Mobility management for highly mobile users and vehicular networks in heterogeneous environments

Karl Andersson, Christer Åhlund
Division of Mobile Networking and Computing
Luleå University of Technology
Skellefteå, Sweden
{karl.andersson,christer.ahlund}@ltu.se

Balkrishna Sharma Gukhool, Soumaya Cherkaoui
Electrical and computer engineering department
University of Sherbrooke
Canada
{balkrishna.sharma.gukhool,s.cherkaoui}@usherbrooke.ca

Abstract—**With the recent developments in wireless networks, different radio access technologies are used in different places depending on capacity in terms of throughput, cell size, scalability etc. In this context, mobile users, and in particular highly mobile users and vehicular networks, will see an increasing number and variety of wireless access points enabling Internet connectivity. Such a heterogeneous networking environment needs, however, an efficient mobility management scheme offering the best connection continuously. In this paper, a mobility management architecture focusing on efficient network selection and timely handling of vertical and horizontal hand-overs is proposed. The solution is based on Mobile IP where hand-over decisions are taken based upon calculations of a metric combining delay and delay jitter. For efficiency reasons, the frequency of binding updates is dynamically controlled, depending on speed and variations in the metric. The dynamic frequency of binding updates helps the timely discovery of congested access points and cell edges so as to allow efficient hand-overs that minimize packet drops and hand-over delays.**

Results show that the overall signaling cost is decreased and changes in networking conditions are detected earlier compared to standard Mobile IP.

Keywords-Mobility management; heterogeneous networking environments; highly mobile users; vehicular networks

I. INTRODUCTION

Mobility support for highly mobile users and vehicular networks is an emerging research area. Besides intra-vehicular and vehicle-to-vehicle communication vehicle-to-infrastructure communication is an important subarea being investigated in current research.

Another important development field is the global deployment of new radio access networking technologies including both dedicated technologies for vehicular communication and more general solutions for wireless communications. The IEEE 802.11p-based Dedicated Short Range Communication, DSRC, [1] is an example of the former type. General solutions include other IEEE standards like the different WLAN standards 802.11a, 802.11b, and 802.11g [2], the mobile WiMAX 802.16e [3] standard, as well as the IEEE 802.20 Mobile Broadband for Wireless Access [4] standard. Besides those IEEE standards, the telecommunications industry is improving the 2.5G and 3G standards offering higher capacities to their customers.

Since many laptops, PDAs, and mobile telephones today are equipped with multiple radio access interfaces the vision of "Always Best Connected" was formulated [5]. By letting the user define its own preferences on e.g. cost, battery consumption, and network capacity the overall idea is to continuously take decisions, based on the user's preferences and varying network conditions, on what radio access network to use and when to perform a possible vertical hand-over. Heterogeneous networking research is focusing on this particular issue.

A heterogeneous networking environment is composed of different access technologies, often both wired and wireless. Vertical hand-over is the process of switching the current point of attachment from one access technology to another. The overall network structure is often organized as an IP overlay network with mobility management functions including support for location management and connection migration.

A number of mobility management schemes for heterogeneous networking environments have been proposed. Some proposals have a focus on network-controlled hand-over schemes. Other proposals let the end-user's terminal take decisions on hand-overs. Interesting combinations also exist.

The design of the mobility management functionality must also define its placement in the layered networking stack. Solutions for mobility management at the network layer, the transport layer, and the application layer have been proposed and standardized. There are also solutions introducing new layers in the layered networking stack. Finally, cross-layer designed solutions also exist.

Currently, the IEEE is working on a standard for media-independent hand-over services under the name of 802.21 [6]. The main idea behind this work is first to introduce an access network neutral function taking care of data-link layer triggers letting upper layer use this information timely. Second, a command service is also letting higher layers send various instructions to the data-link layer. Third, an information service holding information on available networks and other important information pieces will be part of the 802.21 standard.

The rest of the paper is organized as follows: Section II describes the proposed architecture. In section III the evaluation framework is presented. Results from simulations and real-world evaluations are presented in section IV, while section V gives a recommendation based on the results. Section VI discusses the results, section VII contains a survey of related work, while finally, section VIII indicates future work.

