Use of Non-monotonic Utility in Multi-Attribute Network Selection

Farooq Bari and Victor C.M. Leung

Abstract In the past few decades several wide-area and local-area wireless access technologies have emerged. Network convergence across these different access technologies holds a promise of enabling ubiquitous service availability but faces several technical challenges. With anticipated proliferation of multimode IP devices, the optimal selection of a service delivery network among multiple IP-based wireless access alternatives is one of the important issues that is actively studied and discussed in several standardization forums. Use of multi-attribute decision making (MADM) algorithms has been proposed in the past for network selection decisions in a heterogeneous wireless network environment. A direct comparison of these algorithms is difficult as this would require the use of another MADM algorithm. A better approach instead is to ascertain the appropriateness of the algorithm to the problem space. This chapter provides the basis for evaluating the appropriateness of MADM algorithms for network selection. It analyzes the use of MADM algorithms such as TOPSIS, ELECTRE and GRA for network selection and argues that GRA provides the best approach in scenarios where the utilities of some of the attributes are non-monotonic. We propose a novel stepwise approach for GRA that uses multiple reference networks and explain how it works with network selection scenarios.

1 Introduction

In the past few decades several wide-area and local-area wireless access technologies have emerged. While many of these technologies have been successful in deployment and continue to evolve, none of them provides a universal coverage to cater to the mobile lifestyle of today’s user. The desire and expectation to have
ubiquitous broadband connectivity all the time are the driving forces to push the network operators to look at new and innovative ways of service delivery including the possibility of inter-working of these evolving access technologies with the use of multi-mode devices. An example of multimode device is a terminal that supports Institute of Electrical and Electronics Engineers (IEEE) 802.11 wireless local area network (WLAN) technology and Groupe Special Mobile (GSM) wireless wide-area network (WWAN) technology. A major challenge in this new environment is network selection, i.e., identifying the best-suited service delivery network when the user has multiple networks to choose from. Network selection becomes specially challenging for multi mode devices that have an option to get services from different all IP wireless access types. In the case of current GSM only devices, network selection involves a scan for the network identities, i.e., public land mobile network IDs (PLMN IDs) followed by a selection of one of them based on the pre-provisioned information in the terminal about preferred PLMN IDs and forbidden PLMN IDs. In the case of IEEE 802.11 WLANs, the beacon and probe request/response mechanism provides a way for terminals to discover access points (APs) using service set identifiers (SSIDs). Based on this information and signal strength, the terminal can decide which access point (AP) to associate with. Such simplistic approaches, however, are unlikely to yield optimal network selection in inter-worked heterogeneous all IP broadband systems that use different access technologies while delivering a range of services from a variety of operators. This new environment requires consideration of a number of factors such as the QoS capabilities of the network, current network conditions, QoS requirements of the requested service, and the subscription type of the user.

Network selection is an area of active research and a topic of discussion in several standardization forums. IEEE 802.11u Working Group currently has a draft proposal that would enable information exchange for network selection between the network and the terminal [1]. It leverages the protocol being developed by IEEE 802.21 for this purpose which they also plan to use for selecting networks for vertical hand-offs [2]. Similarly 3GPP is looking into the mechanism of network discovery and selection in their work on System Architecture Evolution (SAE) [3]. Along with protocol and architectural aspects of the problem, an essential component in solving the problem of network selection is defining the optimization objective and the algorithm to be used in the selection process. Selection of a non-optimal network creates undesirable results such as poor customer experience or the use of more expensive network. The focus of this chapter is on this aspect of the problem.

1.1 Selection of MADM Algorithm

The requirement to have a consistent service experience for the user requires selection of an optimal delivery network. This issue has special significance for multimode IP devices where services can be delivered over a variety of wireless access technologies under varying network conditions. Several factors related to network
capabilities and quality of service (QoS) conditions influence the network selection decision process, e.g., bandwidth, delay, jitter, and packet loss. This makes it attractive to use deterministic decision-making tools such as multi-attribute decision making (MADM) algorithms [4]. Their use has been previously considered, e.g., for network selection in a heterogeneous wireless network environment [5–8], to derive a ranking of the available networks in terms of their suitability. The highest ranking network is then selected as the best-suited network. The prior work, however, failed to provide a comparison among the MADM algorithms for use in network selection.

Several alternate MADM algorithms can be suitable for solving a decision problem and the decision maker in this situation can be faced with the task of selecting the most appropriate method from among a number of feasible methods. Classification of MADM algorithms into categories can help to eliminate the algorithms in categories that are not well suited to the problem space, but this process does not provide the most suited algorithm. It is conceivable that a suitable MADM algorithm may be selected for a particular decision problem based on one or both of the following criteria.

