A Novel Markov Decision Process-based Solution for Improved Quality Prioritized Video Delivery

Matteo Anedda, *Member, IEEE*, Maurizio Murroni, *Senior Member, IEEE*, Gabriel-Miro Muntean, *Senior Member, IEEE*

***Abstract*—The recent growth in both number of high specification mobile devices and network multimedia data demand, has made difficult supporting access to the multimedia content at high user Quality of Experience (QoE) levels. It is even more challenging to support offering such services adjusted to user specific requirements in current heterogeneous wireless network environments (HWNE). This paper proposes a novel Prioritized Adaptive Real-time Multi-user Access Network Selection framework (P-ARMANS) which employs a Markov Decision Process (MDP) solution to perform improved bandwidth resource allocation. P-ARMANS enables load balancing during multimedia delivery when users with diverse priority and different device screen resolutions access various services. Modeling and simulations show how P-ARMANS distributes content with different types of traffic and load among typical and business users at higher QoE levels compared to a classic no-priority approach.**

***Index Terms*—Network selection, Markov Decision Process, Traffic and performance monitoring, QoS, QoE.**

I. Introduction

THERE is an exponential increase in both the number of wireless mobile devices that exchange traffic in current network environment and amount of traffic generated by these devices, which have become one of the major contributors to the global network traffic growth. For instance the annual global IP traffic has already exceeded the zettabyte (1000 exabytes) threshold is set to reach 3.3 zettabytes per year by 2021. Of this IP traffic, wired communications will be responsible for 37% in comparison with 52% in 2015 and wireless and mobile devices will generate 63%. In particular the video traffic is estimated to reach 82% of the total traffic by 2021, growing from 55% in 2015 , especially as increasing amounts of high quality multimedia content are being exchanged between the diverse latest generation devices [1]. Furthermore, according to Cisco, mobile offload exceeded cellular traffic for the first time in 2015 [2], and since then more than 3.9 exabytes of mobile data is offloaded per month to other types of networks i.e. WiFi or femtocells. However most offload is network operator controlled and network driven and very little flexibility exists for a user driven approach.

Currently various access technologies coexist in the heterogeneous wireless network environment (HWNE), including LTE/LTE-A [3] and IEEE 802.11n/ac [4], and no solution alone can support high quality multimedia content delivery to users all the time, while also making commercial benefits. Diverse solutions try to achieve this goal, including by adapting multimedia content delivery [5], [6] or by employing load balancing in HWNE, where different radio access technologies (RAT) exist [7], [8], [9]. Additionally, increasing device numbers with ever-more complex device characteristics, including higher display resolutions, enable offering rich media services

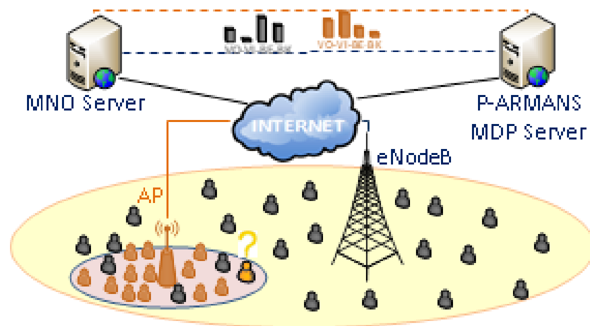

Fig. 1: Load balancing P-ARMANS MDP in a heterogeneous network environment

such as those of Netflix, YouTube, and Amazon Instant Video. These services support delivery of very high quality video content, including 4K video, with impressive details and clarity. In this context, providing high Quality of Service (QoS) levels during remote network delivery of such rich media content and ultimately achieving high user perceived Quality of Experience (QoE) levels is challenging. However, not only rich media content is being exchanged. Considering existing data network traffic mix, some of it has high and other low priority. The high-priority data has specific requirements due to its emergency context (e.g., wired or wireless public safety network control data) or its real-time nature (e.g., streaming applications). The so-called low-priority data delivery can cope with longer latency. Nowadays, within a HWNE, mainly due to economic reasons, there is also a need to consider not only data priority, but also user type. Taking into account heterogeneous communication and user diversity, it should be noted that mobile network operators offer different traffic plans (i.e., standard individuals or business agreements are a non-exhaustive list of these options).

Therefore, in order to support high quality multimedia content delivery in HWNE, there is a need for optimal network selection according to user type, load balancing, data priority and device characteristics (i.e., screen resolution). Despite being a very challenging task, this can be approached by employing multiple attributes decision-making (MADM) methods [10], [11]. This study employs a Markov Decision Process (MDP)-based optimization with the final goal to ensure load balancing and high QoS levels for multimedia delivery considering diversity of users and devices, variable network conditions and data with different priorities.

This paper proposes a novel **MDP-based Prioritized-Adaptive Real-time Multi-user Access Network Selection solution (P-ARMANS)** which supports load balancing and high quality multimedia delivery over HWNEs, as illustrated in

Fig. 1. The proposed P-ARMANS selects dynamically the best candidate network and performs network resource reallocation in order to best balance traffic load and achieve high quality multimedia delivery. P-ARMANS key performance benefits are in terms of better fairness in terms of load balancing and higher QoS and QoE than current state-of-the-art load balancing solutions in HWNE access network selection. Testing results show benefits in terms of multiple quality of service (QoS) metrics including average packet delay reduction (down to 15.3%), average estimated perceived quality increase (up to 15.5%), and average packet loss rate decrease (down to 15.6%) when the proposed MDP-based PARMANS is employed in comparison with when a classic algorithm is used.

The remainder of this paper is organized as follows. Section II summarizes the related works, and section III presents the framework of the proposed P-ARMANS solution, its network load model and the proposed load balancing scheme. Section IV details the MDP-based priority model. In Section V and VI, modeling and simulations considering different scenarios are described and result analysis is provided. Finally, Section VII draws the conclusions.

II. Related Works

This section discusses related works in context of load balancing, prioritized access network selection and MDP.

A. Load Balancing State of the Art

In [12], [13], [14] the authors proposed reputation-based load-balancing network selection strategies for heterogeneous wireless environments. These reputation-based mechanisms select the most appropriate set of networks for the mobile user and a load balancing mechanism to distribute the traffic load among the networks by making use of the different protocols, including TCP, UDP and Multipath TCP (MPTCP). In the above mentioned works the authors introduced a solution based on video delivery content without considering the other type of traffic commonly required by the users that contribute to the overall load balancing. In [15], Araniti et al. proposed to perform radio resource management in dense heterogeneous network environments in order to balance energy saving and delivery performance. Their new HUMANS algorithm offers an additional option of using multicast transmissions in the network selection process during video delivery in order to improve performance. However, the work focuses on LTE microcell and femtocell cases only, without considering available Wi-Fi networks. In [16] Yang et al. presented a novel load balancing scheme conceived for OFDMA LTE cellular and Wi-Fi coexisting networks. The load balancing involves managing of the user equipment (UE) load dynamically. A load balancing algorithm was developed to balance the network load measured by the ratio of the average Wi-Fi load to the LTE cell load as well as by the Jains fairness index. Nevertheless, the proposed load balancing system model does not differentiate the type of mobile client in terms of device characteristics (i.e., screen resolution), application (i.e., type of traffic) and user type (with or without priority). In [17] the authors extended the concept of load balancing to the spatial domain by developing two approaches: Network Load Balancing and Single-Carrier Multi-Link (SCML) for spatial load balancing. Both these methods apply when the device has more than one candidate server and select the server using not only the channel quality from the server to the device, but also the current load on each server. The solutions have merits but also limitations as they do not have any consideration of the type of user, traffic, or devices involved.

B. Prioritization and Radio Access Network Selection

In [18] Wang et al. proposed a priority-based cell selection scheme that divides the cells into different types, decides their priority order, and selects the type of cells with the highest priority. Employing this method, as the number of triggers for cell selection decisions has decreased, when a user moves across cell borders, there is lower probability to trigger cell selection. According to this decision scheme, cell selection (CS) is performed to the cells with the highest priority when a mobile equipment moves across cell border. Thus, the scheme may trigger CS decision just based on user position and no further CS options are considered such as temporary coverage of Wi-Fi, for instance. Liu et al. presented a random-access algorithm considering data network carrying mix traffic, with low and high priority, respectively [19]. The algorithm consists of two dynamically coupled window algorithms, one for the high and one for the low-priority packets. The optimization consists of throughput maximization under no specific delay constraints, and throughput maximization subject to expected delay constraints for the high-priority traffic. Each user continuously observes the channel feedback, from the time a packet is generated to the time that this packet is successfully transmitted. The algorithm limitation includes maximization of throughput with and without delay constraints as the only other QoS parameter. Furthermore, the decision-making process is carried out by the user, thus introducing an additional computational load.

In [20] the authors proposed a solution for sequential dynamic channel selection by different priority users in order to reduce the high priority user access delays. The considered environment contains a number of transmission channels among which each user may select to direct his/her transmission. Time is divided into slots of length equal to the duration of a packet, and the start instants of the slots are identical in all channels. When employing the proposed solution, the access of the high priority users is accelerated significantly at the expense of some throughput reduction.Several studies were carried out for access network selection and multimedia content delivery. Some of these are related to 802.11 WLANs, others involved other heterogeneous wireless networks (HWN) and which include LTE. The latter performs diverse traffic delivery based on QoS Class Identifier (QCI), whereas IEEE 802.11e/n/ac enable QoS differentiation between four different traffic classes (i.e. VO-voice, VI-video, BE-best effort and BK-background), require MAC layer support and assign traffic with static priority. In order to solve this limitation Yuan et al. proposed an intelligent Prioritized Adaptive Scheme (iPAS) to provide QoS differentiation for heterogeneous multimedia delivery over wireless networks [21]. iPAS assigns dynamic priorities to various streams and determines their bandwidth share by employing a probabilistic approach-which makes use of stereotypes. The priority level of individual streams in iPAS is variable and considers service types and network delivery QoS parameters (i.e. delay, jitter, and packet loss rate). In [22], [23] Anedda et al. proposed an adaptive real-time multi-user access network selection (ARMANS) algorithm where overload network problem is accomplished through progressively decreasing the available throughput from high quality to standard quality, starting from users with lower resolution requirements to users with higher resolutions.

