

A mobility-based load control scheme in Hierarchical Mobile IPv6 networks

Sangheon Pack · Taekyoung Kwon ·
Yanghee Choi

© Springer Science+Business Media, LLC 2008

Abstract By introducing a mobility anchor point (MAP), Hierarchical Mobile IPv6 (HMIPv6) reduces the signaling overhead and handoff latency associated with Mobile IPv6. In this paper, we propose a mobility-based load control (MLC) scheme, which mitigates the burden of the MAP in fully distributed and adaptive manners. The MLC scheme combines two algorithms: a threshold-based admission control algorithm and a session-to-mobility ratio (SMR)-based replacement algorithm. The threshold-based admission control algorithm gives higher priority to ongoing mobile nodes (MNs) than new MNs, by blocking new MNs when the number of MNs being serviced by the MAP is greater than a predetermined threshold. On the other hand, the SMR-based replacement algorithm achieves efficient MAP load distribution by considering MNs' traffic and mobility patterns. We analyze the MLC scheme using the continuous time Markov chain in terms of the new MN blocking probability, ongoing MN dropping probability, and binding update cost. Also, the MAP processing latency

is evaluated based on the M/G/1 queueing model. Analytical and simulation results demonstrate that the MLC scheme outperforms other schemes and thus it is a viable solution for scalable HMIPv6 networks.

Keywords Hierarchical Mobile IPv6 · Mobility-based load control · Mobility anchor point · Admission control algorithm · Session-to-mobility ratio

1 Introduction

In wireless/mobile networks, users freely change their points of attachment as they move about. In such environments, mobility management is a key technology for keeping track of the users' current locations and for delivering data correctly. In current cellular networks for voice telephony services, various schemes have been proposed to provide efficient mobility management [1]. However, next-generation wireless/mobile networks will be based on the IP technology and hence they will have different characteristics from the existing cellular networks. Therefore, the design of an IP-based mobility management scheme has become a crucial issue [2].

Mobile IPv6 (MIPv6) [3] is the de facto mobility support protocol in IPv6 wireless/mobile networks. To reduce high signaling overhead incurred in Mobile IPv6 networks when mobile nodes (MNs) perform frequent handoffs, Hierarchical Mobile IPv6 (HMIPv6) [4] introduces the concept of a mobility anchor point (MAP). The MAP handles binding update (BU) requests pertaining to intra-domain handoffs in a localized manner. Without the MAP, each BU request for an intra-domain handoff would have to be handled by the home agent (HA), which would cause the BU traffic to place a burden on the entire network.

The preliminary version of this paper was presented at IEEE Global Telecommunications Conference (GLOBECOM) 2004, Dallas, USA, November 2004.

S. Pack (✉)
School of Electrical Engineering, Korea University,
Seoul, South Korea
e-mail: shpack@korea.ac.kr

T. Kwon · Y. Choi
School of Computer Science and Engineering,
Seoul National University, Seoul, South Korea

T. Kwon
e-mail: tkkwon@snu.ac.kr

Y. Choi
e-mail: yhchoi@snu.ac.kr

In HMIPv6 networks, as the number of MNs serviced by a MAP increases, the MAP becomes a single point of bottleneck. This is because the MAP not only handles binding updates, but also performs encapsulation/decapsulation for every data packet destined for or originated from the MNs. In other words, when many MNs are serviced by a single MAP, the MAP suffers from traffic overload, which results in higher processing latency. Furthermore, since HMIPv6-based mobile networks will be likely to proliferate in the near future, the question of how to control the traffic load at the MAP will soon become one of the most crucial issues.

In this paper, we propose a mobility-based load control (MLC) scheme where the MAP controls its load based on the current load state and the MN's characteristics. The MLC scheme reserves an amount of the MAP capacity for ongoing MNs with higher priority, which enables ongoing MNs to be accepted by the MAP preferentially. Therefore, ongoing MNs can experience significantly reduced handoff latency. More importantly, the MAP maintains the session-to-mobility ratio (SMR) for each MN that it is serving, where the SMR is defined as the ratio of the session arrival rate to the handoff rate (i.e., intra-domain handoff). Using the SMR information, a SMR-based replacement algorithm is employed to distribute the MAP load more efficiently, when no MAP capacity is permitted. Specifically, an MN with low mobility and high session activity is replaced by the SMR-based replacement algorithm, and then the replaced MN does not use the MAP further. Therefore, the replaced MN can eliminate the MAP tunneling overhead, which is more desirable for the replaced MN because of its high session activity. By means of this mobility-based traffic control, the MLC scheme can reduce the amount of binding update traffic transiting over the entire network.

Our major contributions are summarized as follows: (1) the MLC scheme is performed at a single MAP in a distributed manner, and it does not require any overhead for interaction and synchronization between MAPs; (2) the MAP load can be distributed adaptively to the MN's characteristics by employing the SMR-based replacement algorithm; (3) we develop analytical models to evaluate the MN dropping/blocking probability, MAP processing latency, and binding update cost. Also, simulation results are given to validate the analytical results.

The remainder of this paper is organized as follows. We describe the HMIPv6 operations and system model in Sect. 2. We propose and analyze the MLC scheme in Sects. 3 and 4, respectively. Numerical results are presented in Sect. 5 and they are validated by simulations in Sect. 6. Related works are summarized in Sect. 7 followed by the concluding remarks in Sect. 8.

2 System description

In HMIPv6 networks, an MN configures two care-of-addresses (CoAs): a regional care-of-address (RCoA) and an on-link care-of-address (LCoA). The RCoA is an address on the MAP's subnet. An MN is attributed an RCoA when it receives a Router Advertisement (RA) message with a MAP option. On the other hand, the LCoA is an on-link CoA attributed to the MN's interface based on the prefix information advertised by an access router (AR).

Figure 1 illustrates the basic operations in HMIPv6 networks. An MN entering a foreign network first configures its LCoA by the IPv6 address auto-configuration scheme. Also, it receives an RA message containing information on a nearby MAP and configures an RCoA. The MN then sends a local BU message to the MAP. This local BU message includes the MN's RCoA in the Home Address Option field and the LCoA is used as the source address of the BU message, and thus it binds the MN's RCoA to its LCoA.

The MAP then performs a duplicate address detection (DAD) procedure for the MN's RCoA on its link and returns a Binding Acknowledgement (BACK) message to the MN. The BACK message identifies the binding as being successful or otherwise contains an appropriate fault code. After registering with the MAP, the MN must register its new RCoA with its HA by sending a BU message that specifies the binding (RCoA, home address (HoA)). The HoA is recorded in the Home Address Option field, whereas the RCoA can be found in the Source Address field. The MN may also send a similar BU message, which specifies the binding information between the HoA and the RCoA, to its current correspondent nodes (CNs).

Once the MN has successfully registered with the MAP, a bi-directional tunnel is established between them. All packets sent by the MN are tunneled to the MAP. The outer

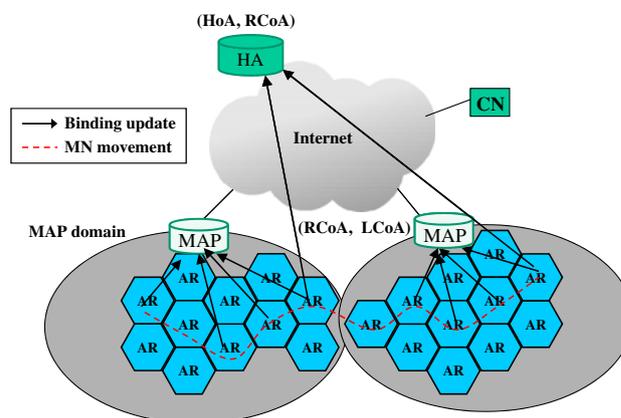


Fig. 1 Basic binding update procedure in HMIPv6

header contains the MN's LCoA in the Source Address field and the MAP's address in the Destination Address field. The inner header contains the MN's RCoA in the Source Address field and the CN's address in the Destination Address field. At the same time, all packets addressed to the MN's RCoA are intercepted by the MAP and tunneled to the MN's LCoA. If the MN changes its current address within the same MAP domain, it only needs to register the new address (i.e., LCoA) with the MAP. The RCoA does not change as long as the MN moves within the same MAP domain. This makes the MN's mobility transparent to the CNs.

In this paper, we focus on the load control at the MAP¹ and assume that there is only one MAP available to each MN. The MAP capacity C is represented by the maximum number of MNs that it can service. First, an MN sends a local BU message to the MAP. If the BU message is accepted by the MAP, the MN will receive a successful BACK message and then send a BU message with its RCoA to the HA. On the other hand, the MN's BU message is rejected by the MAP, the rejected MN registers its LCoA with the HA and then the packets destined for the MN will bypass the MAP. Regarding route optimization, if the local BU message is accepted, the MN's RCoA is notified to the CNs. Otherwise, the MN's LCoA is sent to the CNs.

