A Policy for Group Handover Attempts over Heterogeneous Networks

Nivia Cruz Quental and Paulo André da S. Gonçalves
Centro de Informática – CIn
Universidade Federal de Pernambuco – UFPE
Recife, Brazil
Email: ncq@cin.ufpe.br, pasg@cin.ufpe.br

***Abstract*—The proliferation of heterogeneous wireless networks and devices with multiple wireless link-layer technologies have called attention to the development of efficient vertical handover policies. Research on providing efficient vertical handover includes either the proposal of novel solutions or adaptations into existing horizontal handover schemes. In this work, we propose a policy for Group Vertical Handover (GVHO) attempts. We apply such a policy to an existing GVHO scheme, which handles vertical group handover based on a threshold that limits handover blocking probability. Performance is evaluated through simulation under several scenarios. To provide more realistic situations, we consider channel holding time in our studies. In addition, we study the fraction of blocked nodes and we vary threshold values for blocking probability. We compare our solution to that of the studied GVHO scheme. Results show that our solution reduces the handover latency and the fraction of blocked nodes while maintaining the handover blocking probability under a predefined threshold. In particular, latency is reduced from 11% to 51.5% in some of the scenarios studied.**

Keywords-GVHO; handover; policy of attempts.

I. INTRODUCTION

Load balancing and handover among different Radio Access Technologies (RATs) are the main concerns in Group Vertical Handover (GVHO) studies [1]-[3]. Research on GVHO covers simultaneously issues from Group Handover (GHO) [4]-[6] and Vertical Handover (VHO) [7]-[18].

GHO takes place when two or more Mobile Nodes (MNs) intend to request handover at the same time to the same base station. During GHO, MNs are not necessarily aware of the presence of each other. Thus, GHO procedures must carry out load balancing. To achieve this, criteria such as energy saving, available bandwidth, and type of service may be considered.

The continuity of telephone calls and streaming sessions over heterogeneous networks are covered by the VHO research field. Quality of Service (QoS) and the type of traffic may also define requirements for handover decisions apart from the underlying network technology available. IEEE 802.21 [19] is an example of effort to standardize VHO procedures and facilitate the proposal of new VHO solutions.

Providing support to GVHO has been motivated by the recent popularity of devices such as tablets and smartphones, which are capable of supporting multiple link-layer technologies and handling different kinds of traffic. Additionally, use cases involving users moving in trains and on buses are becoming more common and introduce new challenges. At the IP level, protocols like Proxy Mobile IPv6 [20] manage mobility sessions at the network layer. In this paper, we are particularly interested in the link layer handover.

Research on GVHO may involve the three main handover phases: discovery, decision, and execution [18]. The decision phase interests us the most, since the decision process in GVHO is still an open issue and it may impact the GVHO overall performance. Further, the decision algorithm itself must be associated to an optimized policy for GVHO attempts to guarantee better handover performance. Research on GVHO seeks to provide efficient decision-making techniques with their own policy for handover attempts. Some of them are based on centralized entities [21], distributed algorithms [22], random delays [2], reinforcement learning [23], game theory [2], and optimization problems [3]. We give special attention to Lee *et al.* [3], since it addresses the latency reduction while considering load balancing, support to legacy networks, and handover blocking probability. A reduced GVHO latency means less time spent in the GVHO operation. Load balancing is the consequence of a efficient resource management. Controlling the handover blocking probability means that the probability of the MN having its handover request denied by the target network is limited. Those issues are fundamental for advances in GVHO. The objective of Lee *et al.* [3] is to model GVHO decision as an optimization problem. Latency is minimized given the condition of maintaining the handover blocking probability under a predefined threshold. Although Lee *et al.* [3] present encouraging results, we find optimization opportunities in the policy for handover attempts.

In this paper, we propose a policy of attempts for GVHO. Our policy is based on exponential backoff and uses information from the GVHO scheme itself. We improve on previous experiments [1] by considering channel holding time in performance evaluations, which makes studied scenarios more realistic. In addition, we study the fraction of blocked nodes and several thresholds for the blocking probability. The proposed solution reduces average latency and the fraction of blocked nodes in comparison to results found in [3].

The remainder of paper is organized as follows: we present GVHO concepts in Section II. We present related work in Section III. We detail the GVHO scheme proposed in [3] in Section IV. We present the proposed policy for GVHO attempts in Section V. We present performance evaluation results in Section VI. Finally, we highlight our conclusions in Section VII.

II. GROUP VERTICAL HANDOVER - GVHO

Recently, the concept of handover has evolved to take into account the continuity of communication sessions even among different RATs [24]. Technological evolution has allowed the rising of cheaper gadgets supplied with multiple network interfaces. The appearance of such gadgets, in turn, has encouraged new research in mobility management considering brand-new use-cases. Studies in the Group Vertical Handover (GVHO) area of interest aim at managing different connections taking place at the same time in public spaces with a diverse number of available technologies.

An example of GVHO scenario is illustrated in Figure 1. Suppose an open event, like a music festival where users desire to communicate with friends and transmit multimedia data. In this scenario, users are constantly changing their location. There may be several available RATs and dozens of devices in communication sessions simultaneously. If there are commercial agreements among the telecommunication carriers, it must be possible to maintain a communication session even if a group of users move from one network to another at the same time.

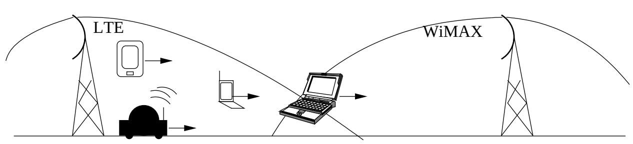

Figure 1. A GVHO scenario.

The integration between heterogeneous networks can be divided in two approaches: loose coupling and tight coupling [7]. In the loose approach, heterogeneous networks are integrated a the IP level, but they operate independently at the link layer. In this case, it is necessary to rely on a gateway to support authentication and accounting. Since handover is done at link layer, routing tables and authentication information must be updated. In order to accomplish this operation, mobility support is added to IPv4 and IPv6 protocols. The Internet Engineering Task Force (IETF) efforts in mobility management are concentrated at the IP level and layers above. The main protocols derived from that effort are Mobile IP [25], Mobile IPv6 Fast Handovers [26], PMIPv6 [20], and Fast Handovers for Proxy Mobile IPv6 (FPMIPv6) [27]. These standards allow the MN to maintain its initial IP address, even when it is out of its home network. The main advantage of loose coupling is the simple adaptation to legacy systems. However, handling handover only at the IP level may not entirely solve the problem of interruption of communication during this operation.

In the tight approach for heterogeneous networks coupling, the network entities of each technology must explicitly collaborate with each other. The authentication, communication management, and accounting are integrated. This approach requires more standardization effort than the loose approach. However, the handover management becomes more effective in terms of the number of lost packets.

