Dynamic and User-Centric Network Selection in Heterogeneous Networks

Xuejun Cai, Ling Chen
Corporate Technology
Siemens Ltd., China
{xuejun.cai, chenling}@siemens.com

Rute Sofia
Corporate Technology
Information and Communications
Siemens AG, Germany
rute.sofia@siemens.com

Yanqi Wu
Beijing Jiaotong University
China
04120242@center.bjtu.edu.cn

Abstract

The use of multiple interfaces to access heterogeneous networks is becoming a strong reality to end-users. Hence one realistic problem is how to select a specific access interface (and consequently network) as well as how to perform smooth and seamless handover among different types of technologies.
In this paper we propose a dynamic and user-centric network selection and decision process which optimizes handover across heterogeneous networks. A Satisfaction Degree Function (SDF) *is used to evaluate, according to user's predefined criteria, available networks and select the best one(s) according to such criteria. The criteria consider incorporating user policies and information from several OSI Layers, including dynamic network status and application requirements. Numerical results show that the proposed network selection process results in the choice of the best network according to the user's choices.*

1 introduction

In recent years, the widespread success of wireless and mobile communications has resulted in the development of a large variety of mobile and wireless technologies, including 2.5G and 3G cellular, satellite, WLAN, WiMAX and Bluetooth. Each technology is tailored to reach a particular market, or a particular type of user with a specific service need. Such diversity offers different choices in terms of bandwidth, security, and the coverage area for the mobile user. For instance, WLANs offer mobility within limited scope and reasonable rates, whereas cellular networks can offer universal network access but with limited bandwidth and high costs.

Moreover, advanced multi-mode mobile terminals enables the users to access information and services from anywhere, at any time over any available network, i.e., the connectivity is 'always on' and seamless. To achieve the seamless handover between heterogeneous networks, there is a basic challenge, namely, the selection of the best network to access both from a network and an end-user's perspective. From the network perspective, the different access network may have different mobility, QoS and/or security characteristics. From the end-user's perspective, different users may have different preference. Some users prefer cheaper access networks, while others may want the network to support high rates. So an optimized network selection and handover decision mechanism is necessary for the users to choose an optimal network that can satisfy their needs.

Within such context we propose a novel method to perform network selection and to optimize handovers in heterogeneous network scenarios based both on network performance parameters (such as bandwidth efficiency), application requirements and end-user's policies.

The remainder of the paper is organized as follows. Section 2 briefly introduces related work. Section 3 describes the network model used in our work. Section 4 then presents the proposed dynamic network selection algorithms for performing vertical handover in overlay networks. Section 5 evaluates the performance of the proposed approach based on the simulation results. The conclusion is given in Section 6.

1-4244-1338-6/07/$25.00/©2007 IEEE

2 Related Works

Handovers performed within the same type of network technology (*horizontal handovers*) often rely on the *Received Signal Strength (RSS)* based algorithms [6] [9]. For *vertical handover* across heterogeneous networks, several handover decision methods are proposed to extend the traditional RSS based methods. In [8], Zahran et al. propose to use an adaptive lifetime-based handover decision algorithm in which the application specific signal strength threshold is introduced. In [2], which relates to an UMTS-WLAN environment, the handover decision is based on both RSS and distance criterion. Though both these two related works provide improvements, they assume that WLAN is the preferred technology. The WLAN is then always selected unless it is considered unusable based on their proposed algorithms. However if the available networks are considered equally important, the mentioned RSS-based methods are not suitable any longer given that the RSS varies significant from different networks due to the different adopted physical technologies. Furthermore, these methods don't consider users policies or preference. Additionally to the RSS, the monetary cost, available bandwidth and other factors may also affect the user's selection.

Related to user's preferences and requirements, several policy based mechanisms have been proposed [1, 3, 7, 10]. In [7], Wang et al. propose a policy-enabled handover decision system in which a cost-based function is used to reflect the user's preference on various network parameters such as cost, performance and power consumption. In [10], a cost-based function is also used in the handover decision algorithms, and a network elimination factor is introduced to exclude those networks that can't meet the *Quality of Service* (QoS) constraints for specific services. In both works, the network selection and handover decision is made based on the cost function which reflects the user's policy and characteristics of different networks. In [1], the authors propose an intelligent Access Selection mechanism which considers the signal quality, charging cost, user's preferences on operator and technology. Though these methods consider user preferences and cost, there are still some limitations to these works. Specifically, the used cost function is considered static rather than dynamic: the network parameters don't change when the mobile terminal is roaming across heterogeneous networks. However, in reality some of the used network parameters would change over time. For instance, in best effort networks the resources are shared among a number of users. Furthermore, the dynamic requirements of certain applications are not considered: there may be times where low bandwidth level can properly fulfill the need of the mobile terminal if no bandwidth consuming applications (such as bulk file transfer) are active in the terminal. While other times the network condition is rather good, but