In HMIPv6 networks, MNs are classified into three types: *new MN*, *ongoing MN*, and *refresh MN*. A new MN refers to an MN performing the initial binding update to the MAP, e.g., when an MN is first turned on. On the other hand, an MN performing an inter-MAP domain handoff is considered as an ongoing MN. Another type is a refresh MN which binding lifetime expires and tries to refresh the binding information. Since the binding refresh operation does not increase the current MAP load, only new and ongoing MNs are considered in the MLC scheme. To distinguish ongoing MNs from new MNs in BU messages, we suggest that a new O flag be added to the existing BU message format (see Fig. 2). If the O flag in the BU message is set, the MN is deemed to be an ongoing MN. Otherwise, the MN is regarded as a new MN.

All packets destined for or originated from MNs go through the MAP in HMIPv6 networks. Therefore, the MAP can measure the MN's SMR as follows. First, the SMR of a new MN is set to a default value, i.e., 1.0. Since the SMR can be changed by session arrivals or intra-domain handoffs, the numbers of session arrivals to the MN and intra-domain handoffs conducted by the MN are counted during a fixed time interval. After then, by its definition, the SMR is calculated by dividing the number of session arrivals into the number of intra-domain handoffs.

¹ That is, load balancing among HAs and security issues are beyond scope of this paper.

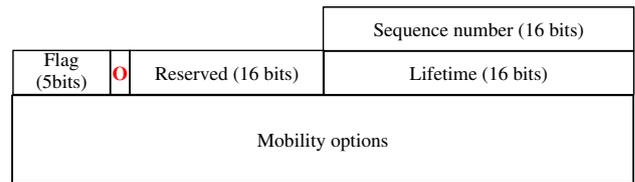


Fig. 2 Modified BU format

To avoid oscillation of the measured SMR, an exponentially weighted moving average (EWMA) scheme can be employed. More details can be found in [5].

3 Mobility-based load control scheme

Figure 3 illustrates the overall procedure of the MLC scheme, which consists of threshold-based admission control and SMR-based replacement algorithms. Here, C_{used} denotes the current number of MNs registered with the MAP. Detailed operations are given in this section.

3.1 Threshold-based admission control algorithm

When a BU message arrives at a MAP, the MAP triggers the threshold-based admission control algorithm. First of all, the MAP needs to determine whether the received BU message comes from an ongoing MN or a new MN using the O flag. Let K be a pre-defined threshold. When the number of MNs serviced by the MAP is less than K , both new MNs and ongoing MNs are admitted. On the other hand, to give higher priority to ongoing MNs [6], when the current MAP load is equal to or greater than K , only ongoing MNs are accepted. This threshold-based admission control algorithm reduces the ongoing MN dropping probability at the cost of increasing the new MN blocking probability [7, 8].

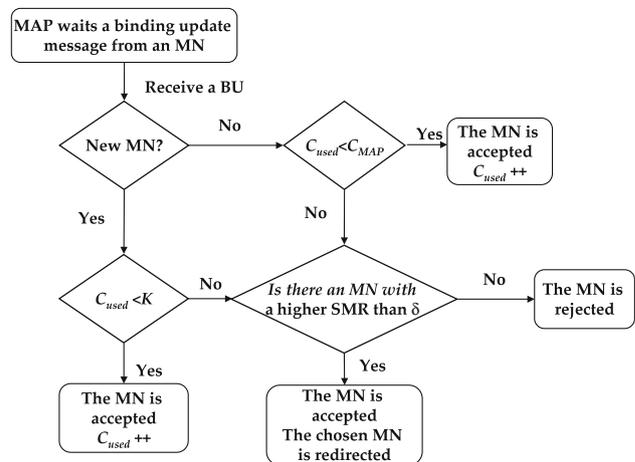


Fig. 3 Mobility-based load control procedure

3.2 SMR-based replacement algorithm

When an MN rarely requests handoffs, it is better not to use the MAP to reduce tunneling overhead at the MAP [9, 10]. On the other hand, for an MN requesting frequent handoffs and thereby resulting in much BU traffic, it would be better for the MAP to intercept and localize the BU traffic. This is the rationale behind the SMR-based replacement algorithm. This algorithm is triggered when there is an MN whose SMR is higher than a pre-determined SMR threshold (δ) and one of the below two cases is satisfied. One is the situation in which a new MN requests a binding update when there is no available capacity for new MNs (i.e., $C_{used} \geq K$). The other is the case where an ongoing MN requests a binding update when there is no remaining capacity for ongoing MNs (i.e., $C_{used} = C$). By the SMR-based replacement algorithm, an MN with a higher SMR than δ is replaced by the MN that newly requests a binding update. In this way, an MN that would have been rejected by the threshold-based admission control algorithm can be admitted by the MAP after all. In the SMR-based replacement algorithm, it is important to determine δ correctly, in order to strike a balance between the dropping/blocking probability and replacement overhead. The issue of how to determine the SMR threshold value will be addressed in Sect. 5.5.

It should be noted that the replaced MN is not forced to be terminated. Figure 4 shows the message flow for the replaced MN. Once an MN to be replaced has been chosen, the MAP sends a BACK message to the MN. In this case, the BACK message contains the status code 130 with the reason “Insufficient resources” [3]. To guarantee the reliable reception of the BACK message by the MN, the existing BACK message format is modified. This is because the current HMIPv6 specification does not define any acknowledgement service for the BACK message itself. To make this possible, we add an A flag to the BACK message (see Fig. 5). Therefore, a BACK message with the

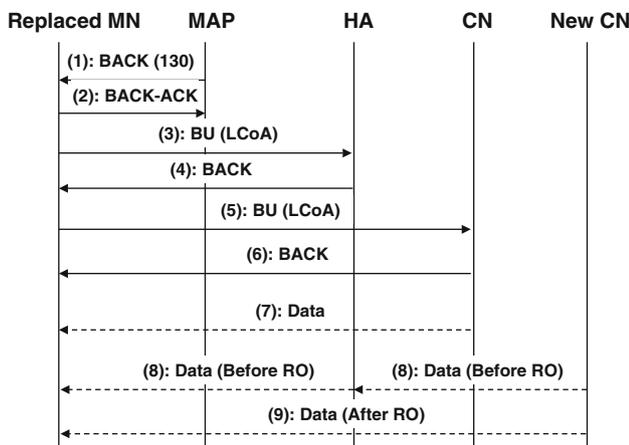


Fig. 4 Message flow of the replaced MN

	Status (8bits)	K	A	Reserved (6 bits)
Sequence number (16 bits)	Lifetime (16 bits)			
Mobility options				

Fig. 5 Modified BACK format

set A flag indicates that another acknowledgement message should be returned upon receipt of the BACK message. Let us refer to the acknowledgement message of the BACK message as a BACK-ACK message. After receiving the BACK message, the replaced MN then sends a BACK-ACK message to the MAP and BU messages with its LCoA to the HA/CNs. Then, the packets destined for the MN will be delivered to the MN bypassing the MAP, which can increase the HA load and HA binding update latency. However, since the replaced MN has large session arrivals and infrequent handoffs (i.e., its SMR is high), it is preferable to reduce the MAP tunneling overhead by choosing the HA for binding update [9]. Consequently, it can be concluded that the SMR-based replacement algorithm can distribute network-wide traffic to the MAPs and HA depending on the MN’s characteristics.

4 Performance analysis

In this section, we evaluate the performance of the MLC scheme using continuous time Markov chains (CTMCs), in terms of the new MN blocking probability, ongoing MN dropping probability, and binding update cost. In addition, we derive the MAP processing latency based on the M/G/1 queuing model.

4.1 Assumptions

To develop CTMCs, we make the following assumptions without loss of generality [11].

- A.1 The inter-session arrival time follows an exponential distribution with rate λ_S .
- A.2 The arrival processes of ongoing and new MNs follow Poisson distributions with rate λ_O and λ_N , respectively.
- A.3 The MAP domain residence times of ongoing and new MNs follow exponential distributions with mean $1/\mu_O$ and $1/\mu_N$, respectively.² The traffic loads of ongoing and new MNs are given by $\rho_O = \lambda_O/\mu_O$ and $\rho_N = \lambda_N/\mu_N$, respectively.

² Although the residence times are typically non-exponential in wireless/mobile networks, the analysis based on the simplified exponential assumption has been widely used [12–14], and provides useful mean value information.

