Vasu *et al. EURASIP Journal on Wireless Communications and Networking* 2012, **2012**:322
http://jwcn.eurasipjournals.com/content/2012/1/322

EURASIP Journal on
Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH **Open Access**

QoS-aware fuzzy rule-based vertical handoff decision algorithm incorporating a new evaluation model for wireless heterogeneous networks

Kantubukta Vasu[1], Sumit Maheshwari[1], Sudipta Mahapatra[1*] and Cheruvu Siva Kumar[2]

Abstract

Next generation networks are envisioned to be heterogeneous in nature with an increase in demand towards ubiquitous services in wireless networks. As various networks have widely different characteristics, it is difficult to maintain the quality of service (QoS) after executing a handoff from one network to another network. Maintaining the QoS, based on applications, during the handoff in an heterogeneous networks needs an intelligent handoff decision mechanism. This article proposes a QoS-aware fuzzy rule-based vertical handoff mechanism that makes a multi-criteria-based decision, found to be effective for meeting the requirements of different applications in a heterogeneous networking environment. The QoS parameters considered are available bandwidth, end-to-end delay, jitter, and bit error rate. A new evaluation model is proposed using a non-birth–death Markov chain, in which the states correspond to the available networks. Simulation results show that compared to other vertical handoff algorithms, the proposed algorithm gives better performance for different traffic classes.

Introduction

Next generation heterogeneous wireless networks require seamless mobility amongst the different access networks while maintaining the required level of quality of service (QoS) for applications, namely, high-speed data services, audio, video and multimedia applications. In such networks, it is necessary to employ efficient mobility management strategies to meet different QoS requirements for various traffic classes while maintaining high or fair utilization of wireless resources. This is achieved with a good mechanism to handle handoff between two dissimilar networks, called vertical handoff. Vertical handoff mechanisms involve three different phases of operations: system discovery, handoff decision process and handoff execution.

In system discovery phase, the system periodically monitors the states of the networks to determine the network to which handoff can be carried out. Several strategies are proposed for implementing the vertical handoff decision process, which is crucial in carrying out handoff, with the help of the available information. A comparison between different vertical handoff algorithms is presented in [1]. The handoff decision generally depends on various parameters including available bandwidth, bit error rate (BER), jitter, average battery lifetime, access cost, transmit power and end-to-end delay (E2EDelay). Smaoui et al. [2] have proposed a new handoff scheme for reducing handoff delay using the concepts of Received Signal Strength and threshold management. Considering reduction of total interference in CDMA, a vertical handoff decision algorithm among the CDMA networks and wireless local area networks (WLANs), is proposed in [3]. A combination of some of the criteria like bandwidth, RSSI and delay is also considered for making a handoff decision [4]. Such a multi-criteria-based handoff offers a number of advantages, especially in the presence of heterogeneous networks. Moreover, the wide variation in the characteristics of the networks involved motivates one to explore the field of fuzzy logic to develop a strategy for implementing multi-criteria-based handoff.

---

*Correspondence: sudipta@ece.iitkgp.ernet.in
[1] Department of Electronics and Electrical Communication Engineering, IIT Kharagpur, Kharagpur, India
Full list of author information is available at the end of the article

Literature reveals that a lot of work has been carried out with a view to improve the QoS in heterogeneous wireless networks [5-8]. Yang et al. [5] propose a new policy-based QoS-supporting system infrastructure and a QoS-aware routing algorithm that is based on the analysis of the basic architecture of an emerging heterogeneous wireless network. The policy-based QoS supporting system allows the network operators to adjust their policies easily to meet specific requirements arising in dynamic networks. However, the QoS requirements of various traffic classes vary according to the type of traffic. As we are moving towards new technologies, admission control needs to deal with many heterogeneous networks and admit new sessions to a network that is most appropriate to supply the requested QoS. Yang and Chen [7] propose a mobile QoS framework for heterogeneous IP multimedia subsystem (IMS) interworking that reduces the service disruption time, supporting the IMS mobility based on the concept of session initiation protocol multicast. In this approach, the mobility of a User Equipment is modelled as a transition in the multicast group membership. To overcome mobility impact on service guarantees, UEs need to make QoS resource reservations in advance at neighbouring IMS networks, which they may visit during the lifetime of an ongoing session.

Guo et al. [9] have proposed a multi-criteria-based approach for making the handoff decision using a fuzzy inference system (FIS) along with a modified Elman neural network. FIS considers bandwidth, velocity and number of users as decision parameters in the handoff process. The fuzzy logic-based vertical handoff decision procedure presented in [10] considers three main input parameters, namely, received signal strength, cost and bandwidth. These parameters are dynamically evaluated and compared to achieve optimal handover. However, the schemes proposed in [9,10] do not consider the QoS requirements of the applications being serviced. A context-aware handover decision using knowledge about context of mobile devices, users and networks, such as user preferences, application requirements, network parameters, link quality for decision making and based on user perceived QoS trigger is presented in [11]. Kim et al. [12] propose a context-based network selection mechanism between WLAN and CDMA networks, where the context information is a combination of grade of service (GoS) and number of handoff attempts. GoS is a function of dropping and blocking probabilities. Handoff trigger is decided based on the RSS and distance. A velocity threshold is used to optimize the system performance. However, the context information used in this is not enough to maintain the required QoS for various kinds of applications, which might be having widely different QoS requirements in terms of data rates, delay bounds and BERs.

In [13-15], fuzzy logic-based vertical handover decision is applied by considering a combination of various parameters such as price, RSS variation, traffic, sojourn time, available network bandwidth, monetary cost, user preferences, dwell time, etc. Fuzzy logic is even applied for interworking between LTE and WLAN where the authors consider the bandwidth, battery life, SNR and network load as system parameters for arriving at a handover decision [16]. As the QoS is a tradeoff among different parameters, maintaining the QoS for different kinds of applications is a critical task in VHO decision making. For example, unlike non-real-time data packets, video services are very sensitive to packet delivery delay, but can tolerate some frame losses and transmission errors. Due to the widely varying requirements of applications in wireless networks, the scenario is changing from best effort to QoS aware networks. However, from the above discussion it is clear that none of the existing approaches take the application-specific QoS parameters into consideration while making the handoff decision.

The various strategies used for executing handoff may in general be classified into Mobile-Controlled Handoff (MCHO), Network-Controlled Handoff (NCHO) and Mobile-Assisted Handoff (MAHO). In MCHO, the mobile node continuously monitors the signal of access points and initiates the handoff procedure when certain handoff criteria are satisfied. Although MCHO has a low complexity in terms of network equipment, latency and loss of packets during inter-subnet handoff can be high. NCHO is a centralized handoff protocol, in which a network takes handoff decisions based on the measurements of the signal quality of a mobile station (MS) at a number of base stations (BS). In MAHO, the handoff process involves feedback from a mobile node reflecting the measured signal strength of the surrounding base stations; but, finally it is the network which makes the handoff decision. Since in heterogeneous wireless access networks only the mobile nodes have specific knowledge about the kind of interfaces they are equipped with, the network dependency on the mobile node is high. Therefore, MCHO with some assistance from the networks is better suited for implementing vertical handoff.

A lot of research has been conducted to model the performance of heterogeneous networks. Queuing theory has been applied to model the performance in several contexts. To predict the heterogeneous environment behavior via a simulation model, we first need to construct an appropriate model to represent the heterogeneous networks. This model is then analysed and simulated using mathematical techniques. Analytical modelling of blocking and packet loss in wireless cellular networks supporting handoff are proposed in [17] where the performance and availability models are developed. In these models, the authors considered the number of virtual channels to represents the states for wireless networks. In [18,19], a similar approach is adopted for evaluating vertical handoff schemes in wireless data networks where the authors propose a performance, availability and performability (a combination of performance and reliability) based model. For comparison between various vertical handoff decision algorithms in heterogeneous wireless networks, Stevens-Navarro and Wong [1] use a Markov model based on a birth–death process. In contrast to other works in the literature, in this article, we develop a simulation model to measure the performance of vertical handoff mechanisms in a heterogeneous wireless network environment. The power of this evaluation model is that the mobility characteristics are taken care of automatically by considering different state, and connection lifetimes in the simulation. Each network is assumed to be associated with different QoS parameter values.

The QoS-aware fuzzy rule-based handoff mechanism proposed in this article assumes MCHO and some assistance from the network. In this, a mobile periodically monitors the available networks and using a fuzzy rule-based algorithm determines the best network for making a handoff decision. This information is then communicated to the current network for executing the handoff. The evaluation model used in the simulation is a non-birth–death Markov chain where a state represents the available networks at any instant of time. The rest of the article is organized as follows: “QoS requirements and vertical handoff strategies” section briefly discusses the QoS requirements in heterogeneous wireless networks and “Implementation issues for vertical handoff” section discusses the implementation issues for vertical handoff; the proposed algorithm is presented in “The proposed QoS-aware FRB vertical handoff decision algorithm” Section, followed by theoretical evaluation of the proposed model in “Evaluation model” section. “Results and discussions” section contains the simulation results and finally “Conclusions” section concludes this article.

QoS requirements and vertical handoff strategies

QoS requirements

It is difficult to maintain a steady customer-base with just a single type of network with the technological advancements and economic changes in the market. So, wireless service providers (WSPs) are adopting a multitude of access technologies, operating on both licensed and unlicensed bands, to serve an increasing number of subscribers. Among the various applications, those involving real-time video are more delay sensitive than non-real-time services such as file downloads. In these kinds of applications, it is likely that numerous types of access networks will coexist to support wireless services with different QoS requirements. To harness the wide variability of coverage, bandwidth and reliability offered by different technologies, operating at different spectrum bands, WSPs are planning to deploy heterogeneous access networks. These heterogeneous networks are able to provide different sets of services with varying QoS requirements. According to the IMT2000 QoS Classes and Requirements (3GPP-TS 23.107), different applications have different QoS requirements as explained in Table 1 [20,21] with the help of linguistic terms.

From Table 1, some simple inferences can be made. For example, BER for conversational type of applications need not be low which means that BER is not a prime requirement in voice-based applications. Whereas for such applications, E2EDelay and Jitter should be low, meaning that these parameters should have low values for better user perceived quality. For streaming kind of applications (including live streaming) Jitter should be low, Bandwidth should be high, and E2EDelay may be low to medium, but should not be high. Similar inferences can be made from this table regarding interactive and background applications.

Vertical handoff strategies

Vertical handoff decision mechanisms need to consider the QoS parameters important for a particular application. Various vertical handover decision mechanisms have been proposed in the past [22] including the following.

Simple additive weighted—SAW

While making a handoff decision, each of the networks involved is assigned a score, which is the weighted sum of all the attribute values. The score of each network is obtained by adding the normalized contributions from

Table 1 QoS requirements and traffic classes where the QoS parameters: BER, E2EDelay, Jitter, Bandwidth; and Traffic Classes: Conversational, Streaming, Interactive, Background

	BER	E2E delay	Jitter	Bandwidth
Conversational	Need not be low	Should be low	Should be low	Need not be high
Streaming	Need not be low	Should be low or medium	Should be low	Should be high
Interactive	Should be low	Medium or low	Need not be low	Need not be high
Background	Should be low	Need not be low	Need not be low	Should be medium at least

each metric r_{ij} multiplied by the importance weight w_j assigned to attribute j. The selected network A^*_{SAW} is

$$A^*_{\mathrm{SAW}} = \mathrm{argmax}_{i \in M} \sum_{j=1}^{N} w_j * r_{ij}$$

where N is the number of parameters and M is number of candidate networks.