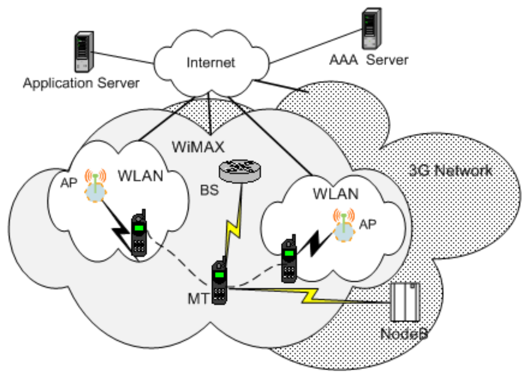


Figure 1. Reference Network Model.

the user's application may still be too bandwidth demanding, exceeding what can be provided by the network. The requirements can then only be properly satisfied by switching to another network.

To overcome the shortcomings mentioned above, we propose a dynamic and user centric network selection process, which utilizes a *Satisfaction Degree Function* to evaluate the available networks and select the best network to be used, from an end-user's satisfaction perspective. The network selection utilizes user-defined policies and cross-layer information including physical, link and application layer.

3 Network Model

To better explain our work, we rely on a generic heterogeneous overlay network composed of several wireless access networks for the explanations provided in the next sections. The mobile terminal is able to access all of these networks with the equipped multi-mode interfaces.

In this work, we rely upon the *loosely coupled* architecture illustrated in Figure 1, where each wireless network is deployed as an access network that is independent from the other networks. In other words, each network provide direct and independent data access to the Internet. The loosely coupled architecture provides therefore a flexible and independent environment for different wireless technologies.

Within such architecture, usually there are three strategies capable of detecting the need for handover [5]: *Mobile-controlled HandOver* (MCHO); *Network-Controlled HandOver* (NCHO), and *Mobile-Assisted HandOver* (MAHO). Under NCHO or MAHO, the network is responsible for the decision for handover, while under MCHO, the mobile terminal makes the handover decision on its own. Among these strategies, MCHO is the one that adjusts better to our work due to several reasons. First and foremost is the low complexity in network equipment. Second, in heterogeneous wireless scenarios, only mobile terminals have the complete perspective on the kind of interfaces they are equipped with. Even if the network is aware of the mobile terminal's interfaces, there may be no way to control another network that the mobile terminal is about to handover to. So in our work MCHO is adopted, i.e. the network selection and handover decision is made by the terminal. The detail of the proposed network selection process is explained in the next section.

4 Dynamic Network Selection Process

Our network selection process is defined as a three-phase process to find the network that can fulfill user's requirements the best. The three phases that compose network selection are: *Network Detection*, *Network Evaluation*, and *Handover Execution* (for both OSI Layer 2 & 3). Network Detection is used to discover available networks and collect their status. The next step is to use the collected information as input to evaluate the available networks and to select the network capable of satisfying best the user's request at an instance in time. We name such network *Always Best Satisfying (ABS)* network. The ABS network provides not only always-on connectivity, but also gives the user the best service according to his preference at any time and any place. Here, a *Satisfaction Degree Function* is used to evaluate the available networks. And the final step is to execute and finish the global handover where both link and network layers are involved.

4.1 Network Condition Detection

When mobile terminals roam among heterogeneous networks, it is important for them to know the availability and the conditions of each network before they make the network selection. The network conditions are classified into two categories: static and dynamic conditions. Static conditions don't change while the mobile terminal is in the same network. Cost and security level associated with a network are such conditions. Static conditions may be obtained by means of, e.g. a AAA server when the mobile terminal is connected to a specific network. Dynamic conditions, on the other hand, may change while a mobile terminal is roaming in the network. Such conditions can only be measured in real time and examples are available bandwidth, signal strength. Because we hope to present an ABS network to the mobile user, the dynamic conditions are also necessary for the network selection. We accomplish this by means of Satisfaction Degree Function which is introduced next.