For all layers of the protocol stack, the proposal of efficient and effective handover procedures for mobility management including handover decisions and optimal resource allocation is a critical need. In this paper, we are particularly interested in the efficiency of GVHO at link layer, when tight coupling takes place. In that context, proposals may involve the three handover phases [18]:

- *Discovery* - Service discovery and network information gathering. A specific criterion is adopted to determine if handover is necessary. According to [28], the gathered information may have a predetermined nature, like user policies and preferences, or time-varying nature, like signal-to-noise ratio, transmission rate, Point of Attachment (PoA) load, battery consumption, Received Signal Strength (RSS), RSS with threshold, RSS with hysteresis, etc.
- *Decision* - One network in a list of candidates is chosen, taking into consideration data collected in the earlier phase. Depending on the network technology, handover may be MN-initiated or network-initiated. The decision technique and the policy for handover attempts strongly impacts the resource management and the overall handover performance. Thus, the decision phase is the focus of this paper.
- *Execution* - Networks and MNs exchange control messages to make channel switching. This phase should minimize service interruption in order to appear imperceptible to the user. This phase is strongly media-dependent. The Hard Handover (HHO) implementation is mandatory for all technologies. HHO takes place when the MN disconnects from its original PoA before making the first contact with its target PoA. Since packages may be lost during that interval, optional Soft Handover (SHO) mechanisms are proposed. SHO mechanisms include Seamless Handover; Entry Before Break (EBB); Multicarrier Handover; Fast Base Station Switch (FBSS), or Fast Cell Selection (FCS); and Macro-Diversity Handover(MDHO) [12].

IEEE 802.21 standard [19], which describes the Media Independent Handover (MIH) can help determining the requirements for discovery and decision phases. MIH intends to be a common mean over the link layer in order to allow different RATs to communicate with each other during handover, abstracting implementation details. MIH is still a relatively new standard and, therefore, it faces challenges such as abstracting wireless technologies in a single interface, incorporation into existing handover schemes, security issues, power management, and storage issues at the information service.

Each RAT must provide its own implementation of MIH and must map the MIH messages to its media-dependent primitives. The main elements of MIH are:

- MIH Function (MIHF) - It detects changes in link layer, controls link state and provide neighborhood information;
- Service Access Points (SAPs) - It defines media-dependent/independent interfaces;

- MIH Users - Entities that make use of MIH services.

A handover scenario with MIH assumes the existence of an information service to help MNs to find neighborhood information. It avoids the MN to waste energy ant time making scanning operations by itself. There are plenty of studies on handover performance adopting MIH as an auxiliary tool for discovery and decision processes [3][29]-[32].

There may be many different decision criteria for GVHO such as available bandwidth, expected QoS, or battery consumption. The type of service (voice or data) is a determinant factor for choosing the most suitable criterion for GVHO decision. Decisions made without network analysis and without considering the MNs in the neighborhood may bring disastrous performance results. Wrong handover decisions may cause MNs to choose the same PoA, overloading it, or to choose an inadequate network for the application in use. The main handover decision approaches found in GVHO research include:

- *Centralized entities* [21][33] - A relay station handles GVHO management, removing complexity from MNs. This approach also reduces the uncertainty level and ensures better performance than decentralized approaches. The main drawback is the lower fault tolerance. Figure 2 illustrates an architecture based on central entity. If the entity suffers a failure, the mobility management would be damaged.

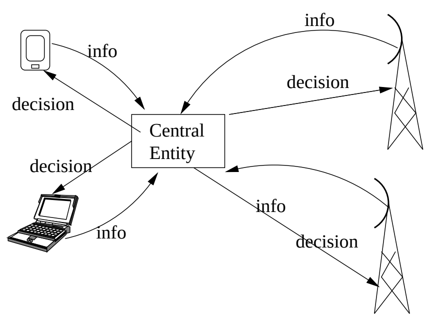

Figure 2. Central-entity-based approach.

- *Distributed algorithms* [22] - The decision algorithm makes use of well-known parallelism and synchronization techniques. Distributed algorithms are usually simple to understand. Figure 3 shows an example of distributed approach. The architecture is fault-tolerant, however, the algorithms are not built to adapt themselves to new scenarios.
- *Random delays* [2] - MNs attempt to handover after a random delay. This procedure minimizes simultaneous handover attempts and is considered a subtype of the distributed algorithm approach. In Figure 4, we present an example of handover that happens in different instants of time for each MN. This approach avoids collision among MNs, distributing handover requests over time.
- *Reinforcement learning* [23] - It employs Artificial Intelligence (AI) techniques to make MNs learn about their surrounding environment as they make handover attempts. This approach does not require message exchange among users; they use the information received from other entities over time. Figure 5 presents this interaction. However, learning algorithms may cause performance issues due to the complex processing.

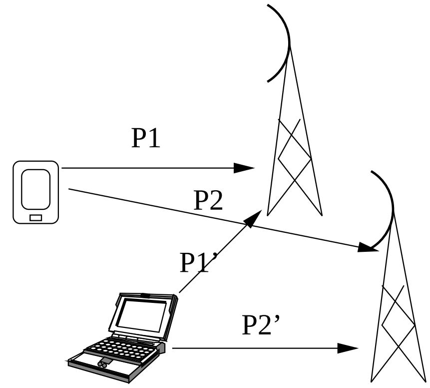

Figure 3. Distributed-algorithm-based approach.

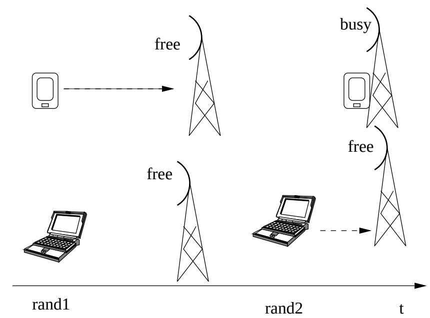

Figure 4. Random-delay-based approach.

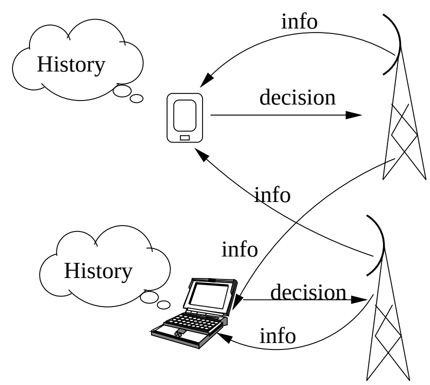

Figure 5. Reinforcement-learning-based approach.

- *Game theory* [2][23] - This approach maps handover scenarios in cooperative or non-cooperative games in which MNs are players interested in getting the best payoff as possible, shown in Figure 6. The payoff may be a larger bandwidth, energy saving, or better security. Nash equilibrium is the desired stable state in which all MNs do not have anymore strategies to obtain better payoffs. The main advantage of this approach is the almost perfect match between a GVHO scenario and the Game Theory competitive models. On the other hand, it is not always possible to model additional parameters.

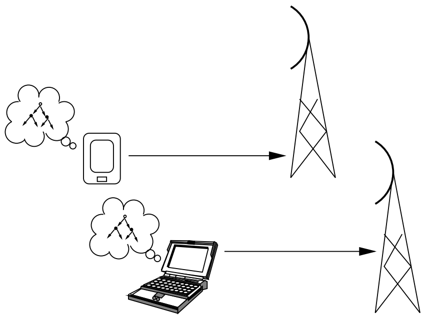

Figure 6. Game-theory-based approach.

