Telecommunication Systems 30:1/2/3, 123–142, 2005

Predictive versus Reactive—Analysis of Handover Performance and Its Implications on IPv6 and Multicast Mobility

THOMAS C. SCHMIDT schmidt@fhtw-berlin.de
HAW Hamburg, Department Informatik, Berliner Tor 7, D-20099 Hamburg, Germany; FHTW Berlin, Hochschulrechenzentrum, Treskowallee 8, D-10318 Berlin, Germany

MATTHIAS WÄHLISCH mw@fhtw-berlin.de
FHTW Berlin, Hochschulrechenzentrum, Treskowallee 8, D-10318 Berlin, Germany

Abstract. Handovers in mobile packet networks commonly produce packet loss, delay and jitter, thereby significantly degrading network performance. Mobile IPv6 handover performance is strongly topology dependent and results in inferior service quality in wide area scenarios. To approach seamless mobility in IPv6 networks predictive, reactive and proxy schemes have been proposed for improvement. In this article we analyse and compare handover performance and frequencies for the corresponding protocols, as they are an immediate measure on service quality. Using analytical methods as well as stochastic simulations, we calculate the performance decreases originating from different handover schemes, the expected number of handovers as functions of mobility and proxy ratios, as well as the mean correctness of predictions. In detail we treat the more delicate case of these rates in mobile multicast communication. It is obtained that performance benefits, expected from simple analysis of predictive schemes, do not hold in practice. Reactive and predictive handovers rather admit comparable performance. Hierarchical proxy environments—foremost in regions of high mobility—can significantly reduce the processing of inter–network changes. Reliability of handover predictions is found on average at about 50%.

Keywords: mobile IPv6, FMIPv6, HMIPv6, multicast mobility, handover performance, handover frequency

1. Introduction

Mobility Support in IPv6 Networks as designed by Johnson et al. [4] has become a proposed standard within these days. Outperforming IPv4, the emerging next generation Internet infrastructure will now be ready for implementation of an *elegant*, transparent solution for offering mobile services to its users.

At first users may be expected to cautiously take advantage of the new mobility capabilities, i.e. by using Home Addresses while away from home or roaming their desktop ‘workspaces’ between local subnets. Major scenarios in future IPv6 networks, though, move towards the convergence of IP and 3GPP devices, strengthening the vision of ubiquitous computing and real-time communication. The challenge of supporting voice and videoconferencing (VoIP/VCoIP) over Mobile IP remains, as current roaming procedures are too slow to evolve seamlessly, and multicast mobility waits for a convincing design beyond MIPv6—see [6] for a detailed discussion on multicast mobility.

Synchronous real-time applications s. a. VoIP and VCoIP place restrictive demands on the quality of IP services: Packet loss, delay and delay variation (jitter) in a constant bit rate scenario need careful simultaneous control. Serverless IPv6 voice or videoconferencing applications need to rely on mobility management for nomadic users and applications (comp. [2,9]), as well as multicast support on the Internet layer. Strong efforts have been taken to improve handover operations in a mobile Internet, both in the unicast and the multicast case. Hierarchical mobility management by Soliman et al. [10], Schmidt and Wählisch [8] and fast handover operations by Koodli [5], Suh et al. [12] both lead to accelerated and mainly topology independent schemes. In addition to specific performance issues and infrastructural aspects, these concepts cause a different eagerness to operate handovers.

The occurrence of handovers is the major source for degradation of mobile network performance and places additional burdens onto the Internet infrastructure. Smoothing their operations and reducing their frequencies thus promises to ease roaming and to lower infrastructural costs. In the present work we focus on the essential performance aspects of handovers as applied in the afore mentioned Internet Drafts. Following a simple reference model, performance is measured both, in theory and in simulations. Furthermore we quantitatively evaluate handover activities with respect to user mobility and geometry conditions.

In addition, Multicast group communication raises quite distinctive aspects within a mobility aware Internet infrastructure. On the one hand, Multicast routing itself supports dynamic route configuration, as members may join and leave ongoing group communication over time. On the other hand, multicast group membership management and routing procedures are intricate and too slow to function smoothly for mobile users. Multicast itself imposes a special focus on addressing. Applications commonly identify contributing streams through source addresses, which must not change during sessions, and routing paths in most protocols are chosen from destination to source. In general the roles of multicast senders and receivers are quite distinct. While a client initiates a local multicast tree branch, the source may form the root of an entire source tree. Within this paper we will concentrate on mobile multicast receiver scenarios, as this is the less complex problem.

This paper is organised as follows. In Section 2 we briefly introduce the current proposals for improved unicast and multicast mobility. Performance aspects of the different handover approaches are evaluated in Section 3. Section 4 is dedicated to our results on handover frequency analysis, derived from analytical models as well as stochastic simulations. Conclusions and an outlook follow in Section 5.

2. Improved unicast and multicast mobility management

2.1. Hierarchical mobility and fast handovers

Two propositions for improving roaming procedures of Mobile IPv6 are essentially around: A concept for representing Home Agents in a distributed fashion by proxies has been developed within the Hierarchical Mobile IPv6 (HMIPv6) by Soliman et al. [10]. While away from home, the MN registeres with a nearby Mobility Anchor Point (MAP) and passes all its traffic through it. The vision of HMIPv6 presents MAPs as part of the regular routing infrastructure. The MN in the concept of HMIPv6 is equipped with a Regional Care-of Address (RCoA) local to the MAP in addition to its link-local address (LCoA). When corresponding to hosts on other links, the RCoA is used as MN's source address, thereby hiding local movements within a MAP domain. HMIPv6 reduces the number of ëvisibleí handover instances, but—once a MAP domain change occurs—binding update procedures need to be performed with the original HA and the CN.