4.2 Satisfaction Degree Function

Based on the collected information (OSI Layers 2 and 3 information as well as user policies), a dynamic and user centric *Satisfaction Degree Function* (SDF) is used to evaluate each available network and find the one that can satisfy the mobile user the best - the ABS network. SDF considers not only the specific network conditions, but also user defined policies and dynamic requirements of active applications. [1] Here we suppose that the network selection is made based on the factors that the mobile user cares about. These factors describe the characteristic of each available network, e.g., cost, bandwidth. The *satisfaction degree* is a value provided by the SDF and is calculated based on cross-layer information including user policies, application specific requirements and dynamic network status. It is denoted as:

$$f = \sum_{i=1}^{N} W_i \cdot S_i,\ i = 1, 2, 3, \ldots, N \,. \tag{1}$$

Where N is the total number of factors. Let W_i represent the weight assigned to the ith factor that affects the user's satisfaction. A higher weight of the ith factor means that the user considers this factor is more relevant. And the sum of all the weights must be equal to 1.

$$\sum_{i=1}^{N} W_i = 1$$

For example, if a mobile user only cares the network access cost, then the weight of the access cost would be set to 1 and all other weights would be set to 0. Let S_i denote the satisfaction degree for factor i, $0 < S_i < 1$. It can be stated as a *Satisfaction Function* of $S_i(V_i, E_i)$ where V_i represents the instantaneous measured value at the specific time and E_i represents the expected value of factor i. E_i is either provided by the end-user or defined by active applications. For each factor, there is a corresponding satisfaction function which may be different from others. For example, for available bandwidth, assuming the end-user would prefer high bandwidth, one simple satisfaction function could be:

$$S_b = \min(\frac{V_b}{E_b}, 1) \tag{2}$$

and for access cost, assuming the end-user would prefer the lowest access cost, one possible satisfaction function could be:

$$S_c = \min(\frac{E_c}{V_c}, 1) \tag{3}$$

The satisfaction function may be defined by the mobile user itself, or defined by service providers. The SDF therefore

[1] While we provide the end-user perspective here because we consider that such is the best perspective in what concerns handovers decisions, the function can be easily applied to the perspective of operators, e.g., by means of keeping agreed user policies beforehand and keeping them somewhere in the network.

provides a global view on the defined factors can then be written as:

$$f = \sum_{i=1}^{N} W_i \cdot S_i(V_i, E_i),\ i = 1, 2, 3, \ldots, N\ . \tag{4}$$

We further distinguish two categories of factors: *application-dependent* and *application-independent* factors. For application-dependent factors the expected values are determined by specific applications that the user runs. For instance, some multimedia applications may have higher bandwidth demand while other applications such as Email, Instant Messaging have lower requirement. The expected value for application-dependent factors is a function of the requirements of all active applications, which can be stated as *Requirements Function*:

$$E_i = r(E_{ij}),\ j = 1, 2, 3, \ldots, M\ . \tag{5}$$

Where M is the total number of the active applications that have expectation on factor i, and E_{ij} represents the expected value of application j for factor i. Taking bandwidth as an example, the requirement function could be stated as $E_b = \sum_{j=1}^{M} E_{ij}$. For 'security' factor, the function could be stated as $E_s = \max(E_{ij})$.

On the other hand, for application-independent factors the expect value has no relevance to specific applications. It can either be provided by the mobile user or a pre-defined default value is adopted instead. For example, the access cost is an application-independent factor.

The previous equations corroborate that the satisfaction degree depends on the weight, expected value and actual measured value for all factors. Whenever there is a change of some values for a specific network, the function f_{SDF} has to be recomputed for each available network, resulting again in the selection of the ABS network. The terminal is then triggered to perform the handover to the ABS network.

Three tables are used to collect and store the required information to compute f_{SDF}: *Network Status table* (NST), *User Policy table* (UPT) and *Application Requirement table* (ART). The NST stores the measured value of all factors for each network. The UPT contains the policy provided by the mobile user, which includes the weight assigned to each factor and the expected value for application-independent factors. The ART is used to store active applications' expected values on those application-dependent network factors such as bandwidth, security and etc.

The network selection process works as follows:

1. We assume that the multi-mode *Mobile Terminal* (MT) is initially in a *Current Network* (CN). The MT periodically checks the signal strength and beacon information through the equipped interfaces to see the availability of all networks. If there is any change, then proceed to step 2;

2. For each available network, the MT needs to measure and collect useful information which includes all factors that the mobile user cares about such as the provided bandwidth, access cost, power consumption, security level. The static network information, i.e., the values related to each network that don't change during the roaming process, are either provided beforehand by the end-user or obtained automatically via a AAA architecture. However, some information is dynamic and may change from time to time, such as available bandwidth. Therefore the MT has to obtain such kind of information periodically, either by means of announcements provided by the network or measuring by itself. After obtaining the required information, the MT should update its NST and proceed to step 5.