Technique for order preference by similarity to ideal situation—TOPSIS

The selected candidate network is the one which is the closest to the ideal solution (and the farthest from the worst case solution). The ideal solution is obtained by using the best values for each metric. If C_i^* denotes the relative closeness (or similarity) of the candidate network i to the ideal solution, the selected network A^*_{TOP} is

$$A^*_{\mathrm{TOP}} = \mathrm{argmax}_{i \in M} C_i^*$$

Multiplicative exponential weighted—MEW

Using this technique, vertical handoff decision can be expressed as a matrix where each row i corresponds to the candidate network i and each column j corresponds to an attribute (Bandwidth, Delay, etc.). The score S_i of network i is

$$S_i = \prod_{j=1}^{N} x_{ij}^{w_j}$$

where x_{ij} denotes attribute j of candidate network i, w_j denotes the weight of attribute j and $\sum_{j=1}^{N} w_j = 1$.

w_j is a positive power for benefit metrics ($x_{ij}^{w_j}$), and a negative power for cost metrics ($x_{ij}^{-w_j}$). Since the score is an upper bound, it is convenient to compare each network with the score of the positive ideal network A^{**}. This network is defined as the network with the best values in each metric. (For a benefit metric, the best value is the largest. For a cost metric, the best value is the lowest.) The value of ratio R_i between network i and the positive ideal is

$$R_i = \frac{\prod_{j=1}^{N} x_{ij}^{w_j}}{\prod_{j=1}^{N} (x_{ij}^{**})^{w_j}}$$

The selected network A^*_{MEW} is obtained as

$$A^*_{\mathrm{MEW}} = \mathrm{argmax}_{i \in M} R_i$$

The vertical handoff decision algorithms considered for comparison needs relative importance of each parameter which is usually given by the set of weights w_j. The analytical hierarchical processing (AHP) method is used to determine the weights [23,24] by comparing a pair metrics with the 1–9 AHP scale. The four traffic classes have different QoS requirements. So, we assigned the different weights according to the importance of parameters in different traffic classes as shown in Table 2. The QoS parameters considered are available bandwidth, E2Edelay, BER and jitter with the corresponding importance weight for each traffic class shown in Table 3 [1].

Table 2 Relative importance of different parameters using AHP

	BER	E2EDelay	Jitter	Bandwidth
Conversational				
BER	1	1/9	1/9	1
E2EDelay	9	1	1	9
Jitter	9	1	1	9
Bandwidth	1	1/9	1/9	1
Streaming				
BER	1	1/5	1/9	1/9
E2EDelay	5	1	1/5	1/5
Jitter	9	5	1	1
Bandwidth	9	5	1	1
Interactive				
BER	1	5	9	5
E2EDelay	1/5	1	5	1
Jitter	1/9	1/5	1	1/5
Bandwidth	1/5	1	5	1
Background				
BER	1	9	9	5
E2EDelay	1/9	1	1	1/5
Jitter	1/9	1	1	1/5
Bandwidth	1/5	5	5	1

Implementation issues for vertical handoff

IEEE 802.21 is responsible for enabling handover and interoperability between heterogeneous network types including both 802 and non-802 networks. It provides information required for handover to and from a range of networks including cellular, GSM, GPRS, WiFi and Bluetooth. The network handover enabling function is a part of the media-independent handover (MIH) function. The general reference model of MIH is as shown in Figure 1

Table 3 Importance weights derived using AHP

	BER	E2EDelay	Jitter	Bandwidth
Conversational	0.04998	0.45002	0.45002	0.04998
Streaming	0.03737	0.11380	0.42441	0.42441
Interactive	0.63593	0.16051	0.04304	0.16051
Background	0.66932	0.05546	0.05546	0.21976

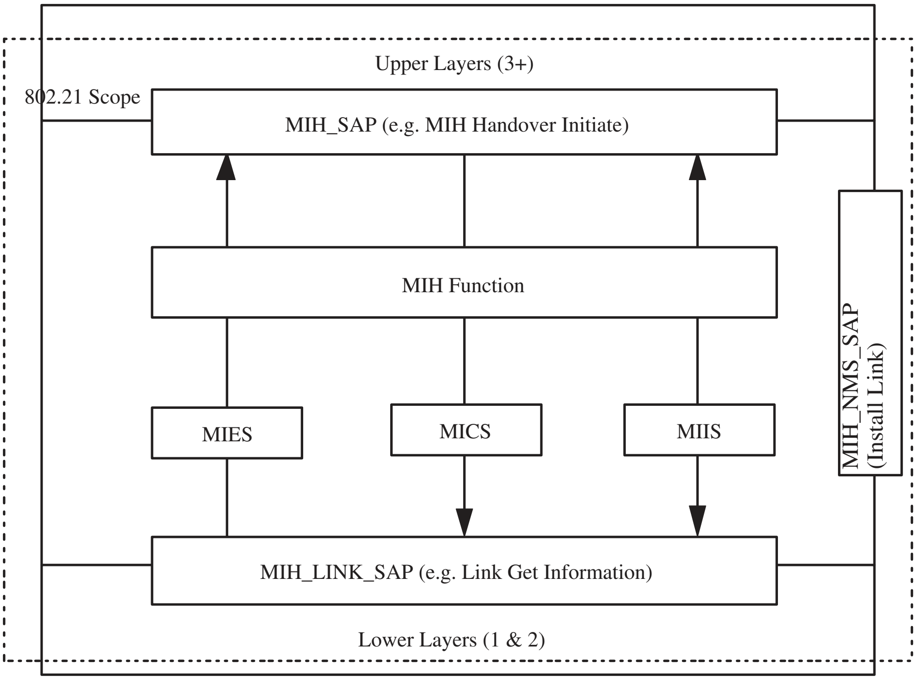


Figure 1 MIH (IEEE 802.21).

[25]. The MIH function consists of three elements, namely, the event service, command service and information service [26,27]. MIH function (MIHF) is used for convergence of multiple heterogeneous access technologies to achieve seamless mobility and provides link layer information to the upper layers of the mobility management protocol stack using service access points (SAPs) [25]. In this MIH, the *MIH_SAP* (*MIH_service_access_points*) allows access from the upper layers, *MIH_NMS_SAP* (*MIH_Network Management system_service*) allows for management and *MIH_LINK_SAP* is used for the link layers. Each or any number of the access networks will interface directly with the MIHF using their own SAPs. Operation of MIH is as follows: When the mobile device notifies that the signal strength of the current network is going below the threshold, the MIH event service informs to MIHF of the mobile device. Then this information is passed on to the MIHF of the access point. The MIH command service in the access point informs the mobile device to initiate handover with a list of possible access points. The mobile device MIHF determines the signal strength and achievable QoS parameters of each access network using the MIH information service. Once this information is obtained, it will be passed on to the handover decision module for deciding the best network. Once the best network is decided, the MIH command service informs the mobile device to commit for handover. Then, mobile IP protocols can be used to switch over to the selected network. A similar kind of standard is proposed from the 3GPP as the generic access network (3GPP TS 43.318), which supports two modes of operation, namely GAN A/Gb mode and GAN Iu mode.

QoS information about the available networks within the range of the current network interface of a mobile node is obtained periodically by using the IEEE 802.21/GAN standard. This information is utilized by our proposed QoS-aware FRB vertical handover decision mechanism. Internet Engineering Task Force IP Performance Metrics Working Group has standardized procedures for performance metrics such as available bandwidth and average delay for Internet services. After the handoff decision is taken, handoff is executed using mobility management protocols, e.g. the Host Identity Protocol.

Assuming the mobile client is able to access any and every network and the mobile client has each and every interface built-in, Figure 2 represents a typical heterogeneous wireless network environment consisting of Universal Mobile Telecommunications System (UMTS), Global System for Mobile Communications (GSM), WLAN and Worldwide Interoperability for Microwave Access (WiMAX) networks. UMTS and GSM networks are connected to the General Packet Radio Service (GPRS) infrastructure via the mobile switching center (MSC). This GPRS infrastructure consists of Serving GPRS Support

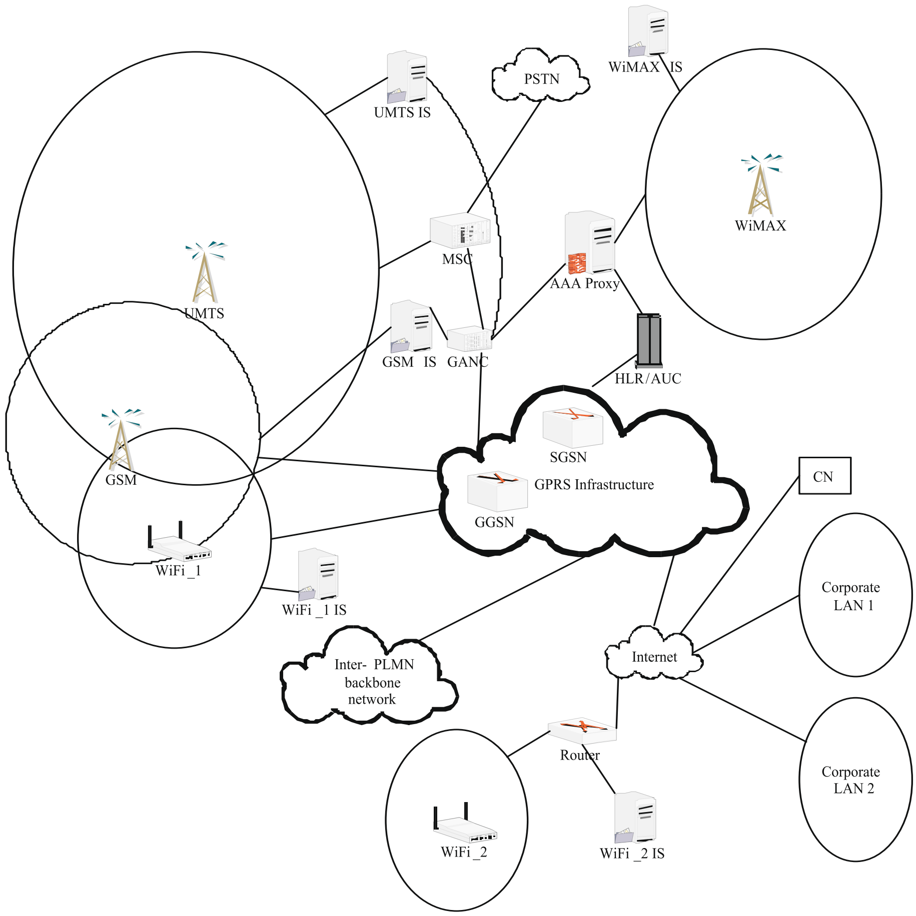


Figure 2 **Typical heterogeneous wireless network environment (deployment mode).**

Node (SGSN) and Gateway GPRS Support Node (GGSN) gateway components. The SGSN is responsible for mobility management and manages the end-user via the Base Station Controller/Radio Network Controller (BSC/RNC) in the radio network. It maintains the connected user's context and integrates with other elements such as the Home Subscriber Server/Home Location Register (HSS/HLR) to manage allowed services and the GGSN to manage access to external IP networks. The GGSN is the IP access point for mobile users to the Internet, corporate Virtual Private Network (VPN) or other IP access networks. Static information about the QoS and energy is maintained by the information servers provided at each network. WLAN infrastructure provided by the operator can directly connect to the GPRS infrastructure via WLAN gateways. A WLAN network can also connect to Internet via a router. A WiMAX network is connected to the GPRS infrastructure via an Access Service Network Gateway (ASN-GW) (data not shown in Figure 2). A critical component of any mobile WiMAX network is the ASN Gateway, which aggregates subscriber and control traffic from base stations within an access network.

The correspondent node can be a streaming server or it can be any other mobile device to which the user is connected. The mobile client is also equipped with MIH client support and generic access network support. The mobile clients access the information about the network through MIH (IEEE 802.21) for non-3gpp networks and through the generic access network controller for 3gpp-based networks. The AAA server provides the authentication, authorization and accounting functions.

While the mobile is moving, if it detects that the signal strength of current network is less than a predetermined threshold value, it scans the available foreign network set. The mobile client needs to obtain the information about the foreign networks via generic access network controller or MIH. Then the information is provided to a QoS-aware FRB module that resides in the mobile client. This module decides the best network based on the information provided from the information servers. The information includes the achievable QoS parameters from the past knowledge. In the proposed scheme for implementing vertical handover, the mobile is switched to the best network by using the mobility management protocols.