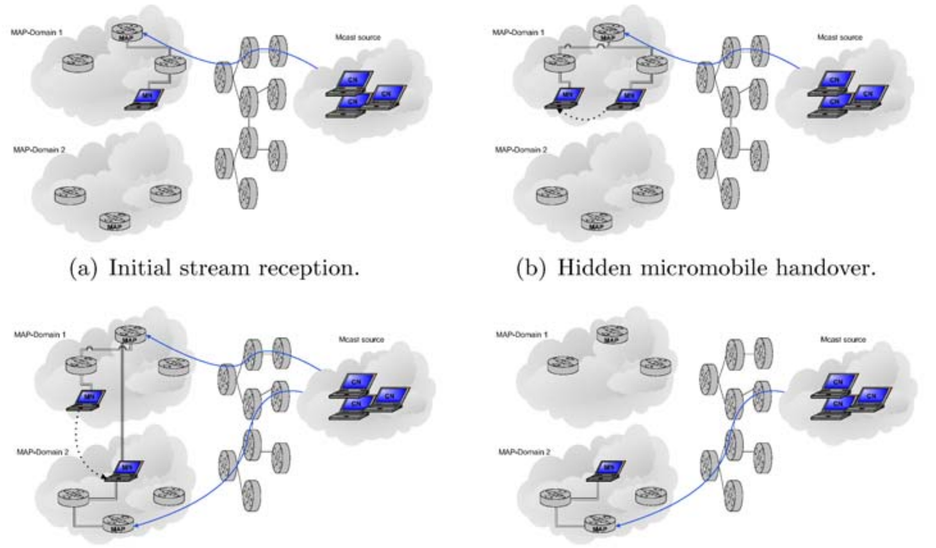
The complementary approach provides handover delay hiding and is introduced in the Fast Handover for MIPv6 scheme (FMIPv6) by Koodli [5]. FMIPv6 attempts to anticipate layer 3 handovers and to redirect traffic to the new location, the MN is about to move to. FMIPv6 aims at hiding the entire handover delay to communicating end nodes at the price of placing heavy burdens onto layer 2 intelligence. A severe functional risk arises from a conceptual uncertainty: As the exact moment of layer 2 handover generally cannot be foreseen, and even flickering may occur, a traffic redirect due to anticipation may lead to data damage largely exceeding regular MIPv6 handover bare any optimization. The significance of this uncertainty has been recently confirmed with empirical studies by Song et al. [11], where even the use of extensive statistical data under fixed geometry condition led to a prediction accuracy of only 72%.

The two multicast mobility approaches introduced below are built on top of either one of these unicast agent schemes. Minor modifications to HMIPv6 resp. FMIPv6 signaling are requested and both proposals remain neutral with respect to multicast routing protocols in use.

2.2. M-HMIPv6—Multicast mobility in a HMIPv6 environment

"Seamless Multicast Handovers in a Hierarchical Mobile IPv6 Environment (M-HMIPv6)" by Schmidt and Wählisch [8] extends the Hierarchical MIPv6 architecture to support mobile multicast receivers and sources. Mobility Anchor Points (MAPs) as in HMIPv6 act as proxy Home Agents, controlling group membership for multicast listeners and issuing traffic to the network in place of mobile senders. Figure 1 visualizes the architecture and handover operations therein.

All multicast traffic between the Mobile Node and its associated MAP is tunneled through the access network, unless MAP or MN decide to turn to a pure remote subscription mode. Handovers within a MAP domain remain invisible in this micro mobility approach. At the event of an inter-MAP handover, the previous anchor



(a) Initial stream reception. (b) Hidden micromobile handover.

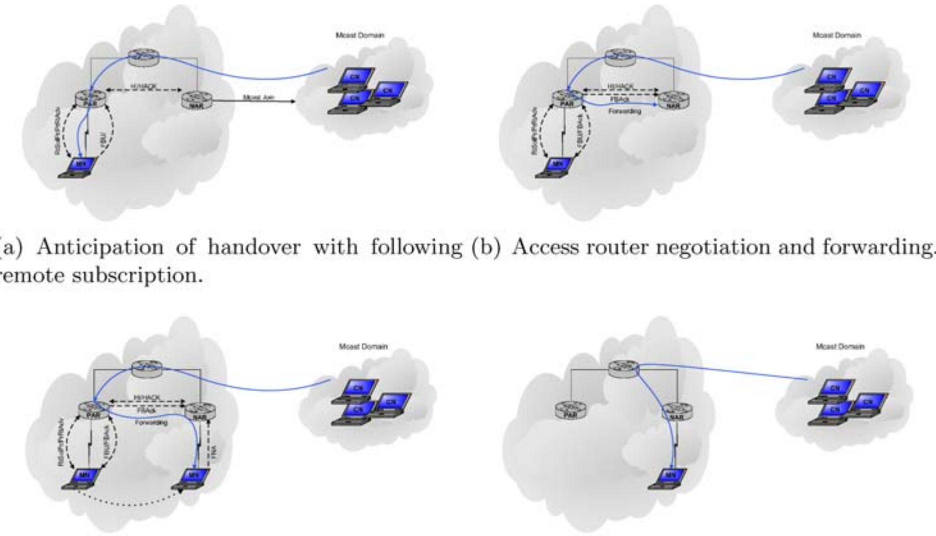
(c) Inter–MAP handover with data forwarding. (d) Termination of forwarding at finalized handover.

Figure 1. Multicast handover for M-HMIPv6 receivers.

point will be triggered by a reactive Binding Update and act as a proxy forwarder. Subsequent to MIPv6 handover continuous data reception is thus assured, while a remote subscription continues within the new MAP domain. A Home Address Destination Option, bare of Binding Cache verification at the Correspondent Node, has been added to streams from a mobile sender. Consequently transparent source addressing is provided to the socket layer. A multicast advertisement flag extends the HMIPv6 signaling.

In cases of rapid movement or crossings of multicast unaware domains, the mobile device remains with its previously associated MAP. Given the role of MAPs as Home Agent proxies, the M-HMIPv6 approach may me viewed as a smooth extension of bi-directional tunneling through the Home Agent supported in basic MIPv6.

The frequency of multicast handovers visible to the network is reduced by this scheme like it is in HMIPv6 for unicast. The relevance of hiding handovers is significantly enhanced in the multicast case, as efforts and costs of multicast routing transitions largely exceed the unicast case (see Section 4 for a detailed discussion). A further effort to decrease handover occurrences recently has been taken by Zhang et al. [13]. The authors introduce negotiation functions to proxy agents, called Dynamic Multicast Agents, to steer handovers according to topological distances. In this sense the authors extend the M-HMIPv6 approach to additionally account for individual paths of the MN—on the price of quite extensive signalling and state records to be kept at proxy agents.



(a) Anticipation of handover with following remote subscription. (b) Access router negotiation and forwarding.

(c) Handover with data forwarding. (d) Termination of forwarding at finalized handover.

Figure 2. Multicast handover for M-FMIPv6 receivers.

2.3. M-FMIPv6—multicast mobility in a FMIPv6 environment

Suh et al. [12] introduced "Fast Multicast Protocol for Mobile IPv6 in the Fast Handover Environments", which adds support for mobile multicast receivers to Fast MIPv6. On predicting a handover to a next access router (NAR), the Mobile Node submits its multicast group addresses under subscription with its Fast Binding Update (FBU) to the previous access router (PAR). PAR and NAR thereafter exchange those groups within extended HI/HACK messages. In the ideal case, NAR will be enabled to subscribe to all requested groups, even before the MN has disconnected from its previous network. To reduce packet loss during handovers, multicast streams are forwarded from PAR to NAR via unicast tunnels established by the FMIPv6 protocol, as is shown in figure 2.