II. Mobility Management for Highly Mobile Users and Vehicular Networks

When designing a mobility management scheme for highly mobile users and vehicular networks some considerations must be undertaken.

First, the point of hand-over control must be decided. Our proposal is built on an end-to-end concept where the user is connected to an arbitrarily number of access networks at each time. Those access networks can both be based on various technologies and be located in different administrative domains (i.e. operated by various service providers). Thus, we have taken a user-centric approach letting the end-user's terminal undertake decisions on hand-over timing and network selection.

Second, the placement in the layered networking stack must be decided. Handling mobility management at the network layer has the advantage of not requiring any change to existing applications. Maintaining a fixed IP address while moving and switching from one subnet to another lets applications work as if they were executed in a stationary environment. Mobility management handling at the transport layer or the application layer is beneficial in various scenarios, but needs changes to existing applications. Thus, we have taken a network layer approach leading us into using Mobile IP (MIP) which is the most common solution for network layer mobility management. MIP is defined both for IPv4 [7] and IPv6 [8].

MIP makes use of a mobility agent located in the home network, a home agent (HA), and, in MIP for IPv4, a mobility agent in the visited network, a foreign agent (FA). The HA is a specialised router responsible for forwarding packets to the mobile node (MN). It is assigned a home address (HoA) in the same subnet as the HA. The FA is responsible for assigning a care of address (CoA) for the MN and forwarding packets for the MN. The HA holds a binding cache with mappings of HoAs to CoAs. The MN can also use a co-located address CoA. In that case, the MN acquires an IP address using regular mechanisms like DHCP and is not dependent on the existence of an FA in the visited network.

Packets are transported from the originating host, the correspondent node (CN), to the HA and then tunnelled through an IP tunnel using IP in IP encapsulation to the MN (possibly via the FA). The MN continually sends binding update (BU) messages to the HA indicating its CoA. The binding cache is updated accordingly and the HA returns binding acknowledgment (BA) messages to the MN. Packets in the direction from the MN to the CN can be sent directly to the CN. In MIPv6 route optimization techniques also exist enabling the CN to send packets directly to the MN. Thus, all packets do not need to travel through the HA.

Although MIP has some drawbacks with handover latencies, introduction of tunnelling overhead, and dependency of mobility agents it is a well known and investigated mobility management scheme gaining interest from both the computer communications and telecommunications industries.

Multihoming is a way of letting a mobile node use more than one point of attachment simultaneously. Using such a mechanism is beneficial from a load balancing viewpoint. Also, multihoming is beneficial for performing smooth vertical hand-overs. There is not yet a standardized version of multihomed MIP, but proposals exist in the literature [9]. By using multihomed MIP with BU messages sent in parallel over more than one access network, the MN indicates to the HA what access network to be used at each time. Along with the existing extensions to MIPv6 of fast hand-overs (FMIPv6) [10] and hierarchical MIP (HMIPv6) [11] most problems with hand-over latencies and packet losses are solved.

Also, the usage of the BU and BA messages to measure delay and interarrival jitter has been proposed [12]. The relative network load, RNL, was defined as the sum of delay and the variance of arrival times of BA messages in the MN. This makes it possible to perform continuously comparisons between access networks and to decide on hand-over timing and target access network.

According to the MIPv4 standard BU messages should be handled by the HA at least once every second. The MIPv6 standard, on the other hand, allows the MN to send BU messages every 50 millisecond.

Using a high frequency for BU/BA messages is good when dealing with highly mobile users and vehicular networks. Some access networks (e.g. 802.11 networks) have sharp cell edges making frequent decisions on possible hand-overs necessary. On the other hand, using a high frequency is increasing the overhead and consuming bandwidth. Possibly, it also increases the monetary cost if the user is charged per amount of downloaded data.

Therefore, we propose a dynamic frequency for sending BU messages following the speed of the user. The faster travelling, the higher BU message frequency is the basic idea. The proposal assumes access to current speed in some way, e.g. through the speedometer in a car or through a GPS system.

A. A policy-based decision model for access network selection

To estimate a particular network's current capacity, the relative network load (RNL) metric is calculated in the MN for all available access networks. The RNL metric represents a quality value for each access network and is based on round trip time, RTT, and RTT jitter values. Such parameters are access technology independent and good indicators on congestion in networks and limitations in bandwidth.