1.1.1 Accuracy of the Results Obtained from an Algorithm

For a variety of reasons, different algorithms, when applied to the same problem under the same assumptions, can result in different rankings of the alternatives. In such scenarios it is not possible to objectively rank the MADM algorithms for their ranking accuracy as it would require the use of another MADM algorithm to get such a ranking. For this reason it has been found difficult to use accuracy of the results as a criterion in selecting a specific type of MADM algorithm.

1.1.2 Appropriateness of Applying the Algorithm to the Problem

Because of differences in the approaches used by different MADM algorithms, a direct comparison among them is difficult. It has been proposed in the past that a method which is capable of solving the decision problem and whose decision-making philosophy reflects the values of the decision maker can be considered to be the best suited. Decision makers in general prefer deterministic algorithms that provide reliable results based on a simple and easy to understand philosophy.

So far it has been difficult to perform such an evaluation of the algorithms because of a rather simplistic assumption about the optimization objectives of the decision maker. Prior studies have ignored a possibly diverse range of optimization scenarios based on service and user types that could exist and hence can help in comparing the suitability of the algorithms. In the remainder of this section we describe different QoS profiles that can be used in the decision process which can then lead to a requirement for support of non-monotonic utilities. The following section describes the concept of non-monotonic utilities for the attributes that can help meet a wider variety of optimization objectives. We propose that this concept be leveraged in assessing the suitability of MADM algorithms to network selection.
1.2 Types of QoS Profiles

To better serve their customers, network operators typically use QoS profiles. The selection of a network is highly dependent upon the type of optimization performed for QoS-related attributes stored in such profiles. The two possible QoS profile types that can be stored in user’s home network are as follows:

- **Overall user QoS profile** that is applicable to all of the services that the user is using; e.g., gold, silver, or bronze profile can indicate the level of QoS that the user is expected to have based on the subscription.
- **QoS profile of an individual service** that is applicable to all subscribers of that service; e.g., VoIP service profile or web-browsing profile.

1.2.1 Service-Based QoS Profile

Based on service-based QoS profile, key service classes can be categorized as follows:

**VoIP** – This is a low bandwidth application that is very sensitive to delay and jitter but can withstand some packet losses. Transport cost factor is considered negligible because of low bandwidth usage. Also because of low bandwidth requirements, total bandwidth and available bandwidth are not significant factors. Since there is some correlation of utilization with jitter and delay, it is preferred to have a low utilization for the selected network.

**Streaming** – Being a multimedia service, a streaming application requires a higher bandwidth than VoIP. Therefore available bandwidth, transport cost, and current utilization are important factors. It is less vulnerable to delay and jitter than VoIP because of ability to buffer longer duration of data before play back. Sensitivity to packet loss is similar to VoIP where some packet loss can be compensated without impact to user experience.

**Web Browsing** – Web-browsing type application is a low QoS service; i.e., the importance of utilization, delay, jitter, and packet loss is low. It does not need a guaranteed bit rate because of spiky nature of web traffic pattern. With statistical traffic multiplexing for such type of traffic, broadband wireless networks can deliver a reasonable customer experience even at lower average data rates. The total bandwidth and allowed bandwidth are therefore less critical but transport cost is considered critical.

1.2.2 Subscription-Based QoS Profile

Based on subscription-based QoS profile the following key subscription classes could be defined.

**Gold Subscription** – This indicates a premier user subscription that would allow the use of the highest level QoS independent of the transport cost.
Silver Subscription – This indicates a medium priority user subscription that would try to balance between the QoS requirements and other factors such as the transport cost.

Bronze Subscription – This indicates a lower priority user subscription where transport cost is significantly important compared with any QoS parameters.

The examples described in the later section have used service-based QoS profile.

2 Use of Non-monotonic Utilities for Attributes in Network Selection

In general, as part of the decision process, the MADM algorithms associate a measure of suitability or appropriateness, hereafter called utility, with the individual attribute’s value. The utility is said to be monotonic if the measure of suitability associated with the attribute shows a monotonic increase or decrease with an increase in attribute value. Otherwise it is said to be non-monotonic. Figure 1 shows a simple decision-making scenario with one attribute, i.e., delay and two networks. The delay attribute is shown with possible monotonic and non-monotonic utilities for different service types. The monotonic utility represents optimization objectives where the network with the least delay value, i.e., Ntwk #1 will be selected for all service types. The non-monotonic utility of delay attribute in Fig. 1 represents the optimization objective of the decision maker where it would like

![Graphical representation of a simple decision-making scenario with one attribute and two networks]

Fig. 1 Graphical representation of a simple decision-making scenario with one attribute and two networks
to use the network closest to the service’s delay requirements but not necessarily the best network, which is the network with the least delay. This type of optimization objective would result in selection of Ntwk#1 for VoIP service and use Ntwk#2 for streaming media and web-browsing services. The decision maker may desire this type of optimization for policy reasons such as load balancing across access networks or for keeping the best networks for services and sessions with higher QoS requirements that it can expect to have. It would be similar to the policy of an airline that decide to fly with some first or business class seats empty and not upgrade people from economy class with the knowledge or the hope that it would be able to get full-fare business or first-class customers at the next stop.