- *Mathematical optimization problems* [3] - Mathematical equations are used to describe the handover decision under predetermined conditions. Figure 7 illustrates that approach. The optimization problem is solved by finding the ideal value for the equation variables. This approach requires a more complex modeling and is more flexible than Game Theory-based models.

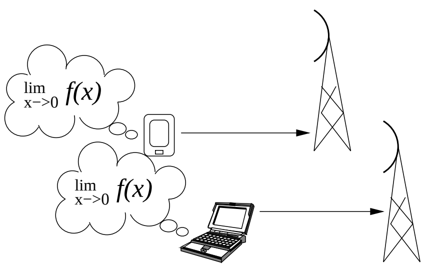

Figure 7. Optimization-problem-based approach.

For any GVHO approach, the MN or the serving PoA may determine if it is possible to request handover in a certain time, or if it is preferable to postpone it, given the network conditions. Policies for handover attempts can influence handover performance, for better or for worse, depending upon the adopted solution.

III. Related Work

In this section, we present a critical analysis of recent research related to GHO, VHO, and GVHO.

A. Group Handover

Chowdhury *et al.* [5] propose a resource management scheme using a dynamic bandwidth reservation policy in mobile femtocellular network deployment. The scheme aims at the vehicular scenario and uses the proximity of new stations and required QoS as information to allocate the corresponding bandwidth only when necessary. Simulation results show a reduction in the handover call drop probability and maintains bandwidth utilization when compared to schemes without QoS criteria and without priority.

Jeong *et al.* [4] propose a specific handover scheme for the IEEE 802.16e standard. It consists on reducing the number of packets necessary to accomplish handover, by means of a group-based channel scan. The MNs form groups and inside each group there is a handover schedule for the MNs. Computer simulations and Markov models were used to compare the scheme performance with the existing scheme in the IEEE 802.16e standard. The authors observed that the proposed scheme reduces blocking probability.

Fu *et al.* [6] highlight a group-based authentication scheme for WiMAX networks. That scheme consists on the PoA sending security context to a group of MNs if a member of this group requests handover. The main objective is to reduce handover latency while maintaining privacy preservation. The metrics evaluated are latency, communication overhead and computation cost. Simulation results show that the two earlier metrics are reduced, however, the latter is increased.

B. Vertical Handover

The state-of-art in VHO schemes can be found in [12]. The authors give more highlights on IEEE 802.16m and 3GPP LTE-Advanced technologies. According to the authors, IEEE 802.16m offers enhancements to link layer performance, such as multicarrier handover, in comparison to IEEE 802.16 legacy. The paper also presents the supported handover procedures besides *Hard Handover*, such as *Seamless Handover* and *Entry Before Break*.

In [7], Park *et al.* propose integration between WiMAX and cdma2000 networks. Their approach takes elements from tight coupling and loose coupling, introducing new messages in link layer and establishing tunnels in the IP layer. Simulations with OPNET measured delay in function of the elapsed time and the speed of nodes. Packet loss ratio is also measured in function of the elapsed time. According to the authors, those metrics are reduced in comparison to a loosely-coupled scheme.

Kim *et al.* [8] present a proposal for the *Hierarchical Mobile IPv6* (HMIPv6). They propose to execute IP-level handover and link-layer handover simultaneously, in order to reduce total time. Results show the reduction of latency and package loss in a intra-domain scenario in comparison to HMIPv6.

Shen *et al.* [9] propose a cost-function-based network selection. The authors consider available bandwidth information, traffic load and RSS. Simulation results show the scheme behavior in different scenarios. The authors conclude that the proposed scheme affects several system parameters, which need to be handled carefully.

Stevens-Navarro *et al.* [10] uses a Markov decision process for VHO having the maximization of the total expected reward per connection as objective. Performance evaluation considers voice and data applications. Numerical results show that the proposed algorithm performs better than the simple-additive-weighting algorithm.

Yeh *et al.* [11] propose the *Fast Intra-Network and Cross-layer Handover* (FINCH). It is a complementary mechanism to Mobile IPv4 for intra-domain mobility management. The main objective is to reduce latency at the IP layer. FINCH uses cross-layering techniques, which allows a more efficient localization and path optimization. The authors use numerical simulation to compare FINCH to Mobile IP, Fast Mobile IP, HMIP, Cellular IP, and HAWAII. The authors observe that FINCH reduces location cost and overall latency.

Gondi *et al.* [13] propose to use network context information during handover. These information include location, required bandwidth, battery status, available network interfaces, and authentication key. The authors give special attention to security in VHO. Experiments are run in a testbed to demonstrate how the proposal can be deployed.

Choi *et al.* [14] propose a new metric for VHO decision: *Interference to other Interferences-plus-Noise Ratio* (IINR). The main objective is to enhance throughput by analyzing the interference among cells in a cooperative fashion. The MN only handover to another cell if there are possibility of throughput gains. The authors use simulation to prove that the proposed scheme increases throughput at the scenarios studied in comparison to schemes that use *Signal to Interference-plus-Noise Ratios* (SINRs) as decision metric.

Koh *et al.* [15] study the fast handover in wireless multicast networks. According to the authors, the message calls to the IGMP protocol can be optimized when introducing *Multicast Handover Agents* (MHAs) at the base stations. Numerical simulations show a reduction of delay in the scenarios studied.

Kim *et al.* [16] propose a common link layer for 3G, WiMAX, and WiBRO networks. Additionally, three decision schemes are presented based on available bandwidth and cost employing neural networks. Simulations measure throughput, cost, and handover success rate and show better results than RSSI-based schemes.

The work in [17] concerns with the TCP throughput during handover in a FPMIPv6 network. The solution includes MIH to make QoS negotiations, preregister, and pre-authentication. Simulations with OPNET show that the proposal reduces TCP overhead in comparison to FPMIPv6.

Zekri *et al.* [18] present a survey on VHO solutions. It highlights the main technical challenges in heterogeneous wireless networks underlying seamless vertical handover. The authors also presents the standards involved and present comprehensively the mobility management process.

C. Group Vertical Handover

In [21], a relay station is used as a centralized entity to coordinate GVHO. The scenario studied is the movement of users in a train. Handover blocking and interruption probabilities are evaluated with the increase of the calls-per-minute ratio. The evaluation compares schemes with and without the relay station. The authors conclude that the proposed scheme reduces handover blocking and interruption probabilities. In this case, the relay station is responsible for executing the policy for handover attempts. The solution has limitations if co-existence with legacy systems is needed. This is due to the need of introducing a new infrastructure with special requirements.

Ning *et al.* [33] propose that a network entity called *Radio Resource Management Center* (RRMC) is responsible for collecting data from the nodes and network candidates. Thus, the RRMC decides which group of nodes may handover to a given network. The decision is based on Fuzzy Clustering, which is used to group nodes with similar characteristics. The policy for handover attempts is totally controlled by that entity. Results show that the solution reduces the blocking probability in comparison to [3], a decentralized scheme. However, it does not give results for the latency and does not make comparisons to another centralized scheme.

Cai *et al.* propose three decentralized algorithms for GVHO in [2]. The first is a Nash equilibrium-based algorithm where the policy for handover attempts is based on the game strategy of each player. The second algorithm adopts random delays, thus using a simpler policy for handover attempts. The third algorithm is a more refined version of the previous one. It considers latency as a basis for delay calculations. Performance evaluations show that latency values under the three algorithms are similar. Handover blocking probability is not considered.