RTT jitter, being the variation in RTT and the mean deviation of the difference in arrival time of two consecutive BA messages compared to sending time of two consecutive BU messages (being equivalent to the variation in transit time of two BU-BA message pairs), is calculated using formulas in RFC 3550 [13]. Instead of using a fixed value of 16 as the

history window we adjust the window dynamically not to use too much history when taking hand-over decisions. This gives the following formulae:

$$RNL_n = \bar{z}_n + c \cdot J_n \qquad (1)$$

$$\bar{z}_n = \frac{1}{h} RTT_n + \frac{h-1}{h} \bar{z}_{n-1} \qquad (2)$$

$$RTT_n = R_n - S_n \qquad (3)$$

$$D_n = R_n - R_{n-1} - (S_n - S_{n-1}) = (R_n - S_n) - (R_{n-1} - S_{n-1}) =$$

$$= RTT_n - RTT_{n-1} \qquad (4)$$

$$J_n = \frac{1}{h} |D_n| + \frac{h-1}{h} J_{n-1} \qquad (5)$$

where S_i and R_i are defined as

S_i = the time of sending BA message i

R_i = the time of arrival of BU message i

h determines the history window for the weighted average calculations. For example, when $h = 5$, the most recent value will contribute to the calculated $\bar{z}_n$ and J_n values with 20%. This way, ping-pong effects could be avoided if the value of h is carefully selected.

c determines the weight of the RTT in comparison to the RTT jitter value. For example, when $c = 5$, the RTT jitter value is contributing five times more to the metric value than the RTT value does.

The variables $\bar{z}$, D, and J are initialized with the following values:

$$\bar{z}_0 = RTT_0$$

$$D_0 = 0$$

$$J_0 = D_1$$

The RNL metric is beneficial to use for its access network independence feature. The fact that no synchronized clocks are needed is also favouring this solution.

Also, the proposed architecture is implementing the IETF's proposed model for policy handling [15] depicted in figure 1. The policy repository (PR), the policy decision point (PDP), and the policy enforcement point (PEP) are in the proposed architecture all located in the MN.
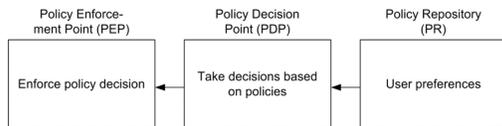

Figure 1.   IETF's model for handling of policies

The overall aim of the policy model is to choose the best available access network taking the end-user's preferences on various variables into account. Therefore, a cost function has been defined using three parameters: monetary cost, power consumption, and network load. The policy function, proposed in [16], is defined as

$$U_j = w_P \ln P_j + w_C \ln C_j + w_L \ln L_j \qquad (6)$$

where

$$w_P + w_C + w_L = 1$$

U_j is the policy value for access network j and is the weighted sum of normalized policy parameters. P_j represents power consumption while C_j is the monetary cost for access network j respectively. Those parameters are dimensionless constants. L_j is the RNL metric for access network j and calculated according to the formulas above. The network selection decision is taken whether to make a handover or not, and in the case of such a decision, to which new access network to transfer the connection to.

Before any handover decision is finally taken, the direction of the vertical handover is determined. Nasser et al. [14] defined an upward vertical handover as roaming to an access network with a larger cell size and lower bandwidth, and a downward vertical handover as roaming to an access network with a smaller cell size and larger bandwidth. Using those definitions, we decide a handover to take place when

$$U_{new} < U_{old} - a$$

for downward handovers where a is a positive hysteresis constant used in order to avoid ping-pong effects between interfaces. We decide upward handovers to take place when

$$U_{new} < U_{old}$$

This asymmetric decision model is used in order to let handovers to access networks with high but unstable capacities wait until the policy value is significantly lower compared to the old access network with lower capacity but better coverage. When doing handovers to access networks with less capacity but better coverage, the handover decision should be executed immediately in order not to loose the connection. This is especially important at high speeds and steep cell edges.