In the case described above only one attribute was considered. For the case of multiple attributes, the overall ranking is either obtained via adding the utility associated with each of the attributes or by comparing the utilities for the attributes individually in the decision process. Prior applications of MADM to the network selection problem [3, 6–8] have generally assumed the use of the best network irrespective of service requirements or user type, but the impact of different optimization objectives and hence the use of non-monotonic utility in the decision process have not been considered. This implies a monotonic utility for all attributes. While some of the decision-related attributes can be considered to have monotonically increasing or decreasing utilities, in reality the overall optimization goals of the decision maker may require a combination of monotonic and non-monotonic utilities for different attributes that are taken into considerations during the decision process for network selection. Associating a monotonic increasing or decreasing utility in general with each of the attributes is therefore a simplistic assumption that would limit the scope of types of optimization available to the decision maker [6, 8].

An example of an optimization objective can be to find the network that along with other factors (such as cost) also has the best QoS characteristics from among the list of available networks. In this case the utility of QoS attributes can be considered to be monotonic. However, under a different deployment scenario, the decision maker may wish to assume a non-monotonic utility for some of the attributes considered in the selection process, e.g., as an optimization objective to distribute network traffic across different access networks by selecting the access network offering a QoS closest to that being requested by the service and not the network that may have the best QoS that far exceeds the service’s QoS requirements.

This chapter analyzes the suitability of several most commonly used MADM algorithms for the problem of optimal network selection, where not all the attributes considered in the decision-making process have a monotonically increasing or decreasing utility. Such network selection scenarios will be quite common in future heterogeneous wireless network environments used for delivery of both real-time and non-real-time services. The algorithms considered in the analysis are TOPSIS [9], ELECTRE [9, 10], and GRA [11]. Among these MADM algorithms, GRA is found to be the most suited for optimization
objectives requiring both monotonic and non-monotonic utilities of attributes. Using a heterogeneous wireless network environment as an example, we demonstrate how the algorithm can be implemented to achieve different optimization objectives and its impacts on the resulting network ranking for network selection.

3 Comparison of MADM Algorithms for Use with Non-monotonic Utilities of Attributes

3.1 TOPSIS

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) is a widely used MADM algorithm that was developed by Yoon and Hwang [9]. The algorithm calculates perceived positive and negative ideal solutions based on the range of attribute values available for the alternatives and selects the best solution as the one with the shortest distance to the positive ideal solution and longest distance from the negative ideal solution. The distances are measured in Euclidean terms. Because of the concept of positive and negative ideal solutions that use Euclidean distances, a standard implementation of TOPSIS requires that the utilities of the attributes under consideration increases or decreases monotonically.

3.2 ELECTRE

The ELECTRE (Elimination et choix traduisant la realite) method [9, 10] performs pairwise comparisons among all alternatives for each one of the attributes separately in order to develop outranking relationship between the alternatives. In its standard implementation, the method first removes the less desirable alternatives and then using a complimentary analysis it selects the best-suited alternatives. Since the comparison is direct among the available alternatives there is no concept in ELECTRE of comparing the alternatives to some reference set of values to see how close the parameters values are to the desired values. The notion of a monotonically increasing or decreasing utility of an attribute is inherent in direct comparison among the alternatives, which makes standard ELECTRE algorithm not well suited for use with attributes having non-monotonic utilities.

3.3 GRA

GRA (gray relational analysis) is another very popular decision-making technique that is based on gray system theory. Originally developed by Deng [11], gray
systems theory has been applied to solve a variety of real-life problems ranging from the fields of business, operations research, and engineering, to social sciences. One of its areas of application has been MADM, decision making where multiple attributes influence the decision process. Unlike other MADM algorithms, GRA uses a reference set of attribute values for comparison with attribute values of the alternatives. It has been applied in the past [6, 7] to solve the problem of network selection in a heterogeneous network environment. The problem of selection of an optimal network in a heterogeneous environment, however, is quite complex and it is possible to apply GRA in several different ways to address the problem. In the prior work, the utility aspects of the algorithm were not explored and a single reference network was constructed, which implied monotonic utilities for all the attributes for all service or user types. Because of its ability to use reference attribute values in the decision process, GRA can be applied where the optimization objectives require non-monotonic utilities for some of the attributes and monotonic utilities for the others. As described in Section 5, such an implementation of GRA would use multiple reference networks. The network rankings in this case could be quite different than if monotonic utilities were considered for all the attributes.

