Load sharing and session preservation with multiple mobile routers for large scale mobile networks

Eun Kyoung Paik*

School of Electrical Engineering and Computer Science,
Seoul National University,
Seoul 151 744, Korea
Now working at Future Technology Lab. KT
Fax: +82-2-526-5759
E-mail: euna@kt.com
*Corresponding author

Hosik Cho

School of Electrical Engineering and Computer Science,
Seoul National University,
Seoul 151 744, Korea
E-mail: hscho@mmlab.snu.ac.kr

Thierry Ernst

Keio University, Jun Murai Lab,
5322 Endo, Fujisawa-shi, Kawagawa 252-8520 Japan,
Now working at INRIA – France
E-mail: thierry.ernst@inria.fr

Yanghee Choi

School of Electrical Engineering and Computer Science,
Seoul National University,
Seoul 151 744, Korea
E-mail: yhchoi@snu.ac.kr

Abstract: By means of NEtwork MObility (NEMO) support, users can organise their various communication devices into a subnet called mobile network. The *Mobile Router* (MR) provides the connectivity to the internet and mobility management transparency for the rest of the mobile nodes in the mobile network. So, the MR must assure reliable communications and a high data rate for the group of nodes behind it. In support of broadband wireless communications, the use of multiple MRs would allow the transfer of large volumes of data to a group of mobile nodes. This paper addresses the protocol issues arising from the use of multiple MRs, and analyses the influence of mobility patterns on load sharing and session preservation. Then, two mobility-aware selection schemes are proposed, which select a MR. Firstly, the best-connected MR selection scheme is designed for regular mobility and is based on the fact that the most recently updated binding remains effective for the longest period of time. It preserves the connection from the mobile network to the internet whether the accessible area is continuous or not, and this even in a fast handoff environment. Secondly, the most-beneficial MR selection scheme is designed for random mobility. The simulation results show that the performance varies depending on the hardware configuration. Based on these results, it was found that application-awareness is critical to achieve good performance.

Keywords: IPv6; load sharing; Mobile IPv6; Mobile Router; MR; multihoming; multi-homing; Network Mobility; NEMO; session preservation; Vehicular Area Networks; VAN.

Reference to this paper should be made as follows: Paik, E.K., Cho, H., Ernst, T. and Choi, Y. (2008) 'Load sharing and session preservation with multiple mobile routers for large scale mobile networks', *Int. J. Wireless and Mobile Computing*, Vol. 3, Nos. 1/2, pp.56–68.

Biographical notes: Eun Kyoung Paik received a BS and an MS in Computer Science from Ewha Womans University and Doctor of Engineering in Electric Engineering and Computer Science from Seoul National University in 1990, 1992 and 2004, respectively. She has worked

on Multimedia Communications for KT from 1992 to till date and worked at IBM Thomas J. Watson Research Center (USA), Keio University (Japan) and Ecole Nationale Superieure des Telecommunications Bretagne (France). She is a Chair of Mobility Working Group at IPv6 Forum Korea. Her research interests are network mobility, mobility management, multihoming, IPv6 and future internet.

Hosik Cho received a BS in Computer Science and Engineering from Seoul National University in 2002. He is now a PhD student in the Multimedia and Communications Lab. at Seoul National University. His research interests are mobile networks, mobile IP, IPv6 and multimedia streaming.

Thierry Ernst received a PhD in Networking from University Joseph Fourier (Grenoble, France) in October 2001 and prepared simultaneously at INRIA and MOTOROLA Labs in France. At the time of writing this paper, he was a Researcher at Keio University, Japan, where he led the Nautilus6 Working Group within the WIDE project. He has setup and chaired the NEtwork MObility (NEMO) Working Group at the IETF. He is also involved with ISO Technical Committee 204 Working Group 16 (TC204 WG16) which is investigating IPv6 communications for vehicles. He now works at INRIA in the IMARA project team.

Yanghee Choi received a BS in Electronics Engineering at Seoul National University, MS in Electrical Engineering at Korea Advanced Institute of Science and Doctor of Engineering in Computer Science at Ecole Nationale Superieure des Telecommunications Paris, in 1975, 1977 and 1984, respectively. He has worked at Electronics and Telecommunications Research Institute (Korea), Centre National D'Etude des Telecommunications (France) and at IBM Thomas J. Watson Research Center (USA), before joining Seoul National University in 1991. He is now president of Korea Institute of Information Scientists and Engineers. He is also Chair of the Future Internet Forum of Korea. He has published over 500 papers in network protocols and architectures.

1 Introduction

AS the number of mobile subscribers increases, the demand for ubiquitous internet communication, that is anywhere, anytime with any device is growing, for example, broadband mobile wireless data communication even on moving vehicles. As indicated in Kellerer et al. (2001), Ernst and Uehara (2002), and Ernst et al. (2003), vehicles represent highly mobile end systems that serve as a service platform for passengers and that use communication services themselves. In addition to this trend, there is a growing tendency for individuals to carry more than one electronic device.

These trends are leading to the development of a new type of mobility, that is, the mobility of an entire network, as illustrated by the following examples:

- An *access network* deployed into a vehicle. This provides the interconnection of the devices belonging to the passengers. In this way, the mobile devices connected to the network can all communicate with the internet.
- A *sensor network* embedded into a vehicle. This allows the interconnection of devices permanently installed inside the vehicle. In this way, fixed devices such as sensors, controllers and local servers, which are connected to the network, can all communicate with the internet.
- A *Personal Area Network (PAN)* carried by an individual. For example, mobile devices such as Personal Digital Assistants (PDAs), mobile phones, radios, portable computers and medical sensors can be organised into a PAN.
- *Nested mobile networks***.** The PAN belonging to an individual rides in a vehicle. The network embedded in a vehicle allows the interconnection of the PAN belonging to the passenger.

As part of the ongoing efforts towards the development of ubiquitous networks, attempts are being made to address all of the issues pertaining to the converging technologies. These efforts include the development of new IP layer protocols, and it is critical that these protocols should guarantee the Quality-of-Service (QoS) that accommodates various requirements, such as a high data rate, low delay and low jitter. The Internet Engineering Task Force (IETF) *NEtwork MObility* (NEMO) Working Group (WG) has been established in order to provide the capability for a particular type of network, which is referred to as a *mobile network*, to maintain its internet connectivity while it changes its point of attachment to the internet.

Mobile networks are connected to the internet using a Mobile Router (MR) via wireless links, and *wireless links suffer from a low data rate and a high error rate*. In spite of the high error rate, connectivity needs to be maintained, in order to enable internet access anywhere and at anytime. The low data rate problem is aggravated as the number of mobile nodes running multimedia applications increases, since they have to share the same connection via a single MR. In addition, if the connection is broken, it is not just the MR that is affected, but all of the nodes inside the mobile network. In such an eventuality, these nodes can no longer communicate with their Correspondent Nodes (CNs) outside of the mobile network. So, the MR must provide reliability and support a high data rate.