The aforementioned works dealt with problems by proposing specific solutions for prioritization and radio access network selection (RANS). However they have either considered a limited set of QoS parameters, have not differentiated the types of traffic, user or device or have used heuristic approaches which do not guarantee accurate solutions.

C. Markov Decision Process in Access Network Selection

Markov Decision Process (MDP) is a rigorous mathematical tool for multi-stage decision making with Markov dependence between the states in different stages. It requires estimates of the transition probabilities associated with each option. MDP procedures are characterized by multi-stage decision problems and are widely used for solving complex issues such as network selection in HWNEs as reported in different works [11]. In [24] Yang proposed a Semi-Markov Decision Process (SMDP) for active state control of a HWN. The author dealt with the problem of a large number of different base stations in idle state and introduced an active state control algorithm based on SMDP with coverage and signal to noise ratio (SNR) constraints. The proposed algorithm properly controls the active state depending on traffic densities without increasing the number of handovers excessively while providing average user perceived rate in a power efficient way in comparison with a conventional algorithm. In this study, active state control has been formulated into a sequential decision problem and one of the most efficient methods for solving sequential decision problem is to employ MDP. In [25] Jakimoski et al. developed an analytical framework using a Markov chain to determine all possible states of the mobile terminals equipped with three interfaces for WLAN, WMAN, and WWAN networks. Such a HWN system is modeled as a Markov chain with appropriate transitions between the states of the mobile terminals. The proposed framework supports a network selection and vertical handover decision algorithm for the three access technologies in order to optimize the performance of the mobile terminal and HWNs. The coexistence of more radio access technologies in the same geographical area offers various advantages. In order to maximize the generated revenue while satisfying the customers increasing demand, the authors of [26] studied a MDP-based radio access technology selection in a cellular/WLAN heterogeneous network environment, where the objective is the maximization of the revenue. Performance is evaluated in comparison with two other policies, namely cellular-first and load balancing. However, most of the previous works introduced priority-based cell selection scheme or occasional accessing requests by public safety-high-priority users. In some cases, QoS differentiation for heterogeneous multimedia delivery over wireless networks (WN) has been reached by assigning dynamic priorities to various streams. Moreover, previous studies do not consider the characteristics of application running at the end-user side and the type of traffic over both IEEE 802.11 and LTE networks. Therefore, this paper proposes the novel P-ARMANS to select the candidate network that best supports the QoS requirements of the transmitted user traffic type (i.e., voice, video, best-effort, and background) and in doing so outperforms existing solutions. Moreover, users are classified by device characteristics, mobility, user-to-user priority and screen resolution. P-ARMANS makes use of MDP to manipulate the optimal transition among the all possible states of the mobile terminal equipped with WLAN and LTE radio interfaces. According to network load, the proposed MDP-based solution assigns throughput shares to each user based on a classification according to its different attributes.

III. P-ARMANS Load System Model

A. PARMANS Framework

The proposed P-ARMANS supports load balancing and high quality multimedia delivery over HWNEs, The framework of the P-ARMANS solution is illustrated in Fig. 2 and consists of three main parts, located at the cloud, client and Mobile Network Operator (MNO), respectively.

The **client** blocks are located at the level of Wi-Fi and LTE-enabled multi-radio user equipment (UE) devices, which can be therefore connected to either an IEEE 802.11 access point (AP) or an LTE evolved Node B (eNodeB). These UE functional blocks are as follows. The *Network Discovery* block provides the user a list of all available wireless connectivity options at their location. The *Device Characteristics* block details information about the device (i.e., screen resolution). Please note that P-ARMANS considers that other device characteristics such as buffer capacity are such set that they do not affect the delivery process. The *QoS Monitor* block measures and sends to the P-ARMANS server unit the quality of the received services in terms of Peak Signal-to-Noise Ratio (PSNR), packet delay and packet loss rate (PLR). PSNR is used for quality assessment as it is one of the most widely used full-reference objective quality metrics. The *Application and User Class Type-ID* block provides to the server unit the current session traffic type. The *user type-ID* is employed to differentiate regular and business users, as the business user traffic has higher priority. Four Wi-Fi classes (i.e., VO, VI, BE, and BK) [4] and four LTE QoS Class Identifier (QCI) options [27] are considered. Information about the *traffic type* in a particular session is periodically reported to the server unit, i.e., at every Transmission Time Interval (TTI). Finally, UEs communicate to the cloud-located server unit their position and speed in order to enable P-ARMANS selection of the best network when moving through HWNE.

The **MNO** blocks collect real-time network characteristics which will be transmitted by the network operator to the P-ARMANS server unit, located in the cloud. Both LTE (i.e.,

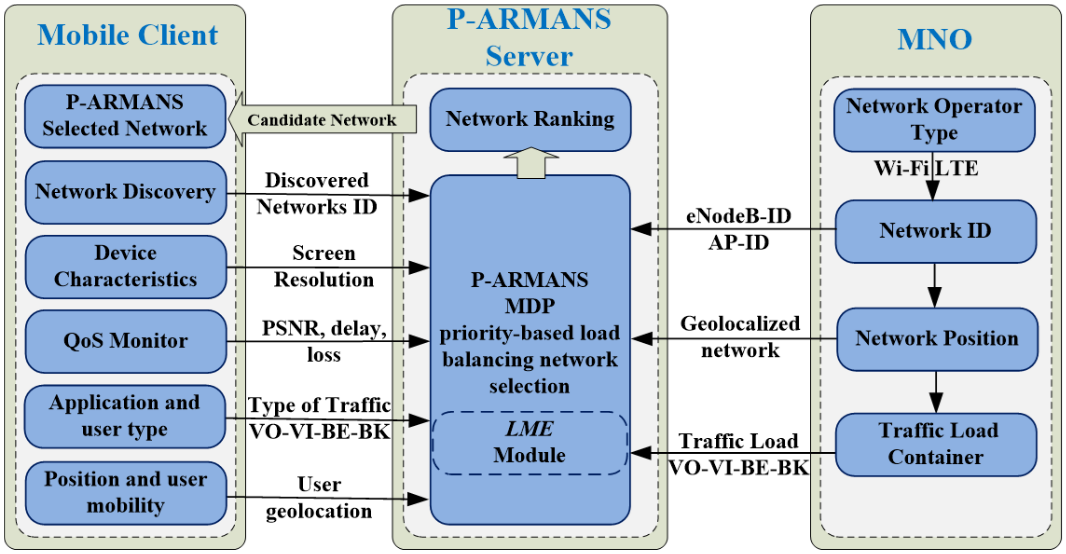

Fig. 2: P-ARMANS scheduling framework

eNodeB) and Wi-Fi (i.e., AP) transmitters are geo-localized and communicate their *network ID* (i.e., eNodeB-ID and AP-ID, respectively), and *position*, respectively. Each transmitter node also communicates the total amount of traffic load and corresponding distribution in terms of different traffic types. For each user, MNOs can keep track of the traffic which can then be used for traffic classification for instance [28].

The cloud-based **P-ARMANS server unit** gathers information from both mobile client and MNO. Data is processed at the server side and sent to the client just after connection and then every time there is a change of the traffic type. The computational cost is linearly related to the number of users and the bandwidth used to transfer this small amount of data is negligible. This unit hosts the P-ARMANS MDP algorithm which computes a network rank list based on estimated delivery quality of multimedia streams. The proposed mechanism generates the network ranking and suggests the best network to each UE either according to the traffic load at MNO, user type and application type running on the mobile UE.

B. Network System Load Model

The generalized network system model considers a HWNE with $l \in \{1, \ldots, L\}$ LTE cells and $w \in \{1, \ldots, W\}$ Wi-Fi APs uniformly distributed within HWNE. It is considered that a LTE base station (BS) eNodeB is located in each LTE cell center. Within the same HWNE, *UEs* are randomly distributed and are able to select either LTE or Wi-Fi connectivity for network access.

Both LTE and Wi-Fi networks employ the same traffic classification based on content $i \in \{VO, VI, BE, BK\}$, where *VO* is voice, *VI* is video, *BE* is best effort and *BK* is background, in order of priority. LTE traffic load is defined as the current total number of resource blocks (RB) for a particular type of traffic scheduled to UEs in a particular *l*-th cell divided to the total number of resource blocks available in that cell. In every TTI, each eNodeB will transmit scheduling information to the UEs.

The LTE load is defined in equation 1:

$$load_l(t) = \frac{\sum_j RB_{l,1}(t)}{RB_{l,max}} = \\ = \frac{RB_{1,VO}(t) + RB_{l,VI}(t) + RB_{l,BE}(t) + RB_{1,BK}(t)}{RB_{l,max}} \tag{1}$$

where $load_l(t)$ represents the number of RBs used within LTE cell l at instant t and is a normalized positive value with the constraint $\sum_i RB_{l,i}(t) \leq RB_{l,MAX}$. $RB_{l,VO}(t)$, $RB_{l,VI}(t)$, $RB_{l,BE}(t)$, and $RB_{l,BK}(t)$ represent the allocated RBs in LTE cell l in the time slot t classified based on type of traffic: VO, VI, BE, and BK, respectively. $RB_{l,max}$ is the maximum available number of RBs in LTE cell l.