It is clear that for optimization scenarios where a utility does not increase or decrease monotonically with an increase or decrease in attribute value, standard implementations of MADM algorithms such as TOPSIS and ELECTRE will have limited applicability. Other simpler compensating MADM algorithms such as SAW (Simple Additive Weighing) [9] and WPM (Weighed Product Method) [9] also have similar limitations because of their inherent assumption about monotonic utilities of attributes. These MADM algorithms do, however, allow assigning different weights to the attributes before they are combined to calculate ranking indices. This general feature of MADM algorithms can be used to apply algorithms such as TOPSIS to network selection for different service types. For example, services that are less sensitive to QoS and more sensitive to transport cost (e.g., web browsing) could have higher weights assigned to the cost attribute and lower weights assigned to the QoS attribute. This would allow alternatives that are closer to the positive ideal solution in transport cost (i.e., lowest in cost) to get more importance in decision making than network alternatives with QoS-related attribute values closest to the positive ideal solution (i.e., best QoS attribute value). However, it is important to note that this type of applicability would provide a different optimization than trying to find the network alternative which QoS attributes are closest to those of the requested service. As stated earlier, GRA favors a selection that gives a closest match to a set of reference data values. This process inherently supports the notion that these reference values do not necessarily need to be the best or the worst values associated with the attributes. In addition, it also has the ability to assign different weights to different attributes. These two tunable aspects of GRA when combined provide a much better mechanism to achieve optimization objectives involving attributes with non-monotonic utilities. The rest of this chapter describes the application of GRA to the problem of network selection with some attributes having non-monotonic utilities.
4 Theory of Gray Relational Space

Gray Relational Analysis (GRA) has sometimes been compared to fuzzy logic. However, GRA is different from fuzzy logic. While both can help in decision making under uncertain conditions, fuzzy logic deals with imprecise information and GRA deals with insufficient or scarce information. GRA is based on the concept of gray relational space (GRS). GRS \((X,Y)\) describes a relationship \(Y\) between reference data values \(X_0\) and sequence of data values \(X\). So if \(y \in Y\), \(x_i \in X, x_0 \in X_0\) such that \(x_0 = x_0(1),...,x_0(n)\) and \(x_i = x_i(1),...,x_i(n)\) then \(y(x_0(k),x_i(k))\) would represent a GRS at point \(k\) provided the axioms documented in [11] are satisfied. In addition, a gray relational grade for a series \(i\) could then be represented as

\[
y(x_0, x_i) = \frac{1}{n} \sum_{k=1}^{n} y(x_0(k), x_i(k)).
\]

A representation of \(y(x_0(k), x_i(k))\) that satisfies all of the axioms in [11] is represented as

\[
y(x_0(k), x_i(k)) = \min_k \min_i \frac{|x_0(k) - x_i(k)| + \xi \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \xi \max_k |x_0(k) - x_i(k)|}
\]

where \(0 \leq \xi \leq 1\) is called a distinguished coefficient and \(y(x_0(k), x_i(k))\) is called the gray relational coefficient (GRC). When applying GRA to ranking of networks while selecting a network, GRC is a measure of how closely a network’s attributes match the reference network’s attributes. In this respect it represents the overall utility that takes into consideration individual values of all the attributes. The higher the value of GRC, the closer would be the candidate network to the reference network. Hence, for the purposes of network selection GRC is equivalent to a utility that takes into consideration all individual attribute values.

5 Application of GRA Adapted to Network Selection with Non-monotonic Utility

For the network selection problem in this chapter, Table 1 provides a typical set of attributes that can be considered in such a decision-making process.

Using the attributes defined above, a candidate network \(NW\) for evaluation by GRA can be represented as follows:

\[
NW = \begin{bmatrix} CB \ TB \ AB \ U \ D \ J \ L \end{bmatrix}
\]

If there are \(N\) alternative networks to be considered in the selection process, they can be represented in the form of a matrix as follows:
A reference access network is needed for application of GRA. In the case of monotonically increasing or decreasing utilities for the attributes, this reference network can be developed by using the maximum or minimum value of the attributes. In this case there will only be one reference network. However, if there are services that have different QoS requirements (VoIP, streaming, web browsing, etc.), or the users are of different categories (e.g., bronze, silver, gold), then the decision maker can use different reference networks for each one of the categories. A reference network in this case can be created based on the information about the user/terminal preferences, e.g., indication of the requested service, or based on the user profile stored in the home network, e.g., the subscribed QoS. These multiple reference networks would result in non-monotonic utilities for some of the attributes. Table 3 shows four different reference networks developed for the example described in Section 6. The reference network $i$ for a particular service or user type can therefore be represented as follows:

$$(\text{NW}_{\text{ref}})_i = ((\text{CB}_{\text{ref}})_i (\text{TB}_{\text{ref}})_i (\text{AB}_{\text{ref}})_i (\text{U}_{\text{ref}})_i (\text{D}_{\text{ref}})_i (\text{J}_{\text{ref}})_i (\text{L}_{\text{ref}})_i)$$