To provide QoS, all network layers from high to low, and all network components from end to end must be coordinated. To provide mobile network with QoS, we propose a load sharing and session preservation architecture with multiple MRs. To use multiple MRs in an advantageous way, we analyse the mobility patterns of NEMO applications, for example embedded access networks in vehicles, and proposed two adaptive MR selection schemes.

Previous approaches to multiple connections used redundant HAs (Deng et al., 2003; Vasilache et al., 2003), multiple routers (Draves and Hinden, 2005) or multiple interfaces (Wakikawa et al., 2002). However, Vasilache et al. (2003) and Deng et al. (2003) did not take into account dynamic wireless channel conditions. Managing the traffic load between MRs requires awareness of the dynamic wireless connection status, since the wireless connection of each MR varies as the mobile network moves. The Router Advertisement (RA) message extension proposed in Draves and Hinden (2005) is aimed at selecting a router. The use of MR advertisement messages is restricted within a mobile network, since MRs act as mobile nodes to the internet. MRs do not advertise route information to the internet but to the nodes within the mobile network. So this proposition can be used only by the nodes within the mobile network. It cannot be used by internet side. The use of multiple interfaces proposed in Wakikawa et al. (2002) made use of multihoming, but it did not consider the spatial advantage that multiple MRs provide.

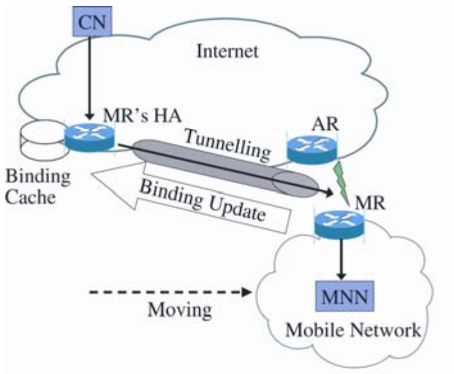
The remaining of this paper is organised as follows: A brief review of NEMO Basic Support is described in Section 2. Section 3 starts by discussing the current issues and problems surrounding the implementation of multiple MRs for load sharing and session preservation. Section 4 describes the mobility-aware protocol design. Section 5 explains two proposed MR selection schemes for session preservation and load sharing. Then, in Section 6, the effects of load sharing and session preservation with multiple MRs are investigated by means of simulation. Finally, we make conclusion on this paper in Section 7, along with a discussion of future work.

2 NEMO Basic Support

Figure 1 shows the protocol operation of NEMO Basic Support (Devarapalli et al., 2005). A mobile network is composed of one or more MRs with communication nodes called Mobile Network Nodes (MNNs) behind the MR. There are three kinds of MNNs (Ernst and Lach, 2007): Local Fixed Nodes (LFNs), Local Mobile Nodes (LMNs) and Visiting Mobile Nodes (VMNs). LFNs never change their point of attachment with respect to the MR, while LMNs and VMNs do. LFNs and LMNs belong to the mobile network, while VMNs do not. All MNNs connect to the internet via the MR. The MR and the Access Router (AR) connect to each other over a wireless link.

The NEMO basic support protocol is based on Mobile IPv6 (Johnson et al., 2004). The main idea of Mobile IPv6 is to separate the two functions of IP addresses: identification and location. So, each mobile node in Mobile IPv6 has two addresses: Home Address (HoA) for identification and Care-of Address (CoA) for location.

Figure 1 Basic architecture of network mobility (see online version for colours)



In NEMO basic support, the Home Agent (HA) and the MR communicate by means of bi-directional tunnelling. When the mobile network moves, the MR sends a Binding Update (BU) to its HA with its CoA. The CoA is a new address that is obtained on each visited link.

The main idea of NEMO basic support is to separate the function of MR interface into two: ingress interface and egress interface. The ingress interface connects to MNNs and manages Mobile Network Prefix (MNP) to identify the mobile network. The egress interface connects to the internet and manages the mobile location by obtaining CoA. So the BU message carries two bindings: HoA to CoA binding for the MR and MNP to CoA binding for MNNs. In this way, the HA can forward packets intended for the MR and MNNs to this CoA.

3 Protocol issues for load sharing and session preservation

When it comes to QoS, the weakest part of mobile networks is the MR. The MR is a critical component of mobile networks since all MNNs' communication pass through it. We face the following problems:

- MR is a single point of failure
- wireless connections to the internet are more sensitive to failure.

When there is only one MR in the mobile network, the traffic of all MNN's has to transit via the MR, so the latter constitutes a single point of failure. The MNNs rely on a single MR for their connection to the internet, and they have to share the same connection. Whenever the MR becomes disconnected, it blocks the communication of all of the nodes in the mobile network. Thus, none of the nodes inside the mobile network can communicate with their correspondents outside the mobile network. In contrast to the issue of host mobility, wherein only one node is affected, NEMO has an impact on all of the nodes located behind the MR. So, providing a reliable connection between the mobile network and the internet is important for the NEMO architecture.

Moreover, the MRs are connected to the internet using wireless technologies. Since all traffic to/from the mobile network must pass through the MR, the wireless connection between the AR and the MR is critical to the communication of the entire mobile network. The wireless link, however, suffers from intrinsic deficiencies as follows (Cao and Li, 2001):

- high error rate and bursty errors
- location-dependent and time-varying wireless link capacity
- scarce bandwidth
- the power constraint of the mobile nodes.

So, it is important to provision QoS for wireless connections, that is the connection between the MR and the internet.

To solve these problems, maintaining multiple connections to the internet, namely *multihoming* (Ng et al., 2007), can help. The importance of multihoming increases as the need for system reliability increases. By maintaining multiple connections, the failure of one connection can be compensated for by another, while in normal operation the traffic load can be distributed between these multiple connections.

When considering load sharing between multiple MRs or recovery from the failure of an MR, the following protocol issues need to be addressed: the physical availability of the MRs, the per-flow MR selection and session preservation. These multiple MRs can be connected to the same link or to different links.

3.1 Physical availability of MRs

With wireless channels, not all MRs may be able to maintain physical connections permanently. In this case, the question arises as to how, the other MRs can become aware of the failure of another MR's connection so as to perform recovery, by redirecting the traffic flow while preserving mobility management transparency. To achieve efficient load sharing and/or session preservation results, a mechanism that notifies the MRs of each other's availability is essential.

Multiple connections can be used either *simultaneously* or *one at a time.*

- When multiple connections are used simultaneously, the mode of operation can be either primary–secondary or peer-to-peer. Simultaneous usage is especially useful for heavy traffic, but there are many implementation issues that need to be addressed. For example, not all of the connections may be maintained physically all of the time. In this case, the question arises as to how the other connections can be apprised of the disconnection of the inoperable one and the traffic flow redirected, while at the same time sharing the traffic load.
- When only one connection can be used at a time, for example, in the case where a single connection has to substitute for all the other candidates, an MR selection policy and switching mechanism is needed for preserving the ongoing session.

Both the above cases require the awareness of connection availability.

In the case of NEMO, connection (i.e. a bi-directional tunnel between the MR and the HA) availability is changed dynamically. Mobility leads to the dynamic creation and expiration of wireless connections between the MR(s) and the AR(s). Therefore, a mobility-aware scheme is required to ascertain connection availability.

