Learning-Based Multi Attribute Network Selection in Heterogeneous Wireless Access

Idris Cinemre¹ · Toktam Mahmoodi¹

Accepted: 17 January 2022 / Published online: 3 February 2022
© The Author(s) 2022

Abstract

The rise of heterogeneous networks including macro, micro, pico, femto, and WLAN presents new challenges in optimizing user's access to the networks. To fully utilize the capacity of such rich field of heterogeneous wireless connectivity, mobile devices should be able to select Radio Access Network (RAN) for their connection depending on their need. The proposed algorithm here is based on an autonomous agent at the mobile node and assisted by integration of distributed cloud services, e.g. the edge cloud. The selection of RAN is based on a Multi Attribute Decision Making process combined with the Reinforcement Learning that reinforces historical data collected from the edge cloud. Mobile agent is responsible for collecting data, executing the selection algorithm, possibly offloading the parts of the execution to the edge cloud, and providing feedback to the edge cloud after termination of the connection. The ultimate aim is to the design of RAN selection algorithm owned by autonomous and intelligent user agents that can significantly improve the user's experience in terms of network coverage, data rate, latency, and battery lifetime. Through extensive simulation scenarios, we demonstrate the stability and precision of our proposed algorithm.

Keywords Heterogeneous networks · Network selection · Reinforcement learning

1 Introduction

With the growing demand for mobile data, mobile operators are increasingly turning to small cells (e.g. femtocells) to deal with the capacity crunch in dense urban areas and to add coverage in areas with low or zero cellular signal levels. To increase the capacity, not only licensed spectrum is exploited, but also service providers are looking to use e.g., WLAN to augment their investment in small cells. Emergence of small cells and their heterogeneity add new dimensions to optimizing users' access to these various Radio Access Networks (RANs). Today's smartphones are equipped with multiple radios and they will soon be capable of connecting to several networks at the same time, opening up enormous capacity and coverage.

✉ Idris Cinemre
idris.i.cinemre@kcl.ac.uk

¹ Department of Engineering, King's College London, Strand, London, UK

These new opportunities cause great challenges including selection of a RAN (or a combination of RANs) at the User Equipment (UE) so as to receive the desired service at a lower cost and with an enhanced battery lifetime. To this end, novel selection techniques should be deployed at the UE. Intelligent and autonomous devices should be able to explore their environment, find the RAN(s), and choose which one(s) to connect to so as to meet their need best. There have been efforts in the autonomous selection of RANs in heterogeneous environments using learning-based techniques [1–3]. Market-based algorithms are also studied for selecting the best available RAN [3, 4]. The proposed technique in [5], designs two entities and that both at the UE side and in the core network; the entity at the UE prioritizes available RANs based on the requirements of running application and the final decision for RAN selection is made at the core network. Combining the ranking algorithm with the Multi Attribute Decision Making (MADM), the network selection technique in [6] deploys the closest algorithm to what we discuss in this paper. In this work, a classification method is applied to build a class which has the similar criteria, and then Analytic Hierarchy Process (AHP) method is applied to determine weights of different classes. The discussed research works in [7–14], all discuss the similar concept of employing multi attribute techniques for addressing network selection problem.

Nonetheless, one of the remaining major challenges of network selection in the largely heterogeneous scenarios of today's mobile network is the consideration of numerous system parameters, which is in fact the common issue in the above mentioned research works. This problem, so far, has been addressed by reducing scenarios to the specific types of RANs or by limiting performance metrics to be optimized. Furthermore, knowledge based on which selection is made and is limited at either of UE or core network.

To this end, we propose adaptive network selection strategies based on an intelligent agent at the UE and based on an entity located at the edge cloud. While mobile agent can run the algorithm at either of the cloud or the UE, it executes the selection algorithm and collects historical information of the access networks from the cloud. Upon termination of the connection, the reports of the users' connection experience turns back to the cloud. The network selection algorithm is modelled by the MADM process combined with the Reinforcement Learning (RL) that reinforces historical data collected from the edge cloud. The algorithm and architecture we presented in this paper is (one way of) realization of the concept we presented in [15]. In contrast with the previous work, our network selection algorithm is fully managed from the UE side, and in addition to the previous works, integration of edge cloud is considered to increase the available data for decision and reduce the complexity.