Niyato *et al.* propose a model for network selection that is based on evolutionary games [23]. The model consider two approaches: a central entity-based approach and a decentralized-based approach that uses a reinforcement learning model. In the first approach, the central entity controls handover attempts. In the second approach, MNs are allowed to infer the best period of time to request a handover. The fraction of MNs choosing the same PoA is the load-balancing metric adopted. They conclude that each approach has its advantages in accordance with the scenario. One drawback is not evaluating the impact of the approaches on latency.

Lei *et al.* [22] present three GVHO schemes. The first scheme schedules simultaneous attempts to random time periods. In the second scheme, MNs select PoAs using a predefined probability as a base. In this case, the policy of handover attempts consists in an immediate attempt. The last scheme requires the network to be responsible for the handover decision. Results show that the last approach is more efficient. However, it may be difficult to adapt it to legacy systems.

Lee *et al.* [3] propose a GVHO scheme, which is based on the solution of an optimization problem. The MN is responsible for the handover decision. The main objective is to minimize latency while limiting the handover blocking probability. Some factors make the scheme in [3] more promising than the other researches:

- it does not require the presence of a relay station.
- it may work together with legacy systems.
- it considers two of the main GVHO metrics: load balancing and latency.

We detail such scheme in Section IV.

IV. Reference GVHO scheme

Lee *et al.* [3] propose an optimization for the total handover latency L, considering the handover blocking probability as follows:

$$\texttt{\textbf{Minimize}}\ L$$
$$\texttt{\textbf{Subject to}}\ P_{HoBlock}(t) \leq P_{HoBlockThreshold}\ ,$$

where $P_{HoBlock}(t)$ is the handover blocking probability in a time t and $P_{HoBlockThreshold}$ is the maximum acceptable value for the handover blocking probability. Latency is calculated as follows:

$$L = N_{HO}.\Delta t, \quad (1)$$

where N_{HO} is the total number of attempts until the MN requests the handover; Δt is the period of time between consecutive attempts. If the MN decides to request in the first attempt, total latency would be Δt. This is because in [3], execution time is also equal to Δt.

Equation (2) presents the calculation of $P_{HoBlock}(t)$. The value of $P_{HoBlock}(t)$ is dependent on the number of candidate networks, their available bandwidth, and the number of participating MNs in GVHO. In [3], it is considered that these values can be obtained by using IEEE 802.21 MIH (*Media Independent Handover*) queries and *ad hoc* communication.

$$P_{HoBlock}(t) = \sum_{k=1}^{K} \sum_{i=C_k(t)}^{M-1} \frac{(i+1-C_k(t)).(M-1)!}{(i+1)!.(M-1-i)!} \times ((P_{sel}^k)^{i+1}.(1-P_{sel}^k)^{M-1-i}), \quad (2)$$

Where:

- M represents the number of participating MNs.
- K represents the number of candidate networks with overlapping areas.
- $C_k(t)$ is the available bandwidth in a time *t* for a network$_k$. The model considers that the available bandwidth is represented by an integer value. Each MN requires one unity for handover;
- P_{sel}^k : The probability of selecting network$_k$.

The Karush-Kuhn-Tucker (KKT) condition is used in optimization problems and it can be applied to (2) to determine the P_{sel}^k value. However, P_{sel}^k can be obtained by using (3), which is simpler than using KKT and induces minor changes in results.

$$P_{sel}^k(t) = C_k(t) / \sum_{k=1}^{K} C_k(t). \quad (3)$$

Now, we can find the $M_{optimal}(t)$ value that ensures the optimization problem condition. This value can be found by setting it initially to one, then increasing it by one unit while the $P_{HoBlock}(t)$ value is still less than or equal to $P_{HoBlockThreshold}$. This procedure is described in Algorithm 1.

Algorithm 1: Find $M_{optimal}$ value

```
M_optimal = 0 ;
repeat
    p = Equation (2) ;
    M_optimal = M_optimal +1 ;
until p ≤ P_HoBlockThreshold;
```

The probability $P_{HO}(t)$ with which a MN can request handover is given by:

$$P_{HO}(t) = M_{optimal}(t)/M. \quad (4)$$

If the MN decides not to request the handover immediately, a new attempt will be made after a constant time interval. The MN requires the number of attempts necessary to have a well-succeeded handover with blocking probability less than or equal to $P_{HoBlockThreshold}$. Algorithm 2 summarizes this process and can also be found in [3].

Algorithm 2: Reference GVHO scheme

```
L = 0;
c_atts = 1;
Mtotal = number of GVHO participants;
Mremaining = Mtotal;
while Mremaining ≤ 0 do
    find Moptimal in function of (2);
    calculate PHO;
    if decision(PHO) then
        choose networkk depending on P^k_sel;
        NHO = c_atts;
        break ;
    else
        L += t_atts(c_atts);
        c_atts++;
    end
    Mremaining = Mremaining - Moptimal
end
L += LHOexec;
```

Where:

- Mtotal is the total number of MNs in GVHO.
- Mremaining is a counter that checks for the end of algorithm.
- `decision()` is a function that returns `true` with probability $P_{HO}(t)$.
- LHOexec is the handover execution time. It is equal to Δt.
- `t_atts()` is a function to calculate the period of time between consecutive attempts. In [3], the return value of this function is always Δt.
- `c_atts` counts the number of attempts. When `decision()` is `true` in the first attempt, the total execution latency is LHOexec.

Function `t_atts()` characterizes the policy for handover attempts. In this case, it is a function that returns a constant value and it is equals to the execution latency L_{HOexec}.

V. The Proposed Policy for GVHO Attempts

Despite of presenting a promising GVHO scheme, the work in [3] lacks a good policy for handover attempts. It is based on a constant delay, which causes a negative impact on the overall GVHO performance as the number of MNs increases. In this section, we present a policy for GVHO attempts that aims at providing reduced handover latency for GVHO schemes like the one proposed in [3]. At the same time, we intend to reduce the latency and the number of blocked nodes, maintaining the blocking probability premise.

In order to enhance performance results, we propose to modify the `t_atts()` function in Algorithm 2. Our proposed solution is exponential backoff-based. It depends upon the `c_atts` counter and the duration of a reference slot time. It is a particular case of random delay. Exponential backoff algorithms have the particularity of keeping the probability of collision and the probability of transmission stable as the number of nodes which are sharing a medium grows [34]. Although our solution is motivated by the performance issues in [3], it is generic enough to be applied in other schemes. Equation (5) shows our modified version of `t_atts()`:

$$t\_atts(c\_atts) = \begin{cases} \textit{random}[0..2^{c\_atts} - 1] \ . \ \textit{timeSlot} \ , \\ \qquad \textit{if } c\_atts \leq \textit{LimBackFactor} \\ \textit{random}[0..2^{LimBackFactor} - 1] \ . \ \textit{timeSlot} \ , \\ \qquad \textit{otherwise} \end{cases} \tag{5}$$

where $random$ picks a uniformly distributed number over the given interval; $LimBackFactor$ is the number of attempts that limits the range of values for $random$; and $timeSlot$ is the duration of a reference time slot, which depends on the target network. This information is obtained via MIH.

