Exploring the Open Problems and Future Trends Concerning Handovers in Heterogeneous Wireless Networks: A Review

Bushra Naeem\textsuperscript{a,b}, Uche A. K. Chude-Okonkwo\textsuperscript{a}, Razali Ngah\textsuperscript{a}, Siti Z. Mohd Hashim\textsuperscript{c}, Wajahat Maqbool\textsuperscript{b,d}, Sadiqa Hashim\textsuperscript{b,d}

\textsuperscript{a}Wireless Communication Centre, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
\textsuperscript{b}Balochistan University of IT, Engineering and Management Sciences, Quetta, Pakistan
\textsuperscript{c}Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
\textsuperscript{d}UTM-MIMOS Centre of Excellence, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

Abstract

The advancements in the telecoms industry have led to the emergence of many new technologies to meet communication requirements worldwide. Consumers have largely been using GSM, CDMA and WiFi. New developments have given birth to more robust communication systems including WiMax and LTE. Lay users are not aware of the technicalities, advantages and disadvantages of these technologies. Therefore, there is a state of great confusion among the users of this generation as to which technology is the best for them to use. Each of these technologies has pros and cons over one another. For example, LTE covers a large area and offers the highest speed, but the user has to pay for it; however, WiFi covers only hotspots but the use is free. Future heterogeneous networks (HetNets) will enable users to achieve seamless connectivity through vertical handovers while providing them the beneficial features of each technology. For HetNets many issues still need to be addressed. This paper explores some of the open problems and future trends related to the handovers in heterogeneous wireless networks. The dynamics with respect to handovers and the network load of cellular architecture in the coming era in the field of wireless communications are also presented. In addition, first a foundation is laid on the IEEE media independent handover and its importance and then the problem of load balancing and its impact on the realization of future HetNets is explained. Finally, an important problem during vertical handovers, known as the ping-pong effect is explained. The open problems for researchers have been identified for the aforementioned issues.

Keywords: Heterogeneous Networks (HetNets); seamless connectivity; vertical handover; load balancing; ping-pong effect

© 2014 Penerbit UTM Press. All rights reserved.

1.0 INTRODUCTION

Future wireless communication networks are seen as the integration of different radio access network (RAN) technologies to achieve best connectivity. Seamless vertical handovers will provide better quality of service (QoS), lower cost, and freedom of movement to the users and will provide service providers with sufficient capacity to support varying user demands [1]. In the 4G wireless environment, mobile users are able to keep connections active while moving from one access point to another [2]. The next generation wireless networks will be based on an infrastructure with the support of heterogeneous access technologies. As the user will be mobile between different networks, there is a chance that a certain hotspot does not have the capacity to sustain the need of users [3]. Moreover, picocells and femtocells, along with conventional macrocells, are being added to the existing cellular architecture. This means that a user can be given bandwidth through heterogeneous cells. In the past few years, many researchers, such as [5–9], have been working on the development of multimode terminals, which consist of a heterogeneous set of transceivers to communicate in different technologies. However, such devices lead to redundant hardware in mobile devices which is not desirable. The concept of vertical handover is studied more widely and expected to enable users to perform a handover to a neighbouring heterogeneous network (HetNet) cell when required. The ability to shift current access links to acquire better QoS in dissimilar wireless communication networks is called seamless vertical handover [10]. Thus, vertical handovers are expected to be the essential mechanism to take advantage of all of the wireless access technologies. In addition, the increasing demand of users for both uplink and downlink communications calls for integration of these heterogeneous technologies. Future generation networks are therefore seen as HetNets. The implementation of such networks faces many concerns which...
need to be addressed. In this paper, we report some of the concerns faced in the handover and implementation of the future generation networks and we present the open problems for the researchers. In the first section, a vision of the IEEE 802.21 Media Independent Handover standard is given, which lays the background for the need for optimized QoS in the HetNet environment. Second, the problem of load balancing in the future generation of networks is explained and the current techniques and the open problem with reference to this problem are also discussed. Third, another important problem in the implementation of HetNets, known as the ping-pong effect is presented. The current techniques to solve this effect and open problems that still remain are presented at the end.

## 2. IEEE STANDARD 802.21(MEDIA INDEPENDENT HANOVER)

In recent years there has been successful deployment of various wireless networks. The notion of HetNets provides users with various options for choosing the best wireless network that fits their requirements. This also provides the network operators with a chance to increase their revenue with more efficient use of radio resources [4]. The latest trends in research and advancement show that much work is being done on the development of multimode terminals and software defined radio (SDR) [5–9].

