Network Selection in Heterogeneous Environment: A Step toward Always Best Connected and Served

K. R. Rao[1], Zoran S. Bojkovic[2], Bojan M. Bakmaz[3]

Abstract – **Service delivery in a heterogeneous wireless environment requires the selection of an optimal access network. In order for users to be always connected through the optimal available network, it is necessary to develop an adequate technique for its selection. This paper deals with network selection concept as a perspective approach to the always best connected and served paradigm in heterogeneous wireless environment. It provides a survey on fundamental aspects of network selection process and mainly focuses on decision making mechanisms. After presenting the related standards, the perspective vertical handover approaches from the open literature are analyzed and compared. The paper concludes with a case study related to the influence of traffic conditions on network selection, which reveals the advantages of the always best connected and served concept.**

Keywords – **Decision making, Handover, Mobility Management, Network selection.**

I. Introduction

Growing consumer demands for access to multimedia services from anywhere at any time is accelerating the technological development toward the integration of wireless heterogeneous networks (HetNets). This next-generation of wireless multimedia systems represents an environment with different radio access network (RAN) technologies that differ in terms of their general characteristics such as coverage, traffic conditions, security, cost and quality of service (QoS). However, differences can exist among networks with the same architecture. For example, two WiFi networks based on the same standard can be different in terms of security and QoS as well as the cost of service.

In this kind of environment, handover management is the essential issue that supports the mobility of users from one system to another. Handover management as one of the mobility management components controls the change of the mobile terminal's point of attachment (PoA) during active communications. Handover management includes mobility scenarios, metrics, decision algorithms and procedures [1].

In homogeneous networks, handover is typically required when the serving PoA becomes unavailable due to user's movement, while the need for vertical (hyper) handover can be initiated for convenience rather than connectivity reasons. Major challenges in vertical handover management are seamlessness and automation aspects in network switching. These specific requirements can refer to the Always Best Connected (ABC) concept [2]. ABC represents a vision of fixed and mobile wireless access as an integral and challenging dimension in developmental paradigm of the next generation wireless networks. It is a strategic goal to define important advancements that happen and are predicted in technologies, networks, user terminals, services, and future business models that include all these issues while realizing and exploiting new wireless networks.

On the other hand, because users can be always connected through the optimal RAN, it is necessary to develop an adequate mechanism for its selection. Since some other parameters must be taken into consideration beside the traditional received signal strength (RSS) and signal to interference and noise ratio (SINR), it is possible that the problem can be pointed out from the aspect of multiple criteria analysis.

After a decade of intensive theoretical and practical application of ABC concept, the logical evolution step toward HetNets development will be a rise of Always Best Connected and Served (ABC&S) paradigm. Concerning this concept, it will no longer be enough for a user to be able to connect anywhere and at any time, but also to be connected in the best possible way to be provided with the best possible service [3].

This paper starts with a synopsis of handover management related standards. After that, definition and systematization of the decision criteria for network selection is provided. Next, the influences of users' preferences together with services' requirements are envisaged. After that, the existing studies on network selection using cost function, multiple attribute decision making, fuzzy logic, artificial neural networks, etc. are systematically discussed. Finally, traffic parameters influences on network selection are envisaged as perspective contribution to the always best connected and served vision of heterogeneous environment.

II. Related Standardization Activities

Network selection is one of the most significant challenges for the next-generation wireless networks. Thus, it draws researchers and standardization bodies attention.

ITU's concept of Optimally Connected, Anywhere, Anytime proposed in M.1645 [4] states that future wireless networks can be realized through the coalition of different RANs. According such a scenario, the heterogeneity of access networks, services and terminals should be fully exploited to enable higher utilization of radio resources. The main

[1]K. R. Rao is with the University of Texas at Arlington, 416 Yates Street, UTA, Box 19016, Arlington, Texas 76019, E-mail: rao@uta.edu

[2]Zoran S. Bojkovic is with the University of Belgrade, E-mail: z.bojkovic@yahoo.com

[3]Bojan M. Bakmaz is with the Faculty of Transport and Traffic Engineering, University of Belgrade, V. Stepe 305, 11000 Belgrade, Serbia, E-mail: bakmaz@ieee.org

objective is to improve overall networks performances and QoS perceived by users.

Third Generation Partnership Project (3GPP) is defining an Access Network Discovery and Selection Function (ANDSF) [5] to assist mobile terminals in vertical handover between 3GPP and non-3GPP networks, covering both automated and manual selection as well as operator and user management.

IEEE 802.21 TG is developing standards to enable handover and interoperability between heterogeneous link layers [6]. This standard defines the tools required to exchange information, events and commands to facilitate handover initiation and handover preparation. IEEE 802.21 standard does not attempt to standardize the actual handover execution mechanism.

The IEEE 1900.4 standard [7] defines the architectural building blocks enabling network-device distributed decision making for optimized radio resource usage in heterogeneous wireless access networks. Specifically, the standard aims to improve the overall composite capacity and QoS of wireless systems in a multiradio environment by defining suitable architecture and protocols to facilitate resource optimization, including dynamic spectrum access control. Radio resource selection is one of IEEE 1900.4 policies, proposed in order to guide terminals in their reconfiguration decisions.

In the network selection scenario, users are always trying to seamlessly access high-quality multimedia service at any speed, any location, and any time through selecting the optimal network. Therefore, ensuring a specific QoS is the objective in the process of network selection. A great number of techniques related to the handover initiation and optimal access network selection are proposed in the open literature. The suggested techniques are using different metrics and heuristics for solving the above mentioned problems. Unfortunately currently proposed vertical handover techniques do not completely satisfy the demands from the covering technologies point of view, as well as for analyzing parameters adequacy, implementation complexity, together with invoking the all entities to the access network selection.

III. Vertical Handover Decision Criteria

Handover criteria represent the qualities that are measured to give an indication of whether or not a handover is needed and select optimal available network. In general, the conventional single criteria-based algorithms cannot react easily to the changing environmental conditions and the possible accumulated human knowledge [8]. Usually, they cannot cope with the different viewpoints and goals of the operators, users, and QoS requirements, which make them inefficient for a multidimensional problem, such as network selection. The network selection is considered to be a complex problem, because of the multiple mix of static and dynamic, and sometimes conflicting criteria (parameters) involved in the process [9].

