P^{(n)}. \tag{5}$$

From the steady state probabilities π, we can derive the average number of active sessions an MR can offer as follows.

$$m = \sum_{i=0}^{n} i \sum_{j=i}^{\min(2i,n)} \pi\,(i, j). \tag{6}$$

4.2 Signal Overhead

If we assume each NEMO node runs RSVP itself and each session could either be an outgoing or an incoming session with an equal probability, then the signal overhead for reservation maintenance in RSVP, denoted as O_{RSVP}, can be represented as Eq. (7).

$$O_{\mathrm{RSVP}} = \frac{(S_{\mathrm{P}} + S_{\mathrm{R}}) \times \frac{1}{2} \times m \times 8}{T}\mathrm{bps} \tag{7}$$

where S_P and S_R are the sizes of RSVP Path and Resv message in octets, respectively; T is the interval of sending RSVP messages in seconds, and m is the average number of active sessions in a NEMO. By replacing m with Eqs. 3 and 6, respectively, we can obtain the equations of O_{RSVP} for the cases of identical requests and multi-class requests.

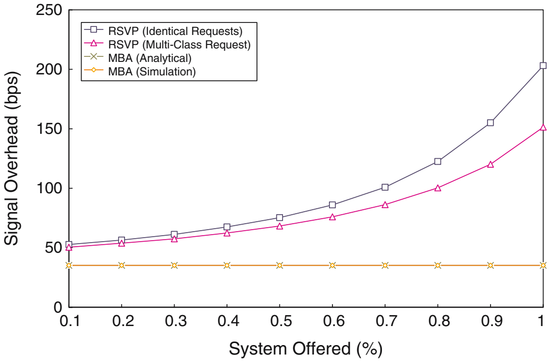
On the contrary, in MBA, an MR aggregates the reservations, and issues a single Tunnel Path/Resv, respectively, to refresh the outgoing/incoming tunnel reservations. Therefore the signal overhead the proposed MBA introduces could be represented as follows in both cases.

$$O_{MBA} = \frac{(S_P + S_R) \times \frac{1}{2} \times 1 \times 8}{T} bps \tag{8}$$

Without security and identification requirements, the default value of S_P is 172 byes, S_R is 92 bytes, and T is 30 s [17]. We apply these default values and obtain the signal overhead over the system load $\rho = \lambda/\mu$ as shown in Fig. 11. From the figure, it is obvious that the signal overhead for establishing a tunnel in MBA is about 35 bps. In addition, we could observe that the signal overhead in RSVP increases as the system load, i.e., average number of active sessions, increases, while the one in MBA retains at the same level, independent of the system load. Furthermore, with the same bandwidth reserved by a NEMO, the NEMO could serve less active sessions for the case of multiple-class requests, and consequently, for RSVP, the signal overhead in this case is comparatively smaller than the one of identical requests.

We also run simulations to verify the above analysis. As shown in the figure, the simulation result of MBA is identical with the analytical one. The simulation will be explained in Sect. 4.4.

Fig. 11 Signal overhead for identical requests



4.3 Blocking Probability

Because we assume the MR uses SRP for tunnel reservations in the above analysis, the blocking probability, p_{block}, for the sessions with identical requests equals to p_{n}, and can be derived from Erlang-B serving model as follows:

$$p_{\text{block}} = p_n = \frac{\frac{\left(\frac{\lambda}{\mu}\right)^n}{n!}}{\sum_{i=0}^{n}\left(\frac{\lambda}{\mu}\right)^i \frac{1}{i!}}. \tag{9}$$

For the sessions with multi-class requests, the equation of p_{block} is more complicate. A multi-class request will be blocked by the MR if it has requested more units than the remaining units of bandwidth. Therefore, the shadowed states, in Fig. 10, represent the ones where a blocking may occur. The shadowed states and equation of p_{block} can be derived as follows.

$$p_{\text{block}} = \sum_{i=0}^{n} p\,(i,n) + \sum_{i=0}^{n-1} \frac{1}{k} p\,(i,n-1) + \cdots + \sum_{i=0}^{n-(k-1)} \frac{k-1}{k} p\,(i,n-(k-1)), \tag{10}$$

where $j/k \times p(i, n-j)$ represents the probability that a new request will be blocked by the MR in state $(i, n-j)$. By summing the blocking probabilities of all shadowed states, we can obtain the equation (10) for the blocking probability, p_{block}. In the equation, we assume the probabilities of the non-existing states are all zeros.

Figures 12 and 13 plot the blocking probabilities for different tunnel reservation policies under various system loads (λ/μ). Because in DRP, the MR will make a new reservation request when it receives a new reservation request, we use DRP to simulate the behavior of RSVP. In both figures, we can easily observe that the proposed MBA scheme is better than the original RSVP protocol. Furthermore, for the proposed MBA scheme, the blocking probabilities for multi-class requests are higher than the ones for identical requests.