3.2 MR selection entity

The MR can be selected by the HA, the MRs themselves, and/or the MNNs:

- To let the HA select which MR to use is the easiest solution. For example, an Internet Service Provider (ISP) can set up a selection policy for load sharing at the HA.
- The second possibility is for the MRs themselves to decide which MR to use. This should require coordination between multiple MRs. For example, the MRs can distribute the load by electing the most available MR for a given flow.
- The MNNs should also be able to select which MR to use, in the case where a user or an application wants to select a particular access technology, ISP or data rate according to his/her preference. This can be implemented at the application layer so that the user can explicitly select the MR, or at the IP layer so that an appropriate IP layer mechanism performs the selection according to the user's preference.
- A hybrid mechanism should also be available, for example, the HA, the MR and the MNNs can be coordinated.

3.3 Criteria for selection

The mechanisms used for selecting an MR from multiple MRs can be listed as follows:

- simple default MR selection
- simple round robin
- the best accessibility to ARs
- minimise the maximum channel utilisation
- isolation of different types of traffic, for example, signal, VoIP, multimedia, etc.
- routing metric
- ISP policy, for example, price.

4 Design of mobility-aware protocol

To achieve session preservation and load sharing described in Section 3, this section begins by analysing the protocol requirements for connection selection in multihomed mobile networks. Then, the characteristics of mobility and access technologies are classified for the purpose of proposing various solutions. There, then follows a description of the selection entity, and design approach to use with multiple MRs.

4.1 Protocol requirements

Based on the listed issues in Section 3, the protocol requirements are analysed as follows:

- *Connection availability*: One of the most critical factors that affect connection availability is mobility. So the proposed scheme should be *mobility-aware*.
- *Connection selection*: To implement a mobility-aware selection protocol, the selection entity (i.e. an entity that selects the connection) needs to be aware of mobility related issues and problems.
- *Selection criteria*: In NEMO, the wireless channel conditions change dynamically, and the MR's connection availability varies according to the network conditions. Thus, the selection scheme needs to be able to decide quickly, before the conditions change. So, a selection algorithm that requires a long computation time is not appropriate. To be able to react to varying connection availability conditions, simple and fast algorithms are required.

4.2 Mobility-awareness

Connection availability depends on the dynamic creation and expiration of wireless connections between the MR and the AR, so classifications of MR mobility and access technology are needed for designing mobility-awareness schemes.

4.2.1 Mobility classification

For the purpose of providing mobility-awareness, mobility can be classified into two categories: *regular* mobility and *random* mobility. If the mobility is regular, we can predict the mobility and make use of this advantage. In the case of irregular and unpredictable mobility, we can select the best connection at each moment.

The mobility characteristics of mobile networks depend on the type of embedded mobile networks. To establish a mobility model for NEMO, we need to consider the various operational constraints. Table 1 gives the mobility patterns of vehicles and pedestrians. As described in Section 1, pedestrians may carry PANs.

4.2.1.1 Movement limitation The movement patterns of vehicles are limited according to the road topology. For example, vehicles such as trains and aircraft can only follow predefined paths, that is, railroads or established flight paths. Bus routes are more flexible than those of trains or airplanes. In the case of buses operating on a highway, however, this flexibility decreases. So, we can apply a certain movement limitation characteristic to public vehicles. In comparison, pedestrians also walk along the road, but the road topology of pedestrians is much denser than that of vehicles. So, the road topology of pedestrians does not limit movement, as much as that of vehicles does.

4.2.1.2 Regularity The movement pattern of public vehicles has a tendency to be regular and iterative. They follow the same route and stop at the same places, for example, train stations or bus stops. In the case of trains or subways, their timing is also strictly regular.

4.2.2 Access technology classification

Connection availability is also affected by access technologies. Table 2 gives the characteristics of various access technologies. Wireless LANs provide a high data rate at low cost, while cellular networks provide a large access area at reasonable cost.

4.3 Selection entity

In NEMO, the HA manages the mobility information, that is, the binding between the HoA and the CoA in the binding cache. In this way, the HA can select a connection with up-to-date mobility information.

In order for it to use multiple connections, a mobile network requires the HA to maintain multiple binding entries for it. From these multiple binding entries for the mobile network, the HA chooses which connection to use by means of a selection algorithm.

MRs and/or MNNs can also participate in selection process. However, to make MR(s) select a connection, additional implementation is needed for MR(s) to know other MRs availability information. To make MNN(s) select a connection, new selection algorithm should be implemented into all mobile terminals. It costs much than to implement a new algorithm only into a HA.

4.4 Design approach: case study

One single selection algorithm cannot always provide good performance with all possible configurations. So different selection algorithms need to be provided according to the multihoming configuration. However, not all of the possible multihoming configurations described in Ng et al. (2007) are practical, so it is necessary to focus on providing solutions for realistic scenarios.

As given in Table 2, various access technologies are possible, so we can select a beneficial access technology. For example, wireless LAN provides a high bandwidth with low cost. However, it is not accessible at all places, since its access coverage is small. For special vehicles such as trains whose mobility patterns are regular as given in Table 1, it is possible to use wireless LAN with session preservation. For the other case, that is, random mobility vehicles, it is practical to make use of heterogeneous access technologies with load sharing schemes.

5 Mobility-aware MR selection schemes

This section explains the proposed MR selection schemes. Since the connection availability differs according to different configuration of MRs and mobility environment, we propose different MR selection schemes for each application scenarios. The first case is a train or subway.

A mobile network in a train or subway allows passengers to access the internet, and its sensors to be organised into a sensor network for its own use. So, public vehicles are one of the most useful NEMO applications.

The second case is cars. With the advances currently being made in the field of telematics, car communication is becoming increasingly popular. To connect to the internet with telematics technologies, NEMO Basic Support is one of the most feasible protocols to use. Car electronic systems such as navigators and audio systems can be organised into a mobile network along with personal mobile devices.

With respect to the above two cases, two mobility-aware selection schemes are proposed in the following sections: the *best-connected selection scheme* and the *most-beneficial selection* scheme.

5.1 Session preservation based on location

The first selection scheme is designed for regular mobility, and is aimed at always-best-connected mobile networks.

5.1.1 Case scenario: multiple MR configuration for public access networks in trains

As described in Table 1, the movement patterns of public trains or subways are severely limited, and we can take advantage of this predictability. Since wireless LANs provide high data rates at low cost, we can install them along the railroad, that is, in a predefined path.

The routers and hosts inside a train or subway are organised into a mobile network. This mobile network is supposed to have multiple MRs as described in Section 4. Each MR is made of a wireless access point, which provides AR functionality for the MNNs. Both the egress and ingress interfaces of MRs are wireless LAN interfaces.