The decision criteria that may be considered in network selection process can be subjective or objective with minimizing or maximizing nature. From the origin point of view, they can be classified in four categories [1], [9]:

- Network-related criteria express the technical characteristics and performance of the RANs in terms of certain performance indicators (technology, coverage, link quality and capacity, throughput, load, pricing, etc.). These criteria are referred to the network and not to the service provisioning.
- Terminal-related criteria include information about the end-user's terminal characteristics and current status (supported interfaces, velocity, location-based information, energy consumption, etc.).
- Service-related criteria concern the QoS offered to the end users through a series of metrics. This is the most commonly used category in network selection problems, since it involves basic QoS metrics of packet switched networks (delay, jitter, error ratio, loss ratio).
- User-related criteria are mostly subjective and express certain aspects of the end user's satisfaction (user's profile and preferences, cost of service, quality of experience - QoE, etc.). Because of the imprecision they bear, these criteria are often expressed in linguistic terms.

Received signal strength is the most widely used criterion because of its measure simplicity and close correlation to the link quality. There is a close relationship between RSS readings and the distance from the mobile terminal to its PoA. Traditional horizontal handover techniques are basically analyzing metrics through the variants of comparing RSS of the current PoA and candidate network PoA. In combination with threshold and hysteresis, RSS metrics represent a satisfying solution for a homogeneous network environment. In a heterogeneous environment, RSS metric is not sufficient criterion for initiating a handover, but in combination with other metrics it can be applied as a compulsory condition.

The following network parameters are particularly important for vertical handover decision (network selection) and because of this, they are most commonly found in the open literature. In terms of network conditions, available bandwidth is the mostly used indicator of traffic performances and transparent parameter for the current and future users of the multimedia services. This is the measure of per user bandwidth allotted by the network operator which is dynamically changeable according the utilization of the network. The maximum theoretical bandwidth is closely related to the link capacity (total number of channels) [10]. Transition to a network with better conditions and higher performance would usually provide improved QoS and it is especially important for delay-sensitive applications.

QoS parameters like delay, jitter, error ratio and loss ratio can be measured in order to decide which network can provide a higher assurance of seamless service connectivity. The QoS level can objectively be declared by the service provider based on ITU-T recommendation Y.1541 [11] and specified parameters. By declaring the QoS level in this way, a complex examination of QoS parameters by users and the additional load of user's terminals and other network elements, can be avoided, while this criterion becomes more transparent to the user. An example of QoS levels mapping is proposed in [12].

When the information exchanged is confidential, a network with high encryption is preferred. The security level concept is similar to service level in QoS management [13]. Security level is a key piece of information within a security profile and is used to determine whether data is allowed to be transferred by a particular network or not.

The cost of service can significantly vary from provider to provider, but in different network environments. In some cases cost can be the deciding factor for optimal network selection, and it includes the traffic costs and the costs of roaming between HetNets. In some context cost of service is in tight relation with available bandwidth and QoS level, but in heterogeneous environment, this criterion is fast time differentiable function depending on many other parameters. Pricing schemes adopted by different service providers is crucial and will impact the decisions of users in network selection [14].

The specified criteria do not represent an exhaustive list and are possible choices that can be used as input information for the decision mechanism. Different approaches may use only a subset of the parameters, or may include additional parameters.

After the definition of the convenient parameters, the question often arises is how to transfer the metrics information from the network entities to the user's multimode terminals. Through the End to End Reconfigurability (E^2R) project, concepts and solutions for a cognitive pilot channel (CPC) were developed [15]. It was concluded that CPC will be able to bring enough information (for example proposed parameters) for network selection to the terminals, when users are preceding either initial connection or handover.

A uniform set of parameters for each candidate RAN has to be provided as input to the decision algorithm to form the basis for the network selection process. Depending on the type of architecture, and protocol in use, and whether it is a centralized or decentralized decision, different information will be available in different form and accuracy levels [9]. For example, for a decentralized approach, the mobile device can collect the network conditions information as statistics, usually represented by mean values of previous sessions, or can estimate network bandwidth, for example, through the use of IEEE 802.21 Hello packets. A mobile station can collect authentication, routing, and network condition (e.g., available throughput, average delay, average packet loss, etc.) information through advertisement Hello packets sent by a gateway node. This information can be collected from the link layer by using the IEEE 802.21 reference model [6]. Another option would be to predict the future network state based on past history. For example, many QoS parameters (e.g., availability, utilization, etc.) can be predicted depending on usage pattern statistics based on location (e.g., home, office, airport, etc.), or time related statistics (e.g., peak/off-peak hours, working days/weekends, holidays, etc.). The accuracy in collecting network state information is very important as the selection of the optimal value network depends on it. However, a trade-off between accuracy and overhead has to be taken into account, as keeping accurate estimates for the more dynamic parameters which depend on their frequency of change and can be data intensive, adding to signaling, processor and memory burden.

Another form of data prediction is the estimation of probabilities. This procedure is related to Markov chains and is most of the times associated with a limited set of decision parameters, such as SINR, blocking probability, and termination probability. One representative approach, related to traffic model with handover call termination probability estimation, is proposed in [16]. Finally, in the presence of imprecise information, network selection can be performed using sequential Bayesian estimation, which relies on dynamic QoS parameters estimated through bootstrap approximation [17]. The bootstrap method is a computer-based, nonparametric approach where no assumptions are made on the underlying population from which the samples are collected and allows estimating the probability distributions of critical QoS parameters from acquired data. This concept can be employed to provide reliable inference from incomplete network conditions information.

Network selection criteria are mainly represented in the form of decision (performance) matrix

$$\mathbf{D} = \left\| x_{ij} \right\|_{m \times n}, \tag{1}$$

where x_{ij} represents performance of i-th RAN (for $i = 1, \dots, m$), related to j-th criteria ($j = 1, \dots, n$). Here, m is a set of available RANs, and n is a set of observed criteria. In order to compare the criteria of different values and different measurement units, normalization is treated as a necessary step for most of the network selection techniques. In normalization process starting matrix (1) moves into normalized matrix

$$\mathbf{R} = \left\| r_{ij} \right\|_{m \times n}, \tag{2}$$

where r_{ij} is defined as normalized performance rating, obtained as

$$r_{ij} = \begin{cases} x_{ij} / \sum_{i=1}^{m} x_{ij}, \\ x_{ij} / \sqrt{\sum_{i=1}^{m} x_{ij}^2}, \\ (x_{ij} - \min_i(x_{ij})) / (\max_i(x_{ij}) - \min_i(x_{ij})). \end{cases} \tag{3}$$

IV. Influence of Services' Requirements and Users' Preferences

Once the decision criteria have been determined, the next step is to define their importance, i.e., weight, of each one of them in the final outcome. Weights are differentiated based on context, since each user or application type may bear different requirements. The users' preferences and services' requirements play an important role in the decision mechanism and they may be used to weight the involved criteria.