The capability of changing the current access link to achieve better QoS in different wireless communication networks is known as seamless vertical handover [10]. In order to achieve such handover, IEEE has proposed standard 802.21 on Media Independent Handover (MIH). The standard is put forward to allow the exchange of entities of different access networks and to assist the handover decisions by defining the set of functional components to be executed. The IEEE 802.21 framework features handover in such a way that handover is not dependent on a particular access network. The mechanism of MIH evidently has the capability to indicate when the link is going down or approaching a break in communication [11–13]. However, the standard defines the overall framework only and the actual implementation of the algorithms is left to the designers [14].

Open problem: It follows from the above literature that efficient protocols for multimode mobile terminals need to be developed to support the required QoS levels.

## 3.0 LOAD BALANCING

One of the open problems in heterogeneous wireless networks is resource management and load balancing. Load balancing is the method of accepting or rejecting a new user request to come to a base station (BS) and forcing users connected to a heavily loaded BS to handover to a lightly loaded one [15].

The authors of [16] propose a semi-distributed and mobile based architecture for load balancing in wireless HetNets. They evaluate and analyse signal overhead and accumulated time as the parameters for distributed design. The basic grid is created with a number of hexagonal cells. The main advantage of this architecture is that it is free from single point of failure [16]. Load balancing is one of the most important techniques in HetNets which can increase their capacity and QoS. The authors of [17] proposed a fuzzy-based radio access technology selection technique to realize load balancing between 3G and WLAN. They set up a simulation model using MATLAB. The simulation parameters used to assess performance were: blocking probability of voice and data traffic, throughput (capacity in bit/s that each network can provide), and network utilization fairness (difference between the amount of resources, i.e. bandwidth used in two networks) [17]. An optimal inter-access system anchor (ISA) scheme for load balancing in heterogeneous wireless networks is presented in [18]. Their proposed mechanism undertakes load status monitoring and evaluation for access gateways. The numerical results show the network utility changing with the number of users in it and that the optimal handoff user number can be obtained corresponding to the maximum of the joint network utility [18]. Wei and Lei in their load balancing algorithm for a heterogeneous biswapped network [19] combine diffusion and dimension exchange where the process of load balancing is divided into four stages. In the first stage the node load is divided until the load vector in the local group achieves a balanced status. In the second stage, a dimension exchange strategy is applied over all inter-group links. Third, the load on each group is again diffused on each of the groups with the same iterative polynomial, similar to stage 1, by flow scheduling on the intra-group edges of each group. In the last stage, first the loads of the two nodes with the inter-group link are summed, and then this sum is diffused to each node as in stage two. This is how they achieve the final load balancing for biswapped networks. Son and Lee [20] present a load balancing mechanism named “soft” load balancing. In this mechanism, the Internet protocol (IP) packets in the downlink are delivered via two heterogeneous access networks, i.e., they are divided into sub-flows according to their air interfaces. Whenever a subscriber is able to work under multiple access networks, the mobile station (MS) is equipped with multiple interfaces to support both access networks. They obtain an optimal load balancing ratio (LBR) to determine the volume of traffic delivered to each network in an overlaid multi-cell environment [20]. The authors of [3] proposed a load balancing scheme for LTE systems that can control handover time based on load status of a cell. Their algorithm acquired the load status of a cell and avoided handover to a congested cell. This increases the network capacity [3]. A recent work proposed a load balancing scheme that minimizes the VHO rate while achieving low call blocking probabilities. The proposed scheme is analysed via user level Markov chains. One drawback of their technique is that, to find the user speed, they only consider the doppler shift to estimate the speed. However, doppler shift may be present due to other reasons, leading to the wrong speed detection [21]. A load balancing algorithm based on entropy weight and grey relation analysis is proposed in [22]. The entropy weight method is used in conjunction with the correlation matrix to evaluate the network performance of each HetNet. Their simulation results show that this algorithm performed better load balancing than traditional algorithms and has reduced the system blocking rate [22].