The units of measurement for the attributes such as cost, bandwidth, and delay will be different. In order to apply the algorithm without having the artifacts related to different units of measurement impacting the results, the attributes will have to be made unit-less before they can be directly compared or combined during the calculations. This process is called normalization. There are several normalization techniques e.g., dividing attribute value with a maximum value for that attribute across all the alternatives. Using these normalized attribute values, an updated matrix is created as follows:
The reference network $i$’s attributes are also normalized and a normalized reference network vector is created as follows:

$$\tilde{N}_{W_i} = \begin{bmatrix} C_{\tilde{B}_1} & T_{\tilde{B}_1} & A_{\tilde{B}_1} & \tilde{U}_1 & \tilde{D}_1 & \tilde{J}_1 & \tilde{L}_1 \\ C_{\tilde{B}_2} & T_{\tilde{B}_2} & A_{\tilde{B}_2} & \tilde{U}_2 & \tilde{D}_2 & \tilde{J}_2 & \tilde{L}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{\tilde{B}_N} & T_{\tilde{B}_N} & A_{\tilde{B}_N} & \tilde{U}_N & \tilde{D}_N & \tilde{J}_N & \tilde{L}_N \end{bmatrix}$$

If the reference attribute values lie outside of the attributes values for the alternatives under considerations, then calculation of the maximum and minimum values to be used in the normalization process should include the reference values as well.

Distance vectors are calculated for attributes of each access network under consideration by taking the absolute difference between the attribute of the reference network and the corresponding attribute of the candidate network. For example, in the case of TB for network $i$, the distance value from reference network $j$ is calculated as follows:

$$|\Delta_{TB}^i| = |T_{\tilde{B}_j} - T_{\tilde{B}_i}|$$

The matrix of distance value for each of the attributes for the $N$ networks under consideration can therefore be created as follows:

$$\Delta_{NW} = \begin{bmatrix} (\Delta_{CB})_1 & (\Delta_{TB})_1 & (\Delta_{AB})_1 & (\Delta_{U})_1 & (\Delta_{D})_1 & (\Delta_{J})_1 & (\Delta_{L})_1 \\ (\Delta_{CB})_2 & (\Delta_{TB})_2 & (\Delta_{AB})_2 & (\Delta_{U})_2 & (\Delta_{D})_2 & (\Delta_{J})_2 & (\Delta_{L})_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (\Delta_{CB})_N & (\Delta_{TB})_N & (\Delta_{AB})_N & (\Delta_{U})_N & (\Delta_{D})_N & (\Delta_{J})_N & (\Delta_{L})_N \end{bmatrix}$$

Gray relational coefficient (GRC) is a measure of similarity of an attribute to its reference value. It is calculated for each of the matrix entries. For example, in the case of TB, it is calculated as follows:

$$GRC_{TB} = \frac{\Delta \min + \xi \Delta \max}{(\Delta_{TB}) + \xi \Delta \max}$$

where $\xi \in [0,1]$ and $\Delta \min$ and $\Delta \max$ can be calculated as follows:

$$\Delta \max = \max_i (\Delta_{CB_i} + \Delta_{TB_i} + \Delta_{AB_i} + \Delta_{U_i} + \Delta_{D_i} + \Delta_{J_i} + \Delta_{L_i})$$

$$\Delta \min = \min_i (\Delta_{CB_i} + \Delta_{TB_i} + \Delta_{AB_i} + \Delta_{U_i} + +\Delta_{D_i} + \Delta_{J_i} + \Delta_{L_i})$$
The next step is to consider the relative importance of each of the attributes in the decision about network selection. For this purpose each of the attribute is assigned a weight “w” such that

\[ W = W_{TB} + W_{AB} + W_{U} + W_{D} + W_{J} + W_{L} = 1 \]