3. If the user decides to change the weights and expected value for specific factors, the MT updates the UPT and goes to step 5.

4. In this step, active applications may add, remove or modify their requirements, i.e., the expected value for their application-dependent factors. If any application has updated the ART, then proceed to step 5.

5. The MT checks its three tables: NST, UPT and ART and then calculates the value of SDF for each available network according to equation 4 and 5. The MT then selects the network that has the highest SDF value as the new network to handover to.

6. The MT triggers the handover to the selected network.

4.3 Handover Frequency Reduction

As mentioned in the previous section, the network selection is based on the instantaneous calculated SDF value for each available network. The handover is triggered immediately if the calculated SDF value of a network is greater than that of the current used network. This mechanism can reflect accurately the change of network status and application requirements, but may result in frequent handovers, e.g., in *ping-pong* scenarios where the conditions of two networks are similar. Furthermore, such frequency of handovers would produce significant signaling overhead both for Layers 2 and 3, thus negatively impacting the end-user experience.

Three algorithms are proposed to deal with the described problem by reducing unnecessary handovers. A comparison of the performance of the first two algorithms is given in section 5. And the final algorithm is left to be explored as future work.

Suppose there are two available networks. Let N_{now} denote the network where the mobile terminal is attached and let N_{other} denote another available network. Let F_{now}

denote the calculated value of f_{SDF} for network N_{now}, and F_{other} denote the calculated value of SDF for network N_{other}.

Algorithm 1 The handover is only triggered when F_{other} is greater than F_{now} and $F_{other} - F_{now} \geq \Delta$, where Δ is a predefined threshold. This simple algorithm reduces the frequency of handovers by introducing a "threshold of satisfaction". In other words, if the degree of satisfaction of the candidate network is too close (lower than Δ) to the one of the current network, then performing the handover does not pay up. An example of such situation is the case of a MT being in the range of two different hotspots belonging to the same provider (thus offering the same conditions) and holding similar network conditions at a specific point in time. Here, the key of the algorithm is, of course, the adequate choice of Δ.

Algorithm 2 The handover is only triggered when F_{other} has been greater than F_{now} for at least D units of time. If such condition is met, the mobile terminal then switches to network N_{other}. We call this mechanism *Lazy Handover* and the duration D *Lazy Period*. One issue is how to select the value of D. We suggest that the *Lazy Period* D should be at least greater than the handover latency, i.e., the time spent for the handover. This algorithm needs to record some history information for each possible network, which may introduce additional overhead.

Algorithm 3 In a similar way to Algorithm 2, the handover provided by the current algorithm is only triggered if the tendency of F_{other} is to be greater than F_{now} particularly in the most recent instants. Such tendency can be easily be tracked with the help of an *Exponential Moving Average* function. This algorithm then requires less state information than the previous one: no need to keep a value for D and less status in what concerns the history of F_{now}.

5 Simulation Results and Analysis

We provide simulation results based upon a heterogeneous network scenario consisting of an 802.11b network (network A) and an 802.11a networks (network B) - these two WLAN technologies have different physical layer parameters. Though in the simulation, only WLAN networks are considered, we believe the proposed network selection mechanism can be easily applied to other wireless technologies. Network A is setup in a way to provide broader coverage than B - thus, Network B is totally included within Network A. The simulation is done with OPNet. The mobile terminal moves around within the coverage area of both networks. Mobile IPv6 is used as the handover management protocol. And it is modified to be able to provide seamless handover during the handover process. For the SDF function, *available bandwidth* (factor b), *charging cost* (factor c) and *security level* (factor s) are considered as the factors that affect the network selection. Then the SDF is written as:

$$f = W_b S_b + W_c S_c + W_s S_s \ . \qquad (6)$$

The satisfaction functions S_b and S_s are computed by the means of equation 2; while the function S_c relies on equation 3. The factor s and c are considered as static parameters while b is considered as dynamic parameter. Furthermore, factor b and s are considered as application-dependent. Factor c is considered as application-independent if end-user is charged based on serving time. Table 1 shows the user policy, and the network characteristics for the mentioned factors b, c and s. Table 2 provides the bandwidth requirements of active applications. We assume all applications have the same security requirement and the access cost is a static application-independent factor. Consequently, in the simulation the satisfaction degree is mainly affected by the satisfaction function of available bandwidth, i.e., the measured bandwidth and the requirements of active applications. In network A, background traffic (CBR) is generated to affect the actual measured available bandwidth. There is no background traffic in Network B. The real-time available bandwidth of each network is estimated periodically based on the measured NAV values. For further detail of bandwidth estimation, please refer to [4].