4.4 Simulation Results

In addition to the mathematical analysis, we also run simulations and compare the results with the ones of analytical model for SRP tunnel reservations. In this subsection, we first explain how we conduct the simulations and then present the utilization of the bandwidth reserved by using the three reservation policies, namely SRP, DRP, and HRP.

Fig. 12 Blocking probability for identical requests

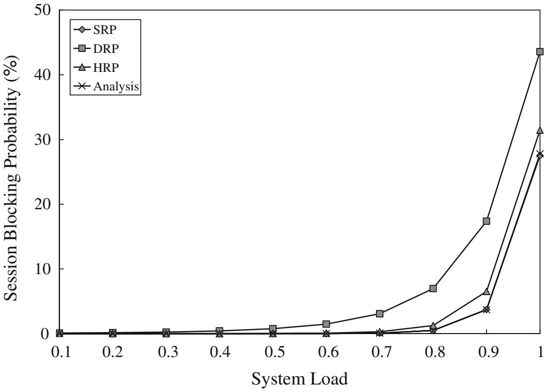
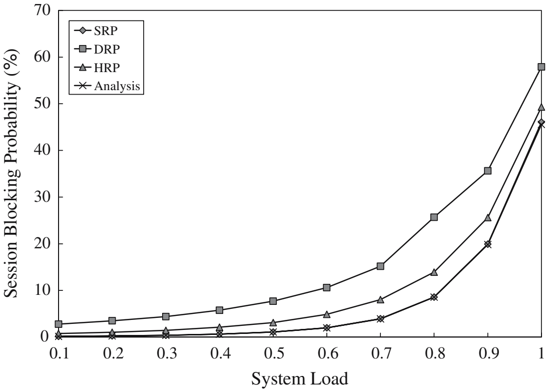


Fig. 13 Blocking probabilities for multi-class requests



In the simulation, we generate background traffic randomly with a mean 40% total bandwidth on the external link of the MR. Similar to the above analytical model, we assume the arrival and departure rates follow the Poisson distribution with a mean of λ and μ, respectively. The total bandwidth of the MR's external link is 10 units and we set $k = 2$ for the multi-class requests.

Figures 12 and 13, respectively, show the blocking probabilities for the identical requests and multi-class requests under various system loads of the MR. In general, the blocking probability increases as the system load increases. SRP has the lowest blocking probability since it reserves maximum available bandwidth in the beginning. DRP has the worst blocking probability because it reserves bandwidth dynamically when the MR (HAoMR) receives a connection request and may fail in competing bandwidth with the background traffic. Furthermore, different values of the background traffic will shift the curves of blocking probabilities horizontally.

On the contrary, as shown in Figs. 14 and 15, DRP always has 100% of bandwidth utilization since it reserves bandwidth only when necessary. SRP, oppositely, has the lowest bandwidth utilization, especially during light system load. On average, the bandwidth utilization of multi-class requests is lower than the one of identical requests.

Fig. 14 Bandwidth utilization for identical requests

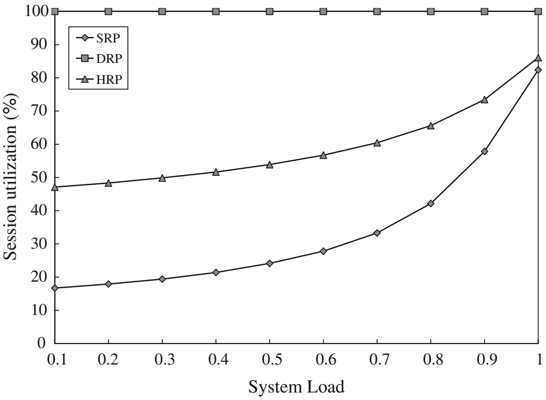
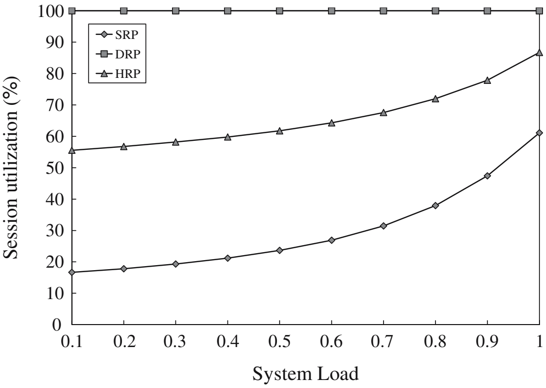


Fig. 15 Bandwidth utilization for multi-class requests



5 Conclusion

In this paper, we propose an MBA scheme to support QoS for NEMO nodes. The proposed MBA scheme makes the MR and HAoMR the resource reservation proxies that handle mobility and reserve bandwidth on behalf of the NEMO nodes. Therefore the MBA scheme can resolve the mobility unawareness and excessive signal overhead in supporting QoS for NEMOs. Mathematical analysis and simulation results show that the proposed MBA scheme can significantly reduce the signal overhead for reservation maintenance. Although we describe the MBA scheme in the context of RSVP, the concept of MBA can work with other resources reservation protocols as well. Furthermore the policies of tunnel bandwidth reservations may affect the blocking probabilities and bandwidth utilization, and thus need further investigation in the future. Furthermore, MBA could incorporate some administration control to drop some ongoing session of NEMO nodes if it demands more bandwidth than whatever the external link can offer. In the future, we will first design the admission control and administration mechanism for MBA, and then implement the proposed MBA scheme.