The new weighted GRC matrix will be as follows:

\[
\text{GRC} = \begin{bmatrix}
W_{CB} \cdot (GRCC_{B})_{1} & W_{TB} \cdot (GRCT_{B})_{1} & W_{AB} \cdot (GRC_{AB})_{1} & W_{U} \cdot (GRC_{U})_{1} & W_{D} \cdot (GRC_{D})_{1} & W_{J} \cdot (GRC_{J})_{1} & W_{L} \cdot (GRC_{L})_{1} \\
W_{CB} \cdot (GRCC_{B})_{2} & W_{TB} \cdot (GRCT_{B})_{2} & W_{AB} \cdot (GRC_{AB})_{2} & W_{U} \cdot (GRC_{U})_{2} & W_{D} \cdot (GRC_{D})_{2} & W_{J} \cdot (GRC_{J})_{2} & W_{L} \cdot (GRC_{L})_{2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
W_{CB} \cdot (GRCC_{B})_{N} & W_{TB} \cdot (GRCT_{B})_{N} & W_{AB} \cdot (GRC_{AB})_{N} & W_{U} \cdot (GRC_{U})_{N} & W_{D} \cdot (GRC_{D})_{N} & W_{J} \cdot (GRC_{J})_{N} & W_{L} \cdot (GRC_{L})_{N}
\end{bmatrix}
\]

Using the GRC matrix thus calculated, the gray relational coefficient for each of the candidate network is calculated as follows:

\[ (GRC_{NW})_{i} = W_{CB} \cdot (GRCC_{B})_{i} + W_{TB} \cdot (GRCT_{B})_{i} + W_{AB} \cdot (GRC_{AB})_{i} + W_{U} \cdot (GRC_{U})_{i} + W_{D} \cdot (GRC_{D})_{i} + W_{J} \cdot (GRC_{J})_{i} + W_{L} \cdot (GRC_{L})_{i} \]

The network with the highest value of gray relational coefficient is considered to be the best network.

6 Evaluation of Using Non-monotonic Utilities in a Heterogeneous Wireless Network Environment

To evaluate the impact of different optimization objectives, we consider a network selection scenario with five networks. For each of these networks, the attribute values to be used in the decision process are shown in Table 2. The table provides the numerical attribute values for these five networks that were used for illustrative purposes in the decision process for network selection. They are representative of listed example network types that a typical user could expect. For example, the cost attribute is derived on the basis of spectral efficiency of the technology and whether the technology runs on licensed or unlicensed spectrum. So the cost is lowest for unlicensed spectrally efficient technology such as IEEE 802.11n and it is highest for licensed and relatively less spectrally efficient technology such as 3G. Also the

| Table 2 Attribute values for alternative networks at the time of network selection |
|-------------------------------|-------|-------|-------|-------|-------|-------|-------|
|                              | CB (%) | TB (mbps) | AB (mbps) | U (%) | D (ms) | J (ms) | L (per10^6) |
| Ntwk#1, e.g., 3G#1           | 100    | 2      | 0.2     | 10    | 400    | 50    | 100    |
| Ntwk#2, e.g., 3G#2           | 100    | 2      | 0.4     | 5     | 200    | 25    | 50     |
| Ntwk#3, e.g., 802.11a        | 10     | 54     | 2       | 20    | 100    | 15    | 15     |
| Ntwk#4, e.g., 802.11n        | 5      | 100    | 5       | 40    | 150    | 30    | 20     |
| Ntwk#5, e.g., 4G             | 30     | 100    | 5       | 20    | 100    | 20    | 15     |
maximum estimated throughput for each of the example technologies has been used for the total bandwidth attribute. The allowed bandwidth has been assumed to be based on operator policy to rate limit its customers differently for different access technologies. Other QoS related attributes such as delay, jitter and loss represent a snapshot of the values that could exist in these networks at the time of decision.

In our case we address the network selection problem for three distinct types of services, namely, VoIP, streaming media, and web browsing. Each of these service types has its distinct set of QoS requirements. In the following subsection we describe how to use an adapted version of GRA for the scenarios under consideration.

6.1 Setting up GRA for Network Selection

The GRA algorithm can be applied in more than one ways to the problem of network selection. However, it is well suited to handling diverse optimization objectives including those requiring non-monotonic utility of attributes. Figure 2 shows three different ways of application of the GRA algorithm. We recommend the third approach as described earlier since it provides the maximum flexibility for tuning the algorithm to different optimization objectives. A two-step process for tuning GRA is proposed below.

6.1.1 Determine Reference Attribute Values for Different Service or User Categories

Based on optimization criteria derived from the QoS requirements for the services or user types, reference networks are created by the decision maker (e.g., user’s home network). It will be toward these reference values that the GRA will try to find a closest match from a given list of alternative networks. This step relates to addressing the non-monotonic nature of utility for some of the attributes under consideration. For example, the VoIP reference network’s attribute values would reflect higher QoS requirements compared to a reference network for web browsing. Creating reference networks is a one-time event and can therefore be provisioned into the decision process.