Due to inevitable unreliability in handover predictions—the layer 2 may not (completely) provide prediction information and in general will be unable to foresee the exact moment of handoff—the fast handover protocol depends on fallback strategies. Fast handover negotiations, which could not be completed by a Fast Binding Acknowledge (fback), are recovered by a final reactive handover part. A complete reactive handover will be performed, if the Mobile Node was unable to submit its Fast Binding Update. The regular MIPv6 handover will take place, if the Mobile Node did not succeed in Proxy Router inquiries. From such fallback positions the mobile multicast listener has to newly subscribe to its multicast sessions, either through a HA tunnel or at its local link. By means of this fallback procedure, fast handover protocols must be recognised as discontinuous extensions of the MIPv6 basic operations.

2.4. *Relevant aspects of handover evaluation*

Both handover approaches introduced above, attempt to assure the continuous reception of real-time data streams for mobile unicast and multicast listeners. M-HMIPv6 uses a reactive handover mechanism, whereas M-FMIPv6 prefers to fulfill a handover prediction, with fallback to a reactive procedure.

Before we evaluate these schemes in a comparative analysis, let us first fix the relevant qualitative aspects:

Handover performance: packet loss, delay and jitter occur, while the Mobile Node switches between networks. Packets are lost, while it is disconnected and no packet buffering takes effect. Delay and jitter are added by the handover procedure, if packets are buffered or transmitted through indirect paths.

Number of performed handovers determines the frequency of changes between networks, while the Mobile Node moves.

Number of processed handovers denotes the frequency of handover procedures processed within the infrastructure. The processed handovers may exceed the actual handover completions, as predictive techniques tend to initiate artificial handover procedures.

Overhead in signalling and traffic distribution depends on the handover protocols in use and places an extra burden to mobility tasks.

Robustness is the expression of dependence of the handover performance on changes in network geometry or rapid movement.

The above aspects summarize the amount of disturbance produced by the roaming procedures, the effort of support requested from the networking infrastructure and the overall price to be paid in overhead costs.

3. Analysis of handover performance

3.1. *Theoretical considerations*

To analize handover performance effects and quantities, we first present simple analytical considerations. They rather intend to derive insights into the different handover mechanisms, than to develop a complete and complex quantitative picture. Throughout this entire section, all buffer effects are neglected, even though the handover protocols under consideration likewise provision for it. Buffers transform packet loss into delay, thereby weakening insight into protocol effects.

In the event of a Mobile Node switching between access networks, it may completely disconnect from the link layer. Thereafter it needs to perform an IP reconfiguration and Binding Updates to its infrastructure. Until completion of all these operations the MN is likely to experience disruptions or disturbances of service, as are the result of packet loss, delay and jitter increases. In general the handover process decomposes into the geometry independent local handoff, the Layer 2 link switching and the IP

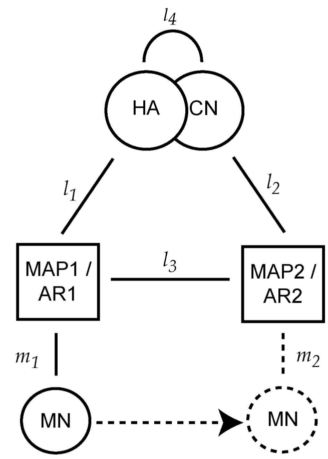


Figure 3. A simple analytical model.

readdressing, and the geometry dependent Binding Update activities. Let t_{L2} denote the Layer 2 handoff duration, $t_{\text{local-IP}}$ the time for local IP reconfiguration and t_{BU} the Binding Update time. Then the following temporal decomposition for handovers holds:

$$t_{\text{handoff}} = t_{L2} + t_{\text{local-IP}} + t_{BU}. \tag{1}$$

Both mobility schemes under consideration attempt to optimize the non-local update procedures by means of proxying and anticipating delay hiding techniques.

To proceed into a more detailed analysis of the handover mechanisms, we consider a simple analytical model (figure 3) in this section: A MN moves from access router 1 (AR1) to access router 2 (AR2) with intermediate 'link' l_3. Based on it we will analytically derive basic properties and present a first quantitative study.

Within this model ARs have been identified with mobility anchor points 1 and 2 for comparative reasons. This simplification merely affects the 'wireless' link dimensions m_1/m_2 to the MN, which are assumed small, but not of one-hop type. Distances l_1 and l_2 to HA or CN must be viewed as possibly large and represent the strongest topological dependence within the model. The distance between the access routers, l_3, should be viewed as a variable, but characteristic geometric entity. As the MN moves between routers, their separation represents the gap to be bridged by forwarding, somewhat the 'size of the step'.

Let t_l denote the transition time of a packet along link l, then the following measures of handover quality for the bi-directional tunneling approach can be easily derived:

Bi-directional Tunneling (MIPv6).

$$\text{Packet loss} \propto t_{L2} + t_{\text{local}-\text{IP}} + t_{m_2} + t_{l_2}$$
$$\text{Additional arrival delay} = t_{l_2} - t_{l_1} + t_{m_2} - t_{m_1}$$

Packet loss remains proportional ($\propto$) with the factor of packet injection rate. Similarly loss and delay for reactive handovers can be calculated, which no longer depend on the distant topology of HA and CN:

Reactive Handover (M–HMIPv6).

$$\text{Packet loss} \propto t_{L2} + t_{\text{local-IP}} + t_{m_2} + t_{l_3}$$
$$\text{Additional arrival delay} = t_{l_3} + t_{m_2} - t_{m_1}$$

To evaluate the more complex operations of the predictive handovers, we introduce t_{Ant}, the anticipation period prior to handover. It should be noted that predictive handovers are optimal, iff anticipation time matches the time needed for á priori router negotiations (excluding fback) exactly. To account for deviations from exact prediction timing, we introduce the difference operators

$$\Delta^{\pm} t = \max\left(\pm t_{\text{Ant}} \mp 2t_{l_3} \mp t_{m_1}, 0\right), \text{ and}$$
$$\Delta t = \Delta^{+} t - \Delta^{-} t.$$

$\Delta^{+} t$ accounts for the time difference of early predictions, i.e. the time a MN leaves later than predicted, $\Delta^{-} t$ calculates the temporal gap resulting from late predictions. Δt finally represents the signed difference operator, combinedly measuring the temporal translation. The packet loss in predictive handovers additionally depends on correctness of the prediction. A misspredicted cell transfer will force the algorithm to fall back onto a reactive handover with inferior performance (3). During a correctly foreseen transition, packets firstly may be lost at previous association, in case of MN leaving too early, and secondly at next association, for the MN reconnecting later than the arrival of forwarded packets:

Predictive Handover (M-FMIPv6, successful prediction).

$$\text{Packet loss} \propto \Delta^{-} t + \max\left(\Delta t + t_{L2} + t_{m_2} - t_{l_3}, 0\right) \qquad (2)$$
$$\text{Additional arrival delay} = t_{l_3} + t_{m_2} - t_{m_1}$$

Predictive Handover (M-FMIPv6, erroneous prediction).

$$\text{Packet loss} \propto \Delta^{+} t + \text{reactive handover loss} \qquad (3)$$

Whereas MIPv6 handover timing is dominated by the update delay with the HA, predictive and reactive handovers solely depend on relative geometry, i.e. the distance of access routers or MAPs. Predictions additionally admit a strong dependence on anticipation time. Figure 4 visualizes the packet loss (2) as function of the variables anticipation time and access router distance in a constant bitrate scenario of one packet per 10 ms. A sharp minimum can be observed for anticipation being roughly equal to doubled router distance with a further descent towards a router distance approaching the L2 handoff

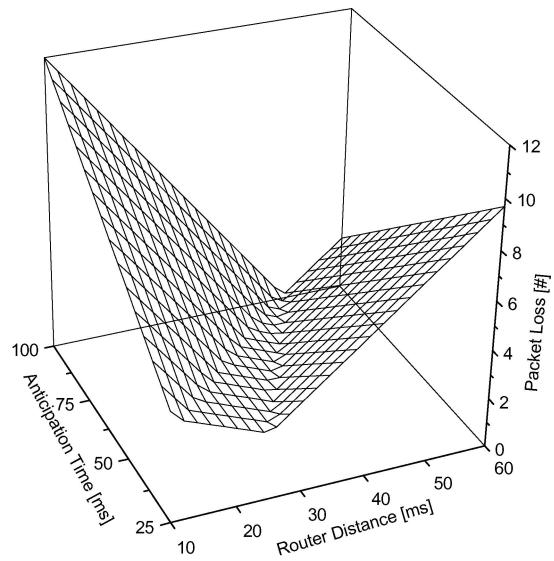


Figure 4. Analytical packet loss in predictive handovers as a function of anticipation time and access router distance.

time (50 ms). When we recall that packet loss in predictive handovers may attain an absolute minimum of zero, if anticipation time exactly matches inter-router negotiations and L2 handoff exactly matches the inter-router packet forwarding time, the shape of the loss function becomes plausible. Of higher interest, though, the steepness of the loss function must be recognized. By observing that predictive handover protocols do not provide mechanisms to adapt to router distances, the effectiveness of the entire scheme is challenged by its parameter sensitivity.

Figure 5 compares predictive and reactive packet loss as functions of the access router distance for selected values of anticipation times. Reactive handover performance admits comparable or lower loss in regimes of close access routers, whereas predictive handovers significantly outperform reactive loss for distant topologies. Again a strong dependence on anticipation time becomes visible.

Considerations so far have been grounded on analytical results obtained from our simplified model. In the subseqent section we will contrast the analysis with empirical simulation results.

3.2. Stochastic simulation of handover schemes

The theoretical model developed in the previous section had been a simplification, as it singled out the predictive from reactive handover mechanisms. Since router negotiations need not complete prior to MN's handover and as predictions may miss completely, any realistic handover scenario will consist of a mixture of (partly) successful predictive and

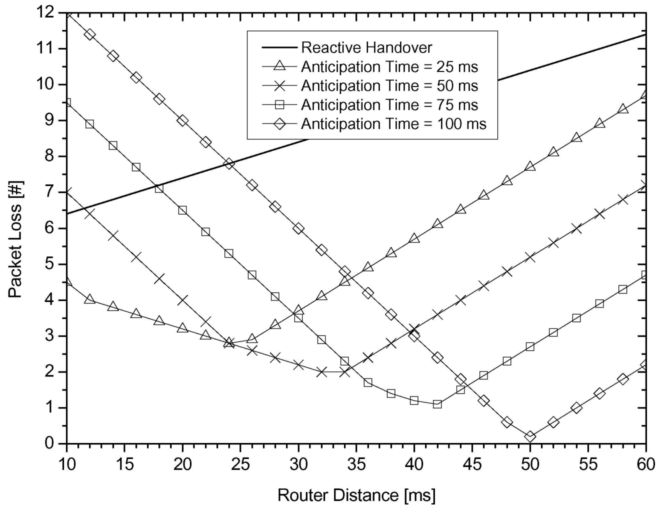


Figure 5. Analytical packet loss in predictive handovers as a function of access router distance.

reactive handovers. Furthermore we did not account for variations in link delays and anticipation timing.

To analyze handover performance empirically, a stochastic discrete event simulation of predictive and reactive handovers has been performed. The MN moves from AR1 to AR2, while the CN continuously sends probe packets at a constant bit rate of one packet per 10 ms. All components operate the M-FMIPv6/M-HMIPv6 protocols. All notation is kept conformal to our previously described model.

In the case of a predictive handover, the MN starts an anticipation timer and submits a Fast BU to AR1, which subsequently negotiates with AR2 and eventually returns an FBU acknowledge (fback) to the MN. At this time the MN may have disconnected from its link to AR1 and thus will have missed fback. As soon as the L2 handoff delay elapsed, the MN reconnects with AR2, submits a fast neighbour advertisement and is ready to receive packets forwarded by AR1.

Following this procedure, we simulated a ‘friendly’ predictive handover, i.e. all probabilities of erroneous predictions as not to arrive in the foreseen subnet, are omitted. For the probability distribution of erroneous predictions see Section 4.2.

Performing a reactive handover the MN leaves the link with AR1 instantaneous to reconnect after the L2 handoff time at AR2. Via AR2 it submits a BU to AR1, which starts packet forwarding subsequently.

For the simulation all temporal entities have been taken to be uniformly distributed random variables with mean $<x>$ and standard deviation σ_x. The L2 handoff delay values were taken from [7], access link delays at the MN at small scale. The anticipation period has been varied on the scale of access router distances, but—as values are of larger uncertainty—chosen with large perturbations. For local IP reconfiguration we

Table 1
Simulation timers representing mean $<x>$ and standard deviation σ_x.