As we are moving towards the paradigm shift towards HetNets we see that a great number of picocells and femtocells are being added to traditional cellular architecture [23–25]. According to Andrews et al. [25], there will be 50 million BSs by 2015. Some researchers also predict that there will be a greater number of BSs than the number of cell phone subscribers in the coming decade [26]. Each of the HetNet cells has different physical properties, such as cell size, number of users that can be accommodated by each cell, and different transmission power. Even though these cells will geographically cover the earth, and all subscribers, the question remains what will be the criteria of serving users from heterogeneous cells? Thus, the current challenge faced by the operators in the near future is expected to be capacity, not coverage. This is depicted in Figure 1.
Open problem: Concluding from the above literature, we perceive the criteria of serving/allocating resources to the users among heterogeneous cells is still an open problem to be addressed.

4.0 PING-PONG PHENOMENON

The ping-pong effect is when a mobile terminal executes handover between two cells back and forth due to natural fluctuations in radio measurements. This happens when a user is mobile or standing at the intersection area of two or more cells. Figure 2 below depicts the area (with grey circles) where the ping-pong effect in intersecting cells takes place.

In future, wireless networks will adopt a micro/pico cellular architecture which will unsurprisingly increase the number of handoffs an MS will make, as the cell size will be smaller [27, 28]. This will make the impact of the ping-pong effect more significant in future HetNets.

Ping pong (repeated handover) may lead to unnecessary handovers which can cause severe degradation in HetNets. These repeated and undue handovers from one access point to another are caused by incorrect decisions. The ping-pong effect may result in loss of packets, loss of energy consumption, service interruption, reduced terminal performance and increased load on the networks [29].

4.1 Ping-Pong Mitigation Techniques

The ping-pong mitigation methods present in the literature can be divided into two broad categories; namely, timing based techniques [22], [30], [31], [37] and [38] and intelligence-based techniques [29], [32–35] and [36].

4.1.1 Timing Based Techniques

Feher et al. reported a ping-pong reduction scheme using sub-cell movement detection [30]. In the reported scheme, whether 1) the mobile terminal (MT) under consideration is stationary or mobile, and 2) the moving terminals are detected using a timing based algorithm, were identified. The user equipment (UE) calculates the difference between the arrival times of system frames from neighbouring cells compared to its inner clock. According to the classification of ping pongs, the decision-rule engine decides either, to invoke a self-optimizing network, perform terminal oriented handover tuning or perform any other action.

Won et al. [31] proposed a vertical handover scheme to alleviate unnecessary handover between 3G cellular and WLAN hotspots. In this study, the ping-pong avoidance algorithm consists of two major schemes. In the first part the single vertical handover threshold is separated, and in the second part the transition of the received signal strength (RSS) by the terminal from the WLAN hotspots based on the least square approach is estimated. This timing approach is used in such a way that the timer is set and when it has expired, the least square is computed.

In another study the handover failure and ping-pong probabilities are analysed. They conclude that the trigger time should, ideally, be both cell specific and UE specific. However, they do not incorporate shadowing and fast fading in the model [22].

One more technique is presented in [37], where a timer is set in a first step to select whether the handover is ping pong or general; if ping-pong handover is detected, then in the next step the handover completion is delayed and the old path between two LTE nodes is kept for a short time.

A semi-active handover algorithm is presented in [38] which is based on the location of WiFi and cellular networks where WiFi is considered to be an attachment network. The speed and direction of the MT using the least square fit model is estimated. The shadow and fast fading model is applied to their RSS and the base station predicts the time during which the MS stays in WiFi. The base station acquires location information from the WiFi APs and a detection algorithm is applied. The dissimilarity found in handover modes restrains the ping-pong effect and reduces the probability of handover failure.

The techniques developed by Feher et al. [30] and Kim et al. [31] are based on timing approaches using different methods: the former compares frame timing, the later scans and collects
RSS for a specific time. However, they do not consider the basic essence of the ping-pong effect, fluctuations in radio measurements and slow and fast fading, which makes the RSS and frame information unpredictable. They do not incorporate any mechanism to deal with the uncertain behaviour of MT signals at a given instant. In [22] no mechanism, such as shadowing and fast fading to deal with uncertain behaviour of MT signal at a given instant, is incorporated. A disadvantage of the technique in [37] is that high timer values can cause dropped calls, moreover maintaining the old path may keep another user from connecting as two paths will be occupied by this user.

4.1.2 Intelligence-based Techniques

Lahby et al. developed an intelligent network selection approach in which they merge two multi-attribute decision making (MADM) methods [33]. In the proposed scheme, two methods are combined, i.e., analytic network process (ANP) and the technique for order preference by similarity to an ideal solution (TOPSIS). ANP is used to find different weights of available networks using individual criteria and TOPSIS is used to rank the available networks. The combined MADM algorithm aims to address the limitations of the ranking abnormality of various networks. It is concluded that the reported technique helps ping-pong occurrence by reducing the number of handoffs, hence less chance of having the ping-pong effect in handover.