The proposed QoS-aware FRB vertical handoff decision algorithm

A method for an intelligent handover decision mechanism among different radio access networks where the available network set is obtained dynamically at the mobile client comprises four steps as explained by the flow chart in Figure 3. The mobile client periodically checks for the signal strength (RSSI) of the current network to monitor the condition for handover. When the RSSI of mobile client is going below the handover threshold, the available networks are examined at the mobile client. In the second step, the QoS parameters of the available networks are obtained either using MIH or GAN or both based on the available network set. In the third step, the best network is decided by using the proposed QoS-aware FRB vertical handoff mechanism based on the application QoS requirements obtained in the second step. Finally, the mobile is switched to the best network from the current network after making a decision using any mobility management protocol.

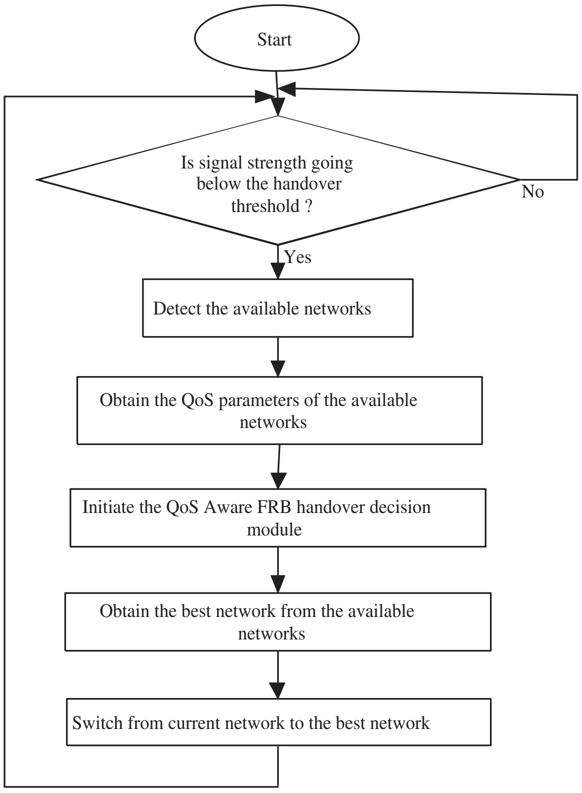


Figure 3 Procedure for making QoS aware fuzzy rule based vertical handover.

Unlike propositional logic where a set membership function takes a value from the set $\{0, 1\}$, fuzzy logic is associated with linguistic variables with membership functions that take values in the interval [0, 1]. A fuzzy set is a set with such a set membership function. The process of taking an observation and creating a fuzzy set from it is called fuzzification. In general, a Fuzzy Logic System is a nonlinear mapping of an input data (feature) vector into a scalar output. Fuzzy logic systems involve a large number of possibilities that lead to a lot many mappings [28]. Fuzzy linguistics descriptions are formal representations of systems made through fuzzy IF-THEN rules. They encode knowledge about a system in statements of the form 'IF (a set of conditions) are satisfied THEN (a set of consequents)' [29,30]. A collection of rules referring to a particular system is known as a fuzzy rule base. Moreover, it is easier to take a crisp decision if the output is represented as a single scalar quantity. The conversion of a fuzzy set to a single crisp value is called defuzzification.

The FRB algorithm proposed here considers four QoS parameters, namely, available bandwidth, E2EDelay, jitter and BER. In heterogeneous networks, handoff between different networks is required to be more efficient while maintaining the QoS requirements for different traffic classes even after the handoff. Therefore, selecting the best network among the available networks is always an important task. The proposed QoS-aware FRB mechanism that gives a decision regarding the network to be selected is depicted in Figure 4. The input QoS parameters considered are the available bandwidth, E2EDelay, Jitter and BER. The FRB system takes the input parameter values of a network and evaluates its handoff score value as shown in Figure 4. The input crisp values are first given to fuzzification module where these input values are converted into membership values using membership functions. The membership functions are considered as triangular functions with three different regions: low, medium and high. The membership values are used for decision making to give an output membership value. The output membership value is converted to a crisp value using a defuzzification process. One of the well-known methods for defuzzification is the centroid method. This study proposes a QoS-aware FRB mechanism which uses the properties of simple fuzzy logic; the FRB is used to evaluate the score value corresponding to the input parameters for each network. The score value is compared to the other networks to obtain the best network to which the handover is to be carried out.

Assume the low region points of parameter P for network N as shown in Figure 5 are as follows $\{Low\_Point1_N^P,\ Low\_Point2_N^P,\ Low\_Point3_N^P\}$; Medium

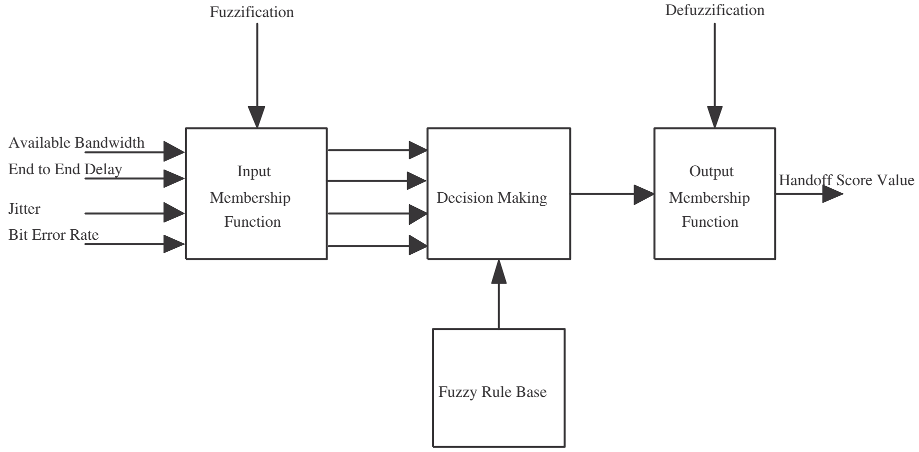


Figure 4 Block diagram of the FRB handoff mechanism.

region points as $\{Med\_Point1^P_N,\ Med\_Point2^P_N,\ Med\_Point3^P_N\}$; and for high region as $\{High\_Point1^P_N,\ High\_Point2^P_N,\ High\_Point3^P_N\}$, where $P \in \{BER, E2EDelay, Jitter, BW\}$ and $N \in \{UMTS, GPRS, WLAN\}$.

In the fuzzification process, if X is the input value for parameter P of network N and falls in the low region then the membership value for low region will be calculated as explained below:

$$\Delta_1 = X - Low\_Point1^P_N;\ \ \Delta_2 = Low\_Point2^P_N - X$$

and slope

$$S_2 = \frac{-1}{(Low\_Point2^P_N - Low\_Point3^P_N)}$$

if $((\Delta_1 \leq 0)$ or $(\Delta_2 \leq 0))$ then the membership value for low region would be

$$Low\_Mem^P_N = 0$$

else

$$Low\_Mem^P_N = min(\Delta_2 \times S_2,\ Max).$$

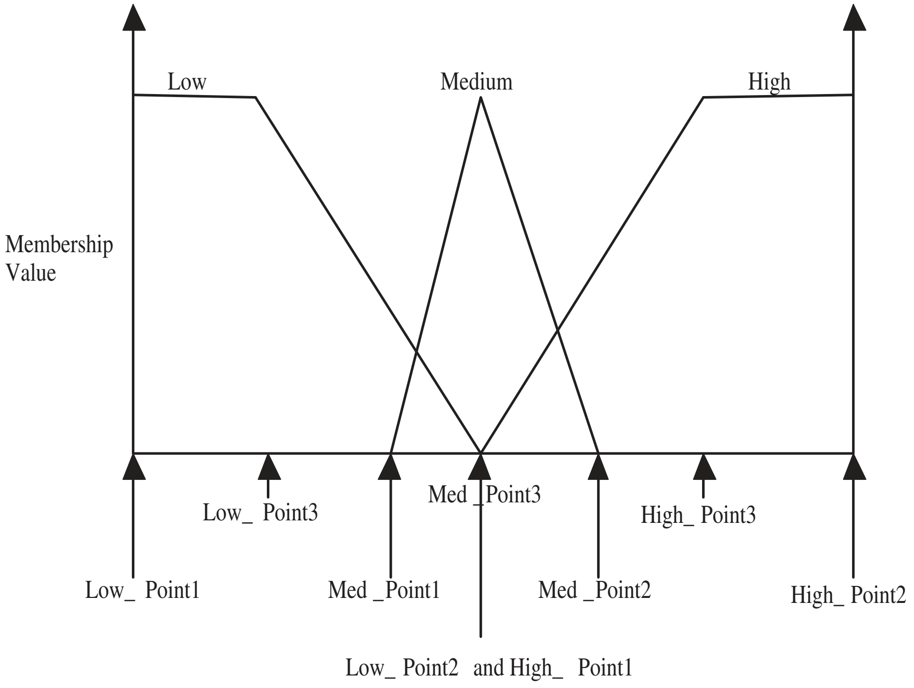


Figure 5 The adopted fuzzy membership function.

If X falls in the medium region then the membership value for medium region will be calculated as follows:

$$\Delta_1 = X - Med\_Point1_N^P;\;\; \Delta_2 = Med\_Point2_N^P - X$$

and slope

$$S_1 = \frac{1}{(Med\_Point3_N^P - Med\_Point1_N^P)};$$

$$S_2 = \frac{-1}{(Med\_Point2_N^P - Med\_Point3_N^P)}$$

if $((\Delta_1 \leq 0)$ or $(\Delta_2 \leq 0))$ then the membership value for medium region would be

$$Med\_Mem_N^P = 0.$$

else

$$Med\_Mem_N^P = min(\Delta_1 \times S_1,\ \Delta_2 \times S_2,\ Max).$$

And if X falls in the high region the membership values for high region would be:

$$\Delta_1 = X - High\_Point1_N^P;\;\; \Delta_2 = High\_Point2_N^P - X$$

and slope

$$S_1 = \frac{1}{(High\_Point3_N^P - High\_Point1_N^P)}$$

if $((\Delta_1 \leq 0)$ or $(\Delta_2 \leq 0))$ then the membership value for high region would be

$$High\_Mem_N^P = 0$$

else

$$High\_Mem_N^P = min(\Delta_1 \times S_1,\ Max).$$

For 4 input parameters and 3 membership regions, there can be a total of 81 ($3 \times 3 \times 3 \times 3 = 81$) possible rules as given in the sample fuzzy rule base in Table 4. For a particular network N, the effective rule set will be maintained by $R\ = [R_1, R_2, R_3, \ldots, R_{81}]$. These rules will be maintained by Boolean value either 0 or 1 and the corresponding membership values for each rule is calculated from the decision module by using max$_$ min product form. The membership values of these rules are the output membership values. These values are defuzzified by using centroid method. Let us assume the center value of each effected output region is $C_N\ = [C_1, C_2, C_3, \ldots, C_{81}]$ and output membership values of each effected rule is $H_N\ = [H_1, H_2, H_3, \ldots, H_{81}]$. The effected areas of output regions is $A_N\ = [A_1, A_2, A_3, \ldots, A_{81}]$. Then using the

Table 4 Sample FRB for illustrating the mechanism

Rule number	BER	E2EDelay	Jitter	Bandwidth	Handoff score
Conversational					
1	Low	Low	Low	Low	High
25	Low	High	High	Low	Low
50	Medium	High	Medium	Medium	Low
81	High	High	High	High	Low
Streaming					
1	Low	Low	Low	Low	Low
25	Low	High	High	Low	Low
50	Medium	High	Medium	Medium	High
81	High	High	High	High	Medium
Interactive					
1	Low	Low	Low	Low	Medium
25	Low	High	High	Low	Low
50	Medium	High	Medium	Medium	Low
81	High	High	High	High	Low
Background					
1	Low	Low	Low	Low	Medium
25	Low	High	High	Low	Medium
50	Medium	High	Medium	Medium	Medium
81	High	High	High	High	Medium

centroid method of defuzzification, the output handoff score value is given by

$$X^* = \frac{\sum A_N \times C_N}{\sum C_N}$$

The handoff score value of given set of input parameters for each network is calculated by using the above proposed method, and then the best network among the available networks is the network whose handoff score value is highest than other networks.