References

1. Devarapalli, V. (2005). Network mobility basic support protocol. *IETF RFC 3963*, January 2005.
2. Perkins, C. (2001). IP mobility support for IPv4. *IETF RFC 2002,* February 2001.
3. Perkins, C. (2004). Mobility support in IPv6. *IETF RFC 3775,* June 2004.

4. Talukdar, A., Badrinath, B., & Acharya, A. (2001). MRSVP: A resource reservation protocol for an integrated services network with mobile hosts. *Wireless Networks, 7*(1), 5–19.
5. Tseng, C., Lee, G., & Liu, R. (2003). HMRSVP: A hierarchical mobile RSVP protocol. *Wireless Networks, 9*, 95–102.
6. Jamoussi, B., Andersson, L., & Callon, R. (2002). Constraint-based LSP Setup Using LDP. *IETF RFC 3212,* January 2002.
7. Awduche, D., Berger, L., & Gan, D. RSVP-TE: Extensions to RSVP for LSP tunnels. *IETF RFC 3209,* December 2001.
8. Kompella, K., & Lang, J. (2004). Procedures for modifying the resource reSerVation protocol (RSVP). *IETF RFC 3936,* October 2004.
9. Baker, F., Iturralde, C., Le Faucheur, F., & Davie, B. (2001). Aggregation of RSVP for IPv4 and IPv6 reservation. *IETF RFC 3175,* September 2001.
10. Braden, R., Zhang, L., Berson, S., Herzog, S., & Jamin, S. (1977) Resource ReServation Protocol (RSVP) version 1 functional specification. *IETF RFC 2205,* September 1997.
11. Terzis, A., Krawczyk, J., Wroclawski, J., & Zhang, L. (2000). RSVP operation over IP tunnels. *IETF RFC 2746,* Jane 2000.
12. Simpson, W. (1995). IP-in-IP encapsulation. *IETF RFC 1853,* October 1995.
13. Braden, R., Clark, D., & Shenker, S. (1994). Integrated services in the internet architecture: An overview. *IETF RFC 1633,* January 1994.
14. Montenegro, G. (2001). Reverse tunneling for mobile IP. *IETF RFC 3024,* January 2001.
15. Gustafsson, E., Jonson, A., & Perkins, E. (2003) Mobile IP regional registration. *Internet-Draft: draft-ietf-mobileip-reg-tunnel-08.txt,* November 2003.
16. Jensen, A. (1953). Markoff chains as an aid in the study of markoff process. *Skandinavisk Aktuarietidskrift, 36*, 87–91.
17. Manner, J., Fu, X., & Pan, P. (2005). Analysis of existing quality of service signaling protocols. *IETF RFC 4094,* May 2005.

Author Biographies

Jui-Tang Wang received his B.S. degree in the Department of Industrial Engineering and Management from National Chin-Yi Institute of Technology, Taichung, Taiwan, in 1995 and his M.S. degree in Computer Science and information Engineering from National Cheng-Kung University, Tainan, Taiwan, in 2000. He is currently a Ph.D candidate in the Department of Computer Science from National Chiao-Tung University, Hsin-Chu, Taiwan. His research interests include Mobile Computing, Network Security and Wireless Internet.

Yuan-Ying Hsu received her B.S. (1999), M.S. (2000), and Ph.D. (2005) degrees in Computer Science and Information Engineering from National Chiao-Tung University, Taiwan. During her Ph.D program, she worked with Prof. Chien-Chao Tseng and Prof. Yu-Chee Tseng, and her background has been focused on wireless application protocols, mobile computing, network management, and wireless ad hoc networks. She has designed and implemented prototypes of a WML/HTML translator, a mobile router, and a personal information system. She also has works in addressing and routing mechanisms in wireless mesh networks. Dr. Hsu joined Telcordia applied research center in Taiwan (TARC-TW) in August 2005. In TARC-TW, she works with the team to implement a seamless roaming testbed which can support audio streaming and VoIP sessions between heterogeneous networks. She also participates in standard making with ITRI on IEEE 802.16j and successfully promotes several contributions to be accepted.

Chien-Chao Tseng is currently a professor in the Department of Computer Science at National Chiao-Tung University, Hsin-Chu, Taiwan. He received his B.S. degree in Industrial Engineering from National Tsing-Hua University, Hsin-Chu, Taiwan, in 1981; M.S. and Ph.D degrees in Computer Science from the Southern Methodist University, Dallas, Texas, USA, in 1986 and 1989, respectively. His research interests include Mobile Computing, and Wireless Internet.