Equation 2 defines the global Wi-Fi traffic $load_w(t)$ in AP w at instant t in terms of client network address translation (NAT) activity calculated in Mbps. Obviously, also in this case, the $load_w(t)$ is characterized by the constraint $\sum_i l_{w,i}(t) \leq l_{w,MAX}$. $l_{w,VO}(t)$, $l_{w,VI}(t)$, $l_{w,BE}(t)$, and $l_{w,BK}(t)$ represent the Wi-Fi load in Wi-Fi AP cell w in the time slot t classified based on type of traffic: VO, VI, BE, and BK, respectively. The load $load_w(t)$ is calculated as the overall load l_w per type of traffic at instant t divided by the maximum available resources $l_{w,MAX}$ UEs can access via the Wi-Fi AP.

$$load_w(t) = \frac{\sum_j load_{w,i}(t)}{l_{w,max}} = \\ = \frac{l_{w,VO} + l_{w,VI} + l_{w,BE} + l_{w,BK}}{l_{w,max}} \tag{2}$$

C. Problem Formulation

Considering that new users login to the network requesting a particular type of traffic and old users logout making available new resources, both LTE and Wi-Fi network loads are time-variable. Thus, performing load balancing is an effective method to adjust fairly the distribution of network traffic among cells, while maintaining certain quality of service (QoS) levels. In order to describe the degree of fairness there is a need for a mathematical model or equation. The average load $load_{avg;i}(t)$ is used to estimate the load average for each type of traffic i among L LTE eNodeBs and W Wi-Fi APs as in equation (3).

$$load_{avg;i}(t) = \frac{\sum_{l=1}^{L} RB_{l;i}(t) + \sum_{w=1}^{W} load_{w;i}(t)}{N} \quad (3)$$

where $\sum_{l} RB_{l;i}(t)$ represents the total number of RBs assigned to the L LTE eNodeBs for i-th content type, $\sum_{l} load_{w;i}(t)$ is the total load assigned to the W Wi-Fi APs for i-th content type, and N=W+L are the total number of networks which have the UE in their area of coverage (i.e. and will be discovered by UE).

Given equation (3), we estimate the load average for each type of traffic i. For each of these average values we calculate the difference between $load_{avg;i}(t)$ and the respective loads classified for type of traffic for each $l \in (1, \dots, L)$ LTE cell and $w \in (1, \dots, W)$ Wi-Fi AP in equation 4 and equation 5, respectively.

$$load_{avg;i}(t) - load_{l;i}(t) \begin{cases} \geq 0 & \textit{no overload} \\ < 0 & \textit{overload} \end{cases} \quad (4)$$

$$load_{avg;i}(t) - load_{w;i}(t) \begin{cases} \geq 0 & \textit{no overload} \\ < 0 & \textit{overload} \end{cases} \quad (5)$$

In case of positive difference, the considered network has a load lower than the average load value $load_{avg;i}(t)$. On the other hand, a negative difference shows an overloaded type of traffic network compared to $load_{avg;i}(t)$. Load balancing is performed considering equation (4) and equation (5) values bounded by lower and upper thresholds:

$$\begin{gathered} Th_1 load_{w;i}(t) < load_{l;i}(t) < Th_2 load_{w;i}(t) \\ 0 < Th_1 < 1 < Th_2 \end{gathered} \quad (6)$$

Th_1 and Th_2 represent load balancing thresholds, and have particular settings in specific scenarios (e.g., Th_1 = 0.9 and Th_2 = 1.1 meaning that there is a need for a convergence value of the term $load_{l;i}(t)/load_{w;i}(t)$ within 10%. This approach provides hysteresis and preliminary simulations showed how ineffective repeated modifications were prevented.

D. P-ARMANS Proposed Load Balancing Scheme

This paper introduces P-ARMANS load-balancing scheme in which the UEs may be prioritized or not and heterogeneous devices have different screen resolutions. P-ARMANS employs a load management entity (LME) which manages the traffic type load and obtains load information from the eNodeB and Wi-Fi AP. It then computes the present cell load state to balance network load dynamically. The P-ARMANS load management framework (LMF) contains current access information from $load_l(t)$ and $load_w(t)$. UEs that have access to the LTE BS are registered in set $load_l(t)$ whereas UEs that access to the Wi-Fi APs are registered in set $load_w(t)$. Given the UEs inputs regarding the application running on the mobile device and the discovered LTE and Wi-Fi networks, the LME

Algorithm 1 P-ARMANS load-based network selection pseudo code

```
Result: Write here the result
t = 0
while While TRUE do
    for each network discovered do
        calculate load_l(t) AND load_avg;i(t)
         if load_avg;i l(t) > load_w;i l(t) then
            then no overload
        else
            l overload
        end
        if load_avg;i(t) > RB_l;i(t) then
            then no overload
        else
            l overload
        end
    end
    for each UEs do
        calculate average received power p̄_UE(t)
         calculate power standard deviation σ²_UE(t)
         calculate mobile index v_UE(t)
         if AI_UE(t) ≥ 0.5 then
            then select LTE
        else
            select Wi-Fi
        end
        update RB_l;i(t) OR load_w;i(t)
    end
end
```

computes UE network ranking and determines the set of best candidate network. There are different supports for mobility in LTE and Wi-Fi networks. Wi-Fi can only support UEs with low mobility, whereas LTE can support both low and high UE mobility [31]. In this paper, a *mobility index* is proposed to describe the UE mobility state. The LME computes the mobility index $v_U E(t)$ according to the received reference signal strength in each UE. A sliding window method is adopted to obtain the UE average received power with window size of T. The average received power and power standard deviation for the user UE are formalized in equations (7) and (8), respectively:

$$\overline{p_{UE}(t)} = \frac{1}{T} \sum_{t=0}^{T-1} p_{UE}(t) \quad (7)$$

$$\sigma_{UE}(t) = \sqrt{\frac{1}{T} \sum_{t=0}^{T-1} [p_{UE}(t) - \overline{p_{UE}(t)}]^2} \quad (8)$$

where t denotes the present time slot, $\overline{p_{UE}(t)}$ denotes the average value of received power in the latest T time slots from transmitters under its radius coverage, and $\sigma_{UE}(t)$ represents the standard deviation of a received power from its mean.

The **mobility index** ($v_{UE}(t)$) is defined as a function of the power standard deviation [16], as shown in equation (9).

$$v_{UE}(t) = e^{\quad UE(t)} \begin{cases} \text{close to 0} & \text{high mobility} \\ \text{close to 1} & \text{low mobility} \end{cases} \tag{9}$$

where $v_{UE}(t)$ is a positive value between 0 and 1. The $v_U E$(t) index close to 0 indicates the UE has a very high mobility and an index close to 1 corresponds to very low user mobility.

The **access index** (AI) at instant *t* measures if the UE within the heterogeneous coverage area has good access to Wi-Fi or LTE for type of traffic *i*. AI is defined in equation (10):

$$AI_{UE;i}(t) = v_{UE}(t)\frac{RB_{l;i}(t)}{load_{w;i}(t)} \begin{cases} \text{close to 0} & \text{LTE} \\ \text{close to 0} & \text{Wi-Fi} \end{cases} \tag{10}$$

In equation (10), higher *AI* values indicate the UE should access the Wi-Fi, otherwise, the UE should access LTE. P-ARMANS algorithm is adopted to adjust the traffic type load in existent LTE and Wi-Fi networks. The resulted $AI_{UE;i}(t)$ value determines user connectivity and the data traffic transfers over Wi-Fi or LTE, respectively. The respective load sets are then updated as in equation (11).

$$\begin{aligned} RB_{l;i}(t+1) &= RB_{l;i}(t \quad 1) + RB_{l;i}(t) \\ load_{w;i}(t+1) &= load_{w;i}(t \quad 1) + load_{w;i}(t) \end{aligned} \tag{11}$$

where *t+1* denotes the load at current state, *t-1* represents the load before load updating, and *t* are the requested resources to update the current load.

IV. P-ARMANS MDP-Based Priority Model

A. Markov Decision Process

As already mentioned, MDP provides a widely used mathematical framework for modeling decision-making in solutions to diverse optimization problems in a stochastic environment. A MDP employs a Markov chain (MC) that includes an agent that makes decisions that affect the evolution of a system over time [29].

MDP is used in a dynamic and non-deterministic environment under conditions of uncertainty. This means that when the rational agent makes a decision, the LME module does not necessarily carry out the action required by artificial intelligence, the system could mistakenly reproduce a different action, or perform the correct action but some exogenous factors unexpected changes the final result. So for every action there is a margin of error to consider. A process is called *markovian* when the decision depends only on the current state of the agent. For each state there are associated possible actions that the rational agent can choose. In a deterministic model each action leads 100% to a predetermined next state.

Oher studies such as [26], [30] suggest that, when compared to other methods, MDP allows to find the solution to a non-linear MDP problem in a very short time. It is undoubtedly more efficient than exploring all possible route combinations starting from the initial state for simple systems. With Petri nets, or more complex systems, such as activity networks, long simulations are often used. On the other hand MDP should not be used for highly complex systems, in which there would be a number of abnormal states, making system modeling difficult/impossible.