The proposed scheme is illustrated by assuming the presence of three different types of networks. These are an UMTS network, a GPRS network and a WLAN. We considered these technologies due to the wide availability of typical parameter values in the literature. But the method and the simulation procedure is transparent to the technology. In the FRB method, the length of the rule set is based on the number of membership regions and the number of QoS parameters considered. The rules are made according to the requirements of applications. So, it is transparent to the kind of technology used. The parameters assumed for these networks, which are based on the standard possible data rates and typical delay values of the networks, are as follows [1]. Bandwidth vector for UMTS: [32, 64, 128, 256, 512, 1024, 2048] kbps, for GPRS: [21, 42, 64, 85, 107, 128, 149, 171] kbps and for WLAN: [1000, 2000, 5500, 11000] kbps; E2EDelay vector for these networks are, respectively, [190, 160, 130, 100, 70, 40, 10], [185, 160, 135, 110, 85, 60, 35, 10] and [160, 110, 60, 10] ms. All networks are assumed to have the same set of jitter and BER vectors. The values for jitter and BER are taken to be [3, 5, 7, 9, 11] and [0.01, 0.001, 0.0001, 0.00001, 0.000001] ms, respectively.

The membership functions for different input and output parameters are as shown in Figure 6. Membership region for BER is assumed to be low if it is in the range $0 - 1e^{-4}$, medium in the range $1e^{-5} - 1e^{-3}$, and high in

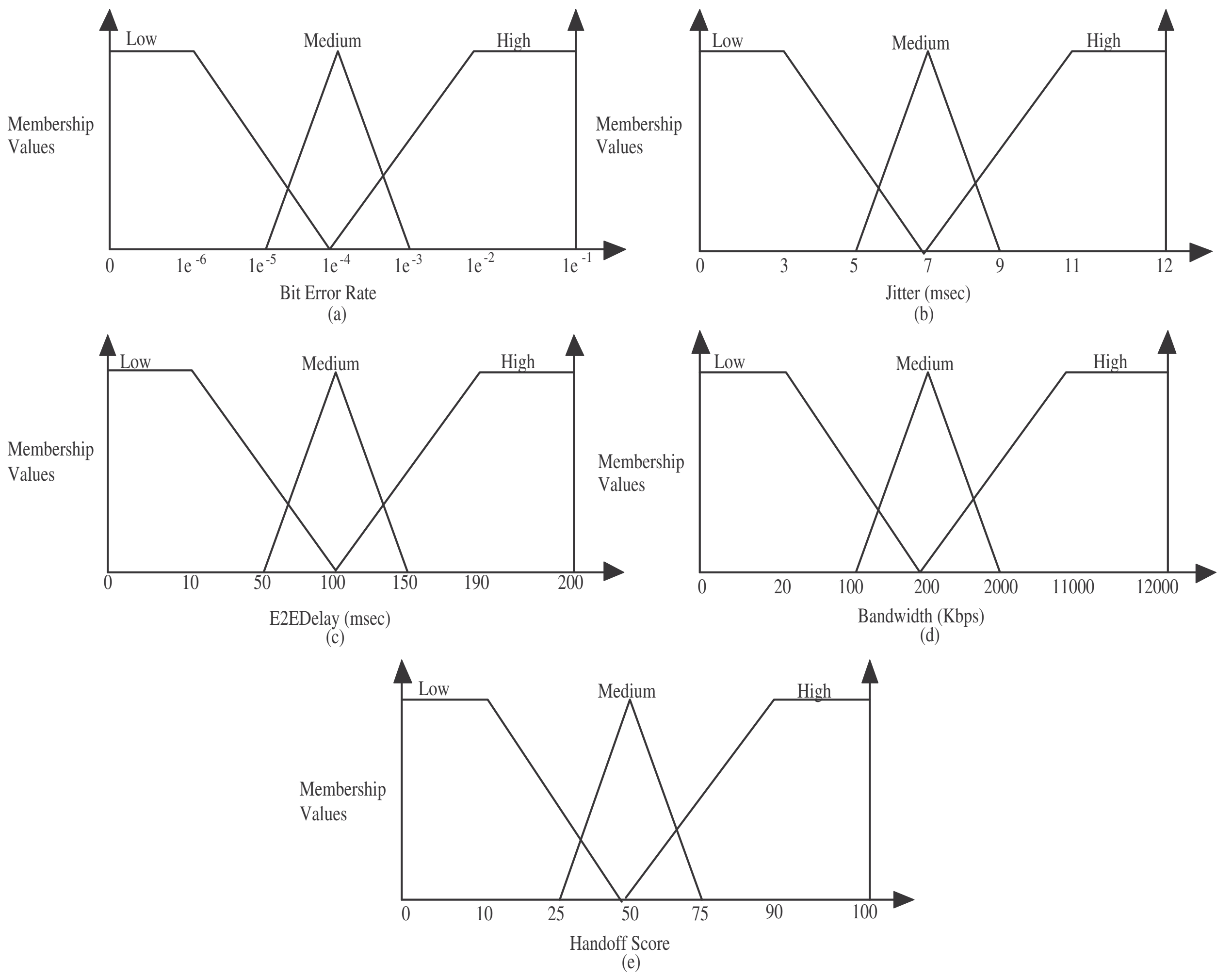


Figure 6 Input and output membership functions. (a) BER. **(b)** Jitter. **(c)** E2EDelay. **(d)** Bandwidth. **(e)** Handoff score.

the range $1e^{-4} - 1e^{-1}$. The parameters' Jitter is assumed to be low if it is in the range 0–7 ms, medium in the range 5–9 ms, and high in the range 7–12 ms. Similarly, E2EDelay is assumed to be in the low region when it is in between 0 and 100 ms, medium in between 50 and 150 ms, and high in between 100 and 200 ms. For bandwidth, we can usually say that 0–200 kbps is a low bandwidth region, medium or enough bandwidth to be in the range of 100–2000 kbps, and the high bandwidth region to be between 200 and 12000 kbps. These membership regions describe more linguistic terms of the input parameters. The basic purpose of considering these ranges is to include at least two different types of networks in each region. For example, the low region includes GPRS and UMTS, the medium region includes GPRS, UMTS and WLAN, and the high region includes UMTS and WLAN.

The fuzzy rules are made as per the requirements of 3GPP QoS classes [20,21]. Eighty-one rules are made for each traffic class (for four input parameters and three membership regions, eighty-one possible combinations of rules can be made for each traffic class). For conversational traffic class, E2EDelay and Jitter should be low. For streaming classes, Jitter should be low and bandwidth should be high. Similarly, for interactive and background traffic classes, BER should be low. Moreover, background traffic needs at least a moderate amount of bandwidth. So, rules satisfying these requirements give the handoff score value in the high region. Rest of the rules for each traffic class have been made as per the requirements of the traffic classes. In Table 4, only four rules are shown for each traffic class for the sake of illustration.

The output is a handoff score value, which can be drawn from the rule base with the help of individual consequents of each rule, and is defined in the range of 0–100 with triangular membership functions of the three regions as shown in Figure 6. The centroid method is used for defuzzification [29]. When handoff is required, a mobile calculates the handoff score value for all the available networks with a set of input parameters by using the FRB scheme proposed above, and selects the best one. The information about the best network is then communicated to the current network to execute the handoff.

Evaluation model

The performance measures extracted from a simulation model must be a good representation of the real network environment to model the heterogeneous environment. Some assumptions must be made about the real network in order to construct the heterogeneous environment model. Figure 7 shows a typical heterogeneous environment where three different networks, that is, a WLAN, a GPRS network and a UMTS network are present. Based on n number of networks, we will get 2^n number of states. Even though we considered three networks for easy understanding of the model, it can be extended to n number of networks. It is assumed that the mobile is equipped with these three network interfaces and hence it is able to access each of the networks. If the mobile is moving in the region of these networks, it can acquire one of these network sets such as only UMTS, {UMTS and GPRS}, {GPRS and WLAN} or {GPRS, UMTS and WLAN}. As there is a

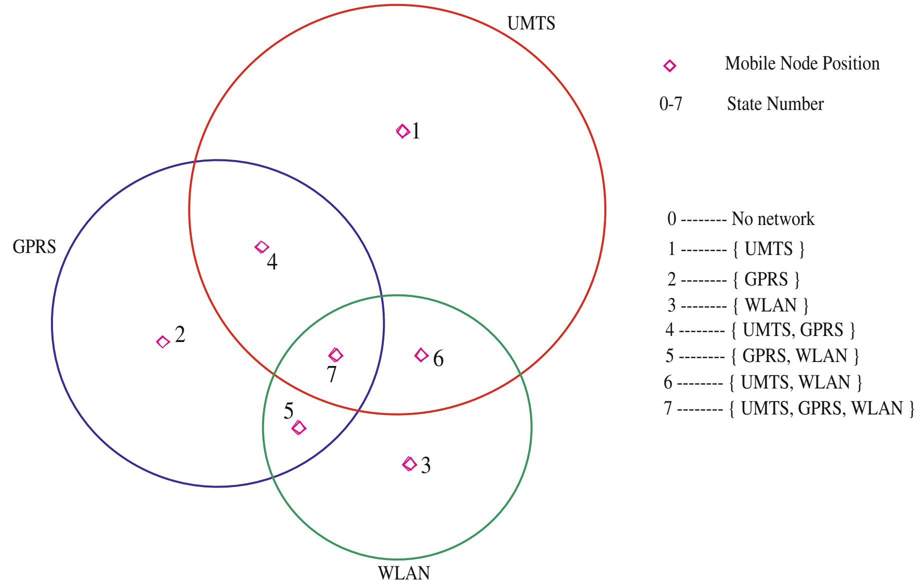


Figure 7 Wireless heterogeneous environment.

possibility that no network is available at a particular time instant so we need to consider this also as a possible scenario. Ultimately, we have eight possible states for three different kinds of networks. We need to transform this real world environment to a logical model. This is the phase where mathematical techniques are used to simulate the heterogeneous environment. Transient behaviour of the heterogeneous environment will be simulated using the proposed model. The steady-state analysis of the model will be used to validate the results.

Among the modelling approaches, Markov models are suitable for modelling of dynamic system behaviour. Of these, a birth–death Markov chain can model the dynamic behaviour of mobiles where a system state is considered in such a manner that the mobile can move between two neighbouring states. For example, a state can represent the available number of transmission channels. Contrary to most of the existing work on vertical handover, in our model, we represent the system using a non birth–death Markov chain where a state corresponds to the available networks, including a state to represent the unavailability of any network. This in fact obliviates the need to consider other parameters such as velocity of mobiles, channel conditions, etc., and in a way simplifies the simulation of a heterogeneous wireless environment. The connection lifetimes of all the states are assumed to be independent identically distributed (i.i.d.) random variables. The primary reason for considering a non-birth–death Markov chain is that a mobile device can be in a region having no network or one or more networks and go from a state to any other state as depicted in Figure 8 where a state represents the available networks at any time.

The simulation model is based on the following assumptions.

(1) A non-birth–death Markov chain is considered for simulation.
(2) A state represents the set of networks available at any instant of the time.
(3) State lifetimes are assumed to follow an exponential distribution with a mean λ.
(4) State transitions are instantaneous and do not incur any waiting delays.
(5) Within a state, connections follow an exponential distribution with a mean μ.

The state transition matrix for the above Markov chain is given by P, with matrix element p_{ij}, $0 \leq i, j \leq 7$ as shown below.

$$\textit{State transition matrix } P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & p_{05} & p_{06} & p_{07} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} & p_{17} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} & p_{27} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} & p_{37} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} & p_{47} \\ p_{50} & p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} & p_{57} \\ p_{60} & p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} & p_{67} \\ p_{70} & p_{71} & p_{72} & p_{73} & p_{74} & p_{75} & p_{76} & p_{77} \end{pmatrix}$$

The state transition probabilities are obtained after normalizing the generator matrix. Using the state balance equations under the equilibrium conditions the generator matrix G for the above Markov chain is derived as shown in (1). Here, λ_{ij} denotes the state transition time from state i to j and μ_{ij} denotes the connection lifetime from state i to j where $0 \leq i, j \leq 7$ (max number of networks).