This paper is structured as follows. In Sect. 2, we elaborate on the two entities involved in network selection, their roles, and their interactions. The details of the network selection algorithm and its different steps are discussed in Sect. 3. We analyze our proposed algorithm thoroughly and through multiple simulation scenarios in Sect. 4.

2 Architecture

The proposed adaptive network selection here is based on a mobile agent, an intelligent agent located at the UE, which collects data from its surrounding wireless environment as well as from the edge cloud. This architecture is depicted in Fig. 1. The ultimate aim is to the design of algorithm for autonomous UEs that can select RAN(s) based on their own need (Quality of Service (QoS) requirements) and prediction/estimation from the

Fig. 1 Heterogeneous wireless access and integration of edge cloud

wireless environment as well as the collaborative observations gathered by other mobile users (assumed to be available at the edge cloud). Hence, the RAN selection algorithm will be divided into two parts. The main algorithm is executed by the mobile agent at the UE, while it interacts with the edge cloud.

2.1 Autonomous Agent at the UE

In the architecture depicted in Fig. 1, the mobile agent is in charge of collecting data from the edge cloud, observing the surrounding wireless, executing the selection algorithm, and sending feedback to the cloud after the connection is terminated. In other words, selection of RAN(s) depends on the observation/estimation from the surrounding wireless as well as the collective experience of previously connected UEs. The selection algorithm at the UE will be designed so as to perform in the absence of access to the cloud while accessing to the shared experience (of other users) in the cloud will improve the selection's efficiency. This algorithm is fully detailed in Sect. 3. Furthermore, the mobile agent can decide to offload execution of the network selection algorithm, partly or fully, to the edge cloud, if resources are available. Such a decision can be made based on the current status of the UE, e.g. its battery level, and its available computing resources.

2.2 Integration of Cloud Services

The new ways of coping with complexity are needed to exploit the characteristics of available RANs and to avoid an excessive need for storage and processing power. Integration of the cloud computing concepts into the mobile environment can overcome the obstacles related to the networking environment and it can improve the mobile computing systems performance including battery lifetime, bandwidth, and capacity [16]. We refer to the edge cloud in this paper as computing & storage resources that are located at close proximity of users and they are also available with short latency, e.g. it could be located at the mobile operators site. The next generation of mobile base stations will indeed feature commodity

servers and include a virtualized platform for general cloud hosting [17]. The cloud-based entity is added mainly to address the challenges of dealing with large data sets of today's heterogeneous networks.

To allow users' mobility, similar model to Follow Me Cloud (FMC) can be used. The FMC model allows end points mobility using different sub-nets that belong to one IP network and it can sustain device communication active during the migration. The migration between the mobile cloud points occurs through a group of OpenFlow switches that are situated at the network edge [18].

3 Network Selection Algorithm

The network selection algorithm at the mobile agent should make decision based on various parameters depending on the application as well as the heterogeneous environment. Hence, we consider the network selection as a Multi Attribute Decision Making (MADM) problem. In this paper, Analytic Hierarchy Process (AHP) is used to analyze weights of each selection criteria. In addition, two techniques will be used for ranking all available access networks: the Total Order Preference by Similarity to the Ideal Solution (TOPSIS) and Distance to Ideal Alternative (DIA). Finally, a combination of MADM and RL will be used to make the decision based on received reward function from the previous selection. Each of these steps is detailed in the following subsections.

3.1 Analysis of Selection Criteria

The AHP aims to solve complex problems by dividing them into sub-problems. The procedure can be summarized in four steps to: (1) determine the decision factors and create the pairwise matrix, (2) normalize the pairwise matrix and obtain normalized decision matrix, (3) calculate the weights of each criterion, and finally (4) calculate coherence ratio and check consistency.