Table 1. User Policy and Network Characteristics

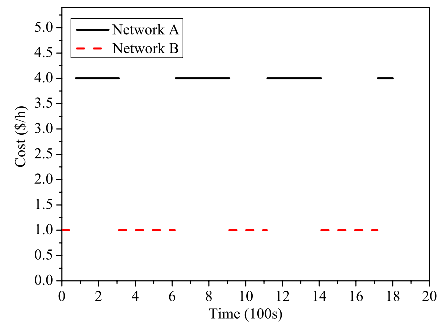
Factors	Bandwidth	Access Cost	Security Level
Weight	0.3	0.5	0.2
Network A	-	\$4/h	3
Network B	-	\$1/h	1
User Expectation	-	\$1/h	-

Table 2. Requirements of Active Applications

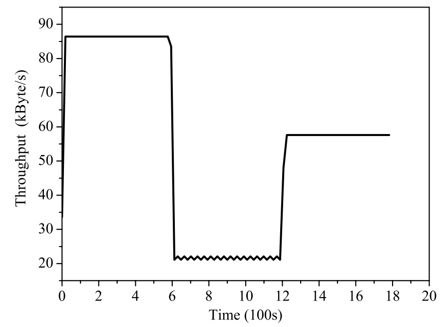
	Duration	Bandwidth (bytes/s)	Security Level
Application 1	1-600(s)	80k	1
Application 2	601-1200(s)	20k	1
Application 3	1201-1800(s)	60k	1

Fig. 2(a) provides the result concerning the selected networks and their corresponding access cost in the simulation process. One network is selected when its calculated satisfaction degree is higher than another network's. In our simulation, the satisfaction degree of each network changes along with the change of background traffic and the requirements of different active application. When the bandwidth requirement decreases, the mechanism would switch to a network that can meet the bandwidth requirement which has less cost. Fig. 2(b) provides result concerning actual throughput. Comparing the results and Table 2 we can see that the proposed network selection process not only does meet the bandwidth requirement but also reduces the overall access cost. So from Fig. 2, we see that the computed satisfaction degree works by choosing the ABS network - the selected network always has the best satisfying degree.

In the simulation, we also evaluate the performance of the proposed algorithms 1 and 2 that reduce handover frequency. Fig. 3 presents the computed satisfaction degree for a ping-pong scenario where, due to the frequent change of the background traffic, the MT switches back and forth between Network A and B. Results provided in Fig. 3(b) refer to the regular operation, i.e., no algorithm is used to reduce the unnecessary handover. We can see that lots of handover occur in it. Fig 3(c) and 3(d) provide results concerning algorithm 1 and 2 respectively. The results show that both algorithms reduce the handover frequency significantly. We also observe that during unstable period algorithm 1 results in a more frequent choice of Network A while algorithm 2 results in a more frequent choice of Network B. This phenomenon is caused by different selection criteria adopted by algorithms 1 and 2. Since the satisfaction degree heavily depends on the selected factors, the performance of the handover algorithms would also depends on these factors. For example, if the available bandwidth is the only factor that end-user cares about, then the satisfaction degree would change according to the traffic characteristics in the network. Hence for different network, such as VoIP or IPTV network, each algorithms may present different results.



(a) Selected Network and Its Cost



(b) MT Throughput

Figure 2. Network Selection Based on SDF.

6 Conclusion

In this paper, we present a dynamic and user-centric network selection method that considers user policies and cross-layer information coming from physical, link and application layers to optimize the handover across mobile heterogeneous networks. The optimization refers not only to a significant reduction in the handover frequency but also in the possibility of allowing the end-user to influence the handover beforehand by means of pre-defined parameters. A *Satisfaction Degree Function* is used to evaluate and select the best network that can fulfill the user's requirements. The numerical results show that proposed mechanism is effective in providing an Always Best Satisfying (ABS) network to the end-user based on user policies and relevant network parameters. However, providing an ABS at any instance in time may result in high frequency of handovers which may degrade the network performance. We therefore propose some algorithms to reduce such frequency and provide a simple evaluation for the two most relevant algorithms. Such results show that the algorithms are successful in balancing the user requirements with the frequency of handovers, thus resulting in a better network selection process. The performance of each algorithm need to be studied further within an analytical model in future work. We also intend to deepen the current investigation in a way that applies the handover selection mechanism to other wireless networks.