Total latency depends directly on the number of attempts, which varies with the return of `decision()`. The exponential backoff approach in `t_atts()` gives to the MN an opportunity for a new handover attempt after a time interval shorter than Δt, or even immediately. When the MN chooses not to request handover, other MNs may request it, reducing concurrency during the next attempts. Thus, MNs finish their handover sooner, decreasing total latency. Additionally, the number of nodes that have their handover blocked also decreases, making the handover more effective.

In [3], the return value of `t_atts()` is constant and equals to the execution latency L_{HOexec}. In that case, latency always grows by a constant factor. It causes a negative effect in the overall handover performance as the number of MN grows, as shown in [3]. Figure 8 presents the behavior of `t_atts()` in function of the number of attempts. We observe that the interval between attempts in (5) is always smaller than the approach in [3]. Analysis of the effect of the proposal in the overall performance is presented in the next section.

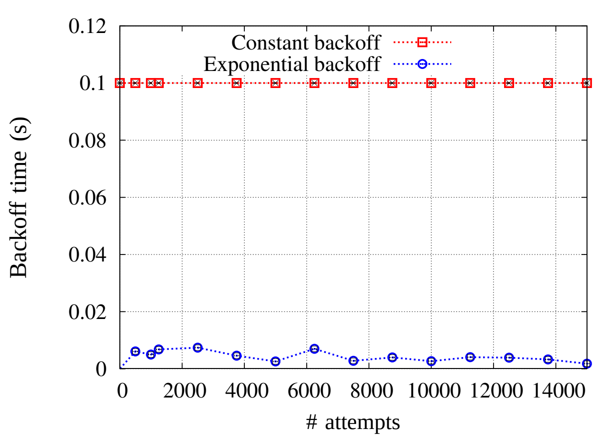

Figure 8. Comparison between the different implementations for the `t_atts()` function.

VI. Performance Evaluation and Comparison

In this section, we extend the experiments made in [1] introducing new parameter values and simulation conditions. The metrics evaluated in this paper are latency and handover blocking probability, as in [3] and [1], and additionally, the fraction of blocked MNs. The fraction of blocked MNs measures the fraction of nodes that decided to request handover and, for lack of bandwidth, had their request denied by the destination network. This metric helps us to evaluate the effectiveness of the GVHO scheme. All metrics are plotted in function of the number of MNs.

The majority of the parameters also follows the work in [3] and [1]. The value of Δt is set to 0.1s. We study scenarios with different values for $P_{HoBlockThreshold}$: 0.01, which is the recommended value according to Telecordia [35]; 0.02, which is a typical value [36][37]; and 0.05, in order to observe the effects of a less conservative parameter. The thresholds of 0.02 and 0.05 has been addressed in the former experiments [3][1] and the threshold of 0.01 is introduced in this paper. The number of MNs varies from 20 to 100, as in [1]. It differs from Lee *et al.* [3], where this number varies from 20 to 65.

In this paper, we introduce the Channel Holding Time (CHT) in the simulations. CHT is the time elapsed while a mobile node occupies a channel in a cell due to new connections or handover in an ongoing call [38]. In [1] and [3] is considered that all nodes leave network as soon as they handover to it. The introduction of the CHT factor give us a more realistic environment for analysis. The CHT modeling usually depends on the call holding time, the cell dimensions, cell residence time, resource allocation strategy, and the network architecture [39]. However, studies has shown that CHT can be approximated to a random variable with exponential distribution [38][39]. We consider 60s as the mean CHT. In other words, in our simulation, a set of nodes arrive, make handover attempts according to the policy adopted ; then, each one remain consuming a unit of bandwidth resource by a time defined by a random variable exponentially distributed with mean 60s.

We maintain the characterization of heterogeneity as the use of different available bandwidths to be compliant with the modeling presented in [3]. The number of available PoAs is 5, considering the following scenarios:

Scenario 1 - All PoAs have 20 bandwidth units.

Scenario 2 - Two PoAs have 15, two PoAs have 17, and one PoA has 20 bandwidth units, respectively.

Scenario 2 is only used in [3] for validating their simulator and in a situation of co-existing individual handover, which is out of the scope of this paper. Nevertheless, we include Scenario 2 in our evaluations. The $FatorLimBack$ parameter is set to 10. This value is based on preliminary experiments. We consider that MNs are switching from an arbitrary network to an IEEE 802.11 area. The parameter $timeSlot$ is set to $9.10^{-6}s$, which is equivalent to the SIFS time slot in IEEE 802.11 standard.

We have implemented the reference scheme and our solution in a discrete-event simulator, which was written in C++. Figure 9 illustrates how our scheduler operates in a given state, when *node 3* has its decision made. In this example, we have a queue of events, ordered by the scheduled time. Initially, we have *n* events of MNs trying to handover at the same time, in time *t = 0*. Then, each event is dequeued and processed according to the Algorithm 2. In this case, *node 3* decided to postpone handover request to time *t3*, enqueuing the corresponding event. In that state, *node 1* already had its event in *t = 0* processed and the decision to retry handover at time *t1* have been put at the queue. There is also another event for *node 2*, which decided to execute handover at time *t2*. Thus we can simulate parallelism in events, since bandwidth allocation will only happen in another event, when the node will in fact execute handover.

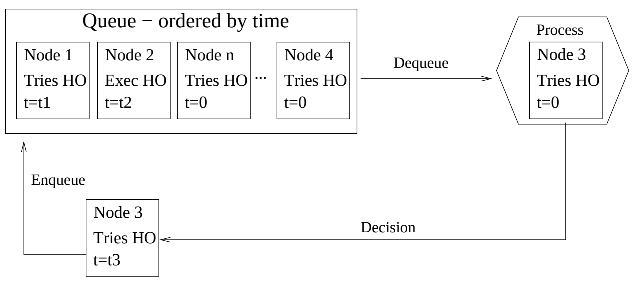

Figure 9. Example of scheduler instance in our discrete-event simulator.

The implementation of the reference scheme in our simulator was validated by the authors of [3]. We consider a group of MNs simultaneously entering a new coverage area and starting handover procedures defined by the GVHO scheme studied.

We represent confidence intervals with 99% of confidence level. Confidence intervals appear imperceptible in Figures 10-14. It is important to point out that we are not interested in evaluating the decision algorithm itself, but the impact of our policy for GVHO attempts on performance.

A. Results for Scenario 1

Figure 10 shows results for handover blocking probability under Scenario 1. The probability increases as the number of MNs grows from 55 for threshold 0.01, from 60 for threshold 0.02, and from 70 MNs for threshold 0.05. Thereafter, the curves are stable. This happens because blocking probability is getting closer to the threshold defined in the optimization problem. Since blocking probability is directly related to the cell utilization [40], it is necessary to limit the number of MNs entering a new cell at the same time in order to maintain the blocking probability under the threshold. When the blocking probability reaches the threshold, the value of $M_{optimal}(t)$ that is calculated in function of (2) can not increase anymore. This leads the remaining MNs to wait for another handover attempt. Thus, the stabilization of the blocking probability curve as the number of MN grows always implies the increase of the average latency. It is important to notice that the curves with and without our solution are similar because the optimization problem conditions are still the same. It means that the application of the proposed solution does not cause damages to the handover blocking probability, despite of the shorter time between attempts.