According to this model, we assume that the decision on whether to connect to Wi-Fi and/or LTE cell (in case both are available) is taken every *t* seconds. Consequently, when only one radio access network (RAN) is available, no decision optimization takes place. The core problem of MDP is finding a policy for the decision maker, i.e., a function that specifies the action (*S*) the decision maker will choose when in state s. Therefore, to identify an optimal policy for deciding which RAN the user UE should connect to, we defined a MDP that associates each state with an action, corresponding transitions and probabilities. In this paper, MDP has been used to define the policy that specifies the action taken by P-ARMANS algorithm when UEs in state *S* request data traffic in an overloaded network scenario. Next sections describe the MDP States, MDP Actions and policy, and the conditions to be verified in order to model moving UEs from one state to another one.

B. MDP States

P-ARMANSs MDP considers eight states, $S_t\ 2\ S$ with *t* = (0, 4K, 2K, FHD, HD, VO, BE, BK). Let s_t be the process describing the evolution of the system state and, let *S* denote the state space. The policy (*S*) of the MDP algorithm is illustrated in Fig. 3 and relate to different traffic types VI, VO, BE and BK, as follows:

S_0 represents the zero state. When all sessions, even the BK sessions are terminated (e.g., the user has no network access or switches off the mobile device), all the allocated resources are returned, and become available to new users (disconnected state);

S_{4K} is a state associated with mobile users holding ultra-high definition screen devices (4K UHD) 2160p (i.e., 3840x2160 pixels), and requesting 4K UHD VI content.

S_{2K} is a state in which users with mobile devices with Quad high definition (QHD) 2K 1440p screen resolution (i.e., 2560x1440 pixels) requesting VI content are placed;

S_{FHD} is a state which corresponds to users with full high definition (FHDi) 1080p (i.e., 1920x1080 pixels) mobile device screen resolution accessing VI content;

S_{HD} is a state associated with high definition (HD) 720p, (i.e., 1280x720 pixels) VI content delivery;

S_t with *t=VO,BE,BK* are states in which users request VO, BE and BK content, respectively.

C. MDP Actions

The actions $a\ 2\ A$ involved in the P-ARMANS MDP algorithm represent transitions from one state to another when various events affect the delivery process e.g., network overload occurs, a user leaves the network, a user changes the desired content. Transitions from one state to another are bidirectional with the unique VI content constraint that a transition from a current state to better one is limited to the native device screen resolutions. For instance, during a VI session a user with a 2K resolution could not make a transition to a UHD state. However, in overloaded network conditions, a 4K user could move to the 2K state and subsequently return to the previous state when enough resources could be again re-allocated. For all actions and VI resolutions we consider there is a solution to generate the data rate associated with the next state following the transition. The set of actions is defined as follows:

- $_{tVI}$ represents the VI user request for VI actions, from the state S_0 to the state s_t with $t = 4K; 2K; FHD$ and according to the native screen resolution device.
- $_{tVI}$ with $t = 4K; 2K; FHD; HD$ represent the VI resource release actions. The $_{tVI}$ actions represent the transition from the current VI t-state to the S_0 state. Furthermore, $1 \quad _{tVI}$ represents the action to remain in the same state when network will update in terms of user and traffic requests.
- $_t$ with $t = 4K; 2K; FHD; HD$ represent the actions to stay at the same state decreasing the available VI data rate from maximum to minimum available.
- $_{t;j}$ represents the action to move from state *t* to state *j*. In particular, we can identify two $_{t;j}$ actions. $_{t;t+1}$ identifies the action to move from one *i*-state to the next *t+1* state. This action is verified when the available resources are not able to satisfy current connected users. The MDP algorithm shares the available resources among users decreasing the data rate while ensuring an acceptable QoS. On the other hand, the action $_{t;t \quad 1}$ = 1- $_{t;t+1}$ characterizes the user action (without screen resolution constraint) to move from i-state to the previous t-1 state with higher data rate provided.

Actions and can be undertaken only by a certain category of users in particular conditions as indicated in section IV.D. Apart from the VI-related actions, this paper considers the user request of other types of traffic and consequently the requests for VO, BE and BK traffic have associated actions $_{VO}$, $_{BE}$, and $_{BK}$, respectively. Furthermore, $_{VO}$, $_{BE}$, $_{BK}$, represent the resources release actions and $1 \quad _{VO}$, $1 \quad _{BE}$ and $1 \quad _{BK}$, are actions to remain in the same state, related to the indicated traffic type, respectively. Depending on device characteristics, all these actions can coexist with each other at the same time. Hence, each user could select a BE download (i.e., $_{BE}$ action) together with a VI request ($_{VI}$ action). Furthermore, the P-ARMANS MDP assumes that requests come sequentially and their treatment is in request order.

D. User Priority Model

Nowadays, users are classified not only according to their device, but also considering their network connection type (i.e., *payment free* or *payment due* networks). In the first case, all users are considered equal, but for the case of *payment due* connectivity, from the wide range of commercial options MNOs offer, this study considers two user profiles: *typical* (TU) and *business* (BU). P-ARMANS algorithm manages and shares the available resources in order to ensure user receive content at QoS levels according to user profile priority. In this context, BU has higher priority than TU.

E. MDP State Transition Probability

The probability that a generic action *a* in a generic state *s* at a generic time *t* will lead to state *s'* at time *t+1*, and is represented by the distribution $P(s^0js; a)$. Some of these probabilities are not defined a-priori, but depend on several factors. In this study, we assume that a generic content request is initialized with a $_{BK}$, action. The probabilities of $_{VO}$, $_{BE}$ and $_{VI}$ actions are defined according to the proposed scenario and the percentage of VO, BE and VI requests, with the constraint P($_{BK}$)+P($_{VO}$)+P($_{BE}$)+P($_{VI}$)=1. The probabilities P($_{VO}$) and P($_{BE}$) are limited to a very small amount of traffic load and have a small influence in comparison with $_{VI}$ actions, hence $P(\; _{VI}) >> P(\; _{VO}) + P(\; _{BE})$.

Particular attention have $_{VI}$ actions. Firstly, whenever a user discovers a single network, P-ARMANS algorithm is disabled in terms of network selection and it can only update its users and requested data rate counters. The discovered network will have TU and BU probabilities $P_{TU} + P_{BU} = 1$ and all $_{VI}$, $_{VO}$, $_{BE}$, $_{BK}$ actions will involve updating the type of traffic and load in the network. In case of medium-high mobility, the algorithm selects the network reducing vertical handover and preferring high network coverage rather than small cell networks. In case of high network load or overloaded network, a random user requesting access network has a probability P_{TU} to be a typical user and a probability $P_{BU} = 1 \quad P_{TU}$ to be a business user. All users are either TU or BU, with the probabilistic constraint $P_{BU} + P_{TU} = 1$. Obviously, P_{TU} and P_{BU} depend on the percentage of TU and BU in the total number of users belonging to a defined MNO. Furthermore, we initially assume that there is a fixed and determined percentage of TU and BU according to MNO statistics on stipulated tariff plans. The transition probability also depends on the number of devices with a particular screen resolution associated to both TU and BU users. In MDP, the conditional probability, introduced as an axiom of probability [31], is considered (i.e., the conditional probability is a measure of the probability of an event given that another event has occurred). Assuming P_{TU} and P_{BU}, we evaluate the joint probability of a 4K, 2K, FHDi or HD state under the condition TU and BU. Hence, $P(4K \setminus TU)$, $P(2K \setminus TU)$, $P(FHDi \setminus TU)$ and $P(HD \setminus TU)$, represent the joint probabilities a request from a user of type TU to fall in the state S_1, S_2, and S_3, respectively. $P(4K \setminus BU)$, $P(2K \setminus BU)$, $P(FHDi \setminus BU)$ and $P(HD \setminus BU)$ are the joint probabilities a request from a user of type BU in state S_0 to fall in the state S_{4K}, S_{2K}, and S_{FHD}, respectively. Hence, the joint probability for a user request to be placed in state s_t is denoted as P(TU or BU $\setminus$ s_t) and is defined as in equation (12):

$$\begin{aligned} P(TU \setminus s_t) = P(s_t jTU)P(TU) \\ P(BU \setminus s_t) = P(s_t jBU)P(BU) \end{aligned} \qquad (12)$$

where t = 4K, 2K, FHD, HD are the considered states. Equation (12) enable introduction of percentage of TU and BU in the initial states S_{4K}, S_{2K}, S_{FHD} and S_{HD}, when $_{tVI}$ actions have occurred. The remaining $_t$ and $_{t;j}$ actions and their probabilities depend upon the political that governs MDP load distribution and retrieval of new resources to satisfy new user requests. For instance, a BU $_t$ action has a probability to be verified equal to $P(\; _t \setminus BU) = 0$ until the total amount of TUs are treated with $_{t;j}$ actions, due the higher BU priority