In certain cases encountered in the literature, the weights of the selection criteria are defined through the derivation and the analysis of questionnaires, which capture the user's overall perception of a service. However, these approaches depend only on user's feedback to determine the relative weights and thus cannot be considered precise, since user's perception and opinion is subjective. Obviously, this approach is adequate when subjective criteria are considered. However, in the case of objective criteria, no accurate results can be guaranteed. For example, some of the existing weighted solutions obtain the weights through questionnaires on users' and services' requirements. Other solutions integrate a user interface in the mobile terminals in order to collect their preferences. An important aspect is to find balance between the cost of involving the user and the decision mechanism. One solution for minimizing the user interaction may be implementing an intelligent learning mechanism that can estimate the users' preferences over time.

A. Subjective Criteria Weights Estimation

In several network selection mechanisms, the use of the Analytical Hierarchy Process (AHP) for subjective criteria weight estimation has been proposed [18]-[20]. This method is considered as a well-known and proven mathematical tool. AHP is defined as a procedure to divide a complex problem into a number of deciding factors and integrate the relative dominances of the factors with the solution alternatives to find the optimal one [21]. AHP criteria weights estimation procedure is carried out in three steps:

(1) Structuring a problem as a decision hierarchy of independent elements.

(2) Compares each element to all the other within the same level through pairwise comparison matrix. The comparison results are presented in a square matrix form

$$\mathbf{A} = \left\| a_{ij} \right\|_{n \times n}, \tag{4}$$

where n is number of observed criteria, and stands that $a_{ij} = 1$, (criteria is compared with itself), $a_{ij} > 1$ (element i is considered to be more important than element j), $a_{ij} < 1$ (element j is considered to be more important than element i), and $a_{ij} = 1/a_{ji}$.

(3) Normalization and calculation of the relative weights using the relation

$$w_j = \sum_{j=1}^{n} a_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}, \quad \sum_{j=1}^{n} w_j = 1. \tag{5}$$

AHP method can be used stand alone for network selection [19], or in combination with other methods [18], [20]. Despite its popularity, the conventional AHP is often criticized for its inability to adequately handle the inherent uncertainty and imprecision associated with the mapping of the decision-maker's perception to exact numbers.

B. Objective Criteria Weights Estimation

As an exact and objective approach, the entropy method was proposed in [22]. This method is based on relation

$$e_j = \left[-1/\ln(m)\right] \cdot \sum_{i=1}^{m} \left[r_{ij} \ln(r_{ij})\right], \quad j \in \{1, \dots, n\}. \tag{6}$$

Deviation within each criterion $d_j = 1 - e_j$, leads to the weight coefficients determination

$$w_j = d_j / \sum_{j=1}^{n} d_j, \tag{7}$$

in the case when user equally prefers all the parameters (objective approach), or

$$w_j = d_j w_j^* / \sum_{j=1}^{n} d_j w_j^*, \tag{8}$$

if the user determines the subjective weights w_j^*.

After objective or subjective criteria weights estimation, in general, normalized matrix (2) moves into weighted matrix

$$\mathbf{V} = \left\| w_j r_{ij} \right\|_{m \times n} = \left\| v_{ij} \right\|_{m \times n}. \tag{9}$$

V. Alternatives Ranking Techniques

The ranking process, through which the optimal choice is pointed out, is based on the input of the previous two steps, as presented in Fig. 1.

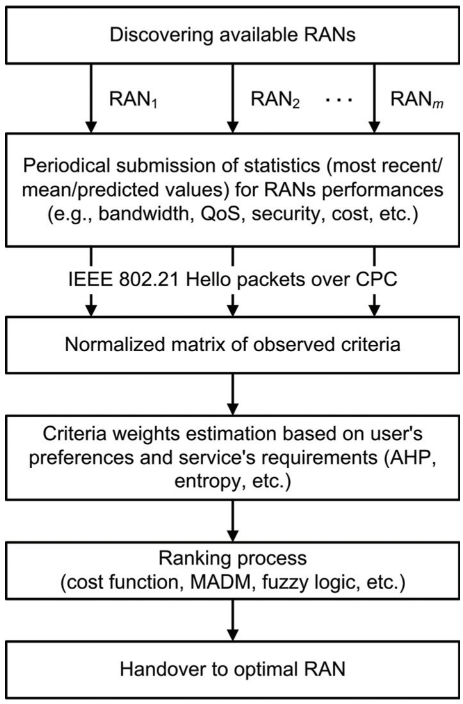

Fig. 1. General model of network selection process

This process can be considered as the core phase of the vertical handover management since it is in charge of evaluating and deciding the most appropriate network choices in order to fulfill both system's and users' requirements, thus providing the desired seamless communication. Due to the different possible strategies and the numerous parameters involved in the process, researchers have tried many different techniques in order to find the most suitable network selection solution.

These techniques are usually called vertical handover decision (VHD) algorithms. Recently, many VHD techniques were proposed in the open literature. Current techniques do not fully comply with requirements related to the technology coverage, adequacy of the analyzed parameters, complexity of implementation and integration of all the entities in the selection of appropriate RAN. Consequently, because of its complexity they are only representing theoretical solutions, which are currently the focus of many researchers, but for a number of deficiencies are not yet applicable in real environment.

Performance analysis of the network selection techniques can be performed through the determination of mean and maximum handover delays, number of handovers, number of handovers failed due to the incorrect decisions, handover failure probability, resource utilization, etc. [22].

Handover delay refers to the duration between the initiation and completion of the handover process. It is related to the complexity of the applied heuristic. Reduction of the handover delay is especially important for delay-sensitive voice and multimedia sessions. Reducing the number of handovers is usually preferred, as frequent handovers can cause wastage of network resources. A handover is considered as superfluous when a mobile terminal coming back to the previous PoA is needed within certain time duration ("ping-pong" effect), and such handovers should be minimized.

A handover failure occurs when the handover is initiated, but the target network does not have sufficient resources to complete it, or when the mobile terminal moves out of the coverage area before the process is finalized. In the first case, the handover failure probability is related to the resource availability (e.g. channel availability) of the target network, while in the second case, it is related to the terminal mobility.