With this configuration (Figure 2), communication for the VMNs can be provided as follows. A passenger with a mobile device can perform mobile communication using Mobile IPv6 before getting onto the train. When the passenger gets onto the train, his or her mobile device receives a RA message from the MR in the train, and in this way the mobile device detects its movement. It configures a CoA with the MNP. From that time on, the mobile device can be served as a VMN in this mobile network. The mobile device does not have to deal with either the problem of the train's mobility or the existence of multiple MRs, since the VMN's CoA is not changed as long as it stays in the same train. When the passenger gets off the train, the mobile device receives a RA message from an AR outside this mobile network, and associates with it, in order to keep on communicating.

5.1.2 Protocol behaviour

This section explains the *best-connected MR selection* scheme for session preservation, which is based on the MR's location. This scheme is based on the characteristics of movement limitation and regularity. Regular mobility can be modelled by linear mobility. Since trains/subways follow predefined railroads and do not move back-and-forth, there is no ping-pong handoff. Thus, we can build a one-dimensional mobility model, as shown in Figure 3. The main idea is as follows:

Figure 2 Multible MR configuration and regular linear mobility model

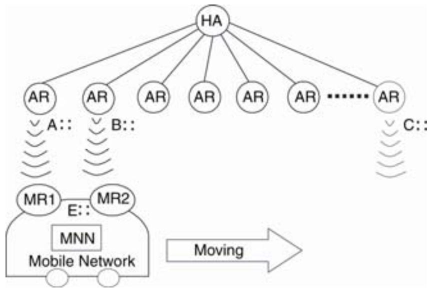
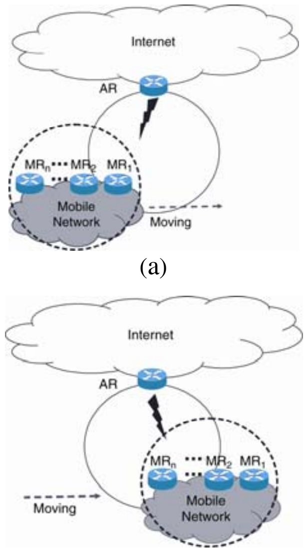


Figure 3 Location-based selection of the best-connected MR (a) before moving and (b) after moving (see online version for colours)



(a)

(b)

An MR that came into the coverage area recently is expected to have a longer connection lifetime.

Therefore, the HA selects the latest binding entry in the binding cache, and forwards packets to the MR that registered this binding (Figure 4).

Figure 4 Best-connected MR selection algorithm

```
for all binding entries of the prefix
    select the  most recent binding entry;
```

Figure 3 shows the selection of the best-connected MR among multiple MRs when a mobile network moves. At first, MR_1 goes into the coverage area of the AR and becomes the best-connected MR (Figure 3(a)). As the vehicle moves, MR_1 leaves the coverage area of this AR and MR_n becomes the best-connected MR (Figure 3(b)). Then, the HA forwards packets to MR_n, so that packet loss can be avoided during handoff.

We assume that the MRs are distributed in an advantageous manner. We control the distance between the MRs, in order to maintain the always-best-connection.

Table 1 Mobility patterns of vehicles and pedestrians

	Pedestrian	*Car*	*Bus*	*Train or subway*	*Aircraft*
Movement limitation	Very low	Medium	Medium	High	High
Road topology density	High	Medium	Medium	Low	Very low
Path predictability (predefined)	Random	Random	Yes	Yes	Yes
Mobility disposition	Random stop-and-start	Random stop-and-start	Regular stop-and-start	Regular stop-and-start	Non-stop
Space regularity	Irregular	Irregular	Regular	Strictly regular	Strictly regular
Time regularity	Irregular	Irregular	Regular	Strictly regular	Regular
Regularity variance	High	High	Medium (Traffic jam)	Low	Low
Network scale (size)	Small	Medium	Medium	Large	Large

For example, if we locate one MR at the right end of the vehicle and the other MR at the left end, the mobile network connection lifetime increases. If we locate the MRs close together, the mobile network connection lifetime decreases.

5.2 Load sharing between heterogeneous access networks

The second selection scheme is designed for random mobility. In the random mobile environment, it is not certain, which access technology will be available. So, the main idea of this scheme is to optimise the use of the available resources dynamically.

5.2.1 Case scenario: multihoming configuration for sensor/access networks in cars

The second case involves cars, which move more randomly than public trains. As described in Table 1, cars move randomly and irregularly. Therefore, in this case, we can make use of the advantages associated with heterogeneous access networks. Since wireless LANs provide high data rates and cellular networks provide wide accessibility, we can set preferences according to these characteristics.

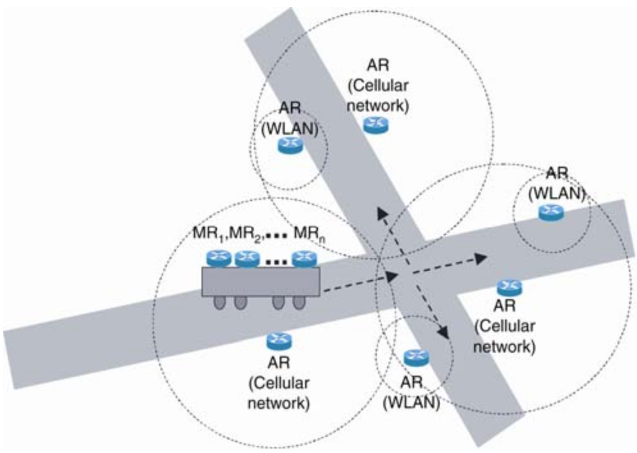
In a car, various components of the car and hand-held devices can be organised into a mobile network. This mobile network may have multiple MRs. For example, one MR can be embedded with a wireless access point, which provides the MNNs with AR functionality. Another MR can be composed of a hand-held device that might have dual interfaces providing support for wireless LAN and cellular networks. The wireless LAN interface would be its ingress interface and the cellular interface would be its egress interface. This kind of configuration will become more feasible when soft interfaces become commercially available.

5.2.2 Basic idea

Terrestrial random mobility can be modelled with the following parameters: the movement *velocity*, *distance* and *direction within a two-dimensional area*. This mobility scenario comes from the case of cars in Table 1. Since their movement is random and the road topology density is not so high, it can be modelled by random global mobility model. For this mobility model, the *most-beneficial MR selection* scheme is proposed.

When a mobile network moves within a global wide area, it can take advantage of the heterogeneous wireless access technologies in order to benefit from different coverage, data rate, price and so on (Figure 5). For example, cellular networks provide wider coverage than wireless LANs, but they do not support as high a data rate as wireless LANs. Table 2 gives the different characteristics of the different access technologies. So, we can provide both a high data rate and high connectivity by supporting heterogeneous access technologies, each with multiple MRs.

Figure 5 Multiple MR configuration and random global mobility scenario (see online version for colours)



5.2.3 Extension to existing message format

The most-beneficial MR selection scheme is based on priority. The BU message is extended to include a priority option, *P*, as shown in Figure 6. In our scheme, the priority bit is set based on the following rules:

- first priority: connectivity
- second priority: data rate.