	Anticipation time	L2 Handoff delay	Wireless link delay	Router distance
$<x>$	25–100 ms	50 ms	2 ms	10–60 ms
σ_x	8.7 ms	3 ms	0.3 ms	0.7 ms

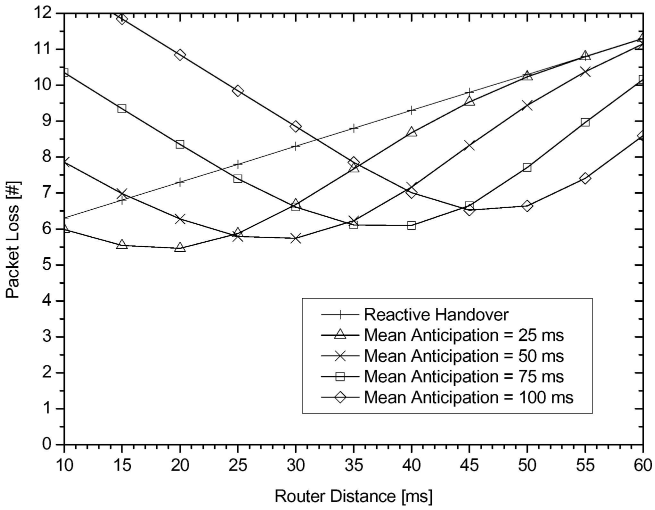


Figure 6. Simulated packet loss in reactive and predictive handovers for various mean anticipation times as a functions of access router distance.

assume rapid movement discovery, e.g. by L2 triggers, and suppression of duplicate address detection, as suggested by MIPv6. In detail we used the parameter set shown in Table 1.

Results

Packet loss as functions of router distance are shown in figure 6 for predictive and reactive handover simulations comparatively. Clearly seen can be the prediction characteristic of minimal loss at a router distance of roughly half the anticipation period. Curves are smoothed as compared to theory (s. figure 5) due to a statistical mixture from varying timer values. Loss yields are significantly enhanced for the empirical results, showing the outcome of reactive handover contributions. Excluded from our theoretical model, the latter account dominantly for packet loss in the regime of low completion probability. Probabilities of successfully completing the anticipation procedure are shown in figure 7 for various anticipation periods as a function of the router distance (‘step size’). Contributions from complete and incomplete handover negotiations are displayed in figure 8. The handover scheme approaches a pure reactive scheme, when anticipation time diminishes compared to the access router distance and predictive accelerations

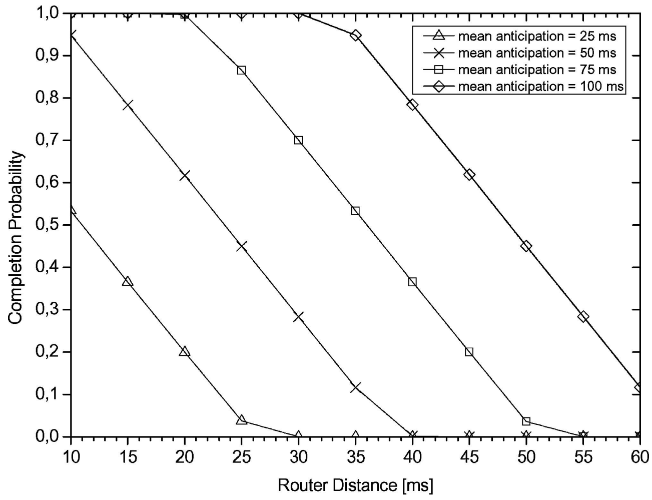
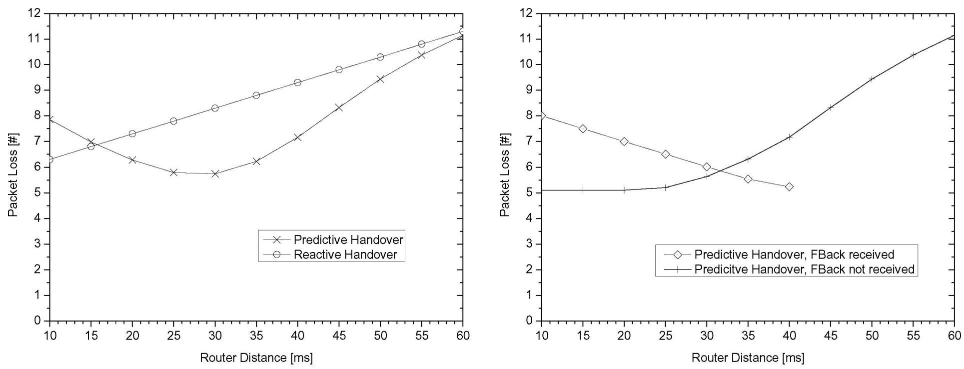


Figure 7. Empirical probabilities of completing the predictive handover negotiations.



(a) (b)

Figure 8. Packet loss during a predictive handover composes from different mechanisms of the scheme: (a) Simulated complete loss for mean anticipation time of 50 ms, (b) contributions from complete and incomplete handover negotiations as functions of access router distance.

cease to be visible. Hence predictive loss rates approach the reactive yields for $t_{\mathrm{Ant}} \ll t_{l_3}$.

In a narrow topology the reactive handover admits equal or better performance than the predictive approach, since update negotiations between access routers or MAPs do only account for small delays. Consequently loss becomes dominant, which is due to inaccuracy of predicting handoff times. Additionally, Layer 2 offtimes cannot be hidden by inter-router forwarding. Note that anticipation timers, Layer 2 handoff and router distances are completely uncorrelated. This regime of close access router topology is well represented by our analytical results of Section 3.1.

3.3. Robustness and overheads

A central goal in designing M–HMIPv6 and M–FMIPv6 has been topological robustness in the sense that handover behaviour remains independent of network geometry. Both protocols fulfill this task in a similar way: Any blocking communication with the Home Agent has been excluded. Signalling in both approaches is dominated by the layer 2 handoff and local router exchanges (compare Section 3.1).

Another important aspect of robustness must be seen in the ability to cope with rapid movement of the Mobile Node. In the case of a mobile multicast listener leaving its association before a handover completed, an M–HMIPv6 device will remain bound with its previously fully established MAP or Home Agent. On the price of a possibly increased delay, forwarding of multicast data is provided independent of handover frequency. On the contrary M-FMIPv6 forwarding will collapse, as soon as a MN admits a handover frequency higher than the signalling completion periods. To see the latter, observe that the fast handover protocol solely establishes forwarding from pAR to nAR, which is useless, if traffic has not reached the previous network. An M-FMIPv6 device then has to fall back onto MIPv6 by establishing a bi-directional tunnel anew. Meanwhile all ongoing services are interrupted.

In addition protocol overheads may be taken into account: Both protocols announce multicast capabilities within regular AR or MAP advertisement. The hierarchical MIPv6 distinguishes cases of handovers between a micro mobile change, which will be accompanied by only one local protocol message, and the inter–MAP handover. The latter is operated by two messages. Handover signalling of fast MIPv6 always requests for seven messages. MIPv6 and multicast routing procedures add on top of both protocols in the event of visible handovers.