B.   Calculation of the registration message frequency

As indicated above, we want to adjust the frequency of sending BU messages according to the speed of the user. Therefore, we propose a frequency of the BU/BA messages according to the formula

$$f = \begin{cases} f_{min} & \text{for } v \leq v_{low} \\ k \cdot v & \text{for } v_{low} \leq v \leq v_{high} \\ f_{max} & \text{for } v \geq v_{high} \end{cases} \qquad (7)$$

so that

$$v_{low} = f_{min}/k$$

and

$$v_{high} = f_{max}/k$$

and where f is the frequency of BU messages (number of messages per second), k is a suitable constant and v is the velocity of the user (measured in m/s). When increasing the speed, the frequency of BU messages should also be increased accordingly. In other words, there needs to be a new RNL value calculated when a certain distance has been covered in order to catch a cell edge timely enough.

III. EVALUATION FRAMEWORK

In order to evaluate the proposed mobility management scheme with respect to responsiveness to changes in network conditions (relative loads, received signal strength, etc.) at different speeds, a scenario with a user traveling 10 kilometers through an area with five IEEE 802.11 access points (APs) placed 2.5 km apart with no overlapping coverage areas was set up using the OPNET Modeler 14.0 PL3 simulation tool [17]. A WiMAX base station was placed in the middle of the area covering the entire area.

Policy values for each access technology were computed and compared according to the scheme presented in Section 2. Each WLAN access point also used by five other stationary nodes. Figure 2 illustrates the scenario. In the evaluations performed, we allowed one BU/BA message to be lost. After two consecutive losses of BU/BA messages, a vertical hand-over was forced immediately. RNL was calculated using RTT and RTT jitter values measured in milliseconds.
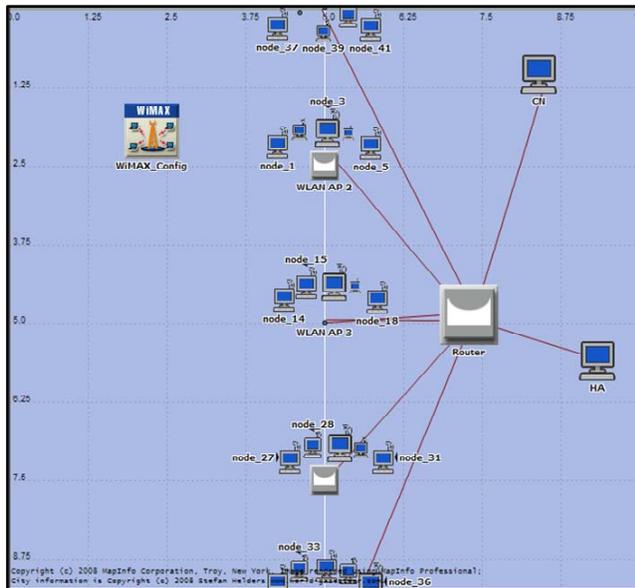

Figure 2. Evaluation topology

Both the MN and the stationary nodes were sending 64 kb/s constant bit-rate data to a common correspondent node (CN).

All such traffic was tunneled through the home agent (HA). Hence, no route optimization was used.

The speed of the mobile node, v, were selected such that

$$v \in \{10, 13.3, 20, 40\} \text{ m/s}$$

The frequencies of the BU/BA messages, f, were selected such that

$$f \in \{1, 2, 5, 10\} \text{ BU/BA messages per second}$$

Results from earlier real-world experiments have shown $h = 5$ being a good value for the history window. Balancing between not causing a hand-over to be executed too fast where RTT and RTT jitter values fluctuate on the one hand, and not delaying hand-over decisions when those values really change on the other hand can though be hard.

Also, different values of the c constant were considered. By setting $c = 1$ we gave RTT and RTT jitter values the same weights when calculating the RNL metric. In order to take variations in the metric itself into account, we increased the value of c and finally used $c = 20$ in the experiments reported in this paper. The cost and energy consumption weights (w_P and w_C) where set to zero on all interfaces in order to evaluate a policy based only on RNL.