Figure 6 NEMO binding update message format extension for load sharing

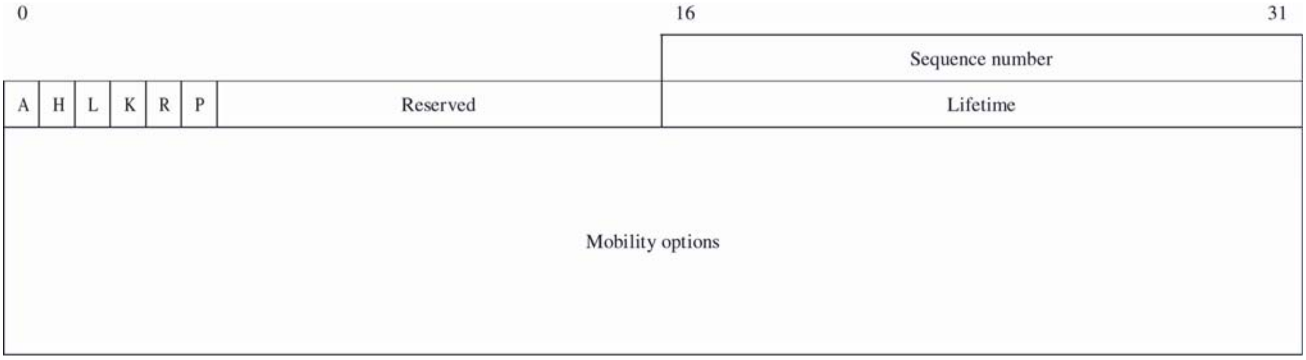


Table 2 Comparison of wireless access technologies

Wireless technologies	*Coverage size*	*Data rate*	*Delay*	*Cost*
Wireless LANs	30 ~ 300 m	11 ~ 54 Mbps	Low	Low
Cellular networks	Around 1 km	30Kbps ~ 2.4 Mbps	Low	Medium
Satellite networks	8000 km ~ almost 1/3 of the Earth's surface	10Kbps ~ 200 Mbps	High	High

If only one MR is connectable, the most-beneficial MR selection scheme chooses this MR, whereas if more than one MR can be connected, the MR that provides the highest data rate is selected (Figure 7).

Table 3 describes each field in the extended BU message. The HA determines the priority of the binding entries having the same MNP, according to the value of the data rate indicated by the priority bit.

Figure 7 Most beneficial MR selection algorithm

```
for all binding entries of the prefix
    select the highest value of P;
```

Table 3 Extended binding update message field description

Field	*Description*
Sequence number	Sequence number for this binding update message
A	Acknowledge
H	Home registration
L	Link-local address compatibility
K	Key management mobility capability
R	MR Flag
P (Priority)	*Priority of this binding update message in the binding cache*
Lifetime	Lifetime of this binding update message

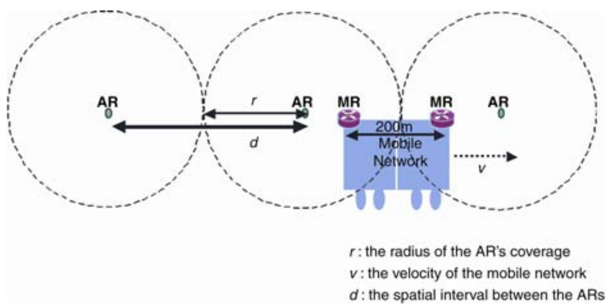
6 Simulations and performance evaluation

We used ns-2 simulator version 2.26 to evaluate the performance of the proposed architecture. To simulate NEMO enhanced schemes in ns-2, we simulated MRs by modifying the mobile nodes of MobiWan (2001) because MRs need to have both egress and ingress interfaces. MRs get their care-of addresses using RA messages received at the egress interfaces. In our simulation, we use one common prefix for the MRs and they periodically advertise their MNP on the ingress interfaces for the sake of the MNNs. When an MR receives a packet from an MNN, the MR tunnels the packet to its HA. When an MR receives a packet intended for an MNN, the MR decapsulates the packet and forwards on its ingress interface.

6.1 Evaluation for session preservation

Figure 8 shows the basic simulation topology for the linear regular mobility case. In this simulation, trains/subways move regularly with two MRs. NEMO is embedded in a 200 m long train, and wireless LAN access points are established at regular intervals. It is assumed that each access point provides AR functionality. Simulations are performed with different values of *r* (the radius of the AR's coverage), *v* (the velocity of the mobile network embedded in the vehicle) and *d* (the spatial interval between the ARs). The session preservation simulation is set up for a frequent handoff environment. The access technology used for our simulation is the wireless LAN.

Figure 8 Basic simulation topology for regular mobility (see online version for colours)



6.1.1 Spatial allocation of multiple MRs

At first, the influence of spatial allocation is investigated.

- *Allocation 1 (benchmark allocation)*: The two MRs are collocated, that is, placed at the same location in the train.
- *Allocation 2 (proposed allocation)*: The two MRs are spatially distributed, that is, placed at each end of the train, thus the distance between the two MRs is 200 m.

The objective of this comparison is to see if the spatial allocation of the MRs affects the session preservation or not. In this simulation, the velocity of the train is 90 km/hr and the AR interval is 500 m. The coverage area of each AR is 125 m in radius. A simple round-robin scheme is used for benchmarking.

Figures 9 and 10 compare the influence of spatial allocation with the benchmark scheme and the best-connected MR selection scheme, respectively. The throughput is evaluated in terms of the number of received packets at the MNNs. The whole throughput can be presented as the area that is outlined by the graph.

Figure 9 Comparison of multiple MR spatial allocation with simple round-robin (see online version for colours)

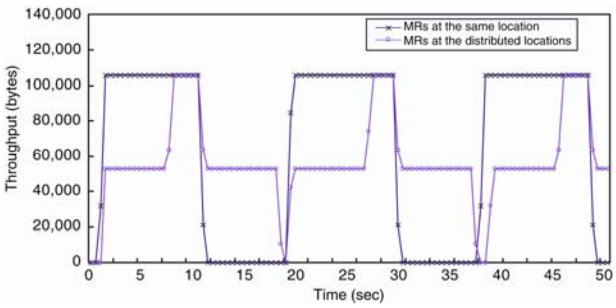
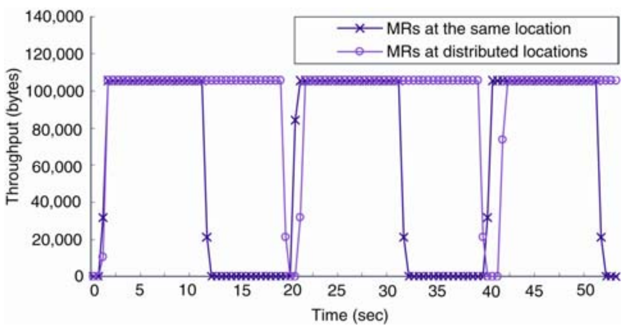


Figure 10 Comparison of multiple MR spatial allocation with the best-connected MR selection (see online version for colours)



In Figure 9, round-robin selection was performed with collocated MRs (allocation case 1) and distributed MRs (allocation case 2). It can be seen that the spatially distributed MRs provide more continuous connectivity. However, the total throughputs of the two spatial allocations do not show a big difference.