Resource utilization is defined as the ratio between the mean amount of utilized resources and the total amount of resources in a system. In the case of efficient channel utilization, the ratio between the mean number of channels that are being served and total number of channels in a system is taken into account.

For efficient network selection strategy the following important issues have to be fulfilled:

- only considerable parameters must be analyzed,
- equilibrium among users' preferences, services' requirements and networks' performance must be achieved,
- technique has to be reliable and transparent to the user,
- algorithm has to minimize handover latency, blocking probability and number of superfluous handovers,
- flexible and suitable implementation in real environment is necessary.

A. Cost Function Techniques

The simplest approach for networks ranking indicates that an optimal RAT is selected based on a cost function, which is a function of selected criteria. Cost function is a measurement of the benefit obtained by handover to a particular network. Usually, the cost of a network can be considered as the inverse of its utility, but the form of this inversion is related with the way to combine multiple attributes.

A general form of cost function for the network selection problem was proposed in [23]. It integrates a large number of attributes, corresponding weights, and network elimination factors through relation

$$CF_i = \sum_k (\prod_j E_{ij}^k) \sum_j f_j^k (w_j^k) N(Q_{ij}^k), \tag{10}$$

where $N(Q_{ij}^k)$ is the normalized QoS parameter Q_{ij}^k, representing the cost in the j-th criteria to carry out service k over network i, $f_j^k(w_j^k)$ is the weighting function of criteria j for service k and E_{ij}^k is network elimination factor of service. The network elimination factor is represented by values of one or infinity, to reflect whether current network conditions are suitable for requested services. For example, if a RAN cannot guarantee the delay requirement of certain real-time service, its corresponding elimination factor will be set to infinite. Thus, the corresponding cost becomes infinite, which eliminates corresponding RAN. The algorithm of the introduced technique is shown in Fig. 2.

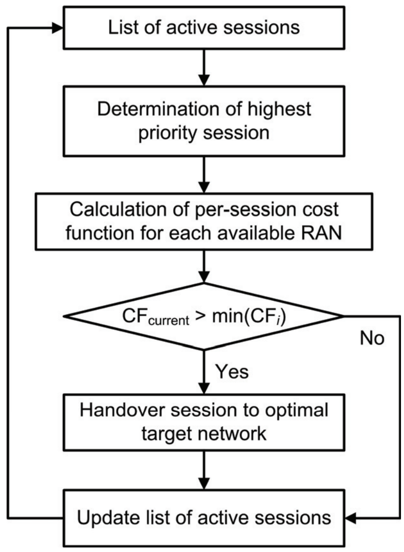

Fig. 2. Example of cost function based vertical handover algorithm

The fundamental benefit of cost function usage and handover independent initiation for different services is reduced failure (blocking) probability. However, weights coefficients estimation techniques are not discussed by the authors. Although cost functions have been widely used, currently, the need for more intelligent and dynamic schemes has been increased since more criteria must be taken into account.

B. Multi Attribute Decision Making Techniques

Multiple Criteria Decision Making (MCDM) problems, although very different, have got some in common characteristics:

- Conflicts among criteria are possible (low cost conditions decrease network's performances),
- Criteria contain non-measurable units (cost, bandwidth),
- The goal is to design an optimal, or to select the best alternative.

Based on third characteristics MCDM can be divided into: Multiple Objective Decision Making (MODM) and Multiple Attributes Decision Making (MADM) algorithms. MODM consists of conflicting set goals that cannot be achieved simultaneously, while MADM deals with the problem of selecting an alternative from a set of alternatives which are characterized in terms of their attributes. In fact, the attributes are parameters or performance factors with an influence on the selection while alternatives are characterized by more attributes with the reliable level of successfulness. Usually the goal is not explicit, i.e., defined and often can be represented as the satisfactions' maximization.

In the network selection problem in heterogeneous environment, the specified candidate network is characterized by attributes like bandwidth, losses, delay, cost of service, etc. The available networks number is finite, while the resolution space is discrete. In this way, the problem can be characterized in MADM category.

MADM algorithms can be divided into compensatory and non-compensatory ones [24]. Non-compensatory algorithms (e.g., dominance, conjunctive, disjunctive or sequential elimination) are used to find acceptable alternatives which satisfy the minimum cutoff. On the contrary, compensatory algorithms combine multiple attributes to find the best alternative. Most MADM algorithms that have been studied for the network selection problem are compensatory algorithms [25], including Simple Additive Weighting (SAW), Multiplicative Exponential Weighting (MEW), Gray Relational Analysis (GRA), Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS), etc. These methods are often called "soft" optimization techniques compared with standard mathematical optimization methods like Linear and Dynamic programming [26], or Game theory [27].

The SAW technique (a.k.a. the Weighted Sum Method) [28] is one of the most widely used MADM methods in the network selection related open literature. The basic concept of SAW in this context is to obtain a weighted sum of the normalized form of each parameter over all candidate networks. Normalization is required in order to have a comparable scale among all parameters. For m available RANs, and n observed parameters, VHD function for i-th RAN is obtained as

$$SAW_i = \sum_{j=1}^{n} w_j r_{ij} . \qquad (11)$$

Application of SAW based VHD function in the process of evaluating the qualitative performance of potential target networks is proposed in [29]. By using the normalization and weights distribution methods, VHD function determines a network quality factor as

$$\begin{aligned} NQ_i = & \frac{w_c(1/C_i)}{\max(1/C_1,...,1/C_m)} + \frac{w_s S_i}{\max(S_1,...,S_m)} \\ & + \frac{w_p(1/P_i)}{\max(1/P_1,...,1/P_m)} + \frac{w_d(D_i)}{\max(D_1,...,D_m)} + \frac{w_f(F_i)}{\max(F_1,...,F_m)} \end{aligned} \qquad (12)$$

where C_i is the cost of service, S_i security, P_i power consumption, D_i network conditions and F_i network performance, while the w_c, w_s, w_p, w_d and w_f, respectively, are weights for each of the criteria which are proportional to the significance of a parameter to the VHD function. These weights are obtained from the user via a user interface. The RAN with the highest NQ_i is the preferred network. If the newly detected network receives a higher Q_i, vertical handover takes place, otherwise, the MT remains connected to the current network. High overall throughput and user's satisfaction can be regarded as major advantages of this technique. Identical approach is proposed in [30], with the difference in terms of observed criteria.