The pairwise matrix $P = [x_{ij}]_{n \times n}$ is shaped by comparison between two different criteria. The Saaty's scale [19] is used for the x_{ij} values, detailed in Table 1, while $x_{ii} = 1$ and $x_{ji} = \frac{1}{x_{ij}}$. Accordingly, P is a $n \times n$ matrix, where n is the number of criteria, e.g. number of considered QoS parameters in network selection. In the second step, P is normalized

Table 1 Saaty's scale for pairwise comparison

Saaty's scale	Relative importance of two criterion
1	Equally important
3	Moderately important with one over another
5	Strongly important
7	Very strongly important
9	Extremely important
2, 4, 6, 8	Intermediate values

(as $P_{norm} = [x_{ij}]$) such that $p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}$ and in the third step weight of each decision criterion is computed based on Eq. (1).

$$W_i = \frac{\sum_{j=1}^n p_{ij}}{n} \quad (1)$$

where $\sum_{i=1}^n W_i = 1$.

Finally, we compute the coherence ratio similar to the [19], which is equal to the Consistency Index (denoted by CI) divided by the random index (denoted by RI). The RI has a fixed value based on the number of considered criteria (from [19], it is 0.9 for our case with four selection criteria), and the CI is computed according to the Eq. (2).

$$\begin{aligned} b_i &= \frac{\sum_{j=1}^n W_i * p_{ij}}{W_i} \\ \lambda_{max} &= \frac{\sum_{i=1}^n b_i}{n} \\ CI &= \left(\frac{\lambda_{max} - n}{n - 1} \right) \end{aligned} \quad (2)$$

The coherence ratio is mainly computed to check that it should not exceed ten percent, which means only inconsistencies less than 0.1 can be ignored. Otherwise, the algorithm should iterate and create the new pairwise matrix for that criterion.

3.2 Ranking Selection Criteria

3.2.1 Total Order Preference by Similarity to the Ideal Solution (TOPSIS)

As its name implies, preferences are sorted based on their similarities to the ideal solution [20] and the idea is that the best option among alternatives must have the farthest distance from negative ideal solution and the shortest distance from positive ideal solution [6]. The TOPSIS process can be summarized in seven steps: (1) construct the decision matrix, (2) normalize the decision matrix, (3) construct weighted normalized decision matrix, (4) determine the negative and positive ideal solutions, (5) measure the distances to ideal solutions, (6) calculate the relative closeness, and finally (7) rank the preference order.

The decision matrix is $D = [y_{ij}]_{m \times n}$, where y_{ij} represents the importance of the criteria j for the alternative i (for example the importance of any QoS parameter for a specific application) and n and m are the number of selection criteria and the number of alternatives consecutively. The rows in matrix D , show selection criteria for different alternatives $A_1, \dots, A_m$.

After shaping the decision matrix, it will be normalized as $D_{norm} = [y_{ij}]$, such that $d_{ij} = \frac{y_{ij}}{\sqrt{\sum_{i=1}^n y_{ij}^2}}$, and the weighted normalized decision matrix $U = [y_{ij}]$ can be shaped by multiplying d_{ij} by the W_i as computed in AHP, i.e. $u_{ij} = W_i x d_{ij}$. In step 4, the negative and positive ideal solutions for desirable (e.g. throughput) and undesirable criteria (e.g. latency) are computed according to Eqs. (3) and (4) consecutively.

$$\begin{aligned} A^+ &= \max_j(u_{ij}) = \{u_1^+, u_2^+, \dots, u_i^+, \dots, u_m^+\} \\ A^- &= \min_j(u_{ij}) = \{u_1^-, u_2^-, \dots, u_i^-, \dots, u_m^-\}. \end{aligned} \quad (3)$$

$$\begin{aligned} A^+ &= \min_j(u_{ij}) = \{u_1^+, u_2^+, \dots, u_i^+, \dots, u_m^+\} \\ A^- &= \max_j(u_{ij}) = \{u_1^-, u_2^-, \dots, u_i^-, \dots, u_m^-\}. \end{aligned} \quad (4)$$

In step 5, distances of alternative A_i to the positive and negative ideal solutions are computed as Eq. (5) and in step 6 relative closeness of alternative A_i is computed as C_i in Eq. (6).

$$\begin{aligned} S_j^+ &= \sqrt{\sum_{i=1}^m (u_{ij} - u_j^+)^2} \\ S_j^- &= \sqrt{\sum_{i=1}^m (u_{ij} - u_j^-)^2} \end{aligned} \quad (5)$$

$$C_j = \frac{S_j^-}{S_j^- + S_j^+} \quad (6)$$

Finally m different alternatives $A_1, \dots, A_m$, are ranked based on their closeness value, C_i , in decreasing order, and the first alternative will be the best choice.