A.4 Let P_δ be the probability that the SMR of an MN is larger than δ . If the SMR of each MN is assumed to be ideally and identically distributed, the probability that there are no MNs whose SMR is larger than δ among k MNs is as follows:

$$P(k, \delta) = \binom{k}{0} \cdot P_\delta^0 \cdot (1 - P_\delta)^k.$$

4.2 Ongoing MN dropping and new MN blocking probabilities

4.2.1 No load control (NLC) scheme

For comparison purposes, we consider a case without any load control scheme, i.e., no load control (NLC) scheme. Figure 6 indicates the state transition diagram for the NLC scheme. The two-dimensional CTMC has a state space $S = \{(i, j) | 0 \leq i, j \leq C, 0 \leq i + j \leq C\}$,

where state (i, j) represents that i new MNs and j ongoing MNs are being serviced by the MAP. Let $p(i, j; i', j')$ be the transition rate from state (i, j) to state (i', j') . The transition rates are given by

$$\begin{aligned} p(i, j; i - 1, j) &= i\mu_N (0 < i \leq C, 0 \leq j \leq C - i) \\ p(i, j; i + 1, j) &= \lambda_N (0 \leq i < C, 0 \leq j \leq C - i) \\ p(i, j; i, j - 1) &= j\mu_O (0 \leq i \leq C - j, 0 < j \leq C) \\ p(i, j; i, j + 1) &= \lambda_O (0 \leq i \leq C - j, 0 \leq j < C). \end{aligned} \tag{1}$$

Then, the steady state probability $\pi(i, j)$ is given by [8]

$$\pi(i, j) = \frac{\rho_N^i \cdot \rho_O^j}{i! \cdot j!} \cdot \pi(0, 0), 0 \leq i, j \leq C, i + j \leq C, \tag{2}$$

where $\pi(0, 0)$ is obtained from the normalization condition, $\sum_{i=0}^C \sum_{j=0}^{C-i} \pi(i, j) = 1$, and it is given by

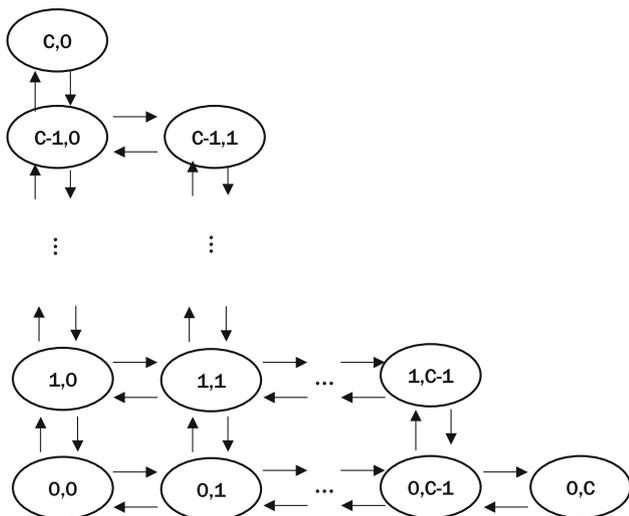


Fig. 6 Two-dimensional CTMC for the NLC scheme

$$\pi(0, 0) = \left[\sum_{i=0}^C \frac{\rho_N^i}{i!} \sum_{j=0}^{C-i} \frac{\rho_O^j}{j!} \right]^{-1}. \tag{3}$$

From the CTMC for the NLC scheme, the new MN blocking probability (P_{NB}) and ongoing MN dropping probability (P_{OD}) are given by

$$P_{NB} = P_{OD} = \sum_{i=0}^C \pi(i, C - i) = \frac{\sum_{i=0}^C \frac{\rho_N^i \rho_O^{C-i}}{i! (C-i)!}}{\sum_{i=0}^C \frac{\rho_N^i}{i!} \sum_{j=0}^{C-i} \frac{\rho_O^j}{j!}}. \tag{4}$$

4.2.2 Threshold-based load control (TLC) scheme

Hereafter, the threshold-based load control (TLC) scheme refers to a load control scheme utilizing only the threshold-based admission control algorithm. In other words, the TLC scheme does not employ any replacement algorithms. The two-dimensional CTMC for the TLC scheme is shown in Fig. 7, where K denotes the threshold value in the threshold-based admission control algorithm. Then, the transition rates are given by

$$\begin{aligned} p(i, j; i - 1, j) &= i\mu_N (0 < i \leq K, 0 \leq i + j \leq C) \\ p(i, j; i + 1, j) &= \lambda_N (0 \leq i < K, 0 \leq i + j < K) \\ p(i, j; i, j - 1) &= j\mu_O (0 \leq i \leq K, 0 < i + j \leq C) \\ p(i, j; i, j + 1) &= \lambda_O (0 \leq i \leq K, 0 \leq i + j < C). \end{aligned} \tag{5}$$

Then, the steady state probability $\pi(i, j)$ can be obtained by an iterative algorithm [15] (see Appendix 1).

In the TLC scheme, a new MN is blocked when the current MAP load is equal to or greater than K . Therefore, the new MN blocking probability is given by

$$P_{NB} = \sum_{i=0}^K \sum_{j=K-i}^{C-i} \pi(i, j). \tag{6}$$

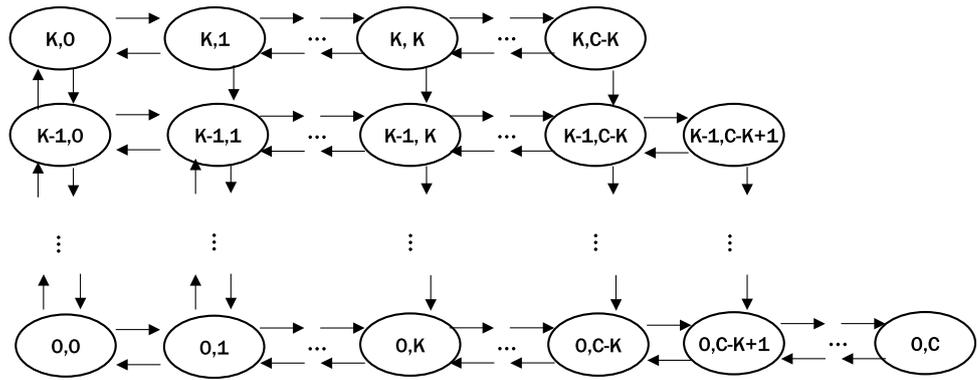
On the other hand, an ongoing MN is dropped when the current MAP load is the same as the MAP capacity, so that its dropping probability is

$$P_{OD} = \sum_{i=0}^K \pi(i, C - i). \tag{7}$$

4.2.3 Mobility-based load control (MLC) scheme

As mentioned before, the MLC scheme consists of the threshold-based admission control algorithm and SMR-based replacement algorithm. The replacement algorithm is performed only when there is no remaining MAP capacity for newly incoming MNs and there is at least one MN whose SMR is greater than δ . Although an ongoing or new MN can be accepted by the replacement algorithm, the total number of MNs at the MAP remains unchanged.

Fig. 7 Two-dimensional CTMC for the TLC scheme



Therefore, the transition diagram of the MLC scheme is the same as that of the TLC scheme shown in Fig. 7. Then, the new MN blocking probability is computed as

$$P_{NB} = \sum_{i=0}^K \sum_{j=K-i}^{C-i} P(i+j, \delta) \cdot \pi(i, j), \tag{8}$$

where $P(i+j, \delta)$ represents the case that there is no MN whose SMR is larger than δ among $(i+j)$ MNs. Similarly, the ongoing MN dropping probability is given by

$$P_{OD} = \sum_{i=0}^K P(C, \delta) \cdot \pi(i, C-i). \tag{9}$$

4.3 Binding update cost

For the binding update cost, we consider binding update traffic incurred by new and ongoing MNs residing in a MAP domain during a session. Let C_H and C_M be the unit binding update cost to the HA and the MAP, respectively. Also, let μ_C be the average subnet crossing rate and then the average subnet crossing rate per session is given by μ_C/λ_S [16]. For a new or ongoing MN, if the MN is blocked (or dropped), it performs a binding update to its HA for every subnet crossing; otherwise, the MN sends a binding update message to the MAP for every subnet crossing. Therefore, the binding update costs for a new MN (C_B^N) and ongoing MN (C_B^O) in the TLC scheme are described by

$$C_B^N = P_{NB} \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_H + C_M \right) + (1 - P_{NB}) \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_M + C_M \right) \tag{10}$$

and

$$C_B^O = P_{OD} \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_H + C_M \right) + (1 - P_{OD}) \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_M + C_M \right), \tag{11}$$

where at least one MAP binding update cost is always counted due to the initial MAP binding update.