IV. RESULTS

To quantitatively evaluate the proposed mobility management scheme, the packet loss rate was studied in detail. Figures 3-6 below summarize the results.
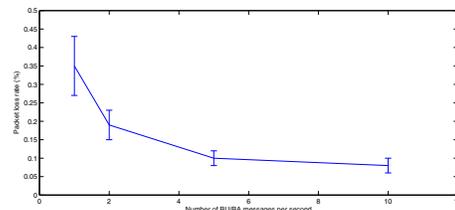

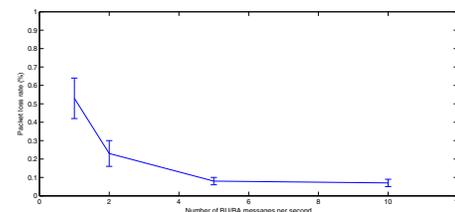Figure 3. Traveling at $v = 10$ m/s

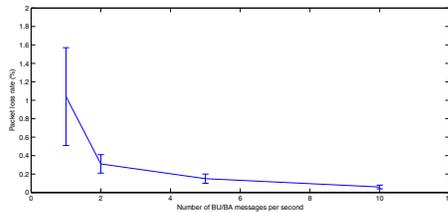Figure 4. Traveling at $v = 13.3$ m/s
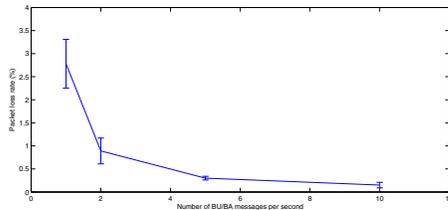
Figure 5.   Traveling at $v = 20$ m/s

Figure 6.   Traveling at $v = 40$ m/s

It is obvious that increasing the BU/BA message frequency is beneficial as the speed of the user increases. Also, the graphs show that a significant improvement wrt. packet losses is achieved for $v = 10$ m/s when increasing f to $f = 2$ BU/BA messages per second, for $v = 13.3$ m/s to when increasing f to $f = 5$ BU/BA messages per second, and finally for $v = 20$ m/s and $v = 40$ m/s when increasing f to $f = 10$ BU/BA messages per second.

Increasing the BU/BA message frequency indeed comes at an increased signaling cost. The associated signaling cost is for each level of frequency of BU/BA messages is summarized in table I below. The signaling cost has been computed using sizes of the BU and BA messages both being 24 bytes long in MIPv6. Additionally two UDP headers of 8 bytes each, two IPv6 header of 40 bytes each, and two data-link layer headers of 34 bytes (a typical value for 802.11 networks) each were added.

TABLE I.   SIGNALING COST AT VARIOUS FREQUENCIES

f (BU/BA messages per second)	Bandwidth used (kb/s)
10	8.5
5	4.2
2	1.7
1	0.8

Users of the proposed mobility management scheme need to take this overhead into account. Therefore, we propose a trade-off such that using our proposed mobility management architecture taking the user's speed into account only to be used for high-speed access networks like WLAN where bandwidth normally is not that limited and where the benefits of our architecture are highest. Access networks with larger coverage and not so steep cell edges may be handled with a fixed, low BU/BA message frequency.

To visualize the results of the experiments further, figure 7 below shows sample values for a user traveling at $v = 40$ m/s using a BU message frequencies of $f = 10$ BU/BA messages per second. Number of packets lost is also plotted as well as policy values for WLAN and WiMAX respectively.

V.   RECOMMENDATION

Taking the results presented in section 4 into account, we recommend using $k = 0.25$, $f_{min} = 1$ BU/BA message per second, and $f_{max} = 10$ BU/BA messages per second in formula (7) giving

$$f = \begin{cases} f_{min} & \text{for } v \leq 4\,\text{m/s} \\ 0.25 \cdot v & \text{for } 4 \leq v \leq 40\,\text{m/s} \\ f_{max} & \text{for } v \geq 40\,\text{m/s} \end{cases}$$

corresponding to a new policy value being calculated every four meter the user is travelling.

VI.   DISCUSSION

The above presented results show that it is possible to adjust the frequency of BU messages according to the speed of the user in order to react quickly to variances in networking conditions while keeping signaling cost down and minimizing packet losses. Also, by assigning a high weight to the RTT jitter component in the network-layer metric proposed, variations in the metric itself is taken into account.

VII.   RELATED WORK

Using Mobile IP as a solution for mobility management is commonly accepted by the research community. Sanmateau et al. [18] propose such a solution with a division of the foreign network into a foreign core network and a foreign access network. Mobile IPv4 is used for macro mobility management, while other protocols may be used for micro mobility management.