MEW (a.k.a. the Weighted Product Method) [28] uses multiplication among attribute values which are raised to the power of the attribute importance, while normalization procedure is not obligatory. VHD function can be represented in form

$$MEW_i = \sum_{j=1}^{n} r_{ij}^{w_j} . \qquad (13)$$

The results obtained in [31] indicate the inaccuracy of the SAW method and the benefits of using MEW as network selection technique.

TOPSIS [32] is one of MADM techniques based on a concept that the chosen alternative should have the shortest distance from the ideal possible solution and the longest from the worst possible solution, where distances are determined with certain values for p ($1 \le p \le \infty$) from Minkowski metric. The Minkowski's row distance p, between two points, $V = (v_1, v_2, \ldots, v_n)$ i $U = (u_1, u_2, \ldots, u_n) \in R^n$, can be defined as

$$L_p = \left(\sum_{i=1}^{n} |v_i - u_i|^p \right)^{1/p} , \qquad (14)$$

where the case $p = 1$ is used very often (Manhattan distance, the first row metric), and $p = 2$ (Euclidean distance). On the boundary case, when $p \to \infty$, Chebyshevs' distance is obtained (L_∞ metrics).

Considering the weighted matrix defined by (9), the ideal and worst solutions are representing the sets

$$\begin{aligned} A^+ &= \{(\max_{1\le i\le m} v_{ij} \,|\, v_{ij} \in V^{\max}); (\min_{1\le i\le m} v_{ij} \,|\, v_{ij} \in V^{\min})\} = \{v_1^+,...,v_n^+\}, \\ A^- &= \{(\min_{1\le i\le m} v_{ij} \,|\, v_{ij} \in V^{\max}); (\max_{1\le i\le m} v_{ij} \,|\, v_{ij} \in V^{\min})\} = \{v_1^-,...,v_n^-\}, \end{aligned} \qquad (15)$$

where V^{max} is the set of larger-the-better criteria and V^{min} is the set of smaller-the-better criteria. Euclidean distances of all alternatives, in relation to the ideal and worst solution, can be calculated from

$$D_i^+ = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^+)^2}\,, D_i^- = \sqrt{\sum_{j=1}^{n}(v_{ij} - v_j^-)^2}\,, \quad i = 1,...,m. \quad (16)$$

Lastly, the ranking of networks can be done through the relative closeness (*RC*) to the ideal solution in the form

$$RC_i = \frac{D_i^-}{D_i^- + D_i^+}\,, \quad RC_i \in (0,1)\,. \quad (17)$$

Network selection algorithm based on TOPSIS method (Fig. 3) is proposed in [33]. The criteria considered in the decision matrix are: available bandwidth *B*, QoS level *Q*, security level *S*, and cost of service *C*. It can be noted that the computational complexity involved in calculating Euclidean distances and entropy weights used in TOPSIS is very low.

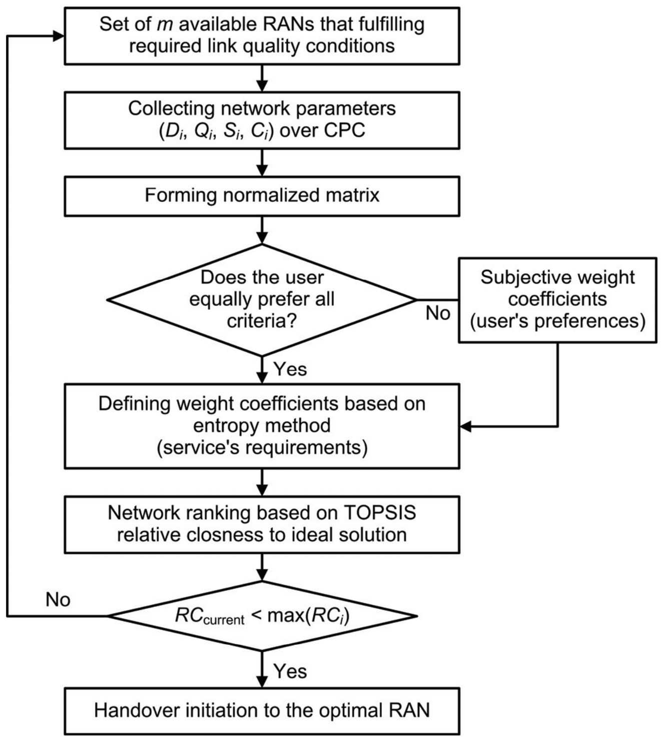

Fig. 3. Network selection algorithm based on TOPSIS method

The proposed solution is evaluated using numerical examples. Through simulation studies, MADM is envisaged as promising tool especially when TOPSIS method is used, because of high sensitivity to users' preferences and the parameter values. This solution is realistic and not very complex to implement in mobile terminals and other network elements.

Researches provided in [34], extend merit function (measure of network quality) obtained as relative closeness to ideal solution (17) with predefined hysteresis whose goal is to minimize the influence of undesirable effect of frequent superfluous handovers and also can be used in admission control mechanism if the need to control or influence the network selection process exists. A handover is considered as superfluous when a mobile terminal back to the previous PoA is needed within certain time duration ("ping-pong" effect). The larger the hysteresis value, the smaller the number of handovers, however there is a longer handover initiation delay. On the other hand, the smaller the threshold value, the shorter the handover initiation delay but the larger the number of handovers. Handover initiation delays lead to an increase in the call dropping probability, especially in the case of highly mobile users. Moreover, frequent handovers can cause an increase in signaling overhead and in the network load. Therefore, the determination of the hysteresis value is very important in terms of mobility performance.

Different from TOPSIS, GRA uses only the ideal solution to calculate the ranking coefficient alternative (network) *i*, given by

$$GRA_i = \left[\sum_{j=1}^{n} w_j \left|v_{ij} - V_j^+\right| + 1\right]^{-1}. \quad (18)$$

In [18] the authors develop a network selection mechanism for an integrated WLAN and 3G cellular systems. The proposed scheme comprises two parts. The first applying an AHP to estimate the relative weights of evaluative criteria set according to users' preferences and service applications, while the second adopts GRA to rank the network alternatives with faster and simpler implementation than AHP. The design goal is to provide the user the best available QoS at any time. This method mathematically presents a complex solution and unnecessarily takes into account a large number of parameters (delay, jitter, response time, bit error rate, etc.) only for 3G and WLAN networks. Processing a large number of parameters leads to the increasing computational time, while the terminal and infrastructure network elements are additionally loaded. Thus, this model is interesting from theoretical point of view, but not adequate for a direct implementation.