6.1.2 Determine Attribute Weights for Different Service or User Categories

To allow further tuning of the GRA algorithm to the optimization objectives, the weights (i.e., the importance) assigned to the attributes used in decision making are adjusted for each type of service. The weight assigned reflect the relative importance of an attribute for that service or user type. For example, the cost attribute would carry relatively higher weight for streaming type service when compared with VoIP service. The process of determining attribute weights is also a one-time event and can be provisioned into the decision process.
Table 3 shows four reference networks that were used in the evaluation. The reference values for the attributes are considered the values toward which the operator would like to optimize while selecting a network from among the alternatives. For example, by adopting the utilization of the lowest utilized alternative network as a reference attribute value, the decision maker could help improve the utilization for under utilized alternative networks and hence have an improved balance of traffic loads across networks. However, this may not be always the optimization criteria and other decision makers may like to have different criteria for selecting this and other attribute values. So the first reference network is created from best values for the attributes from among the alternatives networks. The remaining three reference networks use a combination of best attribute values (for CB, TB, and U) and reference values derived from QoS requirements for the service types (for AB, D,
Table 3 Reference attribute values

<table>
<thead>
<tr>
<th></th>
<th>CB (%)</th>
<th>TB (mbps)</th>
<th>AB (mbps)</th>
<th>U (%)</th>
<th>D (ms)</th>
<th>J (ms)</th>
<th>L (per 10^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best QoS</td>
<td>5</td>
<td>100</td>
<td>5</td>
<td>10</td>
<td>100</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>VoIP</td>
<td>5</td>
<td>100</td>
<td>0.02</td>
<td>10</td>
<td>100</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Streaming</td>
<td>5</td>
<td>100</td>
<td>1</td>
<td>10</td>
<td>400</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Web browsing</td>
<td>5</td>
<td>100</td>
<td>0.1</td>
<td>10</td>
<td>1,000</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

J and L). The entries in Table 3 show that some of the reference values generated from specific service types lie outside the range of attribute values for the network alternatives under consideration. This can potentially cause problems in the normalization process. In order to avoid possible ranking abnormalities, the normalization process is modified by appropriately adjusting the minimum/maximum values used in order to include reference values as well.

The assigned weight for each of the attributes for different service categories considered in the evaluation is shown in Table 4. For example, the importance of transport cost was considered high for web browsing as compared to VoIP. To evaluate the impact of assigned weights, a set of scenarios was also evaluated where only one weight distribution was used for all the different service types.

Scenario 1 uses the best values as the reference for all attributes but uses different weights for different service types. This is shown in the first approach of Fig. 2. Scenario 2 only changes the reference attribute values while keeping the same attribute weights for all service types. This is shown in the second approach of Fig. 2. Scenario 3 calculates ranking when the reference attribute values as well as the attribute weights were changed for different services as shown in approach 3 of Fig. 2. For example, in the case of streaming media type service, it first compares the network attribute values with the reference values specific to streaming media type service to find the degree of match for each of the individual attributes. Then based on the emphasis that should be placed on the degree of match for each of the attributes, a weight is assigned to it as explained earlier.

Table 4 Assignment of attribute weights

Attribute weights used for scenarios 1 and 3

<table>
<thead>
<tr>
<th></th>
<th>CB</th>
<th>TB</th>
<th>AB</th>
<th>U</th>
<th>D</th>
<th>J</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>VoIP</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.15</td>
</tr>
<tr>
<td>Streaming</td>
<td>0.2</td>
<td>0.15</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>Web browsing</td>
<td>0.5</td>
<td>0.05</td>
<td>0.15</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Attribute weights used for scenarios 2

<table>
<thead>
<tr>
<th>VoIP, streaming, web browsing</th>
<th>CB</th>
<th>TB</th>
<th>AB</th>
<th>U</th>
<th>D</th>
<th>J</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2</td>
<td>0.15</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Fig. 3 Results for network selection for three possible configurations of GRA shown in Fig. 2. Configuration 3 is preferred as it provides maximum flexibility to fine tuning the algorithm to optimization objectives.

Figure 3 shows the results for all the scenarios that were evaluated and hence shows the impact of use of multiple reference networks and attribute weights. As explained earlier it is not possible to directly evaluate results of MADM algorithm in terms of accuracy. However, since scenario 1 uses a combination of reference values and attribute weights to distinguish among different service types, it will provide a more balanced approach with results closest to optimization objectives of the decision maker if the intent was to select the network closest in characteristics to the reference network for that service type. This is also apparent by comparing results for scenarios 1 and 3, which show that in this example, for streaming service, using different reference values for different service types actually does impact the
ranking. Similarly a comparison of scenarios 1 and 2 shows that the network rankings are impacted if different distribution of attribute weights is used for each service type as opposed to a single distribution of attribute weights for all services.