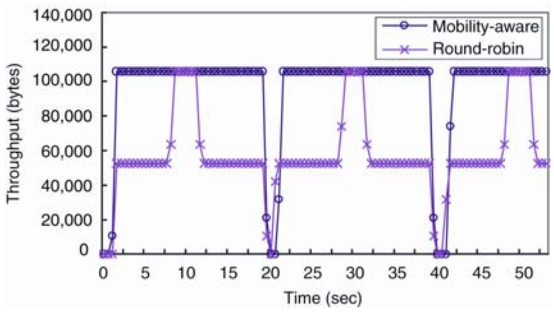
In the case of the best-connected MR selection (Figure 10), the distribution of the MRs improves the throughput. The total throughput of the spatially distributed MR case is 1,01,90,400 bytes, while that of the collocated MR case 68,64,000 bytes. So, the use of spatial distribution provides a 48.46% increase in throughput when the best-connected MR selection scheme is used.

The best-connected MR selection scheme improves the throughput by spatially allocating the MRs, since this scheme is based on the location of the MRs. This effect of spatial allocation increases as the length of the mobile network (i.e. the distance between the MRs) increases and the size of the coverage area decreases.

6.1.2 Mobility-awareness

The purpose of the second simulation is to examine the performance improvement afforded by the mobility-awareness of the best-connected MR selection scheme. We examined the result with the same simulation parameters as those used in Figures 9 and 10. Figure 11 compares the performance of the round-robin selection and the best-connected selection. Both selections are performed with spatially distributed MRs. In Figure 11, the total throughput of the round-robin selection is 61,88,160 bytes, while that of the best-connected selection is 1,01,90,400 bytes. So, the best-connected selection affords a 64.68% increase in throughput.

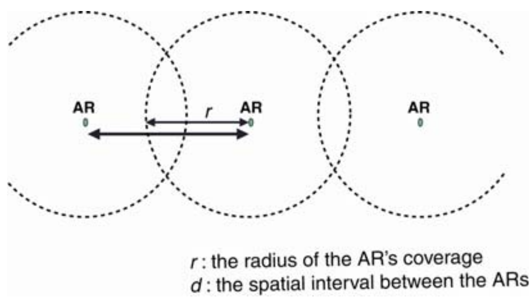
Figure 11 Performance of mobility-aware selection scheme (see online version for colours)



6.1.3 Session preservation in continuously accessible area

For sessions to be preserved, the ARs should be located in such a way that there is no inaccessible area. Even though the ARs are continuously allocated, as shown in Figure 12, mobility causes packet loss. Figure 13 shows that the benchmark schemes (the round-robin selection and the random selection) suffer from packet loss, while the best-connected MR selection scheme preserves all of the packets. In Figure 14, it is shown that the packet loss increases as the velocity increases.

Figure 12 Spatial allocation of ARs and continuous coverage ($d < 2r$)



The packet loss that occurs with the two benchmark schemes in this continuously accessible environment is caused by unawareness of mobility. If the MR moves to another AR's access area while the packets are travelling between the HA and the MR, they are lost. So, the mobility-awareness that the best-connected MR selection scheme provides is beneficial.

Figure 13 Session preservation in a continuously accessible area (velocity of the mobile network = 36 km/hr) (see online version for colours)

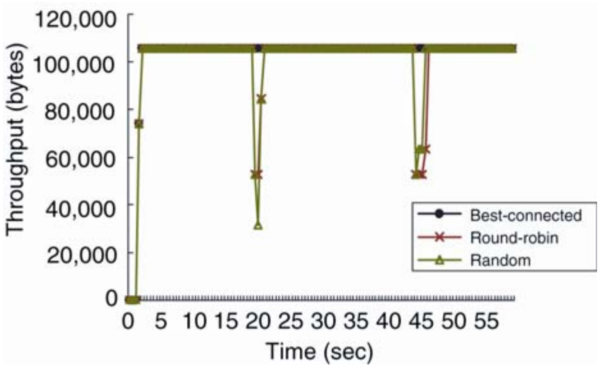
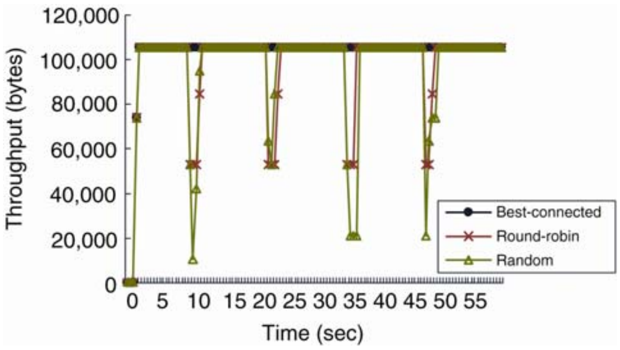


Figure 14 Session preservation in a continuously accessible area (velocity of the mobile network = 72 km/hr (see online version for colours)



6.1.4 Session preservation with intermittently accessible area

The best-connected MR selection scheme provides continuous accessibility even in the case where there are inaccessible areas. So, the use of the best-connected MR selection scheme reduces the number of ARs required.

The simulation configuration for session preservation is set up so as to create a frequent handoff environment (Figure 15). The mobile speed is set to 25 m/sec. The radius of the AR coverage is 175 m and the interval between the ARs is set to 450 m. So there are some areas where internet access is impossible for certain MRs. However, in these areas, at least one of the MRs of the mobile network can access the internet.

In Figure 16, the result of the comparison between the best-connected MR selection scheme and various benchmark schemes is shown. As shown in Figure 16, the best-connected MR selection scheme preserves the session and reduces the amount of packet loss during handoff. In comparison with this scheme, the round-robin and random selections experience more packet loss during handoff. For the entire test period, which lasted 60 sec, the round-robin selection scheme forwarded packets to the failed MR for 18 sec, and forwarded packets correctly for 40 sec. While the benchmark selection schemes suffered from wrong forwarding for 30~67% of the entire test period, the best-connected selection always maintained the connection. The relative performance of the best-connected selection improves as the length of the mobile network increases and the coverage area of the ARs decreases.