C. Fuzzy Logic Techniques

Fuzzy logic (FL) theory [35] utilizes human knowledge by giving the fuzzy or linguistic descriptions a definite structure. In the classical set theory, the membership of elements in a set is assessed in binary terms, which means either belongs or does not belong to the set. On the contrary, the FL theory permits the gradual assessment of membership using a membership function valued within [0, 1].

There are different ways to use the FL in a network selection. Some studies use it standalone as the core of the selection scheme, while some use the FL with recursion (learning techniques), while some combine FL with MCDM algorithms.

An elementary framework for FL based network selection, without combining with any other technique, is proposed in [36]. In the proposed scheme, three input fuzzy variables are considered (the probability of a short interruption, the handover failure probability, and the size of unsent messages). At the beginning of the procedure, the fuzzy variables are fuzzified and converted into fuzzy sets by a singleton fuzzifier. Then, based on the fuzzy rule (IF-THEN) base, the fuzzy inference engine maps the input fuzzy sets into output fuzzy sets by the algebraic product operation. Finally, the output fuzzy sets are defuzzified into a crisp decision point.

Since some dynamic factors change frequently, the recursion can be used to combine the latest information with previous ranking result to obtain the latest rank. The recursion procedure can be a simple recursion without any further operation or certain learning procedure, such as neural networks or learning techniques. Second FL based scheme in [36] proposes simple recursion, which considers the requirements of both operator and user. The rank produced by the fuzzy module is fed back to this module, so that it can produce a new rank when some factors change. The advantage of this method is that the designer is no longer obliged to define rules, which are generated automatically by inserting a series of training data sets to the system.

Network selection solution proposed in [37] represents interesting and promising solution while combining the heuristics of the FL systems and AHP (Fig. 4). In the process of handover initiation, proposed technique uses FL analyzing the criteria such as: received signal strength *RSS*, bandwidth *B*, network coverage *NC* and terminal velocity *V*. The gathered information, depending on their availability, is fed into a fuzzifier and converted into fuzzy sets. A fuzzy set contains varying degrees of membership in a set. The membership values are obtained by mapping the values retrieved for a particular variable into a membership function. After fuzzification, fuzzy sets are fed into an inference engine, where a set of fuzzy rules (81 predefined IF-THAN rules) are applied to determine whether handover is necessary.

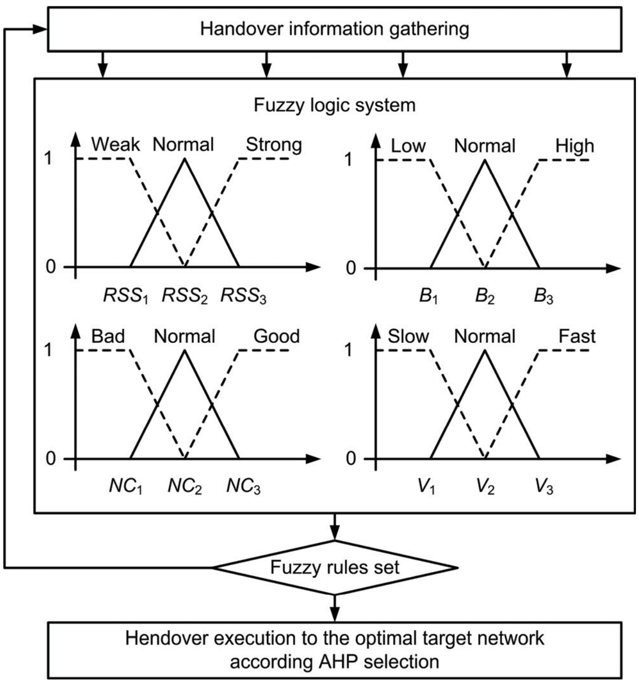

Fig. 4. Network selection technique based on FL and AHP

By application of AHP method and Saaty's scale on criteria such as cost of service, preferred interface, battery status and QoS level, the optimal access network is determined. On the other hand, by applying FL in decision making process the number of unnecessary handovers is reduced, as well as signaling traffic and handover delays. Inflexibility of the impact of user's preferences on the system is basic lack of applied AHP method, which can possibly be exceeded by using some MADM techniques, e.g. TOPSIS.

D. Artificial Neural Networks Techniques

Artificial neural networks (ANNs) can be most adequately characterized as computational models with particular properties such as the ability to adapt or learn, to generalize or to cluster data and which operation is based on parallel processing. ANN consists of a pool of simple processing units which communicate by sending signals to each other over a large number of weighted connections [38]. ANNs are of interest to the researchers because they have the potential to treat many problems that cannot be handled by traditional analytic approaches. Back propagation neural networks are the most prevalent ANN architectures because they have the capability to "learn" system characteristics through nonlinear mapping.

Network selection technique based on the ANN is proposed in [39]. Applied feedforward neural network topology, which consists of input, hidden and output layer, is shown in Fig. 5. The input layer is made of the h nodes representing different criteria for optimal network selection, while the hidden layer consists of the n nodes that represent the available access networks. Output layer is formed by a node that generates the identification of the optimal RAN. For the training process error back propagation algorithm is used. During the simulation, the authors adopted the same VHD function as in [29].

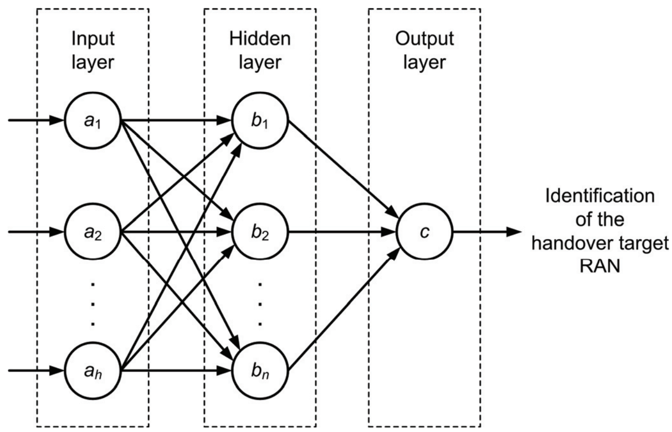

Fig. 5. Example of ANN topology for network selection

In this study, all neurons use sigmoid activation function. Random values, which serve as weights, are generated for all connections from input to hidden (v_{hi}) and from hidden to output layers (w_{ij}). In addition, biases are assigned random values at the hidden nodes (θ_i) and the output node (τ_i). The activation functions at the hidden layer are calculated as

$$b_i = f(\sum_{i=1}^{n} a_h v_{hi} + \theta_i), \quad (19)$$

while activation values at the output layer are calculated as

$$c_j = f(\sum_{i=1}^{n} a_h w_{ij} + \tau_j). \quad (20)$$

In both equations, (19) and (20), $f(x)$ represents the logistic sigmoid threshold function, $f(x) = 1/(1 + e^{-x})$.