Kafle et al. [19] suggest a mobile router based hand-over scheme for long-vehicular-multihomed mobile networks. FMIPv6 is reported to be outperformed with regard to packet losses, signaling-message overhead, and packet-delivery overhead.

Gehlen et al. [20] propose an architecture oriented to the IEEE 802.21 Media Independent Hand-over reference model containing a vehicle gateway connecting to a fixed gateway using the best available communication system at each time and place. The architecture is also handling Quality of Service and security requirements.

An optimal hand-off algorithm for hybrid networks is presented by Majlesi et al. [21]. The proposed algorithm is an adaptive fuzzy logic based algorithm and is reported to decrease hand-off delay and number of unnecessary hand-offs. In particular, the Received Signal Strength averaging window is changed according to the speed of the mobile node.
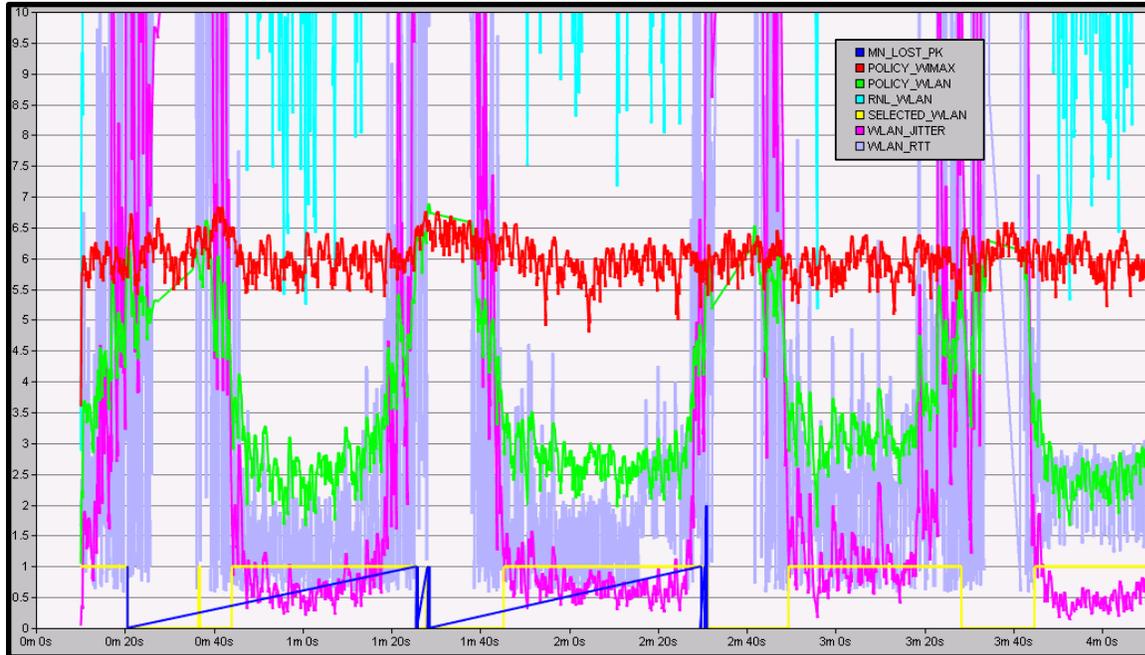
Figure 7. Sample run (v = 40 m/s, f = 10 BU/BA messages/s)
Access network decisions are indicated so that SELECTED_WLAN = 1 when WLAN is selected and
SELECTED_WLAN = 0 when WiMAX is selected. Packet losses are indicated by the MN_LOST_PK curve.

Mäkelä et al. [22] propose using a neural network hand-off algorithm. The number of training points is compared to the number and location of hand-offs.

Finally, Wang et al. [23] propose a Mobile IPv6 based seamless hand-off strategy for heterogeneous wireless networks integrating UMTS and WLAN.

To the best of our knowledge, no previous work has used a combination of a network-layer metric for hand-over decisions with dynamic frequencies for BU messages according to the speed of the mobile node.

VIII. FUTURE WORK

Future studies will include how the event, command, and information services of the emerging IEEE 802.21 standard for media-independent handovers can be integrated and used in our framework.