3.2.2 Distance to Ideal Alternative (DIA)

The DIA method, which belongs to the MADM family, aims to dynamically select the best alternative and identify the ranking of the alternatives with further accuracy [13]. It also deals with ranking abnormalities (we will see this later in the result section). The DIA also performs in seven steps, and while steps 1–5 are the same as TOPSIS, steps 6 and 7 are different. In steps 6, and 7 the Positive Ideal Alternative (PIA) is computed and the selection is made consecutively based on the ranking in Eq. (8).

$$PIA = \left\{ \min(S_j^+), \max(S_j^-) \right\} \quad (7)$$

$$R_j = \sqrt{\left(S_j^+ - \min(S_j^+)\right)^2 + \left(S_j^- - \max(S_j^-)\right)^2} \quad (8)$$

3.3 Reinforcement Learning in the Selection Algorithm

Here, the RL is used as a technique to allow mobile agent to behave optimally, e.g. to maximize its reward, in an environment that agent is not familiar with. We consider each available network as a state, actions can represent either inter or intra technology handover, and reward functions after each action are the level of QoS satisfaction (denoted by Q) as achieved by taking this action. It is assumed that decision at time t depends on current value as well as the two previous values ($t-1$ and $t-2$), and these two historical values are weighted with W_{h1} and W_{h2} . If there is an increasing trend ($Q_{current} \geq Q_{t-1} \geq Q_{t-2}$),

Fig. 2 Network selection procedure based on historical data using RL

Table 2 Parameters of the three wireless access networks

RAN	Available bandwidth (AB) (Mbps)	Delay (D) (ms)	Packet loss (PL)	Jitter (J) (ms)
UMTS	0.1–2	25–50	20–80	4–10
WLAN	1–11	100–150	20–80	10–20
WiMax	1–60	60–100	20–80	3–10

positive reward is received, while a decreasing trend ($Q_{current} \leq Q_{t-1} \leq Q_{t-2}$), results in negative reward. These comparisons will be reversed if they are done over the undesired QoS parameters. For example, available bandwidth is considered as a desired parameter whereas jitter, delay, and packet loss are undesired factors. Detail of how we employ RL in the network selection is explained in the diagram of Fig. 2.

4 Performance Analysis

A simulation model with access networks of three different technologies is designed to examine the performance. They are one cellular network (UMTS) and two different broadband networks (WLAN and WIMAX). These networks are selected because of the availability of their measurements, while they could be substituted by any other access network (e.g. 4G/LTE).

4.1 Simulation Parameters

The ranges of offered QoS parameters by each network are listed in Table 2, and are similar to the network parameters in [6]. In the remainder of this section, referring to the network parameters will be with an array of Net. Param. = {AB, D, PL, J}. We also assume that users' data rate varies depending on the users' distance from the wireless

access point according to the Figure 3, and the data for this plot is from the measurements reported in [21] and [22].

We consider four QoS criteria and five different applications and we investigate performance of our network selection algorithm in five different scenarios, first to show stability of the algorithm and afterwards to show the precision of selection based on the actual conditions of the wireless networks.

4.1.1 Traffic Classes and QoS Parameter

The examined QoS parameters here are available bandwidth, jitter, delay, and packet loss. For clarity and precision, the definition of these parameters, as we used in our performance analysis, are as follows:

- *Available bandwidth* (bps): The bandwidth that network operator allocated to per user and also it can be changed dynamically according to network utilization.
- *Jitter* (ms): The average delay variations on the network link.
- *Delay* (ms): The average packet delay on the network link.
- *Packet loss*: The average number of packet loss during the certain time between RAN and the UE.