For the MLC scheme, the binding update costs are given by

$$C_B^N = P_{NB} \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_H + C_M \right) + \sum_{i=0}^{K-1} \sum_{j=0}^{K-i-1} \pi(i, j) \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_M + C_M \right) + \sum_{i=0}^K \sum_{j=K-i}^{C-i} \pi(i, j) (1 - P(i+j, \delta)) \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_M + 2 \cdot C_M + C_H + C_R \right) \tag{12}$$

and

$$C_B^O = P_{OD} \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_H + C_M \right) + \sum_{i=0}^K \sum_{j=0}^{C-i-1} \pi(i, j) \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_M + C_M \right) + \sum_{i=0}^K \pi(i, C-i) (1 - P(C, \delta)) \cdot \left(\frac{\mu_C}{\lambda_S} \cdot C_M + 2 \cdot C_M + C_H + C_R \right). \tag{13}$$

In the above equations, the first term refers to the binding update cost incurred when the MN is blocked (or dropped). The second term represents the binding update cost when the MN is accepted without any replacement, whereas the third term indicates the binding update cost when the MN is accepted by the SMR-based replacement algorithm. In the third case, the MAP and HA binding update costs to replace an MN should be taken into account. At the same time, the additional binding update cost incurred by the replaced MN, C_R , should be considered. Let μ_C^R and λ_S^R be the subnet crossing rate and the session arrival rate of the replaced MN, respectively. Then, C_R can be calculated by

$$C_R = \frac{\mu_C^R}{\lambda_S^R} \cdot C_H - \frac{\mu_C^R}{\lambda_S^R} \cdot C_M.$$

Consequently, the total average binding update cost is the weighted sum of the binding update costs for the new and ongoing MNs, i.e.,

$$C_B = \omega \cdot C_B^N + (1 - \omega) \cdot C_B^O,$$

where ω is the ratio of new MNs to total MNs.

4.4 MAP processing latency

The MAP processing latency represents the interval from the time that a packet arrives at the MAP to the time that it departs again. To calculate the MAP processing latency, the question of how many MNs are serviced by the MAP should be considered. Let λ_p be the average packet generation rate of an MN. Then, the average packet arrival rate at the MAP, $\bar{\lambda}$, is given by

$$\bar{\lambda} = \lambda_p \sum_i \sum_j (i + j) \cdot \pi(i, j). \tag{14}$$

The MAP processing latency is the sum of the average service time and the queue waiting latency, which are derived by the well-known P-K formula in the M/G/1 queuing model [17]. Therefore, the MAP processing latency is given by

$$\begin{aligned} L &= \bar{X} + \frac{\bar{\lambda} \cdot \bar{X}^2}{2(1 - \rho_p)} = \bar{X} + \frac{\bar{\lambda} \cdot \bar{X}^2}{2(1 - \bar{\lambda} \cdot \bar{X})} \\ &= \bar{X} + \bar{\lambda} \cdot \frac{\sigma_X^2 + \bar{X}^2}{2(1 - \bar{\lambda} \cdot \bar{X})}, \end{aligned} \tag{15}$$

where \bar{X} and \bar{X}^2 are the first moment and the second moment of the service time, respectively. σ_X^2 is the variance of the service time.

5 Numerical results

For the numerical analysis, C is assumed to be 50. As for the threshold value K , the values 40 and 45 are evaluated.

Also, δ is set to 1,000, which represents that a replaced MN has more than 1,000 session arrivals while it performs a subnet crossing. The mean arrival rates, λ_N and λ_O , of the new and ongoing MNs are both set to 1/20 (1/s). To vary the traffic load, the mean service time (i.e., MAP domain residence time $1/\mu_N$ or $1/\mu_O$) is increased from 20 s to 1,180 s, which corresponds to an increase in the offered load, ρ_N or ρ_O , from 1 to 59. Let t_S and t_C be the random variables for inter-session arrival time and subnet residence time, which follow exponential distributions with mean $1/\lambda_S$ and $1/\mu_C$, respectively. Then, P_δ and $P(k, \delta)$ are respectively calculated as

$$P_\delta = \Pr(t_C > \delta \cdot t_S) = \int_0^\infty e^{-\delta\mu_C\tau} \cdot \lambda_S e^{-\lambda_S\tau} d\tau = \frac{\lambda_S}{\lambda_S + \delta\mu_C}$$

and

$$P(k, \delta) = \left(\frac{\delta\mu_C}{\lambda_S + \delta\mu_C} \right)^k.$$

5.1 New MN blocking probability

Figure 8 shows the new MN blocking probability as the new and ongoing MN loads increase. In Fig. 8(a), the new MN load increases from 1 to 59, whereas the ongoing MN load is fixed at 30. On the other hand, only the ongoing MN load is varied from 1 to 59 in Fig. 8(b). As shown in Fig. 8(a), the new MN blocking probability of the TLC scheme is higher than that of the NLC scheme since the TLC scheme reserves a part of the capacity for ongoing MNs. In addition, it can be seen that the blocking probability increases as the threshold value decreases. This is because more capacity is exclusively used for ongoing MNs when the threshold value is small. However, the difference in blocking probabilities of the NLC and TLC schemes does not change significantly even though the new

Fig. 8 New MN blocking probability. **a** New MN traffic load, **b** Ongoing MN traffic load

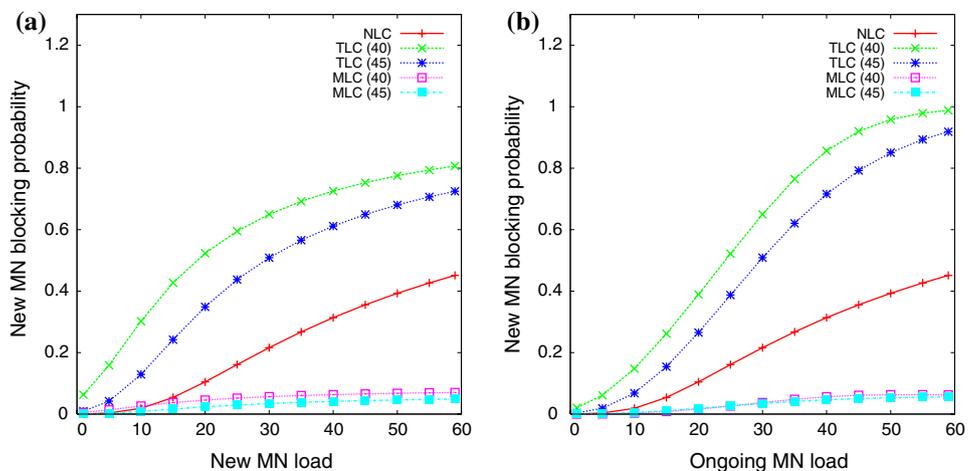
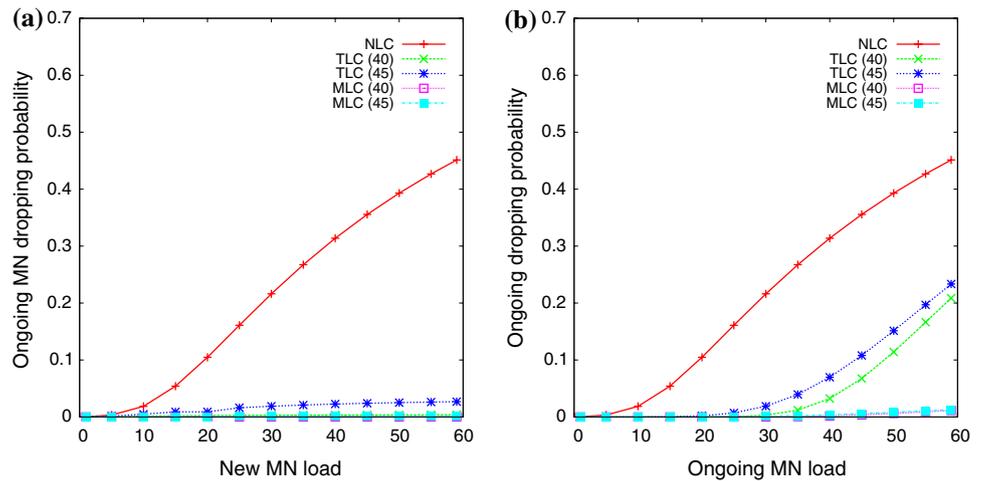


Fig. 9 Ongoing MN dropping probability. **a** New MN traffic load, **b** Ongoing MN traffic load



MN load (or ongoing MN load) increases. On the other hand, it can be found that the MLC scheme has a significantly lower new MN blocking probability than the NLC and TLC schemes. This is because the MLC scheme can accept more MNs, even though there is no MAP capacity available by adopting the SMR-based replacement algorithm.