The simulations have shown high accuracy and reliability of the model while selecting the optimal network. The lacks of the algorithm are reflected to the complexity of the system and to increased handover delays due to the training process.

VI. Traffic Parameters Influence on Network Selection as an Aspect of Always Best Connected and Served

Applying TOPSIS based algorithm proposed in [33], it is possible to evaluate influences of traffic parameters on the network selection process. A software application is developed with source code written in MATLAB environment. Testing model results are shown through representative scenarios for three available RANs, with link capacities, traffic intensities and some others attributes shown in Table I.

TABLE I
INPUT CRITERIA VALUES

RAN_i	s_i	a_i	Q_i	S_i	C_i
RAN_1	20	8 - 15	3	2	2
RAN_2	15	12	3	2	1
RAN_3	18	10	2	3	1

Besides usual criteria (QoS level Q_i, security level S_i, and cost of service C_i), classical Grade of Service (GoS) parameters are pointed out in [34]. The influence of criteria D_i, i.e., network conditions for ith RAN, can be analyzed through link capacity (s_i), the link capacity and traffic ratio (s_i/a_i), as well as the ratio of link capacity and loss (s_i/B_i), link capacity and traffic losses ratio ($s_i/(a_iB_i)$), and finally through the available bandwidth ($s_i - a_i(1 - B_i)$).

In [34] it was concluded that merit functions do not depend on traffic intensity, when network conditions are presented only by link capacity ($D_i = s_i$). By invoking traffic parameters in D_i, i.e., $D_i = s_i - a_i(1 - B_i)$, with the increment of a_1, successive decrease of MF_1 appears, while MF_2 and MF_3 increase successively (Fig. 6). Similar behavior is expected in the case when the network conditions is presented through the link capacity and traffic ratio, $D_i = s_i/a_i$.

The most interesting point is the phenomenon concerning MF_i, when the network conditions $D_i = s_i/(a_iB_i)$ are used. With a sufficient high value a_1, which produces loading per channel like loading in RAN_3 ($a_1/s_1 \approx a_3/s_3$), which is lower compared to the loading per channel RAN_2 (a_2/s_2), the effect of rapid variation for MF_1 and MF_3 exists (Fig. 6, dashed curves). Very similar properties of merit functions is expected also in the case when $D_i = s_i/B_i$. These properties have a significant influence on the network selection process, and provide a step closer to the ABC&S concept.

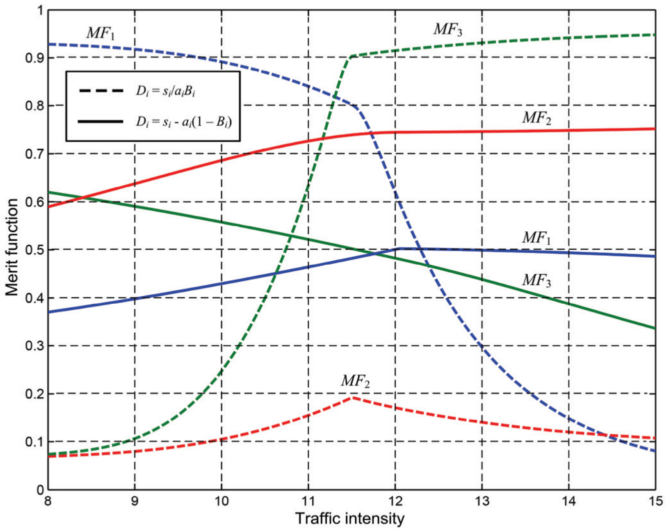

Fig. 6. Influence of traffic intensity variation to the merit functions

VII. Conclusion

Following the principles of heterogeneous networking, a mobile user/terminal may choose among multiple available connectivity alternatives based on the criteria related to networks' performances, users' preferences and services' requirements. The scope of this paper is to address the issue of network selection in heterogeneous wireless environment. The main challenges involved in network selection process are pointed out, and a significant collection of relevant approaches encountered in the open literature is presented. Theoretically acceptable and interesting decision techniques are discussed, emphasizing on the use of tools such as FL and MADM that may lead to more efficient and objective decisions. A qualitative comparison of the analyzed network selection techniques in terms of functionality and complexity of implementation is briefly presented. Motivated by the fact that current proposed network selection techniques in heterogeneous environment require more significant challenges to be overcome before they can successfully be deployed in real systems, we analyzed influence of traffic parameters as unavoidable step toward ABC&S concept realization.

Acknowledgement

This research was partially supported by the Ministry of Education, Science and Technological Development, Republic of Serbia (Project No. TR-32025).

References

[1] M. Kassar, B. Kervella, and G. Pujolle, "An Overview of Vertical Handover Decision Strategies in Heterogeneous