Moreover, we categorized user applications under the five types of Streaming, Background, Interactive, Conversational Audio, and Conversational Video, based on the following QoS requirements:

- *Streaming*: The streaming flows are jitter sensitive and require high bandwidth, but they are tolerable to delay. Video and audio streaming are examples of this class.
- *Background*: The background flows are sensitive to packet loss whereas delay and jitter are not important parameters and the average bandwidth is sufficient for them. The typical background traffic are FTP and Email.
- *Interactive*: For interactive applications, packet loss is the most crucial QoS criterion. A well-known example of this class is web browsing.

Fig. 3 Throughput of networks versus users' distance from their access point; **a** WLAN, **b** WiMax, **c** UMTS

- *Conversational audio*: Flows of this category are delay and jitter sensitive but they are tolerant to packet losses and data rate requirement is relatively low. Examples include audio chats and calls.
- *Conversational video*: The conversation video flows are highly delay and jitter sensitive and require average bandwidth. Video chatting can be an example of this type.

Using the AHP method and Eq. (1), and based on the sensitivity of each application to each QoS criterion, we compute weights for all five applications and for each QoS criterion. The array of weights will shape as Eq. (9).

$$\begin{aligned}
 W_{application} &= [W_{AB}, W_D, W_{PL}, W_J] \\
 W_{streaming} &= [0.2913, 0.0445, 0.0903, 0.5739] \\
 W_{background} &= [0.3063, 0.0540, 0.5857, 0.0540] \\
 W_{interactive} &= [0.1622, 0.0657, 0.7065, 0.0657] \\
 W_{audio} &= [0.0892, 0.4301, 0.0506, 0.4301] \\
 W_{video} &= [0.1177, 0.4217, 0.0389, 0.4217]
 \end{aligned} \tag{9}$$

4.1.2 Setup of the Cloud-Based Information

We assume historical traffic information is available in the cloud—these information are only used in simulation scenario five. For UMTS, we use the traffic pattern of a mobile operator in Istanbul on an occasion of a big event in May 2012. For the WLAN and WiMax, we interpolate the UMTS data. The traffic patterns for these three networks are listed in Table 3. The volume of traffic on WLAN network is considered higher because multiple WLAN access points could potentially be available in an area of event (in scenario five, we assume three WLAN access points).

4.2 Simulation Scenarios

4.2.1 Simulation Scenario One: Stability of TOPSIS and DIA Rankings

The main goal of the first scenario is to compare stability of the TOPSIS and the DIA rankings. In this scenario, there is a single UE that runs conversational video application. User has access to four networks and based on the (fixed) distance of the user to each of these access points, the network parameters are set to:

Table 3 Network utilization by different applications

Network	Streaming	Background	Video	Audio	Interactive
UMTS	0.2096	0.4248	0.0211	0.0855	0.2590
WLAN	0.1883	0.1454	0.1417	0.1728	0.3518
WiMax	0.2383	0.1542	0.0166	0.1189	0.4720

Net. Param.-UMTS = {1, 40, 60, 4}

Net. Param.-WLAN1 = {8, 120, 50, 15}

Net. Param.-WLAN2 = {4, 135, 35, 12}

Net. Param.-WiMax = {20, 80, 55, 3}

Using the two different rankings of TOPSIS and DIA, networks are ranked the same and in the order of WiMax-UMTS-WLAN2-WLAN1. Then, we remove the worst network (in this case WLAN1) and re-run the algorithm. As a result, DIA ranks the three remainder networks similar to the previous rankings (WiMax-UMTS-WLAN2), while TOPSIS ranks them differently (UMTS-WiMax-WLAN2). From this simple setting, it can be seen that TOPSIS may suffer from ranking abnormality while DIA shows more stability. Based on this result, DIA ranking will be our choice for the remainder of simulation scenarios.

4.2.2 Simulation Scenario Two: Further Study of DIA Ranking Stability

In the second scenario, there is a single user that has access to three networks (one access point of each technology) and during ten consecutive runs of simulation, network parameters are chosen randomly from the range given in Table 2—in other words, it is assumed that users are moving randomly within the coverage area of wireless access points. User runs one session of interactive application. Scenario two is designed to examine whether DIA makes reasonable selections according to the application type and network specifications. These selections should reflect the weights in Eq. (9), where packet loss has the highest weight.