Some researchers presented a handoff decision scheme to choose the correct network in a heterogeneous multi-access environment [32]. The RSS is predicted using the LMS algorithm. The predicted RSS, user preference and bandwidth are the input parameters that are passed through a fuzzifier and then fed into the fuzzy inference engine where a set of fuzzy (IF-THEN) rules is applied to obtain fuzzy decision sets. They use the handoff factor to make the final decision; the factor ranges from 0 to 1, in which the higher the value of the handoff factor, the higher the possibility of handoff.

Balarengadurai and Saraswathi [29] proposed a fuzzy logic system for detecting the ping-pong effect attack in low rate wireless personal area networks (WPANs) with a self-reconfiguring protocol for power efficiency. Self-configuring refers to a network protocol that can update its links to maintain strong connectivity with the nodes. The low rate WPAN is self-organized into clusters using an unsupervised clustering method; it applies a fuzzy logic system to master/controller selection for each cluster [29].

Singhrova and Prakash [34] proposed a multiple-parameter-based vertical handover algorithm that uses neuro fuzzy logic, where a supervised learning approach is used for the tuning of the training vector. The input parameters, available bandwidth, speed of MT, number of users, RSS, battery level and coverage area, are fuzzified using the triangular membership function. The fuzzy inference engine then decides on the handover from WLAN to UMTS or vice versa. However, the same rules are not applied in both cases; their algorithm reduces the ping-pong effect hence reducing the number of handoffs in a given time.

A fuzzy-based handover system for avoiding the ping-pong effect in wireless cellular networks is introduced by [35]. They incorporate the Monte Carlo (MC) method to analyse the uncertainty in propagation to find out how the random variations or imprecise data can affect the performance/reliability in the system under consideration [36]. Their method deals with the uncertainty in the handoff process. However, they do not consider all handoff parameters, and only use three parameters, namely, RSS of current AP, neighbouring AP and the distance between the two.

An artificial-neural-network-based handoff decision algorithm to reduce handoff latency is proposed in [36]. A three-layer neural network approach is used and five neural networks for separate technologies including WiFi, GSM, GPRS, UMTS and WiMAX are constructed. The three parameters, i.e., monetary cost, data rate and received signal strength indicator (RSSI), are fed as inputs into their neural networks. Their method outperformed most of the traditional handoff decision methods used in terms of handoff latency and selection of the best network.

It is difficult for handoff algorithms operating in real time to make precise decisions because of the large- and small-scale fading encountered in the mobile environment; for this reason, the fuzzy logic approach is used as it can deal with imprecise or difficult to obtain data [35]. Artificial intelligence and neural networks have been widely used to solve the problem of vertical handovers [38]. These approaches are suitable for pattern classification because of their generalization capability [39]. Neural networks and fuzzy logic overcome the uncertainty and fluctuations of the radio environment by automatic adaptation; that is why they are used as a suitable option for such problems [40]. The neural network is an efficient technique that can respond with an output even though it has never before seen the inputs [41].

Open problem: It follows from the available literature that some researchers realize the problem of the ping-pong effect as an erratic challenge and used artificial intelligence techniques to solve it, whereas other researchers see this problem as a problem that arises because of latency or inappropriate handover execution time. It is now a matter of investigation which technique will best solve this problem for providing seamless connectivity in HetNets.

5.0 CONCLUSIONS

In this paper, some of the issues in the realization and implementation of HetNets were discussed. The open problems in this field were found to assist researchers and industry who are struggling to achieve the paradigm shift towards the future HetNet environment. It was concluded that there is a need to optimize QoS for users in such network environments. The need for new load balancing techniques to cope up with the capacity requirement of users was explained. Lastly, it was established that the ping-pong effect in future HetNets can lead to serious degradation in link quality, especially for mobile users. This study opens new avenues for prospective researchers to further explore the dynamics of future generation wireless networks, currently one of the hottest research topics in wireless communications. This work can provide researchers with an insight into the current challenges faced in the deployment of HetNets so that they can explore further how seamless connectivity can be achieved in future generation wireless networks.

Acknowledgement

The authors would like to thank the Universiti Teknologi Malaysia for supporting the research study.

References

Wireless Networks. EURASIP Journal on Wireless Communications and Networking. 1–17.