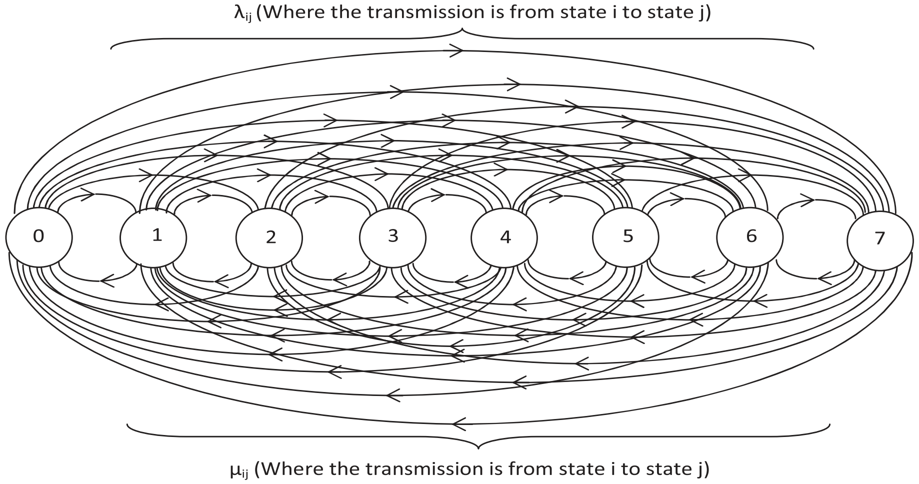


Figure 8 **Non-birth–death Markov chain with state 0 = {no network}; state 1 = {UMTS}; state 2 = {GPRS}; state 3 = {WLAN}; state 4 = {UMTS,GPRS}; state 5 = {GPRS,WLAN}; state 6 = {UMTS,WLAN}; state 7 = {UMTS,GPRS,WLAN}.**

$$G = \begin{pmatrix} -(\sum_{i=1}^{7}\mu_{i0}) & \lambda_{01} & \lambda_{02} & \lambda_{03} & \lambda_{04} & \lambda_{05} & \lambda_{06} & \lambda_{07} \\ \mu_{10} & -(\lambda_{01}+\sum_{i=2}^{7}\mu_{i1}) & \lambda_{12} & \lambda_{13} & \lambda_{14} & \lambda_{15} & \lambda_{16} & \lambda_{16} \\ \mu_{20} & \mu_{21} & -(\sum_{i=0}^{1}\lambda_{i2}+\sum_{i=3}^{7}\mu_{i2}) & \lambda_{23} & \lambda_{24} & \lambda_{25} & \lambda_{26} & \lambda_{27} \\ \mu_{30} & \mu_{31} & \mu_{32} & -(\sum_{i=0}^{2}\lambda_{i3}+\sum_{i=4}^{7}\mu_{i3}) & \lambda_{34} & \lambda_{35} & \lambda_{36} & \lambda_{37} \\ \mu_{40} & \mu_{41} & \mu_{42} & \mu_{43} & -(\sum_{i=0}^{3}\lambda_{i4}+\sum_{i=5}^{7}\mu_{i4}) & \lambda_{45} & \lambda_{46} & \lambda_{47} \\ \mu_{50} & \mu_{51} & \mu_{52} & \mu_{53} & \mu_{54} & -(\sum_{i=0}^{4}\lambda_{i5}+\sum_{i=6}^{7}\mu_{i5}) & \lambda_{56} & \lambda_{57} \\ \mu_{60} & \mu_{61} & \mu_{62} & \mu_{63} & \mu_{64} & \mu_{65} & -(\sum_{i=0}^{5}\lambda_{i6}+\mu_{6}) & \lambda_{67} \\ \mu_{70} & \mu_{71} & \mu_{72} & \mu_{73} & \mu_{74} & \mu_{75} & \mu_{76} & -(\sum_{i=0}^{7}\lambda_{i7}) \end{pmatrix} \quad (1)$$

The state transition time or state lifetime depends on the state from where the mobile is migrating. Let $\lambda_{0i} = \lambda_0,\ \lambda_{1i} = \lambda_1,\ \lambda_{2i} = \lambda_2, \ldots, \lambda_{7i} = \lambda_7$. Connection lifetime or reverse transition time depends on the state it is leaving. Let, $\mu_{i0} = \mu_i,\ \mu_{i1} = \mu_{i2} = \mu_{i3} = \mu_{i4} = \mu_{i5} = \mu_{i6} = \mu_{i7} = \mu_i$. Also, it is assumed that all the state variables are independent identically distributed random variables, then

$$\lambda_0 = \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = \lambda_7 = \lambda$$

$$\mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7 = \mu$$

Hence, the generator matrix G of the non-birth–death Markov chain is given in (2),

$$G = \begin{pmatrix} -7\mu & \lambda_i & \lambda_i & \lambda_i & \lambda_i & \lambda_i & \lambda_i & \lambda_i \\ \mu & -\lambda_i - 6\mu & \lambda_i & \lambda_i & \lambda_i & \lambda_i & \lambda_i & \lambda_i \\ \mu & \mu & -2\lambda_i - 5\mu & \lambda_i & \lambda_i & \lambda_i & \lambda_i & \lambda_i \\ \mu & \mu & \mu & -3\lambda_i - 4\mu & \lambda_i & \lambda_i & \lambda_i & \lambda_i \\ \mu & \mu & \mu & \mu & -4\lambda_i - 3\mu & \lambda_i & \lambda_i & \lambda_i \\ \mu & \mu & \mu & \mu & \mu & -5\lambda_i - 2\mu & \lambda_i & \lambda_i \\ \mu & \mu & \mu & \mu & \mu & \mu & -6\lambda_i - \mu & \lambda_i \\ \mu & \mu & \mu & \mu & \mu & \mu & \mu & -7\lambda_i \end{pmatrix} \quad (2)$$

After finding the generator matrix G as shown in (2), the state transition matrix T is calculated as shown in (3). State transition matrix

$$T = (\frac{G}{q}) + I \quad (3)$$

where $q > \max(|G_{ij}|), 0 \leq i,j \leq 7, I$ is the identity matrix of size 8×8.

An example of the graphical representations for the state and connection lifetime distributions are depicted in Figure 9. In this figure, the state lifetimes are considered to have mean values of 7 and 8 min and the connection lifetimes are assumed to have mean values of 2 and 4 min. Suppose the connections are distributed with a mean value of connection life time as 2 in a particular state, then the sum of all the connection’s lifetimes in that state

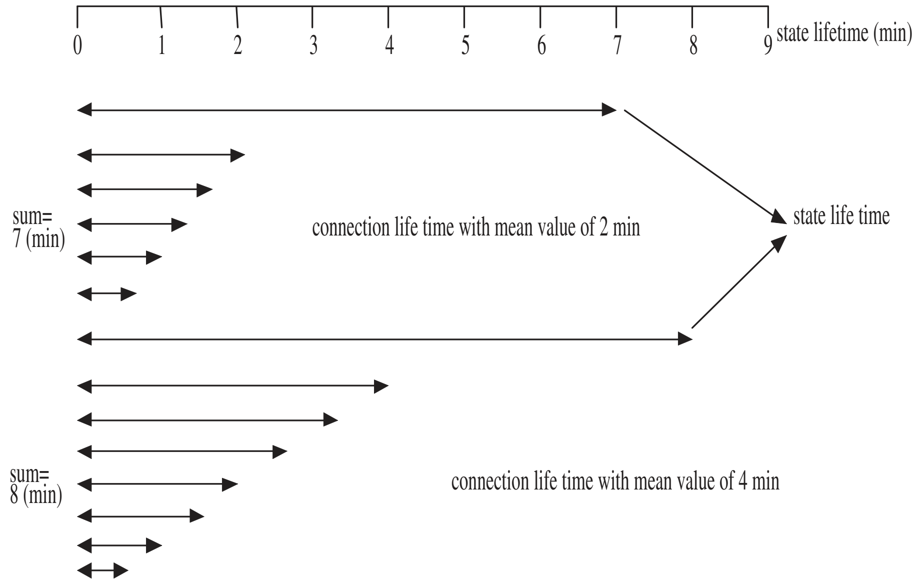


Figure 9 State lifetime and connection lifetime.

should be less than that state lifetime. State change occurs when this sum is more than the state lifetime. To which state it has to go is decided based on the state transition matrix T. We simulated heterogeneous networking environment to evaluate the performance of vertical handover decision mechanisms by using the simulation methodology explained in the following section.

Simulation method

The simulation method using the proposed non-birth–death Markov model explained above is as follows: In each state, the best network is selected from the available networks and connections are assigned to that network. The connection lifetime of the mobile is assumed to follow an exponential distribution. The average connection lifetime (μ) is varied from 1 to 10 min. State lifetime (state changing time) is also assumed to follow an exponential distribution with mean equal to λ_i (where i is 1 for UMTS, 2 for GPRS and 3 for WLAN). The exponential distribution is generally used to characterize an independent set of events. In this article also, we assumed the state and connection lifetimes to follow exponential distributions with different mean values, for illustrating the proposed approach. These lifetime values are chosen to represent the different scenarios. It is assumed that $\lambda_1 = 4$ min and $\lambda_2 = 3$ min, whereas $\lambda_3 = 1$ min [1]. Choice of different mean values for the state lifetimes helps in reflecting different sojourn times within the networks. As coverage area is different for various radio access technologies, we assumed different state lifetimes for different networks. There are eight possible combinations of networks available to the mobile at any instant of the time as explained in Figure 7. The Markov chain for simulation is shown in Figure 8. In the simulation, we assume that the $\lambda_{ij} = \lambda_i$ where i is the network selected from the previous state and $\mu_{ij} = \mu$ as explained by (2). The generator matrix for the Markov chain is obtained by solving the equilibrium conditions [31]. The state changes evolve according to the Markov chain with the adaptive state transition matrix T, given by

$$state\_transition \;\; T = (\frac{G}{q}) + I \qquad (4)$$

where G is the generator matrix, $q > \max(|G_{ij}|), 0 \leq i, j \leq 7$, I is the identity matrix of size 8×8.

A state change occurs only when the sum of the connection lifetimes is more than the state lifetime. The next state is decided based on the state transition matrix T; the state transitions in simulation is better explained by the pseudocode given in Algorithm 1. In this algorithm, $From \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ is the present state and $To \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ is the future state. In each state, different vertical handoff mechanisms including proposed FRB mechanism are evaluated using the available network state parameters given by $\{BER^N, E2EDelay^N, Jitter^N, BW^N\}$ where $N \in \{UMTS, GPRS, WLAN\}$. These network state parameters are obtained randomly from the vector set defined for each network and these network parameter sets are common for all the mechanisms within a particular state.

Algorithm 1 Pseudocode for state transitions insimulation

```
ran ← rand(1);                         ▷ random function which generates a value between 0 to 1
$current\_state = From$;                 ▷ the current state number is from
if $sum\_of\_conn\_lifetime > state\_lifetime$ then
    $sum\_of\_conn\_lifetime = 0$;
    $sum = 0$;
    for To = 0 → 7 do                    ▷ To is the next state number
        $sum = sum + state\_transition(From, To)$;  ▷ state transition matrix T is used to calculate the probability
        if $sum \geq ran$ then
            $next\_state$ ← To;           ▷ next state
            $break$;                      ▷ exit the loop
        end if
        To ← To + 1
    end for
else
    $next\_state$ ← From;                 ▷ next state number is the current state
    $sum\_of\_conn\_lifetime \leftarrow sum\_of\_conn\_lifetime + conn\_lifetime$;   ▷ increment the connection lifetime
end if
```

The QoS parameters considered for each of the four traffic classes are available bandwidth, E2EDelay, BER and jitter, with the corresponding importance weight for each traffic class computed using the AHP technique. The weights are as shown in Table 3 [1]. We compare three existing vertical handoff algorithms: SAW, TOPSIS and MEW with the proposed algorithm. The results are presented in "Results and discussions" section. In the following section, the steady-state analysis that validates the proposed evaluation model is presented.