Figure 15 Spatial allocation of ARs and discrete coverage ($d > 2r$)

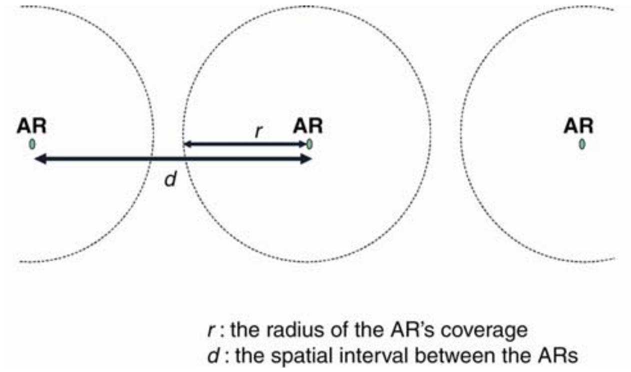
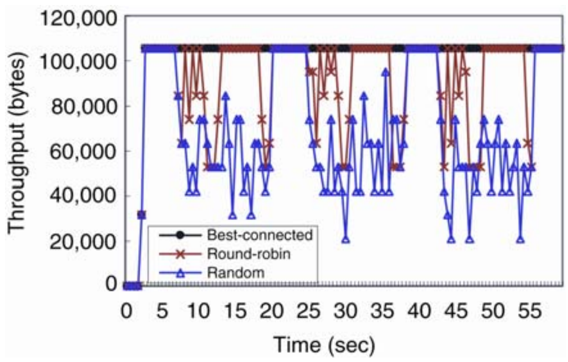


Figure 16 Session preservation with intermittently accessible area (see online version for colours)

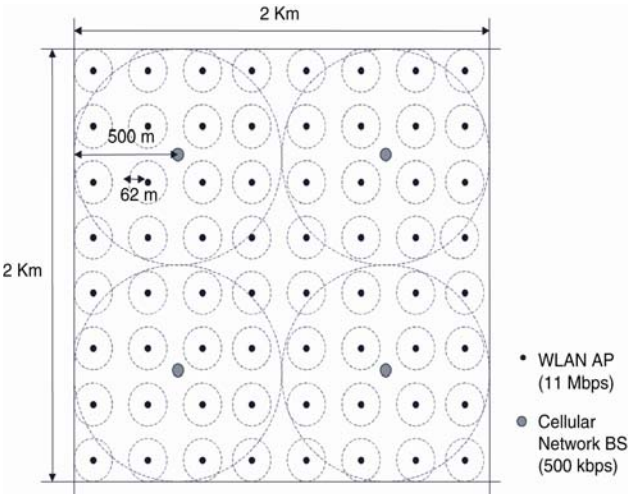


The multiple MR configurations we illustrated for the purpose of session preservation allow for the network to cope with the frail wireless links between mobile networks and the internet. The session preservation scheme results in a gain in performance at the IP layer, as well as at the TCP layer. If the traffic goes to a single MR and the load causes packet dropping and resending, it causes a TCP congestion problem. Thus, it is important to redirect packets.

6.2 Evaluation for load sharing

When a mobile network moves over a wide area across heterogeneous wireless access technologies, its movement and location are unpredictable. To simulate random global mobility, four cellular network Base Stations (BSs) and 64 wireless LAN access points are established in a 2 × 2 km square area, as shown in Figure 17. Each BS and access point is assumed to provide AR functionality. The radius of the access coverage area of the cellular networks is set to 500 m, while that of the wireless LAN to 62 m. So, the cellular networks can be accessed in almost all the areas, whereas the wireless LAN is not accessible in some areas. The data rates of the cellular network and wireless LAN are 500 Kbps and 11 Mbps, respectively. The ingress interfaces of both MRs provide the MNNs with wireless LAN access.

Figure 17 Simulation topology for random mobility



Random mobility is simulated with four parameters: the direction of the *X*-axis, the direction of the *Y*-axis, the distance of movement along the *X*-axis, and the distance of movement along the *Y*-axis. Every 1 sec, the values for the four parameters are randomly selected. The mobile network moves 0~25 m along the *X*-axis and/or *Y*-axis. To simulate FTP traffic, 640 Kbyte-long packets are sent at a Constant Bit Rate (CBR) at 0.05 sec intervals.

6.2.1 *Different capacity of access technologies*

In this section, the effect of adaptive resource selection is shown over heterogeneous networks. Simulations are performed with three different topologies: only wireless LANs, only cellular networks and heterogeneous networks.

Figure 18 shows the throughput of the mobile network that has two MRs, each with a IEEE 802.11b wireless LAN interface, and Figure 19 shows that of a mobile network that has two MRs, each with a cellular network interface. The throughput is defined in terms of the number of received packets. It is shown that the wireless LANs provide a high data rate but low connectivity, while the cellular networks support high connectivity but a low data rate.

Figure 20 shows the performance of the priority-based most-beneficial MR selection scheme. In this simulation, the mobile network has two MRs, one with a wireless LAN interface and the other with a cellular network interface. The result shows that heterogeneous access provides continuous connectivity as well as a high data rate.

Figure 18 Two MRs both with wireless LAN interfaces (see online version for colours)

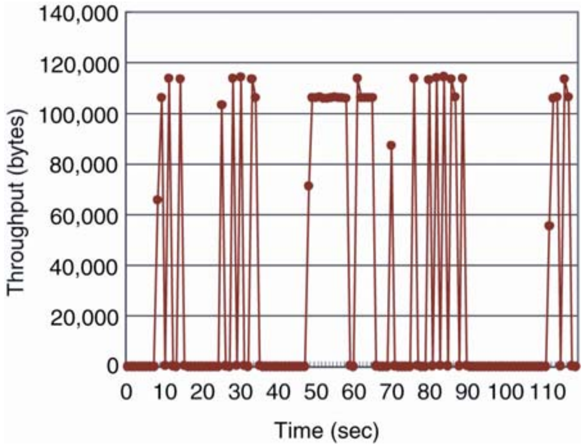


Figure 19 Two MRs both with cellular network interfaces (see online version for colours)

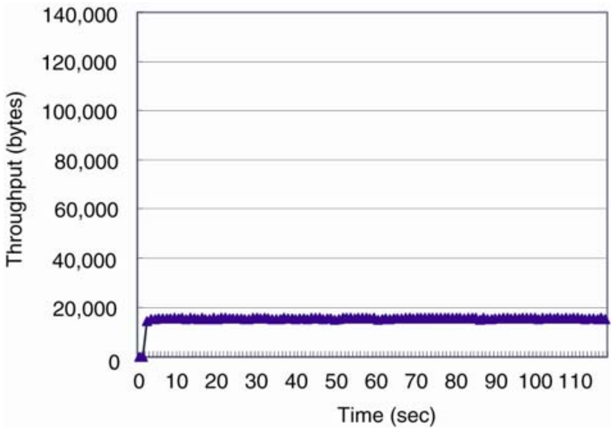
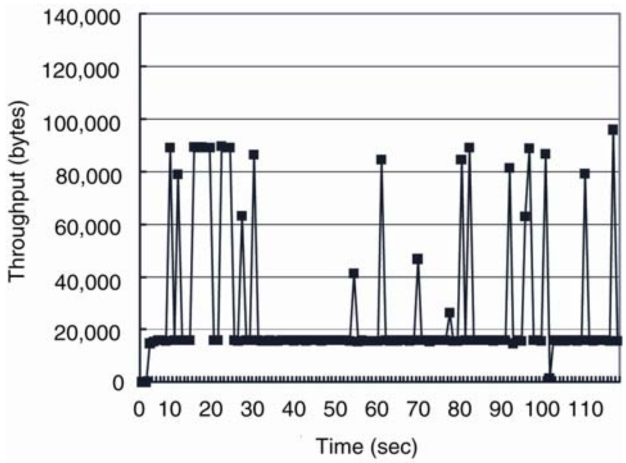