4. Analysis of handover frequencies

4.1. Expected number of handovers

As a Mobile Node moves, handovers potentially impose disturbances and put an extra effort onto the routing infrastructure. Thus the expected frequency of network changes can be viewed as a distinctive measure of smoothness for a mobility scheme. The handoff frequency clearly depends on the Mobile Node's motion within cell geometry. Two measures on quantizing mobility have been established in the literature: The cell residence time and the call holding time. Both quantities fluctuate according to the overall scenery and the actual mobility event.

Let us make the common assumption that the cell residence time is exponentially distributed with parameter η and that the call holding time is exponentially distributed, as well, but with parameter α. Then the probability for the occurrence of a handover from MNs residence cell into some neighbouring can be calculated analytically to

$$\mathcal{P}_{\mathcal{HO}} = \frac{1}{1+\rho}, \quad \text{where } \rho = \frac{\alpha}{\eta}.$$

ρ is known as the *call-to-mobility* factor. For a detailed analysis compare [3]. It can be observed that the handoff probability increases as ρ decreases. Note that all probability distributions are homogeneous in space, e.g. $\mathcal{P_{HO}}$ is independent of the current cell or the number of previously occurred handovers. Spatial scaling can be applied, accordingly.

When comparing Fast MIPv6 and Hierarchical MIPv6 approaches, another distinctive quantity becomes relevant: Whereas FMIPv6 operates handovers at Access Routers, HMIPv6 utilizes MAPs, which form a shared infrastructure. In general one MAP is meant to serve k Access Routers, whereby the expected number of (inter–MAP) handovers reduces in a HMIPv6 domain.

Let us assume MAP regions to be of approximately circular geometry. Then the expected cell residence time changes to

$$\eta_{\mathrm{MAP}}^{-1} = \sqrt{k} \cdot \eta_{AR}^{-1}$$

and the handoff probability transforms into

$$\mathcal{P_{HO}} = \frac{1}{1 + \sqrt{k} \cdot \rho},$$

where k is the ratio of ARs to MAPs.

Now we are ready to calculate the expected number of handovers as a function of the *call-to-mobility* factor ρ and the AR to MAP ratio k

$$\mathcal{E_{HO}} = \sum_{i=1}^{\infty} i \left(\frac{1}{1 + \sqrt{k} \cdot \rho} \right)^{i} = \frac{1}{k \cdot \rho^2} + \frac{1}{\sqrt{k} \cdot \rho}. \tag{4}$$

It can be observed that in highly mobile regimes ($\rho \ll 1$) $\mathcal{E}_{\mathrm{HO}}$ is dominantly a function of the inverse of k, for low mobility ($\rho \gg 1$) of the inverse of $\sqrt{k}$ and attains a singularity at $\rho = 0$. Figure 9 visualizes this expectation value for common values of ρ and k.

4.2. Accuracy of handover predictions

A handover occurance may be predicted correctly and on time. Information on advancing handovers commonly are derived from layer 2 radio interferences and are thus bound to regions of signal overlaps. A mobile node moving within an interference region, though, may retain its network association or perform a handoff different from prediction. It may also decide to reassociate without presensing interferences at all.

Stochastic simulation of walking users

To evaluate the distribution of handover prediction types, we perform a stochastic simulation of motion within radio cells. Bearing in mind lightweight, universal mobility stacks, we restrict modeling assumptions to only include basic geometry and motion. The underlying model combines the following ingredients:

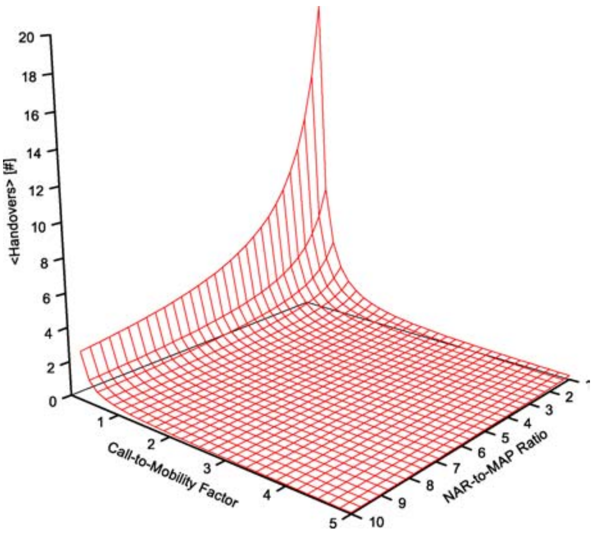


Figure 9. Expected number of handovers as function of the call-to-mobility factor and the ratio of access routers to MAPs.

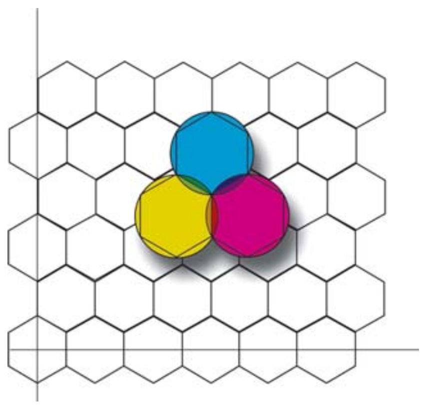


Figure 10. The comb geometry with regions of overlap.

Cell geometry is chosen to be of common honeycomb type, i.e. abutting hexagons completely fill the 2D plane. The ranges of radio transmission are modeled as (minimal) circles enclosing the combs. Thus, regions of prediction are the overlapping circle edges as shown in figure 10. A geometry of coherent Hot Spots is assumed here, where cells—without loss of generality—are identified with individually routed subnets.

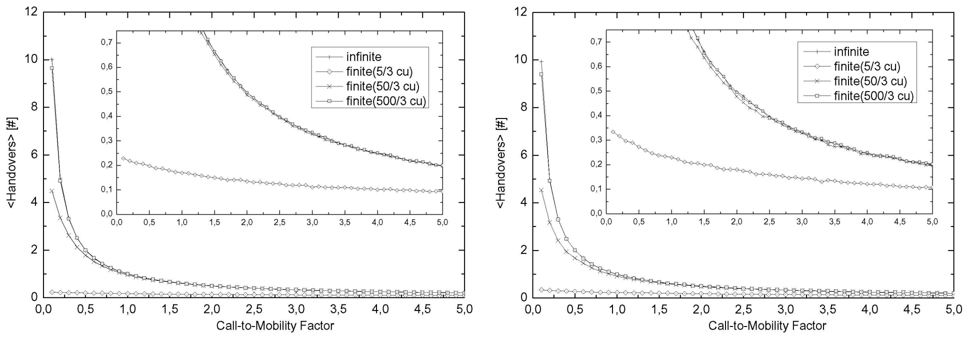
As *Walking models* a Random Waypoint Model and a Random Direction Model [1] are used, where we consider varying sizes of mobility regions, i.e. squares ranging from cell dimension to infinity. Mobile devices move along (piecewise) straight lines within the preset boundaries, possibly coming to rest at their final destination, even if their current call is ongoing. Note that for finite regions the dynamic of both models is ergodic, whereby our simulated motion is equivalent to a walk continuous in time. Predictions are evaluated along the trajectories, distinguishing their correctness according to the outcome in succeeding steps.