Figure 11 shows results for the fraction of blocked nodes under Scenario 1. We can observe that, as the number of

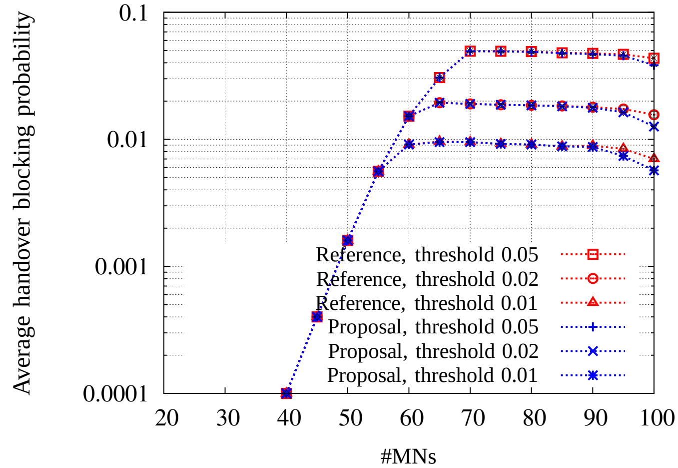

Figure 10. Handover blocking probability *versus* the number of MNs in Scenario 1.

MNs approximates to 100, the number of blocked MNs in the scheme in [3] is greater than the value found in the proposed solution. This is reflected in the blocking probability graph in Figure 10. For all thresholds, the blocking probability is slightly smaller when the number of nodes is between 95 and 100. It is due to the random nature of the attempts, which avoids collision among MNs.

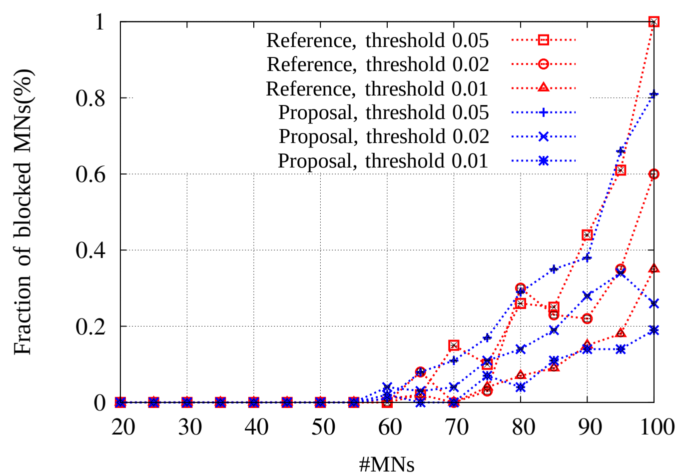

Figure 11. Fraction of blocked nodes *versus* the number of MNs in Scenario 1.

Figure 12 shows results for latency in Scenario 1. With respect to the scheme in [3], we can observe that latency starts growing from 55 MNs for threshold 0.01. For the threshold of 0.02, values start to grow at 60 MNs. Values in that curve are greater than those for threshold 0.05, which starts growing from 70 MNs. As we have stated before, the stabilization of the blocking probability curve observed in Figure 10 implies the increase of the average latency. Also, there is a greater number of handover attempts when we use a lower threshold. It tends to make MNs wait for more time with thresholds 0.01 and 0.02 than those using threshold 0.05. The lower the threshold is, the more conservative is the scheme and the greater is the average latency. We can also observe in Figure 12 the impact of the proposed solution on the latency curve. The curve is much smoother than the curve that does not adopt the solution.

For threshold 0.05, the latency is 11% smaller in the case of 80 MNs and 38% smaller for 100 MNs. For threshold 0.02, latency is 18% smaller for 80 MNs and 50% smaller for 100 MNs. Finally, for the threshold of 0.01, we observe a reduction of 22.5% for 80 MNs and 58% for 100 MNs.

The latency reduction is due to the proposed solution, which makes the delay between attempts more flexible. The exponential backoff also brought randomization to the scheme allowing MNs to try handover again sooner and in different periods of time, eventually reducing the total number of attempts.

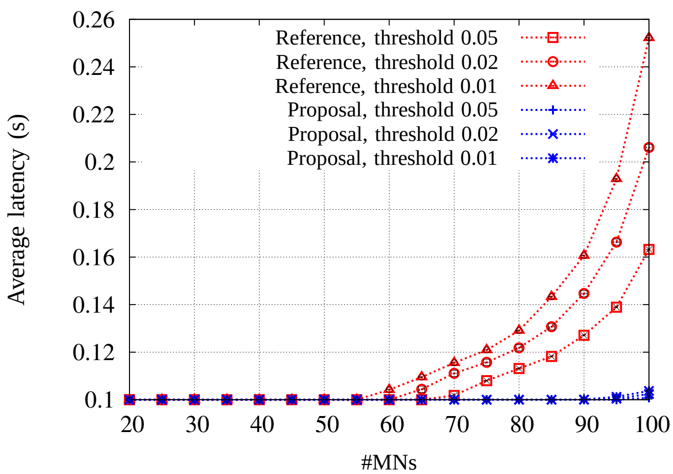

Figure 12. Latency *versus* the number of MNs in Scenario 1.

B. Results for Scenario 2

Figure 13 presents results for the Scenario 2. This figure presents similarities with Figure 10 but the curves stop growing sooner: from 50 MNs for the thresholds 0.01 and 0.02, and from 55 MNs for the threshold 0.05. This anticipation is due to the shorter total available bandwidth in the scenario studied. Thus, handover blocking probability increases faster, but it also gets stable in accordance with the established threshold.

However, we observe that the blocking probability starts to reduce again, from 85MNs. The explanation for this phenomenon is found in Figure 14. Figure 14 presents the fraction of blocked MNs for the Scenario 2. It is important to notice the expressive increase of the number of blocked nodes in the scheme in [3], which causes some nodes leave the concurrency because they were blocked. Thus, for the remaining nodes the blocking probability gets smaller. A similar phenomenon happens to the proposed solution, however, the number of blocked nodes is smaller, because the randomization of attempts makes handover requests less risky. We observe this behavior in all thresholds.

Figure 15 shows results for latency in Scenario 2. As in Scenario 1, the curves for thresholds 0.01 and 0.02 have greater latency values than the one with threshold 0.05. In [3], latency starts growing from 60 MNs for threshold 0.01, from

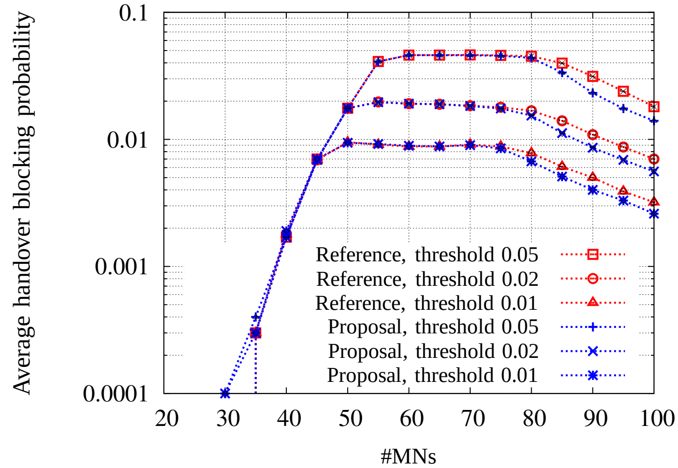

Figure 13. Handover blocking probability *versus* the number of MNs in Scenario 2.