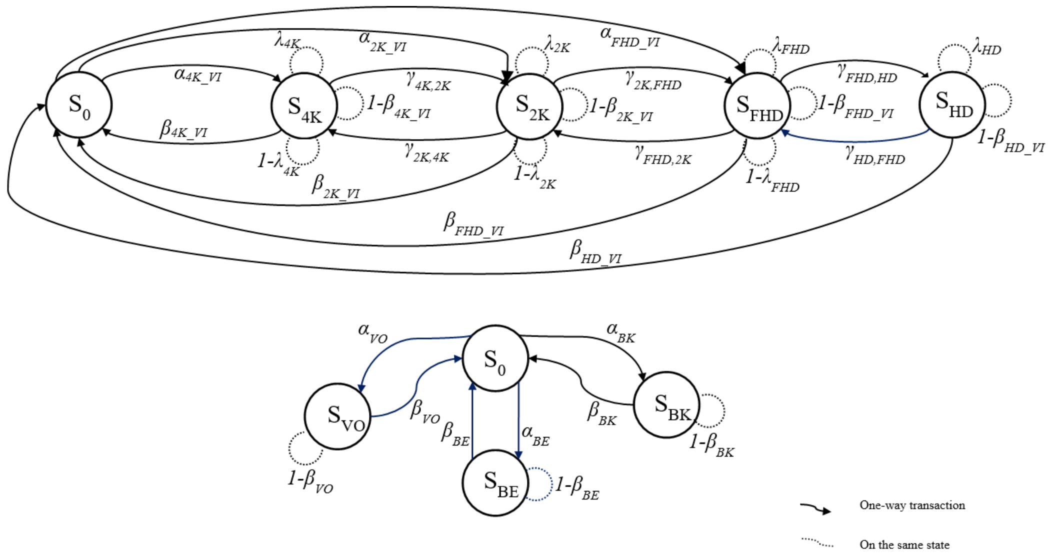

Fig. 3: The policy (S) of the MDP algorithm

TABLE I: Transition Probabilities

Probability	Description
$P(\ _{BK}), P(\ _{BE}), P(\ _{VI}), P(\ _{BK})$	Represent the probability to request traffic type BK, VO, BE, and VI, respectively
$P(\ _{BK})+P(\ _{VO})+P(\ _{BE})+P(\ _{VI}) = 1$	Constraint: when a data connection is established, the sum of the probability to request VO, VI, BE, or BK traffic is equal to 1
$P(\ _{VI}) >> P(\ _{VO}) + P(\ _{BE})$	The probability to request VI traffic type is much greater than the sum of VO and BE type of traffic requests
$P_{BU} + P_{TU}$=1	Constraint: the sum of the probability to be a BU or a TU is equal to 1
P(4K \ TU)	Joint probability of event 4K occurring at the same time that event TU occurs
$P(2K \setminus TU)$	Joint probability of event 2K occurring at the same time that event TU occurs
$P(FHDi \setminus TU)$	Joint probability of event FHDi occurring at the same time that event TU occurs
$P(HD \setminus TU)$	Joint probability of event HD occurring at the same time that event TU occurs
$P(4K \setminus BU)$	Joint probability of event 4K occurring at the same time that event BU occurs
$P(2K \setminus BU)$	Joint probability of event 2K occurring at the same time that event BU occurs
$P(FHDi \setminus BU)$	Joint probability of event FHDi occurring at the same time that event BU occurs
P(HD \ BU)	Joint probability of event HD occurring at the same time that event BU occurs

compared to TUs. P-ARMANS MDP algorithm could be set and proposed in different configurations in such a way that BU $_t$ actions will verify before $_{t;j}$ actions. Both $_{t;j}$ and $_{i;t}$ actions have a probability according to the number of users making VI requests and $_{BK}$ actions. This work assumes $P(\ _{BK}) >> P(\ _{BK})$ equivalent with high traffic load and overloaded networks.

F. MDP Algorithm and Reward Function

P-ARMANS MDP-based algorithm establishes how to assign resources when a new user connection occurs within a heterogeneous wireless environment. P-ARMANS algorithm collects information about the geolocalised user and network conditions under its coverage as shown in Fig. 2 and is explained in III-D.

Typically, when mobile users activate their radio interfaces (i.e., LTE and Wi-Fi in the proposed scheme), it is possible to select different options managed by several applications. For instance, it is possible the automatic switching from LTE to Wi-Fi when a Wi-Fi network is detected or automatic switching from Wi-Fi to LTE when a Wi-Fi connection has been lost or when is under a certain signal strength threshold.

P-ARMANS is deployed at the level of a server unit at the cloud and at a gateway entity placed in between the mobile user and MNO. It dynamically selects the best network for a user according its type of traffic and ensures high QoS level through load balancing. Each not-yet-connected user is placed in state S_0, which represents the starting point of the MDP algorithm. The P-ARMANS process is described next.

n users are placed under the radio coverage of one or more transmitters. Connections and applications running determine BK requests thanks to $_{BK}$, actions. These actions determine a transition from the state S_0 to the state S_{BK}. The BK action represents the transition from the BK traffic state S_{BK}, to S_0 and results in closing connectivity to that particular network. Finally, the action 1-$_{BK}$ means that the user maintains an active connection.

P-ARMANS evaluates the probability of these $1 \quad _{BK}$ actions every TTI, when a new user connects to a network or when a user closes the session. P-ARMANS entity updates the counters of connected users every time any of the above mentioned three cases has been verified.
a user could start a voice session or a BE session by means of $_{VO}$ or $_{BE}$ actions, respectively. These actions determine a transition from the state S_0 to state S_{VO} and S_{BE}, respectively. The $_{VO}$ and $_{BE}$ actions represent the transition from the VO traffic state S_{VO} and BE traffic state S_{BE} to the initial state, S_0. These actions were verified when a user closes the VO or BE sessions. Similar to the BK actions, $1 \quad _{VO}$ and $1 \quad _{BE}$ represent the probability to maintain active a VO and a BE session, respectively. Both actions are analyzed by P-ARMANS entity every TTI, when a new user connects to the network or when it or other users close the session.
a user could start the VI session with $_{tVI}$ actions where $t = 4K; 2K; FHD$. These actions cause a transition from the state S_0 to S_{4K}, S_{2K} and S_{FHD}, respectively. Each user is able to make $_{tVI}$ transitions according to their device screen resolution. In this study the states S_{4K}, S_{2K} and S_{FHD}, represent 4K, 2K and FHD screen resolutions, respectively. Once the P-ARMANS entity recognizes a VI request by a 4K screen resolution user device, follows the $_{4K}VI$ action and transitions from S_0 to S_{4K}. Similarly, transitions to states S_{2K} and S_{FHD}, depending on the user device screen characteristics, respectively. $_{tVI}$ actions with t = 4K, 2K, FHD, HD are associated with closed VI sessions and the correspondent release of resources whereas $1 \quad _{tVI}$ actions maintain users in active mode for the VI traffic sessions.
$_t$ with t=4K, 2K, FHD, HD represent the actions to stay at the same VI state decreasing the available VI data rate from maximum to a guaranteed minimum. Vice versa, 1- $_t$ actions allow the transitions from the minimum to the maximum data rate. It is important to emphasize that the $_t$ and 1- $_t$ actions are available to TUs and BUs and are fully managed by P-ARMANS MDP algorithm.
$_{t;t+1}$ actions represent the transition from the current state to the next worse state in terms of available VI data rate whereas $_{t+1;t}$ actions refer to transition from a current state to an immediately better one. The $_{t;t+1}$ actions make available the saved rate-related resources to other users. These actions are applied by P-ARMANS MDP algorithm to TU users in case of network overload, whereas the BU users are not performing these transitions. When the network load decreases, transitions are performed improving the state level Following this rule, a TU with a screen resolution such that it is originally placed in state S_{4K}, S_{2K} and S_{FHD}3 will be moved if necessary to the state S_{2K}, S_{FHD} and S_{HD}, respectively.

As BUs have priority in comparison with TUs, P-ARMANS MDP-based will apply $_{t;t+1}$ actions to TUs before applying $_t$ actions to BUs. Moreover, MDP will perform $_t$ actions for BUs only when all TUs have been affected by $_{t;t+1}$ actions. Furthermore, we assume that VI content has priority with respect to any other type of traffic. We define $P \ (s; s^0)$ the probability that action in state s at time t will lead to state s^0 at time $t + 1$. We also consider $R \ (s; s^0)$ the immediate reward received after transitioning from state s to state s^0 due to actions , , and . The reward function involves release of resources which are made available to further connections and requests for that particular type of traffic. Table II shows the rewards associated with , , and actions. A value (i.e., +1) is received in two cases: i) after transitioning from the current state to the next state with lower screen resolution video delivery or ii) due to actions through which the devices completely release the requested resources. A negative reward value (i.e., -1) is associated to actions that correspond to resource allocations. Table II summarizes the reward functions involved in P-ARMANS algorithm. The goal of equation (13) is to choose a policy that will maximize the cumulative value of the rewards.

TABLE II: Reward Functions

$P \ (s; s')$	$R \ (s; s')$
$_{2K;4K}$	S_{2K} to S_{4K}
$_{FHD;2K}$	S_{FHD} to S_{2K}
$_{HD;FHD}$	S_{HD} to S_{FHD}
$_{VI}$	$S_{4K;2K;FHD;HD}$ to S_0
$_{VO}$	S_{VO} to S_0
$_{BE}$	S_{BE} to S_0
$_{BK}$	S_{BK} to S_0
$_{VI}$	S_0 to S_{4k}
$_{VO}$	S_0 to S_{VO}
$_{BE}$	S_0 to S_{BE}
$_{BK}$	S_0 to S_{BK}

TABLE III: Simulation Parameters

Parameters	Values
Wi-Fi network	802.11n [4]
LTE Network	LTE Cat. 6 [32]
Number of Wi-Fi AP	3
Number of LTE cell	1
Wi-Fi coverage radius	200m
LTE cell radius	500m
Transmit power of Wi-Fi AP	100mW (20 dBm)
Transmit power of LTE BS	39.8 W (46 dBm)
Wi-Fi/LTE bandwidth	40/20 MHz
Wi-Fi/LTE data rate (DL)	600/300 Mbps
Propagation model	hybrid-indoor and outdoor
User speed	<3, <15, >15 Km/h

$$\sum_{T=0}^{1} m^T R \ _T(s_T; s_T^0) \tag{13}$$

where T represents the action request at time t, m is the discount factor that satisfies $0 \quad m \quad 1$ and is defined as $m = 1=(1 + r)$, and r is the discount rate.