5.2 Ongoing MN dropping probability

Figure 9 indicates the variations of the ongoing MN dropping probability as the new MN and ongoing MN loads increase. Since both the TLC and MLC schemes reserve an amount of MAP capacity for ongoing MNs, the ongoing MN dropping probabilities for both schemes are quite low, regardless of the new MN load (see Fig. 9(a)). However, as shown in Fig. 9(b), when the ongoing MN load exceeds a certain point, the ongoing MN dropping probability in the TLC scheme increases exponentially whereas the MLC scheme does not affected by the increased ongoing MN load. This is because the MLC scheme can reduce the ongoing MN dropping probability in conjunction with the SMR-based replacement algorithm, even in the case where the ongoing MN load is high.

5.3 Binding update cost

To plot the binding update cost, μ_C and λ_S are set to 0.1 and 0.01, respectively. Since δ is set to 1,000, μ_C^R and λ_S^R are assumed to be 0.01 and 100, respectively. The ratio of the number of new MNs to the total number of MNs, ω , can be approximated as

$$\omega \approx \frac{\rho_N}{\rho_N + \rho_O}.$$

The binding update cost is dependent on both the new MN blocking and the ongoing MN dropping probabilities.

If an MN is blocked or dropped by the admission control algorithm, the MN performs binding updates to its HA. Therefore, the blocking (or dropping) of an MN results in a higher binding update cost. Figure 10 shows the relative binding update costs of the TLC and MLC schemes to that of the NLC scheme. As mentioned before, the TLC scheme has a higher new MN blocking probability than the NLC scheme, because it proactively limits the number of MNs based on the threshold value. Therefore, the relative binding update cost of the TLC scheme is larger than 1.0. That is, the TLC scheme results in higher binding update traffic than the NLC scheme. However, as the MN load increases, the binding update cost of the TLC scheme approaches that of the NLC scheme. For the MLC scheme, since the MLC scheme has lower new MN blocking and ongoing MN dropping probabilities, its relative binding update cost to the NLC scheme is lower than 1.0 over wide range of new MN loads.

The effect of C_H/C_M on the binding update cost is also demonstrated in Fig. 10. When C_H/C_M is equal to 1, the relative binding update cost of the MLC scheme is slightly larger than 1.0. This is because the MLC scheme results in additional binding update traffic due to the SMR-based replacement algorithm. However, since the distance from the MN to the HA is typically much longer than the distance from the MN to the MAP, C_H is much higher than C_M . As indicated in Fig. 10, if C_H/C_M is larger than 1, the MLC scheme's relative binding update cost is smaller than 1.0. Furthermore, as the MN load increases, the gain of the MLC scheme becomes apparent.

On the other hand, the relative binding update cost is plotted as a function of the ongoing MN load in Fig. 11. Similar trends to Fig. 10 can be observed in Fig. 11. Namely, if C_H/C_M is equal to 1, the gain of the MLC scheme in the reduction of the binding update traffic is not significant. However, if C_H/C_M is greater than 1, the MLC scheme can reduce the binding update traffic compared

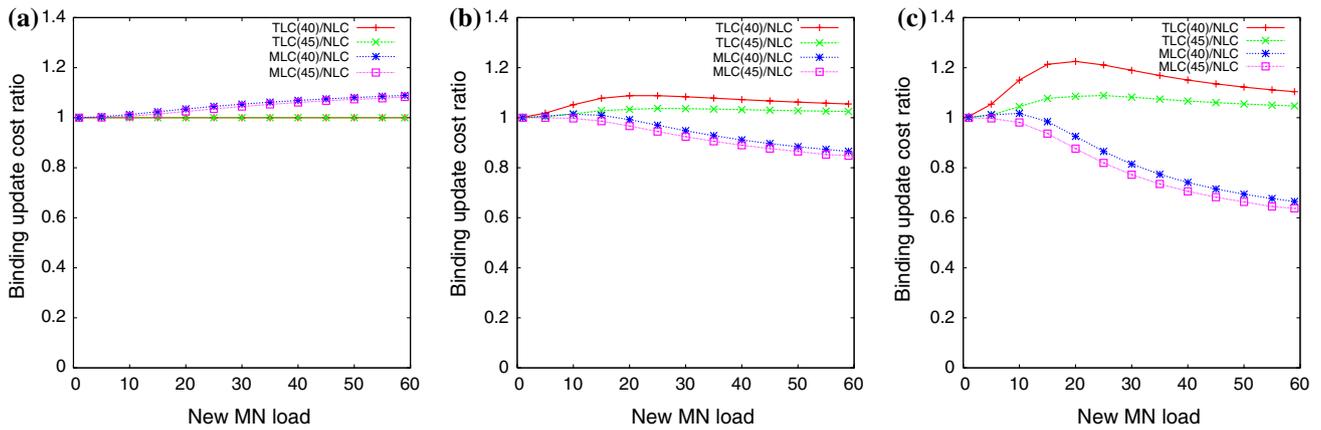


Fig. 10 Binding update cost ratio: New MN load. **a** $C_H/C_M = 1$, **b** $C_H/C_M = 2$, **c** $C_H/C_M = 4$

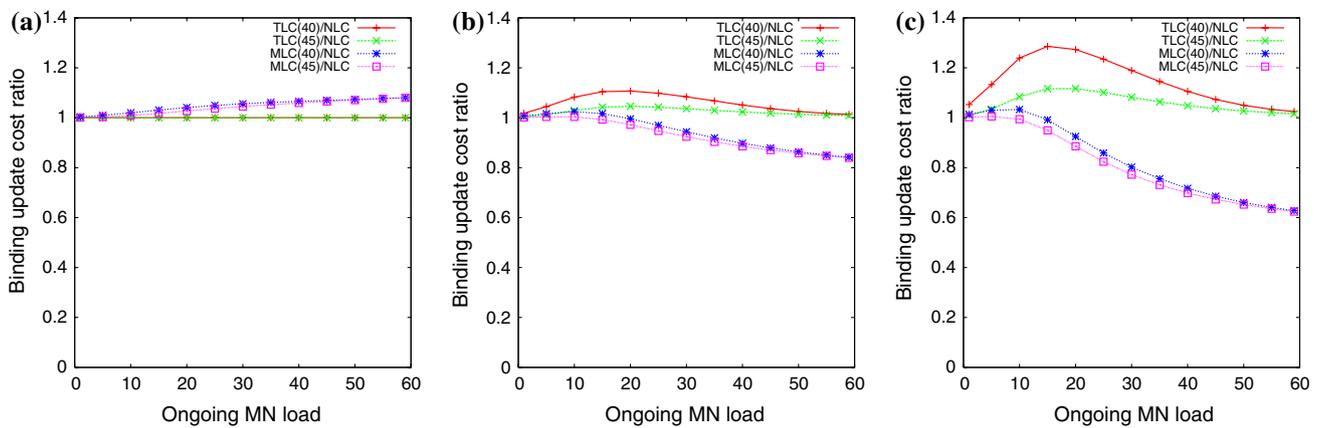


Fig. 11 Binding update cost ratio: Ongoing MN load. **a** $C_H/C_M = 1$, **b** $C_H/C_M = 2$, **c** $C_H/C_M = 4$

with the NLC and TLC schemes. Especially, the binding update cost gain becomes more apparent as the ongoing MN load increases. The TLC scheme shows a higher binding update cost, because of its higher new MN blocking probability. However, since the proportion of new MNs decreases as the ongoing MN load increases, the TLC scheme shows almost the same binding update cost as the NLC scheme when the ongoing MN load is high.

5.4 MAP processing latency

Figure 12 shows the MAP processing latency as a function of the packet arrival rate from an MN. The average service time \bar{X} and variance σ_X^2 of the service time are assumed to be 0.5 ms and 0.01 ms, respectively [18]. As there is no admission control in the NLC scheme, the average number of MNs serviced at the MAP is greater than that in either the TLC scheme or the MLC scheme. Therefore, more packets are concentrated to the MAP in the NLC scheme, which results in longer packet processing latency. In addition, the increasing rate of the MAP processing latency

in the NLC scheme is higher than those in both the TLC and MLC schemes.

5.5 Effect of SMR threshold

The effectiveness of the SMR-based replacement algorithm is highly dependent on the SMR threshold value. Therefore, we investigate the effect of the SMR threshold value on the blocking and dropping probabilities, and introduce an algorithm to determine the optimal SMR threshold value δ^* .