Simulations have been carried out by integrating the equations of motion in a 4D phase space with identified regions of equivalence, as defined by the comb geometry. Velocities are governed by a uniform distribution. Spacial boundaries act as reflectors. Absolute coordinate values cancel out for our integrated quantities measured. It is worth noting that calculation of the mean cell residence time requires the mean path length within a hexagon of edge r, which by the Cauchy–Crofton formula evaluates to $\frac{\sqrt{3}\,\pi}{4}\, r$.

Results

Mean handover occurrences for the different geometries are presented in figure 11. Theoretical results of Section 4.1 reproduce nicely for large mobility regions, whereas small areas of motions are dominated by boundary effects. Values for small regions are attenuated, as tight geometry borders reduce phase space in 'pushing' trajectories into the inner cells, thereby reducing handoff activities. There is no visible dependence on the walking model in use.

Erroneous predictions as the outcome of traversing prediction regions without actually performing the foreseen handover or handovers initiated from outside of prediction areas are shown in figure 12. The graph cumulatively displays the percentage of failures from all predictions in the smallest and infinite mobility region for both walking models. From results varying between 10 and 70% it can be observed that geometry dependence clearly exceeds the influence of the model. Smaller regions of mobility significantly overpredict handovers, as reflecting boundaries are not visible by radio interference. In the regime of high mobility, the boundary-free random waypoint model noticably underestimates erroneous prediction yields. These results can be identified as



(a) Random Waypoint Model (b) Random Direction Model

Figure 11. Mean frequencies of handovers as a function of the call-to-mobility factor for different mobility regions (in units of cell diameters cu).

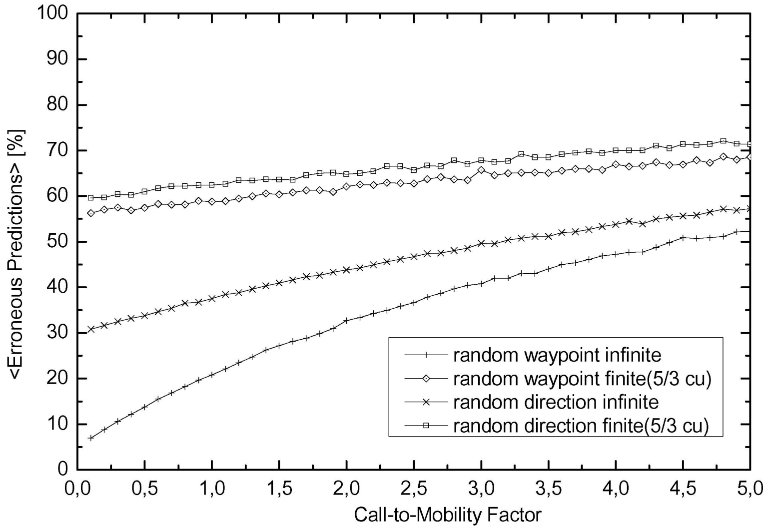


Figure 12. Mean relative yield of erroneous predictions as a function of the call-to-mobility factor for varying transmission ranges.

a model artifact—in the absence of ergodicity the infinite random waypoint model fails to explore large parts of the phase space. Taken as a rough estimate, the average number of erroneous predictions is about equal to the number of correct ones. Thus the average reliability of predictive handover schemes does not exceed 50%.

Previous simulations have been performed under the geometric assumption of layer 2 transmission radii matching comb edge sizes. We now consider different radii for the circular radio transmission areas, simulating a variation in access point densities. Graph 13 compares corresponding results for the random waypoint model with intermediate mobility region (50/3 cell diameter units). The results for systems with optimal radio coverage, i.e. cell radii equal to transmission ranges, show minimal portions of failure predictions. In general a distinctive geometry dependence becomes visible, as well.

To proceed into a more detailed analysis of the sampled predictions, we differentiated the handover events of a simulated trajectory ensemble (model and parameters as in figure 13). Figure 14 visualizes the mean occurrences of correct predictions, false predictions obtained along the path, as the mobile moves contrarily to forecasts derived from radio signal detection, and erroneous predictions generated by terminating movement or call at a position, where a handover is expected. The latter yields *on stop* can be identified as almost mobility independent, resulting in a saturated minimal error rate. Incorrect predictions *on path* as a function of the call-to-mobility factor in contrast scale in correspondence with the correct indicators. It can be concluded from figure 12 that their exact values clearly depend on geometric and walking type conditions.

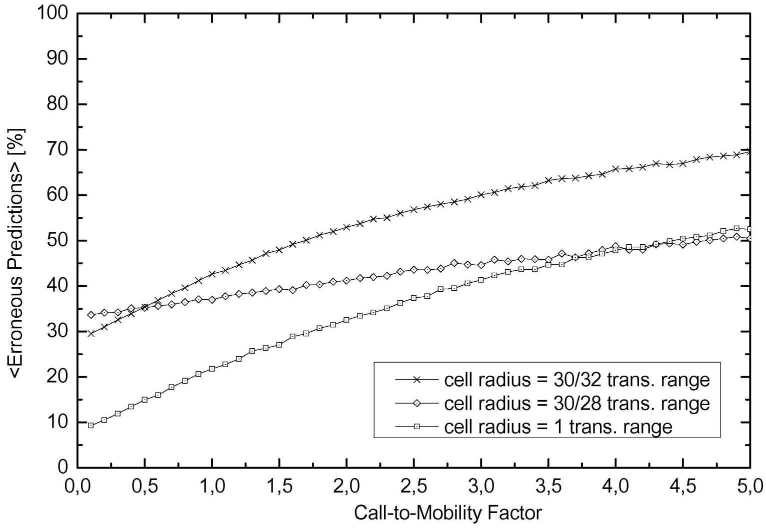


Figure 13. The comb geometry with regions of overlap.