Finally we evaluate the use of non-monotonic utility for a case where there is one physical access network that supports multiple classes of QoS. An example of that kind of network can be WiFi MultiMedia (WMM)-based access network, i.e., an implementation of IEEE 802.11e standard-based access network with four QoS classes. In such a situation we have to select the most optimal QoS class to deliver the requested service. For cases where real-time information about delay, jitter, and packet loss values can not be obtained from the access network, it may be possible to use static or provisioned values from the service level agreement with the network operator. This would allow decision making with only network utilization requiring a real-time update. Even the utilization rate can in some cases be predicted based on, for example, the past seasonal trends. Table 5 represents such an access system that supports five classes of service or five levels of SLAs. The attribute values are for illustrative purposes for use in the decision process for network selection. For example, the cost attribute in this case is derived based on the treatment the packets from that class of service would get. QoS Class 1, for example, gets least preferential treatment relative to other classes and therefore has the lowest cost. The selection of any particular alternative will map to the same physical network but with a different QoS class. Therefore, while the total access network bandwidth will be the same for all alternatives, the values of other parameters (e.g., delay, jitter, packet loss) are different depending upon the QoS class. The allowed bandwidth has been assumed to be based on operator policy to rate limit its customers using different QoS classes on the access network. The reference networks and attribute weights for different service types are the same as for the previous example. The results that used the three possible implementation of GRA are shown in Fig. 4. As explained earlier, scenario 1 uses a combination of reference values and attribute weights to distinguish among different service types. It therefore provides a more balanced approach with results closest to optimization objectives of the decision maker if the intent is to select the network closest in characteristics to the reference network for that service type. It can be seen that depending upon the implementation of the algorithm, a different QoS class can possibly be selected for service delivery. For example, for a VoIP call, the QoS class of service selected by the preferred GRA

<table>
<thead>
<tr>
<th>QoS Class#1/SLA#1</th>
<th>CB (%)</th>
<th>TB (mbps)</th>
<th>AB (mbps)</th>
<th>U (%)</th>
<th>D (ms)</th>
<th>J (ms)</th>
<th>L (per10^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>100</td>
<td>0.1</td>
<td>30</td>
<td>400</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>QoS Class#1/SLA#2</td>
<td>20</td>
<td>100</td>
<td>0.5</td>
<td>20</td>
<td>200</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>QoS Class#1/SLA#3</td>
<td>40</td>
<td>100</td>
<td>1</td>
<td>20</td>
<td>100</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>QoS Class#1/SLA#4</td>
<td>60</td>
<td>100</td>
<td>5</td>
<td>40</td>
<td>50</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>QoS Class#1/SLA#1</td>
<td>60</td>
<td>100</td>
<td>1</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>
Fig. 4 Results for network selection for three possible configurations of GRA shown in Fig. 2. Configuration 3 is preferred as it provides maximum flexibility to fine tuning the algorithm to optimization objectives.

implementation chooses QoS class #3, whereas the other two possible implementations choose QoS class #2 and QoS class #5, respectively. This again shows that using different reference values for different service types actually does impact the ranking and therefore the service delivery network.

7 Conclusion

Network selection is an important problem to be solved for inter-worked heterogeneous all IP wireless systems. Prior research on application of MADM algorithms to the problem of network selection has not compared them to select the most appropriate algorithm. This chapter has presented the decision maker’s optimization
objectives and hence the utilities of attributes as a means to evaluate the algorithms and select the most appropriate one. The need to support non-monotonic utility for attributes in order to handle diverse optimization objectives of a decision maker has been shown and MADM algorithms have been evaluated for handling these objectives. We have shown that many of the commonly used MADM algorithms such as SAW, WPM, TOPSIS, and ELECTRE in their standard form are not best-suited because of assumptions about monotonically increasing or decreasing utilities of the attributes. We have also shown that GRA can easily be adapted to use multiple reference networks so that both monotonic and non-monotonic utilities can be taken into consideration, and is therefore better suited for achieving this type of optimization objectives. The evaluation of adapted GRA in this chapter has also demonstrated that the selection of the best-suited delivery network will be impacted by how the algorithm is used to achieve the optimization objectives. A novel two step process that uses multiple reference values has been proposed and explained through an example, which shows how reference attribute values and attribute weights impact the selection process. The adjustment of these parameters for different service types has been discussed. The decision process proposed in this chapter can be used in a heterogeneous wireless network system environment. Future work on the topic could include research into the usefulness of the approach in terms of its effectiveness for network load balancing, business cost savings, and consistency of customer experience.

References