Figure 4a plot shows the selected network by DIA and Fig. 4b plot shows packet loss of the three networks. Observing from these two figures, the selected network at each time instance has the lowest packet loss rate.

4.2.3 Simulation Scenario Three: DIA for Multiple Users, and Applications

In the third scenario, 10 UEs have access to three networks, while parameters of each network are randomly selected at every run of simulation from the range detailed in Table 2.

Fig. 4 Simulation scenario two shows precision of DIA ranking; **a** selected network by DIA ranking, **b** packet loss (in percentage) of all available networks

Users randomly choose an application in each round of the simulation, and simulation runs for ten rounds. The aim of this scenario is to show DIA ranking works precisely with multiple users that have various preferences (different applications).

Observing the first run of simulation, WiMax network is widely selected by the users. Hence, we further assume that congestion level can change latency in the WiMax network, i.e. initially delay in WiMax is a random value in the range [60–100] ms, and after WiMax is selected by three or more users, then its delay will be a random value in the range [150–250] ms. Interesting observations can be drawn from the results plotted in Fig. 5. For example, although WiMax is the most popular network to be selected, it is not selected by the delay and jitter sensitive applications (conversational video and audio) after serving 30 percent of users. Users running other applications (interactive, background, and streaming) select WiMax even after it is congested. This can be clearly seen in Fig. 5 for all time instances. On the other hand, it can be seen that before being congested, WiMax has been selected repeatedly by users running conversational video and conversational audio (e.g., Fig. 5, Time-4–6–10).

4.2.4 Simulation Scenario Four: Modelling Users Mobility

The fourth scenario simulates users' mobility. In the previous scenarios, all network parameters were randomly defined and they were the same for all users and independent of their location. In this scenario, offered data rate by each network varies for different users and depends on their distances to each actual wireless access point. The rest of parameters are randomly selected similar to the scenario three. There are five users that can walk in the range [0–100] meter in any direction from their current location in each

Fig. 5 Simulation scenario three: selected network by ten users six different time instances of the simulation

round of the simulation. Figure 3 shows how throughput varies depending on users' distance to the access points for our three access networks. Each user has a fixed flow throughout the simulation, i.e. user one runs background traffic, user two runs conversational video, user three runs conversational audio, user four runs interactive application and user five runs streaming.

Observing the result in Fig. 6a, various interesting conclusions can be drawn. For example, user one with background traffic, for which data rate is the only important criterion, chooses WLAN throughout the simulation except the 6th round. At the same time we can see in Fig. 6b that user one experiences the lowest WLAN data rate at the 6th round of simulation—the WLAN data rate for user one in the first eight rounds of simulation are {35.42, 41.72, 43.08, 32.25, 27.85, 9.55, 20.61, 31.42} Mbps. Moreover, user five with streaming flow chooses WLAN throughout the simulation but the first time stamp despite the higher throughput—in the first simulation round, WLAN data rate for user five is 30.68 Mbps Vs. the WiMax data rate that is 19.74 Mbps. This selection is mainly due to a higher jitter in the first round of simulation. Data rate of all five users throughout the simulation, and over the three different networks are plotted in Fig. 6b, which can be compared against the selected networks in Fig. 6a.

Fig. 6 Simulation scenario four; **a** selected networks, **b** throughput as experienced, by different users throughout simulation

4.2.5 Simulation Scenario Five: Using the Historical Information from the Cloud

Simulation scenario five integrates the historical knowledge available at the edge cloud into the network selection, by using the RL. We assume five access points are available three of which are WLAN and the other two are WiMax and UMTS. We use the historical information in Table 3 and split the WLAN traffic between three access points as follows (order of data here is the same as Table 3).