Wireless Networks", *Computer Communications*, vol. 31, no. 10, pp. 2607-2620, 2008.
[2] E. Gustafsson, A. Jonsson, "Always Best Connected", *IEEE Wireless Communications*, vol. 10, no. 1, pp. 49-55, 2003.
[3] Z. Ji, I. Ganchev, and M. O'Droma, "An iWBC Consumer Application for 'Always Best Connected and Best Served': Design and Implementation," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 462-470, 2011.
[4] ITU-R Rec. M.1645, "Framework and Overall Objectives of the Future Development of IMT-2000 and System beyond IMT-2000", June 2003.
[5] 3GPP TS 24.312 V11.3.0, "Access Network Discovery and Selection Function (ANDSF) Management Object (MO)", June 2012.
[6] IEEE Std. 802.21-2008, "Media Independent Handover Services", Jan. 2009.
[7] IEEE Std. 1900.4-2009, "Architectural Building Blocks Enabling Network-Device Distributed Decision Making for Optimized Radio Resource Usage in Heterogeneous Wireless Access Networks", Feb. 2009.
[8] D.E. Charilas and A.D. Panagopoulos, "Multiaccess Radio Network Environments", *IEEE Vehicular Technology Magazine*, vo. 5, no. 4, pp. 40-49, 2010.
[9] R. Trestian, O. Ormond, and G. Muntean, "Game Theory - Based Network Selection: Solutions and Challenges", *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1212-1231, 2012.
[10] D. He et al., "A Simple and Robust Vertical Handoff Algorithm for Heterogeneous Wireless Mobile Networks", *Wireless Personal Communications* vol. 52, no. 2, pp. 361-373, 2011.
[11] ITU-T Rec. Y.1541, "Network Performance Objectives for IP-Based Services", Dec. 2011.
[12] B. Bakmaz et al., "Mobile IPTV over Heterogeneous Networks: QoS, QoE and Mobility Management", *Proc. 18th ERK2009*, Portoroz, Slovenia, pp. 101-104, 2009.
[13] B. Bakmaz, M. Bakmaz, and Z. Bojkovic, "Elements of Security Aspects in Wireless Networks: Analysis and Integration", *International Journal of Applied Mathematics and Informatics*, vol. 1, no. 2, pp. 70-75, 2007.
[14] D. Niyato, E. Hossain, "Competitive Pricing in Heterogeneous Wireless Access Networks: Issues and Approaches", *IEEE Network*, vol. 22, no. 6, pp. 4-11, 2008.
[15] Y. Ji et al., "CPC-assisted Network Selection Strategy", *Proc. 16th IST Mobile and Wireless Communications Summit*, Budapest, Hungary, pp. 1-5, 2007.
[16] W. Shen, Q.-A. Zeng, "Cost-Function-Based Network Selection Strategy in Integrated Wireless and Mobile Networks", *IEEE Transactions on Vehicular* Technology, vol. 57, no.6, pp. 3778-3788, 2008.
[17] E.H. Ong, J.Y. Khan, "Dynamic Access Network Selection with QoS Parameters Estimation: A Step Closer to ABC", *Proc. 67th IEEE VTC*, Marina Bay, Singapore, pp. 2671-2676, 2008.
[18] Q. Song, A. Jamalipour, "Network Selection in an Integrated Wireless LAN and UMTS Environment Using Mathematical Modeling and Computing Techniques", *IEEE Wireless Communications*, vol. 12, no. 3, pp. 42-48, 2005.
[19] Y. Wei, Y. Hu, and J. Song. "Network Selection Strategy in Heterogeneous Multi-Access Environment", *The Journal of China Universities of Posts and Telecommunications*, vol. 14, Supplement 1, pp. 16-20, 2007.
[20] P. Zhang et al., "A Novel Network Selection Mechanism in an Integrated WLAN and UMTS Environment Using AHP and Modified GRA", *Proc. 2nd IEEE IC-NIDC 2010*, Beijing, China, pp. 104-109, 2010.
[21] T.L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, NY, McGraw Hill, 1980.
[22] X. Yan, Y. A. Şekercioĝlu, and S. Narayanan, "A Survey of Vertical Handover Decision Algorithms in Fourth Generation Heterogeneous Wireless Networks", *Computer Networks*, vol. 54, no. 11, pp. 1848-1863, 2010.
[23] J. McNair, F. Zhu, "Vertical Handoffs in Fourth-generation Multinetwork Environments", *IEEE Wireless Communications*, vol. 11, no. 3, pp. 8-15, 2004.
[24] F. Bari, V.C.. Leung, "Automated Network Selection in a Heterogeneous Wireless Network Environment", *IEEE Network*, vol. 21, no. 1, pp. 34-40, 2007.
[25] L. Wang, G. Kuo, "Mathematical Modeling for Network Selection in Heterogeneous Wireless Networks - A Tutorial", *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 271-292, 2013.
[26] L. Pirmez et al., "SUTIL - Network Selection based on Utility Function and Integer Linear Programming", *Computer Networks*, vol. 54, no. 13, pp. 2117-2136, 2010.
[27] I. Malanchini, M. Cesana, "Modelling Network Selection and Resource Allocation in Wireless Access Networks with Non-Cooperative Games", *Proc. 5th IEEE MASS*, Atlanta, GA, USA, pp. 404-409, 2008.
[28] C.-H. Yeh, "A Problem-based Selection of Multi-attribute Decision-making Methods", *International Transactions in Operational Research*, vol. 9, no. 2, pp. 169-81, 2002.
[29] N. Nasser, A. Hasswa, and H. Hassanein, "Handoffs in Fourth Generation Heterogeneous Networks", *IEEE Communications Magazine*, vol. 44, no. 10, pp. 96-103, 2006.
[30] R. Tawil, G. Pujolle, and O. Salazar, "A Vertical Handoff Decision Scheme in Heterogeneous Wireless Systems", *Proc. 67th IEEE VTC,* Marina Bay, Singapore, pp. 2626-2630, 2008.
[31] Q.-T. Nguyen-Vuong, Y. Ghamri-Doudane and N. Agoulmine, "On Utility Models for Access Network Selection in Wireless Heterogeneous Networks", *Proc. IEEE NOMS 2008*. Salvador, Bahia, Brazil, pp. 144-151, 2008.
[32] C.L. Hwang, K. Yoon, *Multiple Attribute Decision Making: Methods and Applications*, NY, Springer-Verlag, 1981.
[33] B. Bakmaz, Z. Bojkovic, and M. Bakmaz, "Network Selection Algorithm for Heterogeneous Wireless Environment", *Proc. IEEE PIMRC 2007*, Athens, Greece, pp. 1007-1010, 2007.
[34] B. Bakmaz, Z. Bojkovic, and M. Bakmaz, "Traffic Parameters Influences on Network Selection in Heterogeneous Wireless Environment", *Proc. 19th IWSSIP 2012*. Vienna, Austria, pp. 306-309, 2012.
[35] L.A. Zadeh, "Fuzzy Sets", *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965.
[36] S. Kher, A.K. Somani, and R. Gupta, "Network Selection Using Fuzzy Logic", *Proc. IEEE BroadNets 2005*, Boston, MA, USA, pp. 876-885, 2005.
[37] M. Kassar, B. Kervella, and G. Pujolle, "An Intelligent Handover Management System for Future Generation Wireless Networks", *EURASIP Journal on Wireless Communications and Networking*, Article ID 791691, 2008.
[38] B. Kröse and P. van der Smagt, *An Introduction to Neural Networks*, University of Amsterdam, 1996.
[39] N. Nasser, S. Guizani, E. Al-Masri, "Middleware Vertical Handoff Manager: A Neural Network-based Solution", *Proc. IEEE ICC'07*, Glasgow, Scotland, pp. 5671-5676, 2007.