The new MN blocking and ongoing MN dropping probabilities for different SMR threshold values are illustrated in Fig. 13. The new and ongoing MN loads are both set to 30. As shown in Fig. 13, the ratio θ of the SMR threshold value to the average SMR is used as the x-coordinate. From Fig. 13(a), it can be seen that the new MN blocking probability is negligible when θ is less than 10. However, the blocking probability drastically increases when θ exceeds 10. A similar trend can be found in Fig. 13(b). Consequently, the new MN blocking and

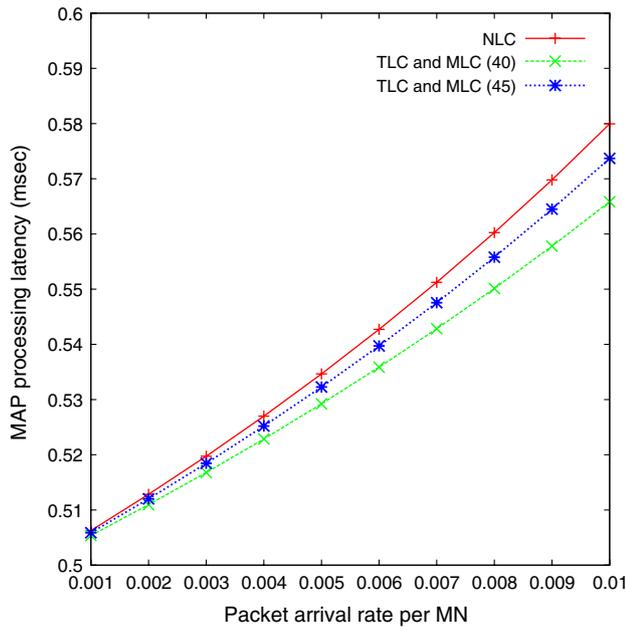


Fig. 12 MAP processing latency versus packet arrival rate

ongoing MN dropping probabilities can be represented as functions of θ .

From these results, we can devise an algorithm to find a suitable SMR threshold given the upper bounds for the new MN blocking probability and ongoing MN dropping probability. Algorithm 1 shows the proposed algorithm. Let T_{NB} and T_{OD} be the given upper bounds for the new MN blocking probability and ongoing MN dropping probability, respectively. α is a coefficient used to determine δ^* , based on the average SMR. In Algorithm 1, a number of threshold values are evaluated and these values are successively augmented by η , which is a step size. Once these repeated calculations have been done, this algorithm determines the largest δ meeting the upper blocking and dropping probabilities as δ^* . This is because a smaller δ results in more frequent

replacements even though it gives lower P_{NB} and P_{OD} . Note that the new MN blocking probability and ongoing MN dropping probability can be changed as MNs arrive or handoffs occur. Therefore, Algorithm 1 can be performed periodically to obtain the optimal performance.

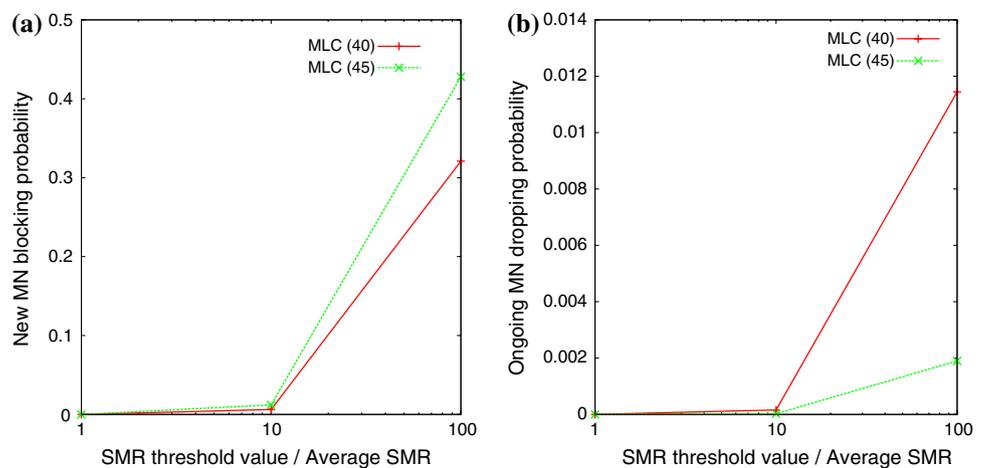
Algorithm 1 Determination of SMR threshold value

- 1: calculate the average SMR (SMR_A);
- 2: determine the average new MN and ongoing MN loads;
- 3: $\alpha \leftarrow 1$;
- 4: $\eta \leftarrow 1$;
- 5: $\delta \leftarrow \alpha \times SMR_A$;
- 6: calculate the new MN blocking probability (P_{NB}) and the ongoing MN dropping probability (P_{OD});
- 7: **while do**
- 8: **if** $P_{NB} < T_{NB}$ and $P_{OD} < T_{OD}$ **then**
- 9: $\alpha \leftarrow \alpha + \eta$;
- 10: $\delta \leftarrow \alpha \times SMR_A$;
- 11: calculate P_{NB} and P_{OD} ;
- 12: **else if** $P_{NB} > T_{NB}$ or $P_{OD} > T_{OD}$ **then**
- 13: **break**;
- 14: **end if**
- 15: **end while**
- 16: $\delta^* \leftarrow (\alpha - \eta) \times SMR_A$;

6 Simulation validation

To validate the analytical results, we have developed an event-driven simulator using C programming language and carried out comprehensive simulations. To obtain more accurate results, the simulation is conducted for 100,000 s and repeated 20 times. The simulation topology consists of 7 MAPs (six neighbor MAPs for each MAP) and the wrap around model is used to eliminate boundary effects [19]. In our simulations, new MNs are generated by a Poisson

Fig. 13 Effect of SMR value. **a** New MN blocking probability, **b** Ongoing MN dropping probability



process with rate λ_M . For the mobility model, we use the random-walk mobility model, where the routing probability for each neighbor MAP is identical. Since we assume a hexagonal MAP configuration, the routing probability is $1/6$. After deciding the direction of movement, the MN stays in a given MAP domain during t_R , which is the MAP domain residence time following a Gamma distribution and its probability density function (pdf) is

$$f_R(t) = \frac{b^k t^{k-1}}{\Gamma(k)} e^{-bt}, \tag{16}$$

where b is equal to $k\lambda_m$ and $\Gamma(k)$ is the Gamma function defined as $\int_0^\infty t^{k-1} e^{-t} dt$. The mean and variance of the Gamma distribution are $1/\lambda_m$ and $1/k\lambda_m^2$, respectively.

On the other hand, the MN's session arrival process follows a Poisson distribution with rate λ_S and the session length (in numbers of packets) follows a Pareto distribution with shape parameter a and scaling parameter k . The mean session length is $ak/(a-1)$. The pdfs of the session arrival process and session length process are respectively given by

$$f_A(k) = \frac{e^{-\lambda_S} \cdot \lambda_S^k}{k!} \tag{17}$$

and

$$f_D(t) = \frac{a}{k} \left(\frac{k}{t}\right)^{a+1}. \tag{18}$$

Tables 1 and 2 summarize the simulation and analytical results (including minimum, average, and maximum values) on the new MN dropping and ongoing MN dropping probabilities, respectively. As shown in these

tables, the analytical results are consistent with those obtained from the simulations. Specifically, the differences between the analytical and simulation results are less than 3–5% for most new MN loads.

7 Related works

Load control (or overflow control) at the mobility agent (e.g., visitor location register (VLR) and home location register (HLR)) has been extensively investigated in cellular networks [20–23]. For the purpose of load control, when the capacity of a mobility agent is full, an existing entry in the overloaded mobility agent is replaced with a new entry (i.e., a newly incoming mobile user). How to decide the entry to be replaced is the key issue for the load control in cellular networks. However, unlike VLR and HLR, since the MAP is in charge of data tunneling as well as mobility management, data traffic pattern should be considered for the load control at the MAP.

Several load control schemes have been reported for IP-based mobile systems, but they focused on the issue of load distribution among multiple HAs [24–27]. However, most mobility-related signaling and data traffic is processed at the MAP rather than the HA in HMIPv6 networks. Accordingly, the load control at the MAP is a more important issue than that at the HA.