TABLE IV: Type of Traffic Characteristics

	VO	VI	BE	BK
Traffic	Opus or G.711	VP9 and H.264	Pareto distribution traffic model	Pareto distribution traffic model
Protocol	RTP/UDP/IP	HTTP/TCP/IP	RTP/UDP/IP	RTP/UDP/IP
Encoding Data-rate	Table V	Table V	128 Kbps	100 Kbps
Packet size	100 Bytes	1024 Bytes	512 Nytes	512 Bytes

TABLE V: Encoding Video Data Rate

State S_i	Resolution	$VP9$ $H{:}264$ AVC (Mbps)		Audio (Kbps)
		Max	Min	
2160p (4K)	3840x2160	13.0	10.4	128
1440p (Quad HD)	2560x1440	7.549	5.816	128
1080p (Full HD)	1920x1080	3.676	2.804	128
720p (HD Ready)	1280x720	2.674	1.229	128

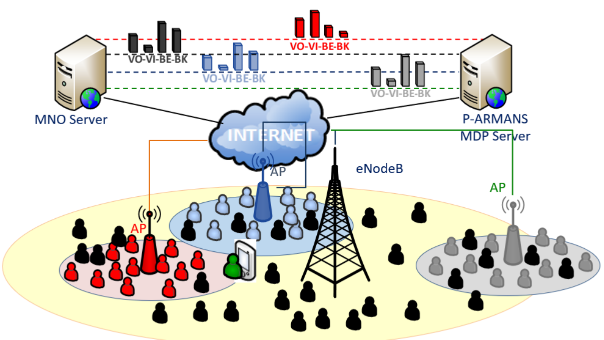

Fig. 4: Scenario 1

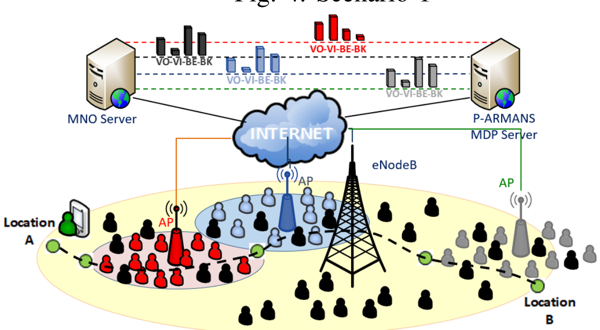

Fig. 5: Scenario 2

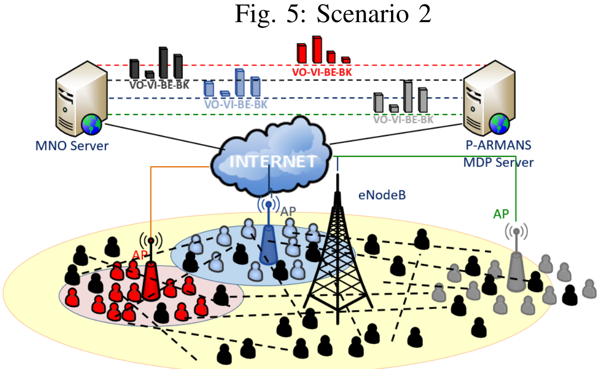

Fig. 6: Scenario 3 and Scenario 4

TABLE VI: Type of Users and Traffic Characteristics

Typical User (TU)			Business User (BU)		
4K	2K	FHDi	4K	2K	FHDi
15%	25%	60%	20%	25%	55%
User with P(TU \ S_i)			User with P(BU \ S_i)		
12%	20%	48%	4%	5%	11%

TABLE VII: P-ARMANS MDP Action Sequence & Data Rate

Step	Step 1	Step 2	Step 3	Step 4
Actions	3- 2- 1	3;4- 2;3- 1;2	4- 3- 2	3- 2- 1
User Type	TUs	TUs	TUs	BUs
User No.	141 - 173	174 - 208	209 - 279	80 - 307
TU Rate	6.04 - 4.70	4.69 - 3.66	3.65 - 2.31	2.31
BU Rate	6.51	6.51	6.51	6.50 - 5.08
Avg Rate	6.21 - 5.05	5.06 - 4.23	4.22 - 3.15	3.14 - 2.86

V. Simulation-Based Testing Setup

A. Simulation Network Environment

The proposed method is scalable and may be applied for load balancing management among different eNodeBs and APs at the same time. We consider a HWNE where cooperate one eNodeB and three APs with the topology illustrated in Fig. 4. It includes one cloud-located P-ARMANS server unit communicating with a scalable number of n clients, via one LTE eNodeB [28] and one IEEE 802.11n Wi-Fi Access Point (AP) [27]. In this work we consider a coverage radius of 200 meters and a transmitted power no higher than 100 mW for the 802.11n. The expected bandwidth is 40 MHz and the maximum data rate in downlink according to the standard and MCS could reach 600 Mbps.

The LTE has a radius coverage of about 500 meters, a transmitted power equal to 39.8 W and a bandwidth 20 MHz, as defined in the standard. The propagation model is a mixed configuration named hybrid-indoor and outdoor. This model allows the user performance evaluation when inside a building or outdoor, in free space and urban environment. Three user speed ranges are considered in the proposed scenarios: static user (i.e., speed<3 Km/h), slow moving users (i.e., 3<speed<15 Km/h), and high speed users (i.e., speed>15 Km/h). P-ARMANS MDP algorithm has been modeled and simulated using NS-3 network simulator and Table III details the simulation parameters.

B. Content Type

Four types of multimedia traffic were used for both IEEE 802.11n and LTE: VO, VI, BE, and BK. In Table IV the characteristics of the four traffic classes are detailed. The most common audio and video codecs ITU-T G.711[1] and H.264/MPEG-4 AVC[2] and VP9[3] , respectively are considered. The encoding data-rates specified in standards and typical packet sizes of 100 bytes, 1024 bytes, 512 bytes and 512 bytes are set for VO, VI, BE and BK traffic, respectively.

Table V presents the maximum and minimum encoding data rates associated with each state Si. These values are recommended for H.264 AVC and VP9 web browsing streaming with HTML5 Video and FLASH video file format. The typical audio 128 Kbps quality, the most common format for audio streaming or storage, is used.

C. Simulation Scenarios

In this work we propose a percentage of BUs among UEs to supply user priority managed by MDP. Table VI shows how TUs have a probability equal to 80% and BUs have a probability equal to 20% to request network resources. Moreover, Table VI includes assume pre-defined percentages of the three higher screen resolution 4K, 2K and FHDi devices, both for TUs and BUs and joint probabilities computed according to equation (12). For instance, there is a 4% probability of BU random access with a 4K device VI request. Network load considered a 5% of users generating BK, BE, and VO traffic only and the remaining 95% of users being involved in high-quality VI requests. The number of mobile devices has been progressively increased in steps of 1 every 2 seconds in order to collect statistics in relatively stable conditions, and increasing the overall offered load. According to Table VII, random TU and BU accesses after 200 seconds and without disconnection cases, consist of 100 users: 80 of these are TUs and 20 are BUs. First 100 users produce a total amount of traffic load around 613 Mbps. According to Table IV we assumed that both TUs and BUs are receiving the maximum data rate given its native screen resolution. This assumption is valid when the network load is between 1% and 95%. Considering 140 connected users at the same time, network load reaches a total amount of traffic load equal to 860 Mbps. Starting from random user number 141, P-ARMANS MDP algorithm starts a decision process based on load balancing and user priority. The deployed P-ARMANS MDP can be summarized as follows:

1) $_t$ are the first actions applied to TUs in the sequence $_{FHD-2K-4K}$ $_t$ actions involve a reduction data rate from maximum to minimum, making available to new users resources previously occupied.
2) $_{t;t+1}$ actions in the sequence $_{2K;4K-2K;FHD-4K;2K}$ are applied to TUs. TUs passed from the native state to the state S_{t+1} at the maximum data rate.
3) TUs suffer a further reduction of data rate with the actions $_{HD-FHD-2k}$.
4) The last step involves BUs. The actions $_{FHD-2K-4K}$ were applied to BUs to decrease the data rate from maximum to minimum, making it available for new users while preserving their screen resolution state.

The proposed MDP ensures a gradual resource re-allocation and a growing number of users perform $_{FHD-2K-4K}$ actions as detailed in step 1. These actions allow an increase in the number of users connected to request VI contents, from 141 to 173. Step 2 forced TUs to decrease the original data rate via $_{FHD;HD-2K;FHD-4K;2K}$ actions. Thanks to these actions the maximum number of connected users increases to 208. Step 3 applies $_{HD-FHD-2K}$ actions and a further VI data rate reduction for TUs. New users requesting VI content are served and the user count increases to 279 users. No other actions involve TUs as in step 4 MDP reduces the available BU's data rate via $_{FHD-2K-4K}$ actions. The number of BUs and TUs managed by MDP after 4 steps reaches 307 users. Given the two networks considered and their characteristics, no more users requesting VI content were further accepted, but more VO, BE, BK requests were accepted. Table V details the VI data rate [Mbps] assigned to each user type in each step. In particular the data rates assigned to TUs, BUs and on average are indicated when the number of users increases. Following the actions in step 1, the TUs start from the average data rate of 6.04 Mbps and it has been limited to 4.70 Mbps. During step 4, there is only a data rate reduction for BUs.