Also, we plan to perform detailed studies on applications of our mobility management framework for Voice over IP and multimedia services with Quality of Service (QoS) requirements.

ACKNOWLEDGMENT

REFERENCES

[1]   grouper.ieee.org/groups/scc32/dsrc

[2]   www.ieee802.org/11

[3]   www.ieee802.org/16

[4]   www.ieee802.org/20

[5]   E. Gustafsson and A. Jonsson, Always best connected, IEEE Wireless Communications, Volume 10, Issue 1, pp. 49-55, February 2003

[6]   www.ieee802.org/21

[7]   C. Perkins (ed.), IP Mobility Support for IPv4, IETF, RFC 3344, August 2002

[8]   D. Johnson, C. Perkins, and J. Arkko, IP Mobility Support in IPv6, IETF, RFC 3775, June 2004

[9]   C. Åhlund, R. Brännström, and A. Zaslavsky, M-MIP: Extended Mobile IP to Maintain Multiple Connections to Overlapping Wireless Access Networks, In Lecture Notes in Computer Science, Volume 3420/2005, pp. 204 – 213, April 2005

[10] R. Koodli (ed.), Fast Handovers for Mobile IPv6, IETF, RFC 4068, July 2005

[11] H. Soliman, C. Castelluccia, K. El Malki, and L. Bellier, Hierarchical Mobile IPv6 Mobility Management (HMIPv6), IETF, RFC 4140, August 2005

[12] C. Åhlund, R. Brännström, and A. Zaslavsky, Traffic load metrics for multihomed mobile IP and global connectivity, In Telecommunication Systems, Volume 33, Numbers 1-3, pp. 155 – 185, December 2006

[13] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, RTP: A Transport Protocol for Real-Time Applications, IETF, RFC 3550, July 2003

[14] N. Nasser, A. Hasswa, and H. Hassanein, Handoffs in fourth generation heterogeneous networks, In IEEE Communications Magazine, Volume 44, Issue 10, pp. 96 – 103, October 2006

[15] R. Yavatkar, D. Pendarakis, and R. Guerin, A Framework for Policy-based Admission Control, IETF, RFC 2753, January 2000

[16] G. Nyberg, C. Åhlund, and T. Röjmyr, SEMO: A Policy-based system for Handovers in Heterogeneous Networks, In Proceedings of The International Conference on Wireless and Mobile Communications (ICWMC 2006), Bukarest, Rumania, July 2006

[17] OPNET Modeler, www.opnet.com

[18] A. Sanmateau, L. Morand, E. Bustos, S. Tessier, F. Paint, and A. M. Sollund, Using Mobile IP for provision of seamless handoff between heterogeneous access networks, or how a network can support the always-on concept, In Proceedings of EURESCOM Summit, Heidelberg, Germany, November 2001

[19] V. P. Kafle, E. Kamioka, and S. Yamada, CoMoRoHo: Cooperative mobile router-based handover scheme for long-vehicular multihomed networks, IEICE Transactions on Communications, Volume E89-B, Number 10, pp. 2774 – 2785, October 2006

[20] G. Gehlen, E. Weiss, S. Lukas, C.-H. Rokitansky, and B. Walke, Architecture of a Vehicle Communication Gateway for Media Independent Handover, In Proceedings of the 3rd International Workshop on Intelligent Transportation (WT2006), pp. 205 – 209, Hamburg, Germany

[21] A. Majlesi, and B.H. Khalaj, An adaptive fuzzy logic based handoff algorithm for hybrid networks, In Proceedings of the 6th International Conference on Signal Processing, Volume 2, pp. 1223 – 1228, August 2002

[22] J.-P. Mäkelä, and K. Pahlavan, Performance of Neural network Handoff algorithm Under Varying Mobile Velocities, In Proceedings of the 2005 Finnish Signal Processing Symposium (FINSIG'05), Kuopio, Finland, August 2005

[23] Y.-H. Wang, C.-P. Hsu, and J.-Y. Lai, A Mobile IPv6 Based Seamless Handoff Strategy for Heterogeneous Wireless Networks, In Proceedings of the Fourth International Conference on Computer and Information Technology (CIT'04), pp. 600 – 605, September 2004

[24] www.hybrinet.org