Steady-state analysis

The steady-state probabilities of finite Markov chains can be determined using the following approaches for the solution of a linear system of the form $xA = 0$. (1) Direct numerical methods; (2) Iterative numerical methods; (3) Techniques that yield closed-form results. Each type of numerical method has merits of its own. Whereas direct methods yield exact results, iterative methods are generally more efficient, both in time and space. Disadvantages of iterative methods are that for some of these methods convergence is not assured in general and determination of suitable error bounds for termination of the iterations is not always easy. Since iterative methods are considerably more efficient in solving Markov chains, they are commonly used for larger models. For models with less than a few thousand states, direct methods are reliable and accurate. Though closed-form results are highly desirable, they can be obtained only for a small class of models that have some structure in their coefficient matrix [31]. In this article, the solutions for steady-state probabilities are obtained using the direct method. Considering the state probabilities as $\pi_i, 0 \leq i \leq 7, \lambda$ as the state lifetime, and μ as the connection lifetime within a state, the following can be inferred using global balance equations of the states.

$$-\pi_0(7\mu) + \mu\left(\sum_{i=1}^{7}\pi_i\right) = 0 \tag{5}$$

$$-\pi_k(k\lambda + (8-k-1)\mu) + \lambda\sum_{i=0}^{k-1}\pi_i + \mu\sum_{i=k+1}^{7}\pi_i = 0 \tag{6}$$

$$\sum_{i=0}^{7}\pi_i = 1 \tag{7}$$

From (7) we have $\sum_{i=1}^{7}\pi_i = 1 - \pi_0$ and substituting this in (5) we get,

$$-\pi_0(7\mu) + \mu(1-\pi_0) = 0$$

$$-\pi_0(8\mu) + \mu = 0$$

$$\pi_0 = \frac{1}{8} \tag{8}$$

Solving (6) for π_k,

$$-\pi_k(k\lambda+(8-k-1)\mu)+\lambda\sum_{i=0}^{k-1}\pi_i+\mu\left(1-\pi_k-\sum_{i=0}^{k-1}\pi_i\right)=0$$

$$-\pi_k(k\lambda+(8-k-1)\mu+\mu)+(\lambda-\mu)\left(\sum_{i=0}^{k-1}\pi_i\right)+\mu=0$$

$$\pi_k=\frac{\mu+(\lambda-\mu)\times(\sum_{i=0}^{k-1}\pi_i)}{\mu+k\lambda+8\mu-k\mu-\mu}$$

$$\pi_k=\frac{\mu+(\lambda-\mu)\times(\sum_{i=0}^{k-1}\pi_i)}{8\mu+k(\lambda-\mu)} \tag{9}$$

These state probabilities can be used to calculate performance measures such as the mean state lifetime, the mean connection time and the average state change time (which is the mean handoff delay time) for these systems using a Markov Reward Model (MRM). MRMs have long been used in Markov decision theory to assign cost and reward structures to states of Markov processes for optimization. With MRMs, rewards can be assigned to states or to transitions between states of a continuous time Markov chain. MRMs provide a unifying framework for an integrated specification of model structure and system requirements. The reward rates are defined based on the system requirements. Let the reward rate r_i be assigned to state $i \in S$. Then, a reward $r_i\ t_i$ is accrued during a sojourn of time t_i in state i. Once the model structure has been defined so that the infinitesimal generator matrix is known, the basic equations can be written depending on the given system requirements and the structure of the matrix. Consider an example of Markov reward assignment to calculate the average connection lifetime within states and average state processing time per connection as shown in Table 5. Utilization measurements describe availability of the network at the steady state. Here, Markov reward assignments are presented as an example for performability analysis of steady-state measures. The average lifetime within state, mean state processing delay per connection, expected number of state changes and utilization are calculated as follows.

$$\textit{Average lifetime within state} = E[z]$$

$$=\sum_{i=0} r_i\pi_i = \mu(\pi_7+\pi_6+\cdots+\pi_1)$$

$$=\mu\sum_{i=0}^{k-1}\pi_i \tag{10}$$

Table 5 Markov reward (r_i) assignments

State i	Average lifetime within state	Average state processing time[1]	Utilization measurements
7	μ	3n	1
6	μ	2n	1
5	μ	2n	1
4	μ	2n	1
3	μ	1n	1
2	μ	1n	1
1	μ	1n	1
0	0	0	0

[1] n delay units assumed for a network.

Mean state processing delay or mean handoff delay per connection,

$$E(t) = [\,3\pi_7+2(\pi_6+\pi_5+\pi_4)+1(\pi_3+\pi_2+\pi_1)]\,n \tag{11}$$

where n is the number of delay units for a network.

If λ is average state lifetime, total number of connections is N. Assuming $n = 1$ delay unit and using Little's law, the expected number of state changes $E(n)$ is calculated as

$$E(t) = \lambda \times E(n) \times N$$

$$E(n) = \frac{1}{\lambda} \times E(t) \times N$$

Hence,

$$E(n) = \frac{1}{\lambda}\times(3\pi_7+2(\pi_6+\pi_5+\pi_4)+1(\pi_3+\pi_2+\pi_1))\times N \tag{12}$$

Utilization,

$$U=\sum_{i=0} r_i\pi_i = \pi_7+\pi_6+\pi_5+\pi_4+\pi_3+\pi_2+\pi_1 \tag{13}$$

Results of steady-state analysis are presented in "Results and discussions" section.

Results and discussions

The simulation is carried out using Matlab and the results obtained are plotted in Figures 10, 11 and 12, respectively, for average E2EDelay, availability and average available bandwidth. For conversational, streaming and interactive traffic classes, average E2EDelay and availability are obtained for various vertical handoff algorithms. Availability is defined as the probability that the mobile is in any state other than state '0'. The mean value is obtained by taking the average over 10,000 connections. Average available bandwidth and average E2EDelay values are obtained with 95% confidence level where the margin of error is 1.86. State '0' is assumed to be associated with a bandwidth of zero, delay of 500 ms, and BER and Jitter of zero values.

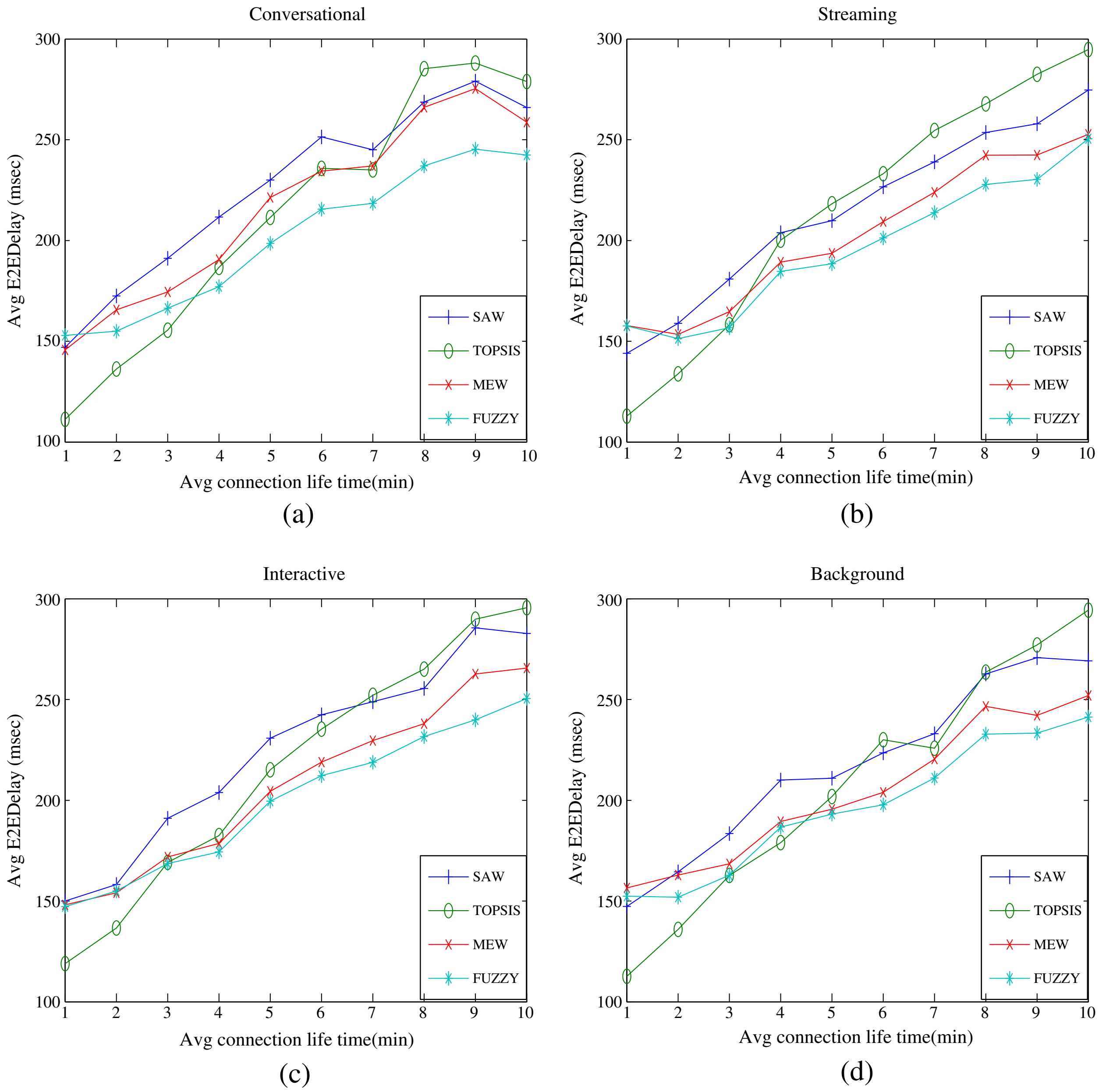
State changes evolve according to the Markov chain with the adaptive state transition matrix T following (4) as explained in "Simulation method" section. A mobile can be in any state as the state changes. In any state the best network is selected from the available networks. The best network will be decided based on handoff score value calculated for the networks. The handoff score value is calculated using various handoff decision algorithms. The average connection lifetime is varied from 1 to 10 min. Connections are distributed within a state using exponential distribution with a certain mean value. To see the effect of various traffic characteristics, the average connection lifetime is varied from 1 to 10 min. The fuzzy rules are effected differently for different types of applications based on their QoS requirements. As the best network will be selected from the networks available in any state, average state lifetime is decided based on the network selected from the previous state. Different set of parameters are assumed for different networks as explained in Algorithm 1. To incorporate the effect of a change in network parameters and a change in scenario, the average lifetime is assumed to be different for different networks. For UMTS and GPRS networks, the state lifetime is assumed to be 4 and 3 min, respectively, whereas for the WLAN, the state lifetime is assumed to be 1 min. In any state, for each network the bandwidth, delay, jitter and BER values are assigned randomly from the possible values for that network.

From Figure 10 it is observed that if average connection lifetime increases, average E2EDelay also increases. Up to a connection lifetime of 3–4 min, the average E2EDelay performance is less with the proposed FRB approach (FUZZY) compared to TOPSIS, because the average connection lifetime is less than the state lifetime. However, for background kind of applications, up to 5 min of average connection lifetime, the performance of TOPSIS is better than FUZZY. As the average connection lifetime increases beyond 4 min, the state changes occur and now the proposed mechanism gives a better performance compared to other vertical handover mechanisms. Till 4 min of average connection lifetime, the average E2EDelay is within the range of 120 to 220 ms with all the vertical handover mechanisms; the rate of change in E2EDelay is small for all the approaches. Above 4 min of average connection lifetime, even though the rate of change of E2EDelay is more, the rate of change is less for the FUZZY mechanism compared to other vertical handover mechanisms.

From Figure 11 it is observed that availability decreases as the average connection lifetime increases. For ubiquitous support, a network should be available at all the times. The availability is more using the proposed FUZZY logic-based approach compared to SAW, TOPSIS and MEW. Moreover, availability performance is better for conversational and interactive traffic classes compared to traffic generated by streaming applications. The availability is moderate for all the traffic classes using SAW and MEW whereas it is poor for TOPSIS. Till 4 min of average connection lifetime the availability does not fall less than 65 and 70% for TOPSIS, and SAW, respectively, in conversational and streaming kind of applications. Whereas for MEW, it does not fall less than 75% and for FUZZY it is about 80% while the availability does not fall below 75% for both of these mechanisms in streaming kind of applications. In interactive and background kind of applications, till 4 min of average connection lifetime, the availability does not fall below 65 and 75% for TOPSIS and SAW mechanisms. Whereas, the availability does not fall below 75% for MEW and FUZZY mechanisms. From Figure 11, it is observed that for conversational and interactive kind of applications, the availability for FUZZY does not fall below 80% while for streaming and background kind of applications it does not fall below 75%.