VI. Simulation-Based Testing Results

Four scenarios have been analyzed and are illustrated in Fig. 4, 5, and 6. The number of users considered includes 62 BUs and 245 TUs with a total number of UEs equal to 307.

Fig. 7 shows a first test result where the difference between P-ARMANS MDP and ARMANS are outlined. The horizontal trends for BU and TU aggregate throughput is specific to ARMANS. In ARMANS, when the number of users exceeds 185, new users are admitted without increasing the overall radio resources. The flat throughput trends identify the maximum capacity to be shared among all users, both for BU and TU. On the other hand, P-ARMANS' MDP shares the wireless resources among users while ensuring higher number of connected devices and high QoS. MDP step 1, 2, and 3 act when the users number 141, 154, 164, 174, 177, 193, 209, 250, 263, and 280 require VI content. Limited resources involve actions $_{VI}$, $_t$, and $_{t;t+1}$ during step 1, 2, and 3 for TUs and when no further action could be applied to TUs, MDP acts with actions $_t$ on BUs when users 288 and 296 need VI content. According to the results of this test, we propose 4 different detailed testing scenarios as follows.

[1]G.711, Pulse Code Modulation (PCM) of voice frequencies, https://www.itu.int/rec/T-REC-G.711

[2]Video encoding settings for H.264 excellence, http://www.lighterra.com/papers/videoencodingh264

[3]The WebM project—VP9 video codec summary, https://www.webmproject.org/vp9

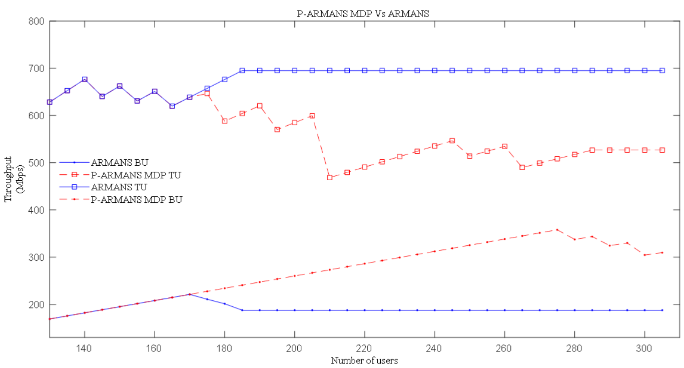

Fig. 7: Throughput for TUs and BUs evaluated with ARMANS and P-ARMANS MDP

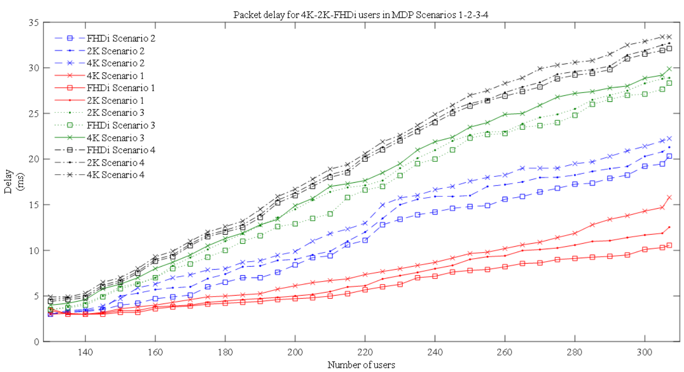

Fig. 8: Packet delay for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP

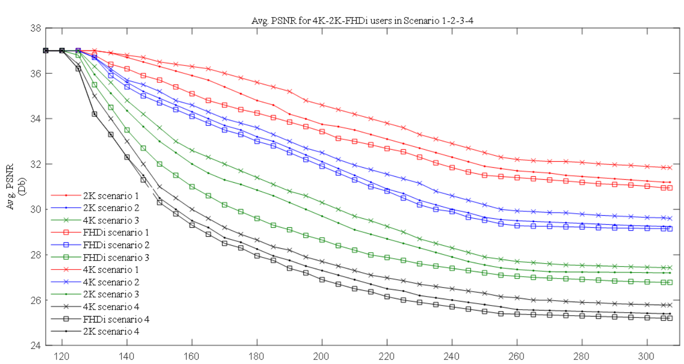

Fig. 9: Average PSNR for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP

In **Scenario 1**, TUs do not move and losses are mainly due to the amount of traffic requested by the TUs and BUs connected to the AP and LTE. Each node generates traffic classified in four different types of traffic, with different device characteristics and screen resolutions. The overall number of users and traffic gradually increases in order to overload both networks. The network hosts a new user every two seconds and this process is repeated for 614 seconds during which the number of UEs increases up to complete saturation. This procedure is repeated for all scenarios.

In **Scenario 2**, one TU user moves on a path towards and then away from the AP and LTE at a constant speed of 1 m/s.

TABLE VIII: Simulation Results - Evaluation

Scenario No.	1	2	3	4	Benefits
Avg. Delay (ms)					=13.28
FHDi	6.199	11.128	16.224	19.184	15.30%
2K	6.974	12.140	16.951	19.500	13.07%
4K	7.924	13.059	17.630	20.119	12.37%
Std Dev	4.57	8.16	11.45	13.20	
Avg. PLR (%)					=6.89
FHDi	5.814	10.093	17.819	21.103	15.56%
2K	5.934	10.842	18.806	21.565	12.79%
4K	6.340	11.591	19.889	22.029	9.71%
Std Dev	2.075	8.20	14.2	15.25	
Avg. PSNR (dB)					
FHDi	34.031	33.467	33.033	27.925	
2K	32.032	31.64	31.466	27.464	
4K	29.547	29.966	28.838	27.219	

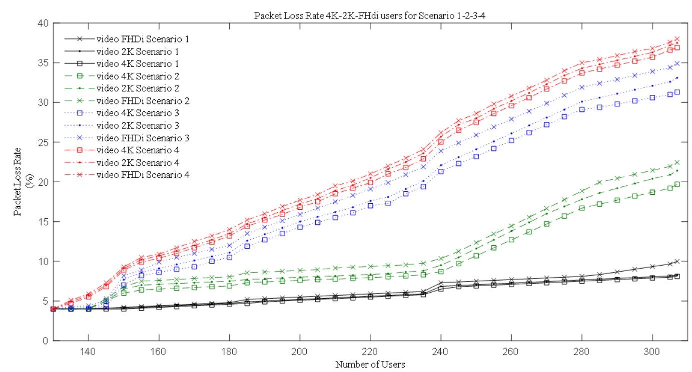

Fig. 10: Packet loss rate for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP

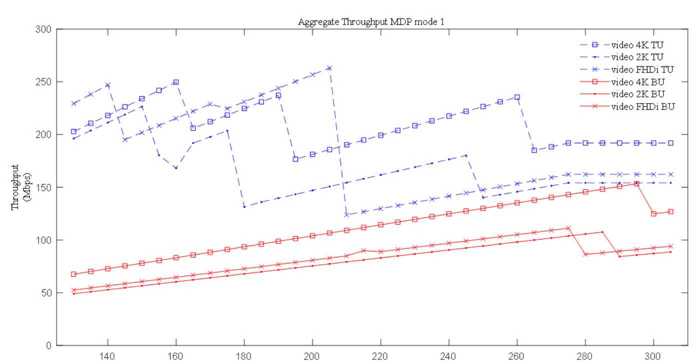

Fig. 11: Aggregate throughput for 4K, 2K and FHDi TUs and BUs evaluated with P-ARMANS MDP

The loss is due to variable received power with increased or decreased distance from the AP and/or LTE, and time variation of traffic load. The other settings are as in scenario 1.

In **Scenario 3** the users are both mobile and non-mobile, and their number progressively increases. Performance was evaluated for a static TU placed within LTE and AP area.

Scenario 4 is similar to scenario 3, but the P-ARMANS load balancing with resolution priority is simulated. P-ARMANS with priority progressively decreases data rate from 4K to FHDi. Each scenario focuses on a TU (i.e., 4K, 2K, FHDi) and its performance evaluation for MDP in cases 1, 2 and 3, respectively. Simulation-based testing analyzed QoS in terms of packet delay, packet loss rate, average PSNR and aggregate throughput comparing P-ARMANS MDP with ARMANS load balancing in each of the four testing scenarios.

Fig. 8 shows in detail the packet delay calculated for a TUs with a different screen resolution when the number of users is progressively increased. All the curves are characterized by a monotonically increasing trend with the increasing number of users. Scenario 1 is characterized by static users within a LTE and Wi-Fi area and the packet delay is evaluated for the users from 130 to 307. For all scenarios, the packet delay for 2K screen resolution users is higher than for the FHDi case and lower than for the 4K. Higher data rate requirements due to higher screen resolution result in higher packet delays when receiving multimedia video content. Measured packet delay for the maximum number of users considered, falls within the range 10.55-15.8 ms for FHDi and 4K screen resolutions, respectively. Scenario 2, similar to scenario 1, involves a TU moving with a constant speed of 1 m/s and three different screen resolution devices. Speed determines an increase in the packet delay in the range of 20.32-22.25 ms when the maximum number of users and traffic load is reached. Scenario 3 considers initially mobility and non-mobility users and progressively increasing the number of mobility users. A generic TU is influenced by the growing number of mobility users and represent the worst case study to evaluate QoS in terms of packet delay. In this case a 4K TU receives packet with a maximum delay of 29.89 ms, whereas FHDi and 2K TUs receive video content with a packet delay of 28.33 ms and 28.90 ms, respectively. Finally, in scenario 4 we evaluate the case in which no-priority and MDP were considered in turn. This case results in higher delay for all screen resolutions. Simulated packet delay falls within 31.12-33.40 ms for FHDi and 4K screen resolutions, respectively.