$$\text{WLAN1} = \{0.2560, 0.0946, 0.1266, 0.2304, 0.2924\}$$

$$\text{WLAN2} = \{0.0983, 0.1469, 0.1502, 0.2028, 0.4018\}$$

$$\text{WLAN3} = \{0.2103, 0.1947, 0.1482, 0.0857, 0.3611\}$$

The five active users, each run a different application, and their mobility model and how network parameters vary, are the same as scenario four. In this scenario, in addition to ranking, we use RL to benefit from the historical information of network utilization collected from the edge cloud. The RL is also integrated into the selection algorithm, i.e., selected networks by the user in previous rounds affect its decision in the next round. Hence, the ranking by DIA for network selection depends not only on the current value but also on the two previous values. The previous values are weighted (with $W_{h1} = 0.8$, $W_{h2} = 0.6$), and integrated in the decision according to the algorithm in Fig. 2. After ranking is completed by the DIA, the collected data from the cloud is integrated with weight $W_{h0} = 0.6$.

Further assumption is that users are always in the coverage area of all access networks, but when they get to the cell edge (or the coverage edge) of any RAN, a fixed/low value of the data rate is assumed for that RAN to punish the RL algorithm for selecting that network. Figure 7b shows the input to RL algorithm, i.e. $Q_{current}$. Figure 7a shows the selected network by users at each iteration of the simulation.

From the presented results in Fig. 7, it can be seen that users with background and streaming flows keep their initial selection throughout the simulation and users with conversational video and audio flows only handover once, and then stay with the same network for the rest of the simulation time. The most interesting conclusion here is that combining DIA with RL that also benefits from the historical traffic pattern can significantly reduces number of handover.

5 Conclusions

This paper discusses the design of a radio access network selection algorithm based on an autonomous mobile agent integrated with access to the edge cloud. A combination of Multi Attribute Decision Making (MADM) and Reinforcement Learning (RL) are used for selecting the best available network at each UE. Through extensive simulations, we show that our selection algorithm is stable, it improves the QoS satisfaction of the users, and it reduces the number of handovers (changes in the selected network). The concept of network selection, as discussed in this paper, is motivated and in fact enabled, by the advancements in three major areas of mobile communications including, software defined radio, small cell networks and mobile cloud computing. We assume mobile devices have multiple radio interfaces and software defined radio techniques allow the mobile to explore their surrounding wireless spectrum. The small cell mobile networks provide a rich platform of heterogeneous RANs—RANs with different technologies and hence different latencies, data rate and range of coverage. Finally, the existence of mobile cloud computing infrastructure

Fig. 7 Simulation scenario five; **a** selected networks, **b** throughput as experienced, by different users throughout simulation