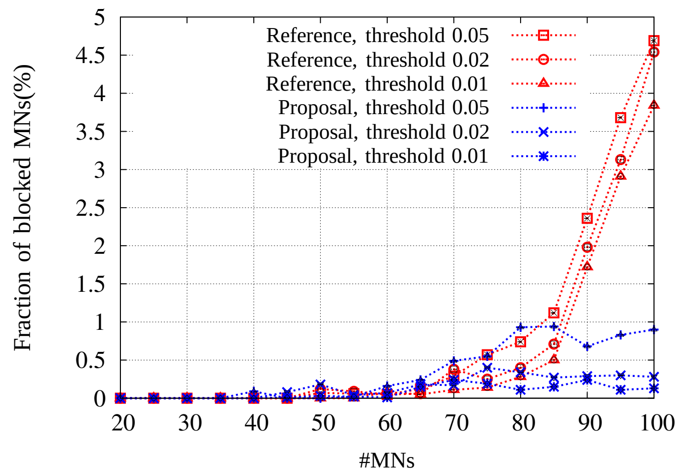

Figure 14. Fraction of blocked nodes *versus* the number of MNs in Scenario 2.

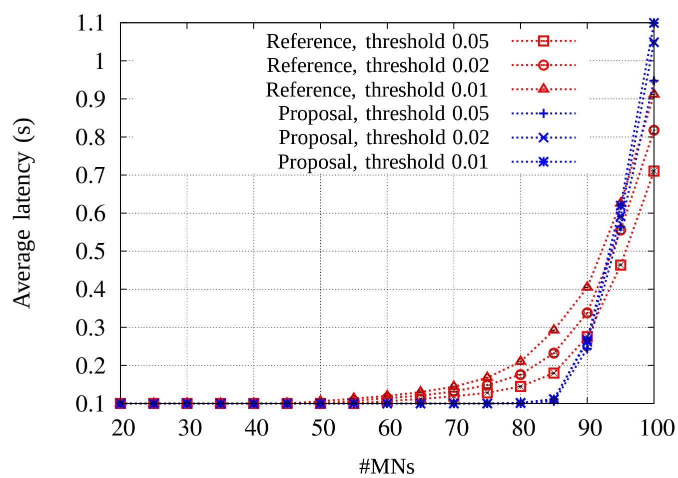

Figure 15. Latency *versus* the number of MNs in Scenario 2.

65 MNs for threshold 0.02, and from 70 MNs for threshold 0.05. Greater latency values are expected because the total available bandwidth is shorter than in Scenario 1.

Figure 15 also shows that for the threshold 0.05, latency has a reduction of 30% for 80 MNs. For the threshold 0.02, we observe a reduction of 42% for 80 MNs. In the threshold of 0.02, the latency is 51.5% smaller for 80 MNs. We also notice that from 95 MNs, our solution presents a greater latency than that one in [3]. The latency for 100MNs with our solution is 25%, 22%, and 17% greater than the scheme in [3] with the thresholds of 0.05, 0.02, and 0.01, respectively. It happens because, since our solution has a smaller percentage of blocked nodes, as shown in Figure 14. The remaining nodes, instead of being blocked as in [3], wait for more time for a handover opportunity and consequently, they increase the average latency.

VII. Conclusion and Future Work

In this paper, we have proposed a policy for GVHO attempts. Our solution uses exponential backoff in order to allow a better distribution of handover attempts over time. Performance evaluations have shown that our proposal makes it possible to reduce handover latency and the percentage of blocked nodes during handover. In particular, results have shown that latency was reduced up to 51.5% in accordance with the scenarios evaluated. Our future efforts will focus on including MIH queries in the solution design and including the information gathering phase in performance evaluation. Although this solution is well-suited to resolve performance issues in the scheme presented in [3], we are also interested in studying the impact of our solution on other GVHO schemes.

References

[1] N. C. Quental and P. A. S.Gonçalves, “A Policy for Group Vertical Handover Attempts,” in Proc. of the 13th International Conference on Networks (ICN), Nice, 2014, pp. 154–159.

[2] X. Cai and F. Liu, “Network Selection for Group Handover in Multi-access Networks,” in Proc. of the IEEE International Conference on Communications (ICC), Beijing, 2008, pp. 2164–2168.

[3] W. Lee and D. Cho, “Enhanced Group Handover Scheme in Multi-Access Networks,” IEEE Transactions on Vehicular Technology, vol. 60, no. 5, 2011, pp. 2389–2395.

[4] H. Jeong, J. Choi, H. Kang, and H. Youn, “An Efficient Group-Based Channel Scanning Scheme for Handover with IEEE 802.16e,” in Proc. of the 26th IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA), Fukuoka, 2012, pp. 639–644.

[5] M. Chowdhury, S. H. Chae, and Y. M. Jang, “Group Handover Management in Mobile Femtocellular Network Deployment,” in Proc. of the 4th International Conference on Ubiquitous and Future Networks (ICUFN), Phuket, 2012, pp. 162 – 165.

[6] A. Fu, S. Lan, B. Huang, Z. Zhu, and Y. Zhang, “A Novel Group-Based Handover Authentication Scheme with Privacy Preservation for Mobile WiMAX Networks,” IEEE Communications Letters, vol. 16, no. 11, 2012, pp. 1744–1747.

[7] S. Park, J. Yu, and J. Ihm, “A Performance Evaluation of Vertical Handoff Scheme between Mobile-WiMax and Cellular Networks,” in Proc. of 16th International Conference on Computer Communications and Networks (ICCCN), Honolulu, 2007, pp. 894 – 899.

[8] D. Kim, H. Shin, and D. Shin, “A network-based handover scheme for hierarchical mobile ipv6 over ieee 802.16 e,” in Proc. of the 10th International Conference on Advanced Communication Technology (ICACT), vol. 1. Gangwon-Do: IEEE, 2008, pp. 468–472.

[9] W. Shen and Q. Zeng, “Cost-function-based network selection strategy in integrated wireless and mobile networks,” IEEE Transactions on Vehicular Technology, vol. 57, no. 6, 2008, pp. 3778–3788.

[10] E. Stevens-Navarro, Y. Lin, and V. Wong, “An MDP-Based Vertical Handoff Decision Algorithm for Heterogeneous Wireless Networks,” IEEE Transactions on Vehicular Technology, vol. 57, no. 2, 2008, pp. 1243–1254.

[11] J. Yeh, J. Chen, and P. Agrawal, “Fast intra-network and cross-layer handover (finch) for wimax and mobile internet,” IEEE Transactions on Mobile Computing, vol. 8, no. 4, 2009, pp. 558–574.