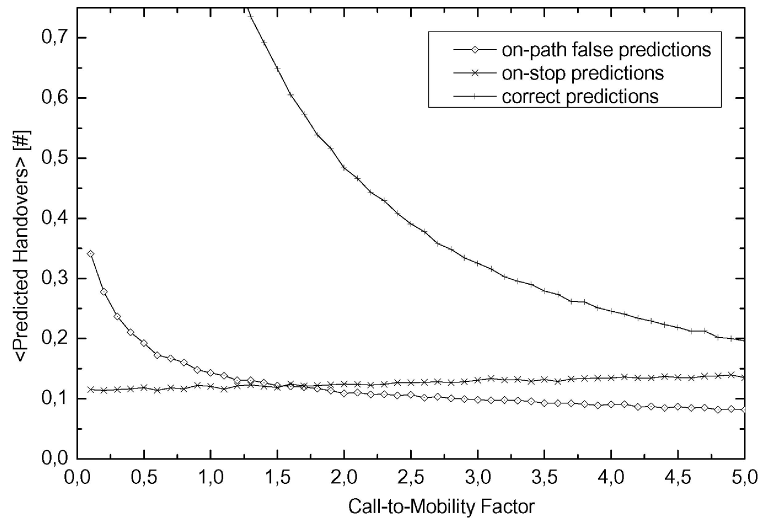


Figure 14. Detailed view on handovers: Correct predictions, failures along the path and on stop.

4.3. Immediate implications

A common goal in designing HMIPv6 and FMIPv6 has been to approach seamless handovers in mobile IPv6 networks. As was shown in the previous Section 3, the predictive scheme of FMIPv6 can lead to faster roaming operations, but the reactive HMIPv6 procedures admits a comparable timing. In scenarios of significant mobility, i.e. $\rho \leq 1$, this advantage may be easily compensated by reducing the number of attained handovers within an HMIPv6 environment up to an order of magnitude. High prediction error rates, as observed from our simulations, place an additional burden onto the infrastructure, since any handover forecast will lead to a negotiation chain between access routers.

This burden notably increases in the case of multicast communication. A preparatory roaming procedure in M-FMIPv6 will initiate a remote multicast subscription, causing multicast routers to erect new branches for the corresponding distribution trees. In combining the results of Sections 4.1 and 4.2 we face scenarios, where the same (high) mobile movement leads to three handovers in a M–HMIPv6 domain, but about 40 handover processings under the regime of M-FMIPv6.

5. Conclusions and outlook

Mobility is one of the most challenging and demanded developments in IP networks today. Mobile multicasting—as needed in many group conferencing scenarios—places an even stronger challenge on today's IP concepts and infrastructure. In this paper we presented the current approaches and elaborated a comprehensive comparison of predictive and reactive handover behaviour in mobile IPv6 environments. Special focus had been dedicated to multicast receiver protocols built upon these unicast operations.

At first major performance aspects have been studied analytically and in stochastic simulations. It was derived that predictive schemes could not approve expectations to significantly outperform reactive handovers. This was due to a narrow parameter band of optimized performance, which is barely met in practice. In contrast handover anticipation raises the issue of erroneous predictions, which are punished by a significant performance degradation.

At second we analyzed common predictive, reactive and proxy mobility schemes w.r.t. their eagerness for handovers. Starting from simple, fundamental assumptions a quantitative study of expected handover occurrences was derived. Anticipation reliability was simulated using common mobility models. A large yield of incorrect predictions, about 50% was obtained, generated a high burden of needless handover processings and erroneous change predictions.

The 'nervousness' of predictive handovers performed at access routers could be shown to reduce significantly in the presence of Mobility Anchor Points established within the hierarchical MIPv6 approach. This smoothing effect gains additional importance by observing an instability of fast handovers in the case of Mobile Node's rapid movement.

The perspective of these results may give rise to further improvements on the smoothing of roaming procedures within the realm of seamlessness, attained at mobile and infrastructure nodes.

Acknowledgment

We would like to thank Otfried Cheong for his invention of Ipe and helpful hints and discussions on geometric affairs. This work has been supported in part by the

German Bundesministerium für Bildung und Forschung within the project MUSICAL.

References

[1] C. Bettstetter, Mobility modeling in wireless networks: Categorization, smooth movement, and border effects, ACM Mobile Computing and Communications Review 5(3) (2001) 55–67.
[2] H.L. Cycon, M. Palkow, T.C. Schmidt, M. Wählisch and D. Marpe, A fast wavelet-based video codec and its application in an IP version 6-ready serverless videoconferencing system, International Journal of Wavelets, Multiresolution and Information Processing 2(2) (2004) 165–171.
[3] Y. Fang and I. Chlamtac, Analytical generalized results for handoff probability in wireless networks, IEEE Transactions on Communications 50(3) (2002) 396–399.
[4] D.B. Johnson, C. Perkins and J. Arkko. Mobility Support in IPv6. RFC 3775, IETF, June 2004.
[5] R. Koodli, Fast handovers for mobile IPv6. RFC 4068, IETF, July 2005.
[6] I. Romdhani, M. Kellil, H.-Y. Lach, A. Bouabdallah and H. Bettahar, IP mobile multicast: Challenges and solutions, IEEE Comm. Surveys & Tutorials 6(1) (2004) 18–41.
[7] T.C. Schmidt and M. Wählisch, Roaming real-time applications—mobility services in IPv6 networks, Zagreb, May 2003. TERENA Networking Conference. URL http://www.rz.fhtw-berlin.de/projekte/mipv6/terena03.pdf.
[8] T.C. Schmidt and M. Wählisch, Seamless multicast handover in a hierarchical mobile IPv6 environment (M-HMIPv6), Internet Draft—work in progress 03, individual, April 2005. URL ftp://ftp.rfc-editor.org/in-notes/internet-drafts/draft-schmidt-waehlisch-mhmipv6-03.txt.
[9] T.C. Schmidt, M. Wählisch, H.L. Cycon and M. Palkow, Global serverless videoconferencing over IP, Future Generation Computer Systems 19(2) (2003) 219–227.
[10] H. Soliman, C. Castelluccia, K. Malki and L. Bellier, Hierarchical mobile IPv6 mobility management (HMIPv6). Internet Draft—work in progress 04, IETF, December 2004.
[11] L. Song, D. Kotz, R. Jain and X. He, Evaluating location predictors with extensive Wi-Fi mobility data, in: *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, Vol. 2, (March 2004) pp. 1414–1424.
[12] K. Suh, D.-H. Kwon, Y.-J. Suh and Y. Park, Fast multicast protocol for mobile IPv6 in the fast handovers environments, Internet Draft—work in progress (expired) 00, IETF, February 2004.
[13] H.-K. Zhang, B. Shen and B.-Y. Zhang, Mobile IPv6 multicast with dynamic multicast agent, Internet Draft—work in progress 00, IETF, June 2005.