From Figure 12, it is observed that the average available bandwidth is more for SAW compared to TOPSIS, MEW and FUZZY approaches for the streaming and background traffic classes. Although MEW and FUZZY give nearly the same bandwidth performance, compared to the other two methods their performance is moderate. TOPSIS gives poor performance in all the cases. Till 3 min of average connection lifetime, the average available bandwidth is 1750–2250 kbps for SAW and 700–1200 kbps for FUZZY. However, above 3 min of average connection lifetime, the average available bandwidth maintains steady value for FUZZY while for SAW it decays rapidly to 1400 kbps. Till 3–4 min of average connection lifetime, the average available bandwidth is 900–1400 kbps for MEW in streaming, interactive, and background kind of applications while it is 1300–1700 kbps in conversational traffic. Whereas, the average available bandwidth is 1100–500 kbps in TOPSIS for all kind of applications. Above 4 min of average connection lifetime the average available bandwidth nearly maintains steady value for MEW, and TOPSIS.

From the results obtained using different vertical handoff techniques, it is concluded that the proposed FUZZY rule-based technique gives better performance for different types of applications. The QoS-aware FRB approach will make a clear decision regarding implementing handoff among the networks. The fuzzy membership regions help in making a clear distinction among the parameter values of the networks and the fuzzy rule base is used to compute the handoff score value. The SAW, TOPSIS and MEW mechanisms follow simple additive or multiplicative approaches. These mechanisms require information about the relative importance of each of the QoS parameters as explained in Table 3. It is usually given by a set of weights. In these approaches, the parameter values of networks are weighted by these values. These weights are



(a)

(b)

(c)

(d)

Figure 10 Average E2EDelay (ms) to average connection life-time (minutes) in comparison with SAW, TOPSIS, MEW, FUZZY. **(a)** Conversational. **(b)** Streaming. **(c)** Interactive. **(d)** Background.

derived based on the QoS requirements of various traffic classes. These techniques are devoid of any intelligence to implement QoS-aware handoff.

After calculating the steady-state probabilities using Equations (8) and (9), it is verified that $1 - \pi_0 = \pi_7 + \pi_6 + \pi_5 + \pi_4 + \pi_3 + \pi_2 + \pi_1$ and $\pi_k = 0.125$ for given values for λ and μ, where $0 \leq k \leq 7$. Utilization measurements (13) carried out using the MRM with reward assignments given in Table 5) are plotted in Figure 13a. Utilization is equivalent to the steady-state value of availability. In transient analysis, as the average connection lifetime increases, the total number of state changes reduces. This is because the average lifetime within a state also increases. This will increase the relative ratio of the number of times a mobile is in state 0 to total number of times it stays in all the states. Hence, the availability decreases with an increase in the average connection lifetime. The availability analysis validates the proposed simulation model. Average lifetime within a state is plotted in Figure 13b. As average lifetime within

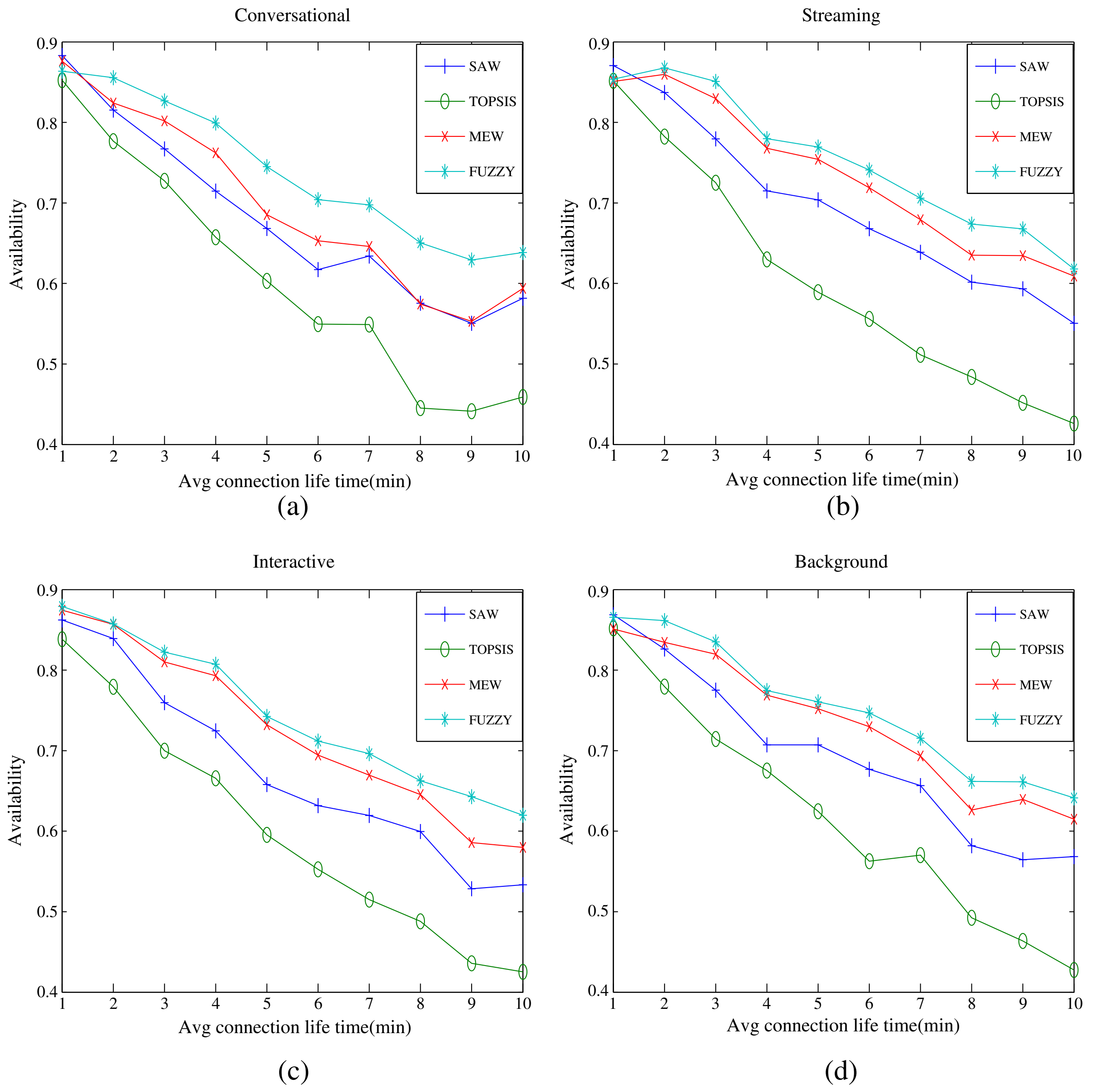


Figure 11 Availability to average connection life-time (min) in comparison with SAW, TOPSIS, MEW, FUZZY. (a) Conversational. **(b)** Streaming. **(c)** Interactive. **(d)** Background.

a state is proportional to mean connection lifetime as shown in (10), it increases linearly with an increase in average connection lifetime. Mean handoff delay, taken as the mean state processing delay per connection, is also plotted in Figure 13c. In this figure, it is assumed that each network contributes to one unit of delay. The mean handoff delay at steady state as calculated in (11) does not depend upon the connection duration and so it gives a constant value. Mean number of state changes, plotted in Figure 13d, is around 3,750 for 10,000 connections and a λ value of 4 min. Mean number of state changes at steady state is a fixed value as explained in (12).

Conclusions

This article proposes a QoS-aware FRB algorithm that makes a multi-criteria-based decision considering the available bandwidth, E2EDelay, jitter and BER of the networks for different traffic classes. The algorithm is simulated by considering the presence of three different types of networks, such as UMTS, GPRS and WLAN networks.

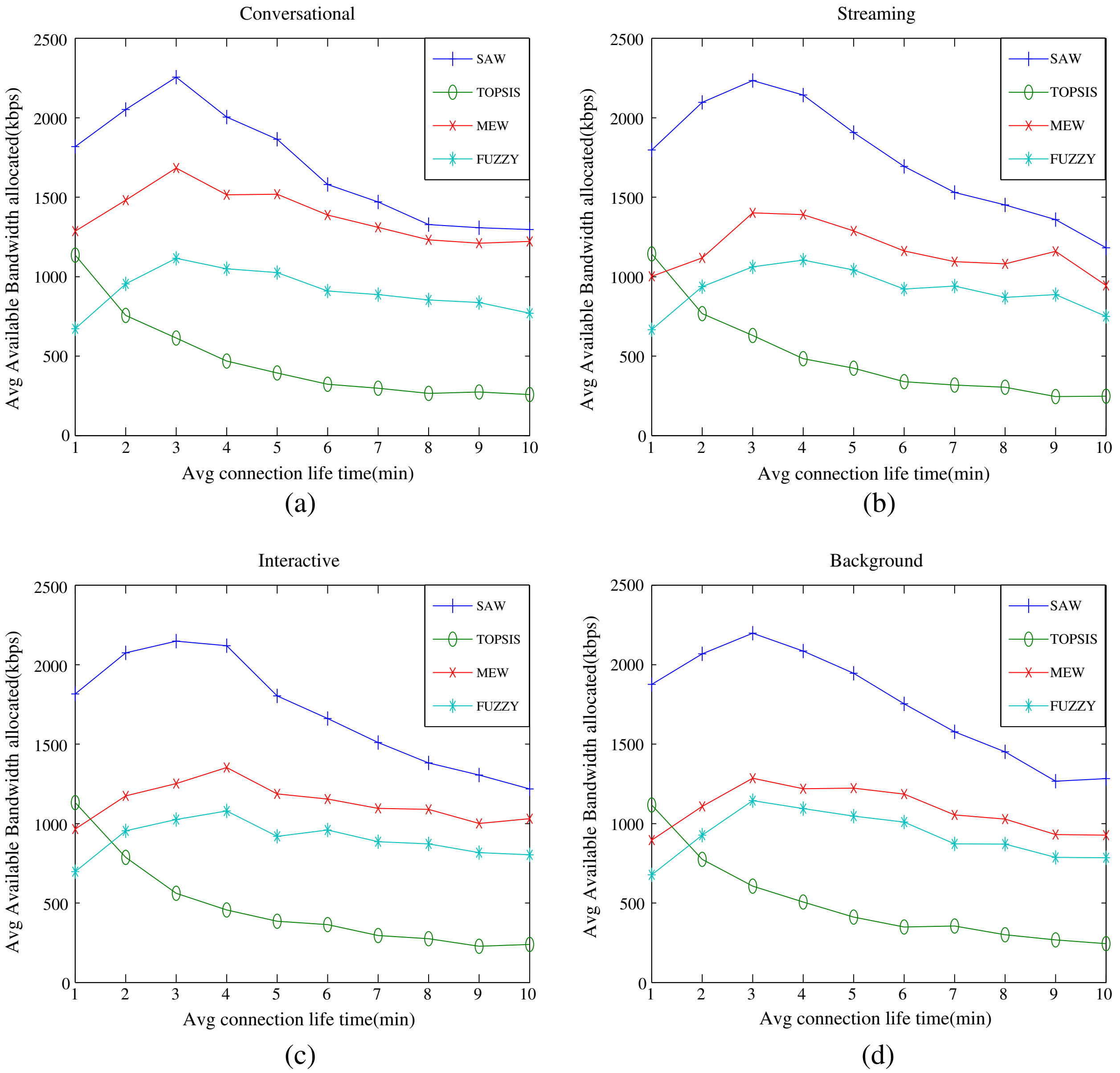


Figure 12 Average available bandwidth (kbps) to average connection life-time (min) in comparison with SAW, TOPSIS, MEW, FUZZY. **(a)** Conversational. **(b)** Streaming. **(c)** Interactive. **(d)** Background.