[12] R. Y. Kim, I. Jung, X. Yang, and C.-C. Chou, “Advanced Handover Schemes in IMT-Advanced Systems,” IEEE Communications Magazine, vol. 48, no. 8, 2010, pp. 78–85.

[13] V. Gondi and N. Agoulmine, “Low latency handover and roaming using security context transfer for heterogeneous wireless and cellular networks,” in Proc. of the IEEE Asia-Pacific Services Computing Conference (APSCC). Hangzhou: IEEE, 2010, pp. 548–554.

[14] H. Choi, “An optimal handover decision for throughput enhancement,” IEEE Communications Letters, vol. 14, no. 9, 2010, pp. 851–853.

[15] S. Koh and M. Gohar, “Multicast Handover Agents for Fast Handover in Wireless Multicast Networks,” IEEE Communications Letters, vol. 14, no. 7, 2010, pp. 676–678.

[16] T.-S. Kim, R. Oh, S.-J. Lee, C.-H. Lim, S. Ryu, and C.-H. Cho, “Cell selection and trigger point decision for next generation heterogeneous wireless networking environment: Algorithm & evaluation,” in Proc. of the 5th International Conference on New Trends in Information Science and Service Science (NISS), Macao, 2011, pp. 247–254.

[17] I. Kim and Y. Kim, “Performance Evaluation and Improvement of TCP Throughput over PFMIPv6 with MIH,” in Proc. of the 12th IFIP/IEEE International Symposium on Integrated Network Management (IM), Dublin, 2011, pp. 997–1004.

[18] M. Zekri, B. Jouaber, and D. Zeghlache, “A Review on Mobility Management and Vertical Handover Solutions over Heterogeneous Wireless Networks,” Computer Communications, vol. 35, no. 17, 2012, pp. 2055–2068.

[19] K. Taniuchi, Y. Ohba, V. Fajardo, S. Das, M. Tauil, Y. Cheng, A. Dutta, D. Baker, M. Yajnik, and D. Famolari, “IEEE 802.21: Media Independent Handover: Features, Applicability, and Realization,” IEEE Communications Magazine, vol. 47, no. 1, 2009, pp. 112–120.

[20] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, “Proxy Mobile IPv6,” RFC 5213, aug 2008, [Accessed Nov. 11, 2014]. [Online]. Available: http://tools.ietf.org/html/rfc5213

[21] L. Shan, F. Liu, L. Wang, and Y. Ji, “Predictive Group Handover Scheme with Channel Borrowing for Mobile Relay Systems,” in Proc. of the International Wireless Communications and Mobile Computing Conference (IWCMC), Crete Island, 2008, pp. 153–158.

[22] S. Lei, T. Hui, and H. Zheng, “Group Vertical Handover in Heterogeneous Radio Access Networks,” in Proc. of the 72nd IEEE Vehicular Technology Conference Fall (VTC), Ottawa, 2010, pp. 1–5.

[23] D. Niyato and E. Hossain, “Dynamics of Network Selection in Heterogeneous Wireless Networks: an Evolutionary Game Approach,” IEEE Transactions on Vehicular Technology, vol. 58, no. 4, 2009, pp. 2008–2017.

[24] S. Lee, K. Sriram, K. Kim, Y. Kim, and N. Golmie, “Vertical Handoff Decision Algorithms for Providing Optimized Performance in Heterogeneous Wireless Networks,” IEEE Transactions on Vehicular Technology, vol. 58, no. 2, 2009, pp. 865–881.

[25] D. Johnson, C. Perkins, and J. Arkko, “Mobility Support in IPv6,” RFC 3775, june 2004, [Accessed Nov. 11, 2014]. [Online]. Available: http://tools.ietf.org/html/rfc3775

[26] E. R. Koodli, “Mobile IPv6 Fast Handovers,” RFC 5568, july 2009, [Accessed Nov. 11, 2014]. [Online]. Available: http://tools.ietf.org/html/rfc5568

[27] H. Yokota, K. Chowdhury, R. Koodli, B. Patil, and F. Xia, “Fast Handovers for Proxy Mobile IPv6,” RFC 5949, september 201, [Accessed Nov. 11, 2014]. [Online]. Available: http://tools.ietf.org/html/rfc5949

[28] K. P. C. Viho, A. Ksentini, and J.-M. Bonnin, “Resource Management in Mobile Heterogeneous Networks: State of the Art and Challenges,” INRIA, Tech. Rep., 2008.

[29] S. J. Bae, M. Y. Chung, and J. So, “Handover Triggering Mechanism Based on IEEE 802.21 in Heterogeneous Networks with LTE and WLAN,” in Proc. of the International Conference on Information Networking (ICOIN), Barcelona, 2011, pp. 399–403.

[30] M. Q. Khan and S. H. Andresen, “PoA Selection in 802.11 Networks Using Media Independent Information Server (MIIS),” in Proc. of the 26th International Conference on Advanced Information Networking and Applications Workshops (WAINA), Fukuoka, Japan, 2012, pp. 454–459.

[31] Y.-H. Liang, B.-J. Chang, and C.-T. Chen, “Media Independent Handover-based Competitive On-Line CAC for Seamless Mobile Wireless Networks,” Journal Wireless Personal Networks, 2011.

[32] R. Silva, P. Carvalho, P. Sousa, and P. Neves, “Heterogeneous Mobility in Next Generation Devices: An Android-Based Case Study,” Mobile Multimedia Communications, vol. 77, 2012, pp. 316–330.

[33] L. Ning, Z. Wang, Q. Guo, and K. Jiang, “Fuzzy Clustering based Group Vertical Handover Decision for Heterogeneous Wireless Networks,” in Proc. of the IEEE Wireless Communications and Networking Conference (WCNC), Changai, 2013, pp. 1231–1336.

[34] B. Kwak, N. Song, and L. Miller, “Performance Analysis of Exponential Backoff,” IEEE/ACM Transactions on Networking, vol. 13, no. 2, 2005, pp. 343–355.

[35] Transport Systems Generic Requirements (TSGR): Common Requirements. GR-499, Issue 2, Telecordia Report 2009.

[36] M. K. Karray, “Evaluation of the Blocking Probability and the Throughput in the Uplink of Wireless Cellular Networks,” in Proc. of the International Conference on Communications and Networking (ComNet), Tozeur, Tunisia, 2010, pp. 1–8.

[37] C. Chekuri, K. Ramanan, P. Whiting, and L. Zhang, “Blocking Probability Estimates in a Partitioned Sector TDMA System,” in Proc. of the 4th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIAL-M), Boston, USA, 2000, pp. 28–34.

[38] H. Khedher, F. Valois, and S. Tabbane, “Channel Holding Time Characterization in Real GSM Network,” in Proc. of the 14th International Symposium on Personal,Indoor and Mobile Radio Communication (PIMRC), Beijing, 2003, pp. 46–49.

[39] Y. Zhang and B.-H. Soong, “Channel Holding Time in Hierarchical Cellular Systems,” IEEE Communications Letters, vol. 8, no. 10, 2004, pp. 614–616.

[40] G. Haring, R. Marie, R. Puigjaner, and K. Trivedi, “Loss Formulas and Their Application to Optimization for Cellular Networks,” IEEE Transactions on Vehicular Technology, vol. 43, no. 3, 2001, pp. 664–673.