Figure 20 Two MRs each with a different access technology



6.2.2 *Heterogeneous access with multiple MRs*

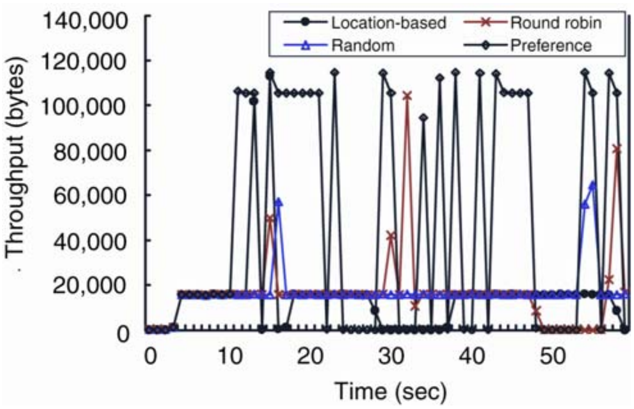
To obtain the benefits of heterogeneous technology access, a mobile network with two MRs is examined: one MR for wireless LAN access and the other for cellular network access. Proposed most-beneficial selection scheme selects a bi-directional tunnel between the HA and the MR, based on the data rate information that the MRs send to the HA.

As shown in Figure 21, the most-beneficial MR selection scheme provides a high data rate when wireless

LAN access is possible, and maintains the connectivity when wireless LAN access is impossible. The total throughputs were 8,56,200 bytes, 9,73,203 bytes, 10,21,148 bytes and 29,27,642 bytes for the best-connected MR selection, round-robin selection, random selection, and the most-beneficial MR selection schemes, respectively. So the most-beneficial MR selection scheme provides about 2 times more throughput than the other selection schemes in this simulation.

Both MRs' ingress interfaces provide the same wireless LAN access. Therefore, the MNNs do not suffer from the overhead associated with heterogeneous access network handoff delay.

Figure 21 High-data-rate MR selection for random mobility (see online version for colours)



7 Conclusion and future work

In this paper, we illustrated the protocol issues, and proposed the use of multiple MRs. The use of multiple MRs allows for a more powerful architecture by offering multiple connections to the internet. Among the many possible multiple MR configurations, realistic configurations were focused, based on case studies involving configurations for in-vehicle networks, since vehicles are coming to represent an important platform for mobile communications. Then, the mobility characteristics of vehicles, namely their movement limitation, and regularity are defined.

Based on these mobility characteristics, two mobility-aware multihoming schemes were proposed for different mobility environments and hardware configurations. Firstly, the MR selection scheme was proposed for session preservation, which takes the regular moving pattern of mobile networks into consideration. This scheme is based on the rationale that the most recently updated binding is likely to have the most stable wireless connection to the AR. By simulation, it was demonstrated that this MR selection scheme shows continuous connectivity, not only when the accessible area is continuously located, but also when it is intermittently located. Thus, this scheme also reduces the number of ARs required. While the benchmark schemes suffer from fast handoff, the proposed MR selection scheme shows stable connectivity with high velocity.

Secondly, the random and global mobility behaviour of cars in a heterogeneous access environment was investigated. In this case, the load sharing MR selection scheme was proposed to obtain the dual benefits of the high data rate of wireless LAN and the connectivity of cellular networks. The secondly proposed MR selection scheme reassigns the most beneficial MR, for example, an MR that provides high data rate, dynamically. This scheme can be useful with caching. Mobile networks can download and cache large volume of data when it is able to access a wireless LAN, while maintaining a continuous connection to cellular networks for signalling or VoIP.

The simulation results with different configurations show that the session preservation and load sharing schemes are influenced by the application mobility behaviour and different characteristics of wireless access technologies. Our best-connected MR selection scheme shows good performance with trains or subways whose mobility is regular. On the other hand, the most-beneficial MR selection scheme made profits from different access technologies.

As NEMO is based on IPv6 and provides for these new demands, we are developing a Linux-based testbed for the purpose of evaluating NEMO Basic Support with nested and multihoming configurations. As a future work, we will experiment these enhanced schemes on this testbed. The aim is to show how the multihoming may enhance the overall connectivity of hosts and mobile networks. We are investigating the application of NEMO to large public vehicles, small private vehicles and embedded PANs to demonstrate the deployment of IPv6 mobility.

References

Cao, Y. and Li, V.O.K. (2001) 'Scheduling algorithms in broad-band wireless networks', *Proceedings of the IEEE*, Vol. 89, No. 1.

Deng, H., Huang, X., Zhang, K., Niu, Z. and Ojima, M. (2003) 'A hybrid load balance mechanism for distributed home agents in mobile IPv6', *Proceedings of PIMRC*.

Devarapalli, V., Wakikawa, R., Petrescu, A. and Thubert, P. (2005) 'Network Mobility (NEMO) Basic Support Protocol', Request for Comments (RFC) 3963, Internet Engineering Task Force (IETF), January.

Draves, R. and Hinden, R. (2005) 'Default router preferences and more-specific routes', Request for Comments (RFC) 4191, Internet Engineering Task Force (IETF), November.

Ernst, T. and Lach, H. (2007) 'Network mobility support terminology', Request for Comments (RFC) 4885, Internet Engineering Task Force (IETF), July.

Ernst, T., Mitsuya, K. and Uehara, K. (2003) 'Network mobility from the internetCAR perspective', *Journal of Interconnection Networks*, Vol. 4, No. 3, pp.329–343.

Ernst, T. and Uehara, K. (2002) 'Connecting automobiles to the internet', *International Workshop on ITS Telecommunications (ITST)*, pp.257–262.

Ernst, T. (2001) MobiWan: NS-2 extensions to study mobility in Wide-Area IPv6 Networks, http://www.inrialpes.fr/planete/pub/mobiwan/.Network Simulator – ns-2 Home page, http://www.isi.edu /nsnam /ns/.

Johnson, D., Perkins, C. and Arkko, J. (2004) 'Mobility support in IPv6', Request for Comments (RFC) 3775, Internet Engineering Task Force (IETF), June.

Kellerer, W., Bettstetter, C., Schwingenschlögl, C., Sties, P., Steinberg, K-E. and Vögel, H-J. (2001) '(Auto) mobile communication in a heterogeneous and converged world', *IEEE Personal Communications*, pp.41–47.

Ng, C., Ernst, T., Paik, E.K. and Bagnulo, M. (2007) 'Analysis of multihoming in network mobility support', Request for Comments (RFC) 4980, Internet Engineering Task Force (IETF), October.

Vasilache, A., Li, J. and Kameda, H. (2003) 'Threshold-based load balancing for multiple home agents in mobile IP networks', *Telecommunications Systems*, Baltzer, Vol. 22, Nos. 1–4, pp.11–31.

Wakikawa, R., Uehara, K. and Murai, J. (2002) 'Multiple network interfaces support by policy-based routing on mobile IPv6', *Proceedings of the International Conference on Wireless Networks (ICWN)*, July.