In [28], a threshold-based MAP load control scheme was proposed. In this scheme, if a MAP receives a BU message and the number of MNs serviced by the MAP exceeds a predefined threshold, the MAP forwards the BU message to the

Table 1 New MN blocking probability: analytical results (A) versus simulation results (S)

New MN load	NLC				TLC (45)				MLC (45)			
	A	S (min)	S (avg)	S (max)	A	S (min)	S (avg)	S (max)	A	S (min)	S (avg)	S (max)
10	0.0187	0.0188	0.0194	0.0199	0.1292	0.1295	0.1301	0.1306	0.0087	0.0080	0.0085	0.0091
20	0.1048	0.1047	0.1053	0.1058	0.3488	0.3567	0.3572	0.3576	0.0236	0.0240	0.0244	0.0250
30	0.2161	0.2198	0.2204	0.2210	0.5089	0.5122	0.5127	0.5134	0.0344	0.0347	0.0352	0.0355
40	0.3138	0.3086	0.3092	0.3099	0.6117	0.6235	0.6240	0.6244	0.0413	0.0417	0.0423	0.0427
50	0.3931	0.3944	0.3949	0.3955	0.6805	0.6880	0.6884	0.6889	0.0460	0.0465	0.0472	0.0478

Table 2 Ongoing MN dropping probability: analytical results (A) versus simulation results (S)

New MN load	NLC				TLC (45)				MLC (45)			
	A	S (min)	S (avg)	S (max)	A	S (min)	S (avg)	S (max)	A	S (min)	S (avg)	S (max)
10	0.0187	0.0188	0.0194	0.0199	0.0048	0.0041	0.0049	0.0055	0.0003	0.0001	0.0002	0.0004
20	0.1048	0.1045	0.1053	0.1057	0.0090	0.0084	0.0093	0.0098	0.0007	0.0002	0.0007	0.0012
30	0.2161	0.2199	0.2204	0.2211	0.0188	0.0185	0.0192	0.0195	0.0010	0.0007	0.0010	0.0014
40	0.3138	0.3084	0.3092	0.3099	0.0226	0.0226	0.0231	0.0236	0.0012	0.0010	0.0013	0.0015
50	0.3931	0.3942	0.3949	0.3955	0.0252	0.0253	0.0258	0.0262	0.0014	0.0011	0.0014	0.0019

next candidate MAP. In [7], the MAP limits the number of new MNs in order to give high priority to ongoing MNs by adopting admission control algorithms. In [29], the MAP load distribution is achieved by using a dynamic MAP selection scheme. On the other hand, multi-level HMIPv6 has been introduced to mitigate the load at the MAP [30].

The existing load control schemes have the two representative limitations. First, the coordination between MAPs is not practical in a small foreign network where multiple MAPs cannot be deployed due to high cost. Moreover, the coordination between MAPs requires additional signaling overhead. Second, the existing schemes do not take into account MNs' characteristics, e.g., the handoff frequency and traffic pattern. Therefore, even though the load control schemes can reduce the MAP load, they may degrade the performance of individual MNs.

8 Conclusion

In this paper, we have proposed a MLC scheme, which consists of the threshold-based admission control algorithm and the SMR-based replacement algorithm. By combining these two algorithms, the MLC scheme significantly reduces the new MN blocking and ongoing MN dropping probabilities, compared with the TLC and NLC schemes. In addition, when the MLC scheme is used, the reduced MAP processing latency and binding update cost can be obtained. Unlike the existing load control schemes, the MLC scheme does not require any interaction among MAPs, which makes the MLC scheme to be implemented with minimal overhead. Moreover, the MLC scheme controls the MAP load adaptively to MNs' characteristics, i.e., session and mobility rates. Hence, overall network load can be distributed more efficiently. Supporting scalable mobile Internet service is one of the most crucial issues in future wireless/mobile networks. Therefore, in our future works, we will investigate how to implement and deploy the MLC scheme on a global-scale testbed.

Acknowledgments This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2008-314-D00354).

Appendix 1

Derivation of $\pi(i, j)$ in the TLC/MLC schemes

As shown in Fig. 7, the Markov chain for the TLC scheme (and the MLC scheme) has a state space, $S = \{(i, j) | 0 \leq i \leq K, 0 \leq i + j \leq C\}$, and the following balance equations are established:

$$\begin{aligned}
 (1) \quad & i = 0 \text{ and } j = 0: \\
 & (\lambda_O + \lambda_N) \cdot \pi(0, 0) = \mu_N \cdot \pi(1, 0) + \mu_O \cdot \pi(0, 1); \\
 (2) \quad & i = 0 \text{ and } 0 < j < K: \\
 & (\lambda_O + \lambda_N + j\mu_O) \cdot \pi(0, j) = \lambda_O \cdot \pi(0, j - 1) + \mu_N \cdot \pi(1, j) \\
 & \quad + (j + 1)\mu_O \cdot \pi(0, j + 1); \\
 (3) \quad & i = 0 \text{ and } K \leq j < C: \\
 & (\lambda_O + j\mu_O) \cdot \pi(0, j) = \lambda_O \cdot \pi(0, j - 1) + \mu_N \cdot \pi(1, j) + \\
 & \quad + 1)\mu_O \cdot \pi(0, j + 1); \\
 (4) \quad & i = 0 \text{ and } j = C: \\
 & C\mu_O \cdot \pi(0, C) = \lambda_O \cdot \pi(0, C - 1); \\
 (5) \quad & 0 < i < K \text{ and } j = 0: \\
 & (\lambda_O + \lambda_N + i\mu_N) \cdot \pi(i, 0) = \lambda_N \cdot \pi(i - 1, 0) + (i + 1)\mu_N \\
 & \quad \cdot \pi(i + 1, 0) + \mu_O \cdot \pi(i, 1); \\
 (6) \quad & 0 < i < K \text{ and } 0 < j < K - i: \\
 & (\lambda_O + \lambda_N + i\mu_N + j\mu_O) \cdot \pi(i, j) \\
 & \quad = \lambda_N \cdot \pi(i - 1, j) + \lambda_O \cdot \pi(i, j - 1) \\
 & \quad + (i + 1)\mu_N \cdot \pi(i + 1, j) + (j + 1)\mu_O \cdot \pi(i, j + 1); \\
 (7) \quad & 0 < i < K \text{ and } j = K - i: \\
 & (\lambda_O + i\mu_N + j\mu_O) \cdot \pi(i, j) \\
 & \quad = \lambda_N \cdot \pi(i - 1, j) + \lambda_O \cdot \pi(i, j - 1) \\
 & \quad + (i + 1)\mu_N \cdot \pi(i + 1, j) + (j + 1)\mu_O \cdot \pi(i, j + 1); \\
 (8) \quad & 0 < i < K \text{ and } K - i < j < C - i: \\
 & (\lambda_O + i\mu_N + j\mu_O) \cdot \pi(i, j) \\
 & \quad = \lambda_O \cdot \pi(i, j - 1) + (i + 1)\mu_N \cdot \pi(i + 1, j) \\
 & \quad + (j + 1)\mu_O \cdot \pi(i, j + 1); \\
 (9) \quad & 0 < i < K \text{ and } j = C - i: \\
 & (i\mu_N + j\mu_O) \cdot \pi(i, j) = \lambda_O \cdot \pi(i, j - 1); \\
 (10) \quad & i = K \text{ and } j = 0: \\
 & (\lambda_O + K\mu_N) \cdot \pi(K, 0) = \lambda_N \cdot \pi(K - 1, 0) + \mu_O \cdot \pi(K, 1); \\
 (11) \quad & i = K \text{ and } 0 < j < C - K: \\
 & (\lambda_O + K\mu_N + j\mu_O) \cdot \pi(K, j) = \lambda_O \cdot \pi(K, j - 1) + (j + 1)\mu_O \\
 & \quad \cdot \pi(K, j + 1); \\
 (12) \quad & i = K \text{ and } j = C - K: \\
 & (K\mu_N + (C - K)\mu_O) \cdot \pi(K, C - K) \\
 & \quad = \lambda_O \cdot \pi(K, C - K - 1).
 \end{aligned}$$

Each balance equation can be represented as a form of $\pi(i, j) = A\pi(i - 1, j) + B\pi(i + 1, j) + C\pi(i, j - 1) + D\pi(i, j + 1)$, where A , B , C , and D are constants. Then, the steady state probability can be computed using an

iterative algorithm and the normalization condition, $\sum_{i=0}^K \sum_{j=0}^{C-i} \pi(i,j) = 1$. In the iterative algorithm, ε is a sufficiently small value and $|S|$ represents the cardinality of S .

Iterative algorithm

Step 1: t is set to 0.

Step 2: For all $(i, j) \in S$, $\pi^t(i, j)$ is initialized as $1/|S|$.

Step 3: For all $(i, j) \in S$, $\pi^{t+1}(i, j) \leftarrow \pi^t(i, j)$.