The article also proposes a new evaluation model using a non-birth–death Markov chain with state parameters corresponding to the available networks. The assumption of a state to represent the available networks in a heterogeneous networking environment is more realistic. This assumption also circumvents the need to consider other system level parameters, such as the mobility patterns or velocity of mobiles during simulation.

It is likely that in heterogeneous wireless networks, numerous types of access networks will prevail to support wireless services that have varied QoS requirements. From the requirements of IMT2000 QoS classes [20,21], the conversational, streaming and interactive traffic classes expect less delay. Applications like conversational, interactive video conferencing and live streaming require more network availability, less E2EDelay with tolerable bandwidth. Results obtained using the proposed technique show better performance for E2EDelay and network availability. Average available bandwidth is also obtained for streaming and background classes. The available bandwidth performance is moderate while satisfying the E2EDelay and availability requirements. Hence, it is concluded that the proposed QoS-aware FRB algorithm gives a good QoS performance for delay

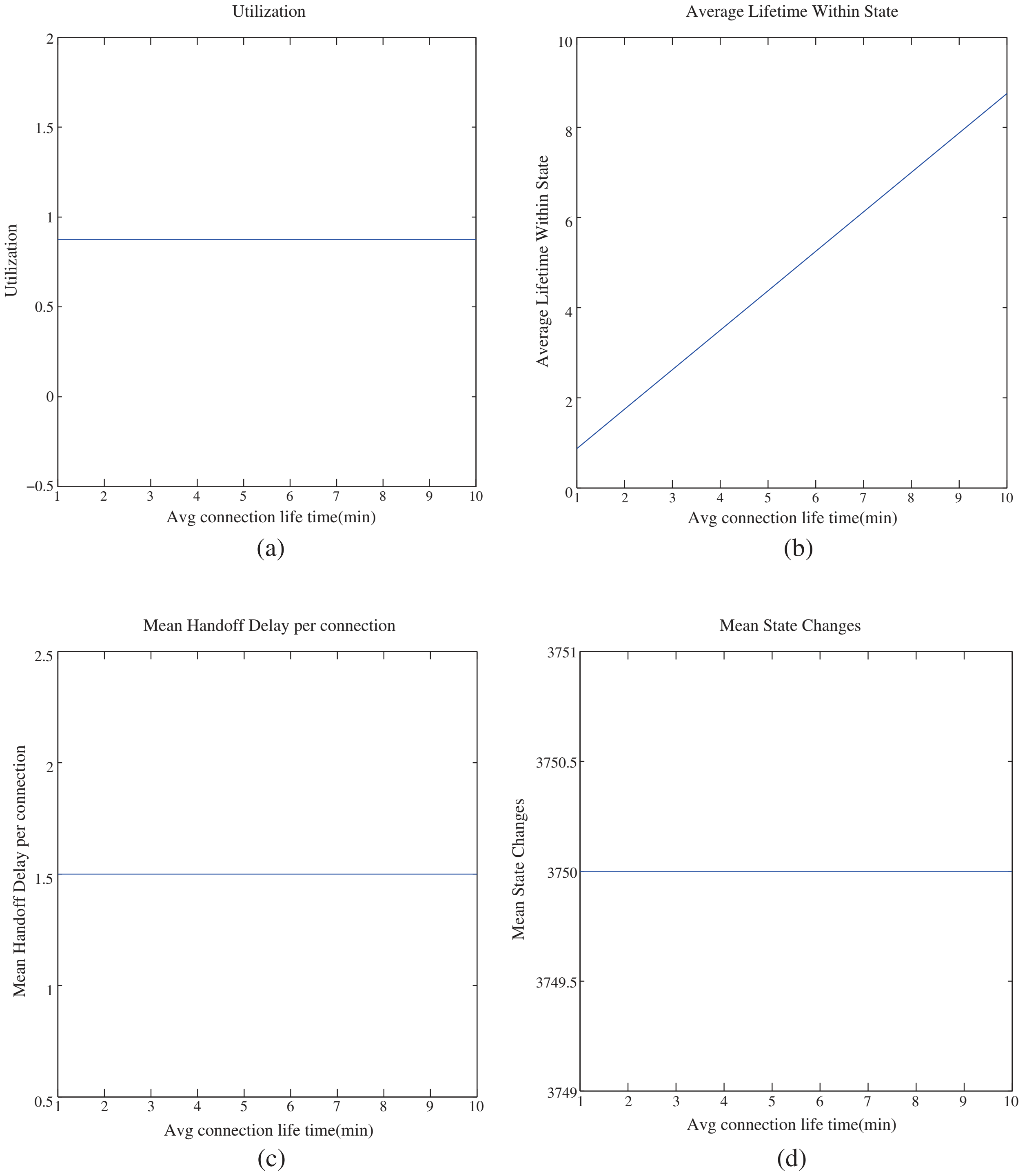


Figure 13 Steady-state parameters to average connection life-time (minutes). (a) Utilization. **(b)** Average state lifetime (min). **(c)** Mean handoff delay per connection (min). **(d)** Mean number of state changes.

sensitive applications like conversational, interactive and live streaming applications. It is demonstrated that the proposed evaluation model using a non- birth–death Markov chain with states representing the available networks can be used for comparing different vertical handoff mechanisms.

Competing interests

The authors declare that they have no competing interests. However, they have filed patent at an Indian office in relation to some aspects of the work for which they have already got the filing date.

Acknowledgements

This study was carried out under the Vodafone Essar-sponsored research project on Wireless Internet at IIT Kharagpur, India.

Received: 6 April 2012 Accepted: 5 September 2012
Published: 26 October 2012

References

1. E Stevens-Navarro, V Wong, in *IEEE 63rd Vehicular Technology Conference, 2006. VTC 2006-Spring*, vol. 2 Comparison between vertical handoff decision algorithms for heterogeneous wireless networks. (Melbourne, Australia, 2006), pp. 947–951
2. I Smaoui, F Zarai, L Kamoun, in *ITI 5th International Conference on Information and Communications Technology, 2007. ICICT 2007*, vol. 1 Vertical handoff management for next generation heterogeneous networks. (Cairo, Egypt, 2007), pp. 19–25
3. S Xie, M Wu, Vertical handoff algorithm in heterogeneous networks for reducing interference. J. Electron. (China). **26**, 71–79 (2009)
4. N Shenoy, S Mishra, in *Heterogeneous Wireless Access Networks*, ed. by E Hossain Vertical handoff and mobility management for seamless integration of heterogeneouswireless access technologies. (Springer, New York, 2009), pp. 1–33
5. K Yang, Y Wu, HH Chen, QoS-aware routing in emerging heterogeneous wireless networks [quality-of-service-based routing algorithms for heterogeneous networks]. IEEE Commun. Mag. **45**(2), 74–80 (2007)
6. E Tragos, G Tsiropoulos, G Karetsos, S Kyriazakos, Admission control for QoS support in heterogeneous 4G wireless networks. IEEE Netw. **22**(3), 30–37 (2008)
7. SR Yang, WT Chen, SIP multicast-based mobile quality-of-service support over heterogeneous IP multimedia subsystems. IEEE Trans. Mob. Comput. **7**(11), 1297–1310 (2008)
8. K Samdanis, A Aghvami, Scalable inter-area handovers for hierarchical wireless networks. IEEE Wirel. Commun. **16**(6), 62–68 (2009)
9. Q Guo, J Zhu, X Xu, in *IEEE International Conference on Communications, 2005. ICC 2005*, vol. 4 An adaptive multi-criteria vertical handoff decision algorithm for radio heterogeneous network. (Seoul, Korea, 2005), pp. 2769–2773
10. A Singhrova, N Prakash, in *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology, Mobility '07* A review of vertical handoff decision algorithm in heterogeneous networks. (ACM, New York, 2007), pp. 68–71. doi:10.1145/1378063.1378075
11. BS Ghahfarokhi, N Movahhedinia, A context-aware handover decision based on user perceived quality of service trigger. Wirel. Commun. Mob. Comput. **11**(6), 723–741 (2011). doi:10.1002/wcm.854
12. JS Kim, E Serpedin, DR Shin, K Qaraqe, Handoff triggering and network selection algorithms for load-balancing handoff in CDMA-WLAN integrated networks. EURASIP J. Wirel. Commun. Netw. **2008**, 136939 (2008). http://jwcn.eurasipjournals.com/content/2008/1/136939
13. A Ezzouhairi, A Quintero, S Pierre, in *Canadian Conference on Electrical and Computer Engineering, 2008. CCECE 2008* A fuzzy decision making strategy for vertical handoffs. (Ontario, Canada, 2008), pp. 000583–000588
14. Q He, in *2nd International Conference on Networking and Digital Society (ICNDS), 2010*, vol 2 A fuzzy logic based vertical handoff decision algorithm between WWAN and WLAN. (Wenzhou, China, 2010), pp. 561–564
15. T Yang, P Rong, in *6th International Forum on Strategic Technology (IFOST), 2011*, vol. 2 A fuzzy logic vertical handoff algorithm with motion trend decision. (Harbin, China, 2011), pp. 1280–1283
16. A Aziz, S Rizvi, N Saad, in *International Conference on Intelligent and Advanced Systems (ICIAS), 2010* Fuzzy logic based vertical handover algorithm between LTE and WLAN. (KLCC 2010), pp. 1–4
17. KS Trivedi, S Dharmaraja, X Ma, Analytic modeling of handoffs in wireless cellular networks. Inf. Sci. **148**(1-4), 155–166 (2002). http://www.sciencedirect.com/science/article/pii/S002002550200292X
18. Gowrishankar, GN Sekhar, PS Satyanarayana, Analytic performability model of vertical handoff in wireless networks. J. Comput. Sci. **5**, 1549–3636 (2009). doi:10.3844/jcssp.2009.445.450
19. Gowrishankar, H Ramesh Babu, G Raju, P Satyanarayana, in *The Fourth International Conference on Wireless and Mobile Communications, 2008. ICWMC'08* Performability model of vertical handoff in wireless data network. (Athens, Greece, 2008), pp. 309–314
20. 3GPP: IMT -2000 QoS classes. TSG 17, 3rd Generation Partnership Project (3GPP)
21. 3GPP: QoS concepts and architecture. TS 23.107, 3rd Generation Partnership Project (3GPP) 2009
22. E Stevens-Navarro, U Pineda-Rico, J Acosta-Elias, Vertical handover in beyond third generation (3G) wireless networks. Int. J. Future Generation Commun. Netw. **1**, 1–8 (2010)
23. K Teknomo, Analytic hierarchy process (AHP) tutorial. http://people.revoledu.com/kardi/tutorial/AHP/index.html
24. TL Saaty, *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process. The Analytic Hierarchy Process*, vol. 6 (RWS Publications, Pitsburgh, 2000)
25. J Stein, Survey of IEEE802.21 media independent handover services, http://www.cse.wustl.edu/jain/cse574-06/ftp/handover/index.html
26. E Stevens-Navarro, Y Lin, V Wong, An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks. IEEE Trans. Veh. Technol. **57**(2), 1243–1254 (2008)
27. *Towards a Media Independent Handover Based Approach to Heterogenous Network Mobility* (Performance Engineering Lab, Derry, Northern Ireland, 2007)
28. J Mendel, Fuzzy logic systems for engineering: a tutorial. Proc. IEEE. **83**(3), 345–377 (1995)
29. S Rajasekaran, G Pai, *Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications* (Prentice-Hall of India, 2004), http://books.google.co.in/books?id=HDg7wCP_bmUC
30. A Celikyilmaz, I Türkşen, *Modeling Uncertainty with Fuzzy Logic: With Recent Theory and Applications*. Studies in Fuzziness and Soft Computing (Springer, 2009), http://books.google.co.in/books?id=XNHzqfvMTbwC
31. G Bolch, *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications* (Wiley-Interscience Publication, Wiley-Interscience, 2006), http://books.google.co.in/books?id=8Mei8w6YUHYC

doi:10.1186/1687-1499-2012-322
Cite this article as: Vasu *et al.*: **QoS-aware fuzzy rule-based vertical handoff decision algorithm incorporating a new evaluation model for wireless heterogeneous networks.** *EURASIP Journal on Wireless Communications and Networking* 2012 **2012**:322.