allows our proposed algorithm to improve its efficiency with little extra computing and battery cost.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Wang, Y., & Zhang, K. (2011). Decision tree based unsupervised learning to network selection in heterogeneous wireless networks. In *IEEE consumer communications and networking conference (CCNC '11)*, Las Vegas, Nevada, USA (pp. 1108–1109).
2. Tabrizi, H., Farhadi, G., & Cioffi, J. (2011). A learning-based network selection method in heterogeneous wireless systems. In *IEEE global communications conference (GLOBECOM '11)*, Houston, Texas, USA (pp. 1–5).
3. Paris S., Martignon F., Filippini I., et al. (2012). A truthful auction for access point selection in heterogeneous mobile networks. In *IEEE international conference on communications (ICC '12)*, Ottawa, Canada (pp. 3200–3205).
4. Niyato, D., & Hossain, E. (2009). Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach. *Transactions on Vehicular Technology*, 58(4), 2008–2017.
5. Kaloxylou, A., Modeas, I., Georgiadis, F. G., et al. (2009). Network selection algorithm for heterogeneous wireless networks: From design to implementation. *Network Protocols and Algorithms*, 1(2), 27–47.
6. Lahby, M., Cherkaoui, L., & Adib, A. (2013). A novel ranking algorithm based network selection for heterogeneous wireless access. *Journal of Networks*, 8(2), 263–272.
7. Sasirekha, V., & Ilanzkumaran, M. (2013). Heterogeneous wireless network selection using FAHP integrated with TOPSIS and VIKOR. In *International conference on informatics and medical engineering (PRIME)*, Tamilnadu, India (pp. 399–407).
8. Song, Q., & Jamalipour, A. (2005). A network selection mechanism for next generation networks. In *IEEE international conference on communications (ICC '05)*, Seoul, South Korea (pp. 1418–1422).
9. Godor, G., & Detari, G. (2007). Novel network selection algorithm for various wireless network interfaces. In *Mobile and wireless communications summit*, Budapest, Hungary (pp. 1–5).
10. Fu, J., Wu, J., Zhang, J., et al. (2007). Novel AHP and GRA based handover decision mechanism in heterogeneous wireless networks. In *Information computing and applications*, Tangshan, China (pp. 213–220).
11. Bari, F., & Leung, V. (2007). Multi-attribute network selection by iterative TOPSIS for heterogeneous wireless access. In *IEEE consumer communications and networking conference (CCNC '07)*, Las Vegas, Nevada, USA (pp. 808–812).
12. Lahby, M., Cherkaoui, L., & Adib, A. (2007). A hybrid approach for network selection in heterogeneous multi-access environments. In *IFIP International conference on new technologies, mobility and security (NTMS)*, Paris, France (pp. 1–5).
13. Tran, P., & Boukhatem, N. (2008). The distance to the ideal alternative (DIA) algorithm for interface selection in heterogeneous wireless networks. In *Proceedings of the sixth ACM international workshop on mobility management & wireless access (MobiWac '08)*, Vancouver, British Columbia, Canada (pp. 1–8).
14. Sgora, A., Vergados, D., & Chatzimisios, P. (2010). An access network selection algorithm for heterogeneous wireless environments. In *IEEE symposium on computers and communications (ISCC '10)*, Riccione, Italy (pp. 890–892).
15. Mahmoodi, T. (2013). Network selection for autonomous mobile devices. In *IEEE wireless on-demand network systems and services (WONS)*, Banff, Canada (pp. 125–127).
16. Chun, B.-G., Ihm, S., Maniatis, et al. (2011). CloneCloud: Elastic execution between mobile device and cloud. In *ACM sixth conference on computer systems (EuroSys '11)*, Salzburg, Austria (pp. 1–14).
17. IBM and Nokia siemens networks announce worlds first mobile edge computing platform. Retrieved from <https://www-03.ibm.com/press/us/en/pressrelease/40490.wss>. Accessed 11 Nov 2017.
18. Bifulco, R., Brunner, M., Canonico, R., et al. (2012). Scalability of a mobile cloud management system. In *ACM mobile cloud computing (MCC '10)*, Helsinki, Finland (pp. 17–22).
19. Saaty, T. L. (2012). *Decision making for leaders: The analytic hierarchy process for decisions in a complex world*. RWS Publications.
20. Hwang, C., & Yoon, K. (1981). *Multiple attribute decision making: Methods and applications. A state of the art survey in lecture notes in economics and mathematical systems* (Vol. 186, pp. 58–191). Springer.
21. Eira, J. P., & Rodrigues, A. J. (2009). Analysis of WiMAX data rate performance. Instituto de Telecomunicações/Instituto Superior, Technical University of Lisbon, Lisbon, Portugal (pp. 1–5).
22. IEEE 802.11g: The new mainstream wireless LAN standard (2003).

Idris Cinemre was awarded a postgraduate scholarship by the Republic of Turkey's Ministry of National Education and completed his MSc in Mobile & Personal Communications with distinction at King's in 2013. He worked as a research assistant in Turkey between 2014–2021, and also enrolled at Karadeniz Technical University in 2014 to pursue a PhD in Telecommunications with a focus on Visible Light Communication. In 2019, he has received the same scholarship again, and joined King's in October 2020 as a PhD student in Telecommunication, concentrating on the functional split in 5G.

Toktam Mahmoodi received her B.Sc. degree in electrical engineering from Sharif University of Technology, Iran, in 2002, and her Ph.D. degree in telecommunications from King's College London, United Kingdom, in 2009. She is currently head of the Centre for Telecommunications Research in the Department of Engineering, King's College London. Her research interests include mobile communications, network intelligence, and low-latency networking.