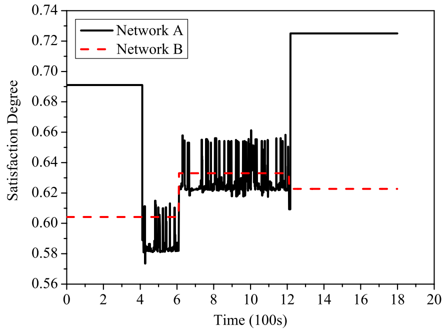
Acknowledgment

The authors would like to thank Huang Zheng, Mu Feng and Tang Wenzhi, the work could have not been done without their help.

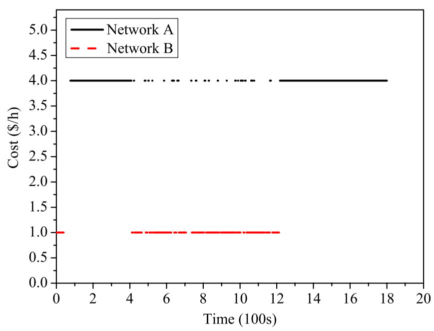
References

[1] E. Adamopoulou, K. Demestichas, A. Koutsorodi, and M. E. Theologou. Intelligent access selection in heterogeneous networks - simulation results. In *Proc. 2nd International Symposium on Wireless Communication Systems 2005*, pages 279–283, 2005.

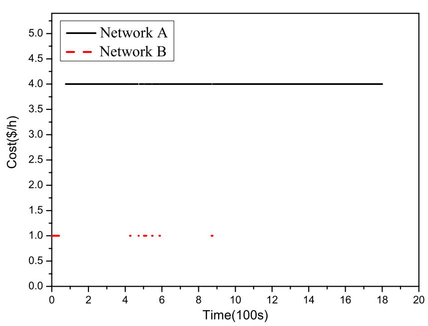
[2] H. Bing, C. He, and L. Jiang. Performance analysis of vertical handover in a umts-wlan integrated network. In



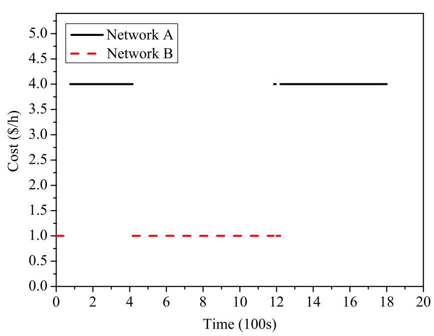
(a) Satisfaction Degree



(b) No Optimization



(c) Optimized Algorithms 1



(d) Optimized Algorithms 2

Figure 3. Optimized Algorithms to Reduce Handover Frequency.

Personal, Indoor and Mobile Radio Communications, volume 1, pages 187–191, 2003.

[3] A. Calvagna and G. D. Modica. A cost-based approach to vertical handover policies between wifi and gprs. *Wirel. Commun. Mob. Comput.*, 5(6):603–617, 2005.

[4] C. Guo, Z. Guo, Q. Zhang, and W. Zhu. A seamless and proactive end-to-end mobility solution for roaming across heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 22(5):834–848, 2004.

[5] N. D. Tripathi, J. H. Reed, and H. F. Vanlandingham. Handoff in cellular systems. *IEEE Personal Communications Magazine*, pages 26–37, 1998.

[6] R. Vijayan and J. M. Holtzman. A model for analyzing handoff algorithms [cellular radio]. *IEEE Transactions on Vehicular Technology*, 42(3):351–356, 1993.

[7] H. J. Wang, R. H. Katz, and J. Giese. Policy-enabled handoffs across heterogeneous wireless networks. In *Mobile Computing Systems and Applications*, pages 51–60, 1999.

[8] A. H. Zahran, B. Liang, and A. Saleh. Signal threshold adaptation for vertical handoff in heterogeneous wireless networks. *Mob. Netw. Appl.*, 11(4):625–640, 2006.

[9] N. Zhang and J. M. Holtzman. Analysis of handoff algorithms using both absolute and relative measurements. *IEEE Transactions on Vehicular Technology*, 45(1):174–179, 1996.

[10] F. Zhu and J. Mcnair. Optimizations for vertical handoff decision algorithms. In *IEEE WCNC*, volume 2, pages 867–872, 2004.