Step 4: For all $(i, j) \in S$, compute $\pi^{t+1}(i, j)$ by balance equations, $\pi^{t+1}(i, j) = A\pi^t(i-1, j) + B\pi^t(i+1, j) + C\pi^t(i, j-1) + D\pi^t(i, j+1)$.

Step 5: For all $(i, j) \in S$, if $|\pi^{t+1}(i, j) - \pi^t(i, j)| < \varepsilon$, the iteration is terminated. Otherwise, $t \leftarrow t + 1$ and go to Step 3.

References

- Akyildiz, I., Mcnair, J., Ho, J., Uzunalioglu, H., & Wang, W. (1999). Mobility management in next-generation wireless systems. *Proceedings of the IEEE*, 87(8), 1347–1384.
- Akyildiz, I., Xie, J., & Mohanty, S. (2004). A survey of mobility management in next-generation all IP-based wireless systems. *IEEE Wireless Communications*, 11(4), 16–28.
- Johnson, D., Perkins, C., & Arkko, J. (2003). Mobility support in IPv6. *IETF RFC 3775*, June 2003.
- Soliman, H., Castelluccia, C., Malki, K. E., & Bellier, L. (2005). Hierarchical Mobile IPv6 mobility management (HMIPv6). *IETF RFC 4140*, August 2005.
- Pack, S., Nam, M., Kwon, T., & Choi, Y. (2006). An adaptive mobility anchor point selection scheme in Hierarchical Mobile IPv6 networks. *Elsevier Computer Communications*, 29(16), 3066–3078.
- Hong, D., & Rappaport, S. (1986). Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures. *IEEE Transactions on Vehicular Technology*, 35(3), 77–92.
- Pack, S., Lee, B., & Choi, Y. (2004). Load control scheme at local mobility agent in mobile IPv6 networks. In *Proceedings of World Wireless Congress (WWC) 2004*, May 2004.
- Fang, Y., & Zhang, Y. (2002). Call admission control schemes and performance analysis in wireless mobile networks. *IEEE Transactions on Vehicular Technology*, 51(2), 371–382.
- Pack, S., & Choi, Y. (2004). A study on performance of Hierarchical Mobile IPv6 in IP-based cellular networks. *IEICE Transactions on Communications*, E87-B(3), 462–469.
- Hwang, S., Lee, B., Han, Y., & Hwang, C. (2004). An adaptive Hierarchical Mobile IPv6 using mobility profile. *Wiley Wireless Communication and Mobile Computing*, 4(2), 233–245.
- Ma, W., & Fang, Y. (2004). Dynamic hierarchical mobility management strategy for mobile IP networks. *IEEE Journal on Selected Areas in Communications*, 22(4), 664–676.
- Ortigoza-Guerrero, L., & Aghvami, A. H. (1999). A prioritized handoff dynamic channel allocation strategy for PCS. *IEEE Transactions on Vehicular Technology*, 48(4), 1203–1215.
- Zhuang, W., Bensaou, B., & Chua, K. (2000). Adaptive quality of service handoff priority scheme for mobile multimedia networks. *IEEE Transactions on Vehicular Technology*, 49(2), 494–505.
- Huang, Y. (2002). Determining the optimal buffer size for short message transfer in a heterogeneous GPRS/UMTS Network. *IEEE Transactions on Vehicular Technology*, 52(1), 216–225.
- Lin, P., Lai, W., & Gan, C. (2004). Modeling opportunity driven multiple access in UMTS. *IEEE Transactions on Wireless Communications*, 3(5), 1669–1677.
- Xiao, Y., Pan, Y., & Li, J. (2004). Design and Analysis of Location Management for 3G Cellular Networks. *IEEE Transactions on Parallel and Distributed Systems*, 15(4), 339–349.
- Kleinrock, L. (1975). *Queueing systems, Volume 1: Theory*. New York: Wiley.
- Wang, W., & Akyildiz, I. (2001). A cost-efficient signaling protocol for Mobility Application Part (MAP) in IMT-2000 systems. In *Proceedings of ACM Mobicom, 2001*, October 2001.
- Zeng, H., Fang, Y., & Chlamtac, I. (2002). Call blocking performance study for PCS networks under more realistic mobility assumptions. *Telecommunication Systems*, 19(2), 125–146.
- Lin, Y. (2000). Overflow control for cellular mobility database. *IEEE Transactions on Vehicular Technology*, 49(2), 520–530.
- Lin, Y. (2001). Eliminating overflow for large-scale mobility databases in cellular telephone networks. *IEEE Transactions on Computers*, 50(4), 356–370.
- Hung, H., Lin, Y., Peng, N., & Yang, S. (2001). Resolving mobile database overflow with most idle replacement. *IEEE Journal on Selected Areas in Communications*, 19(10), 1953–1961.
- Luo, Y., Pan, Y., Li, J., Xiao, Y., & Lin, X. (2003). A new overflow replacement policy for efficient location management in mobile networks. In *Proceedings of IEEE ISDCS 2003*, May 2003.
- Jue, J., & Ghosal, D. (1998). Design and analysis of a replicated server architecture for supporting IP-host mobility. *Cluster Computing*, 1(2), 249–260.
- Deng, H., Huang, X., Zhang, K., Niu, Z., & Ojima, M. (2003). A hybrid load balance mechanism for distributed home agents in mobile IPv6. In *Proc. IEEE PIMRC 2003*, September 2003.
- Vasilache, A., Li, J., & Kameda, H. (2003). Threshold-based load balancing for multiple home agents in mobile IP networks. *Telecommunication Systems*, 22(1-4), 11–31.
- Fritsche, W., & Guardini, I. (2006). Deploying home agent load sharing in operational mobile IPv6 networks. In *Proceedings of ACM MobiArch 2006*, December 2006.
- Bandai, M., & Sasase, I. (2003). A load balancing mobility management for multilevel Hierarchical Mobile IPv6 networks. In *Proceedings of IEEE PIMRC 2003*, September 2003.
- Taleb, T., Suzuki, T., Kato, N., & Nemoto, Y. (2005). A dynamic & efficient MAP selection scheme for mobile IPv6 networks. In *Proceedings of IEEE Globecom 2005*, December 2005.
- Pack, S., Nam, M., & Choi, Y. (2006). Design and analysis of optimal multi-level hierarchical mobile IPv6 networks. *Wireless Personal Communications*, 36(2), 95–112.

Author Biographies



Sangheon Pack received the B.S. (2000) and Ph.D. (2005) degrees from Seoul National University, both in Computer Engineering. Since March 2007, he has been an assistant professor in the School of Electrical Engineering, Korea University, Korea. From 2005 to 2006, he was a postdoctoral fellow in the Broadband Communications Research (BBCR) Group at University of Waterloo, Canada. From 2002 to 2005, he was a recipient of the Korea

Foundation for Advanced Studies (KFAS) Computer Science and Information Technology Scholarship. He has been also a member of Samsung Frontier Membership (SFM) from 1999. He received a student travel grant award for the IFIP Personal Wireless Conference (PWC) 2003. He was a visiting researcher at Fraunhofer FOKUS, Germany, in 2003. His research interests include mobility management, multimedia transmission, and QoS provision issues in next-generation wireless/mobile networks.



Taekyoung Kwon an associate professor in the School of Computer Science and Engineering, Seoul National University (SNU) since 2004. Before joining SNU, he was a post-doctoral research associate at UCLA and at City University New York (CUNY). He obtained B.S., M.S., and Ph.D. degree from the Department of Computer Engineering, SNU, in 1993, 1995, 2000, respectively. During his graduate program, he was a visiting student at IBM T.

J. Watson research center and University of North Texas. His research interest lies in sensor networks, wireless networks, IP mobility, and ubiquitous computing.



Yanghee Choi received B.S. in Electronics Engineering at Seoul National University, M.S. in Electrical Engineering at Korea Advanced Institute of Science, and Doctor of Engineering in Computer Science at Ecole Nationale Supérieure des Télécommunications Paris, in 1975, 1977, and 1984, respectively. He has been with Electronics and Telecommunications Research Institute during 1977–1979 and 1984–1991 as director of Data Com-

munication Laboratory, and Protocol Engineering Center. He worked also at Centre National D'Etude des Télécommunications, France for 1981–1984, and at IBM Thomas J. Watson Research Center during 1988–1989 as a visiting scientist. Since 1991, he is professor at School of Computer Science and Engineering, Seoul National University. He was chairman of SIG on Information Networking, editor-in-chief for the society journal, and is now president-elect of Korea Information Science Society. He is now director of Computer Network Research Center at Seoul National University, and leads Multimedia and Mobile Communications Laboratory. He was Associate Dean of Research Affairs of Seoul National University. He is also chair of the Future Internet forum of Korea.