Fig. 9 presents average PSNR values for different screen resolutions and for all scenarios. All the curves are characterized by a monotonically decreasing trend with the increasing number of users. Scenario 1 is characterized by static TUs and growing number of users in an overloaded network case. Increasing the number of mobile users results in each user being more affected by the others. With respect to scenario 2, a further PSNR reduction has been noted. FHDi screen resolution devices have a PSNR reduction of 2.17 dB, whereas 4K devices loose 2.36 dB. FHDi, 2K and 4K have PNSR values equal to 27.43, 27.2 and 26.78 dB, respectively. Finally, scenario 4 considers the worst case with mobile users and no MDP. In this case, the highest PSNR reduction is 1.8 dB and is associated with a 2K screen resolution user (i.e., from 27.2 dB to 25.4 dB), whereas 4K and FHDi devices experience PSNR decreases of 1.58 dB (i.e., from 26.78 dB to 25.20 dB) and 1.65 dB (i.e., from 27.43 dB to 25.78 dB), respectively.

In Fig. 10, simulated PLR, as experienced at the application layer due to, for instance, excessive packet delays, is characterized by curves with a monotonically increasing trend with the increase in number of users. In scenario 1 a static TU is slowly affected by PLR within the range of 130-180 users where the trend is almost flat. In this range, PLR increased by up to 0.7%, whereas for 180-235 users a further increment of 1.1% is noted, for a global value of 5.8%. For 235-280 users a PLR around 7.5% was recorded for 2K, 4K and FHDi resolutions. Above 280 users, for 4K screen resolution, higher PLR of about 10% is recorded in comparison with 2K (i.e., 8.25%) and FHDi (i.e., 8.1%). Mobility affects high data rate VI traffic, as a significant higher PLR has been encountered in scenario 2 compared to scenario 1. After the first traffic increase and with more than 150 connected users, a growing PLR affects 4K, 2K, and FHDi videos with values of 2.8%, 2.4%, and 1.8%, respectively. A second and significant PLR increase is observed around the 240 connected users and the curves change from almost flat to a severe rise. At the maximum number of considered users, PLR is equal to 22.45%, 21.4% and 19.7% for FHDi, 2K and 4K screen resolution devices, respectively. Scenario 2 has a high impact on PLR when overload occurs with 307 users, with an increase of 12.46%, 13.15%, and 11.45% for FHDi, 2K and 4K screen resolution devices, respectively. In scenario 3 PLR is influenced by the presence of many users with mobility. PLR progressively increases and the simulated values are equal to 34.9%, 33.1%, and 31.3% for 4K, 2H and FHDi screen resolution devices, respectively. Finally, in scenario 4 the P-ARMANS algorithm without MDP optimization achieves higher PLR but has a similar curve than P-ARMANS tested in scenario 3. Fig. 11 illustrates throughput variations in the same conditions.

Table VIII summarizes the benefits introduced by P-ARMANS MDP analyzed in different scenarios and compared to ARMANS algorithm. Packet delay, PSNR, and PLR are compared in terms of average values for each scenario and screen resolution devices. Average packet delay, progressively increase when adding mobility to the users, starting from a scenario with static users and finally evaluating a scenario where all users are moving at a constant sped of 1 m/s. At the same time, average PSNR progressively decreases due to the increment of interferences produced by a growing number of mobile users. The results show benefits when comparing scenario 3 with scenario 4. P-ARMANS MDP has an average packet delay reduction of 12.37%, 13.07%, and 15.30% for 4K, 2K, and FHDi screen resolutions, respectively. In terms of PSNR a delta variation equal to 3.93 is achieved when P-ARMANS manages the access network selection and load balancing in comparison with existing solutions. Finally, MDP algorithm provides advantages also in terms of PLR with percentages of 9.71%, 12.79%, and 15.56% for 4K, 2K, and FHDi screen resolutions, respectively.

VII. Conclusions and Future Work

This paper proposes a novel P-ARMANS MDP-based algorithm to improve user perceived quality in highly loaded network conditions. P-ARMANS performs network selection based on several user parameters such as screen resolution, type of traffic requested, user position, and user mobility. In each network, P-ARMANS evaluates the total amount of traffic and the relative load for different traffic types. Next, it suggests the best network, minimizing the delay and PLR, maximizing throughput and achieving load balancing in HWNE. Testing which considered three screen resolution devices: 4K, 2K and FHDi involved P-ARMANS delivering high-quality video content, which represents the largest component of radio resources requested. A basic two-level priority model was employed among users: business users with a guaranteed data rate and typical users without any privileges. Compared to existing solutions, P-ARMANS provided important benefits in terms of packet delay, PLR, aggregate throughput and estimated user QoE. Noteworthy is that by employing error concealing solutions further improvements in terms of user perceived quality can be obtained. Future work will study how user perceived quality can be further improved, including by employing error concealing solutions. Additional future work considers deploying a MDP-based P-ARMANS adaptation algorithm at the client in an MPEG-DASH-based approach and architecture and assessing its benefits in comparison with that of a sender-based solution as introduced in this paper.

ACKNOWLEDGMENT

The research activities described in this paper have been conducted within the R&D project "Cagliari2020" partially funded by the Italian University and Research Ministry (grant # MIUR_PON04a2_00381). The support of the Science Foundation Ireland (SFI) Research Centres Programme under Grant Number 13/RC/2094 (Lero - the Irish Software Research Centre) is gratefully acknowledged.

REFERENCES

[1] S. Pizzi, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and G.-M. Muntean, "A unified approach for efficient delivery of unicast and multicast wireless video services," in IEEE Transactions on Wireless Communications, vol. 15, no. 12, 2016, pp. 8063–8076.

[2] Cisco visual networking index: Global mobile data traffic forecast update, 20172022 white paper. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html

[3] LTE. [Online]. Available: https://www.3gpp.org/technologies/keywords-acronyms/98-lte

[4] 802.11-2012 - IEEE standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks. [Online]. Available: https://ieeexplore.ieee.org/document/6178212

[5] L. Zou, T. Bi, and G.-M. Muntean, "A DASH-based adaptive multiple sensorial content delivery solution for improved user quality of experience," IEEE Access, vol. 7, pp. 89 172–89 187, 2019.

[6] I. Ghergulescu, A. Moldovan, and C. H. Muntean, "Energy-aware adaptive multimedia for game-based e-learning," in IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting, 2014, pp. 1–6.

[7] F. Zhou, L. Feng, P. Yu, and W. Li, "Energy-efficiency driven load balancing strategy in LTE-WiFi interworking heterogeneous networks," in 2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), pp. 276–281.

[8] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," in IEEE Transactions on Wireless Communications, vol. 12, no. 6, 2013, pp. 2706–2716.

[9] H. H. Yang, G. Geraci, and T. Q. S. Quek, "Energy-efficient design of MIMO heterogeneous networks with wireless backhaul," in IEEE Trans. on Wireless Communications, vol. 15, no. 7, pp. 4914–4927.

[10] J. D. Martnez-Morales, U. Pineda-Rico, and E. Stevens-Navarro, "Performance comparison between MADM algorithms for vertical handoff in 4g networks," in 2010 7th Int. Conference on Electrical Engineering Computing Science and Automatic Control (CCE), pp. 309–314.

[11] R. Trestian, O. Ormond, and G.-M. Muntean, "Performance evaluation of MADM-based methods for network selection in a multimedia wireless environment," in Wireless Networks, vol. 21, no. 5, 2015, pp. 1745–1763. [Online]. Available: https://doi.org/10.1007/s11276-014-0882-z

[12] T. Bi, R. Trestian, and G.-M. Muntean, "Reputation-based network selection solution for improved video delivery quality in heterogeneous wireless network environments," in 2013 IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1–8.

[13] T. Bi, Z. Yuan, R. Trestian, and G.-M. Muntean, "URAN: Utility-based reputation-oriented access network selection strategy for HetNets," in 2015 IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting, pp. 1–6.

[14] T. Bi, R. Trestian, and G.-M. Muntean, "RLoad: Reputation-based load-balancing network selection strategy for heterogeneous wireless environments," in 2013 IEEE Int. Conf. on Network Protocols (ICNP), pp. 1–3.

[15] G. Araniti, P. Scopelliti, G.-M. Muntean, and A. Iera, "A hybrid unicast-multicast network selection for video deliveries in dense heterogeneous network environments," in IEEE Transactions on Broadcasting, vol. 65, no. 1, 2019, pp. 83–93.

[16] Z. Yang, Q. Yang, F. Fu, and K. Sup Kwak, "A novel load balancing scheme in LTE and WiFi coexisted network for OFDMA system," in 2013 Int. Conf. on Wireless Comms. & Signal Processing, pp. 1–5.

[17] K. Azarian, R. Patwardhan, C. Lott, D. Ghosh, R. Gowaikar, and R. Attar, "On spatial load balancing in wide-area wireless networks," in 2012 IEEE Wireless Communications and Networking Conference (WCNC), pp. 2230–2235.

[18] L. Wang, W. Huang, Y. Fan, and X. Wang, "Priority-based cell selection for mobile equipments in heterogeneous cloud radio access networks," in 2015 Int. Conf. on Connected Vehicles and Expo (ICCVE), pp. 62–67.

[19] M. Liu and P. Papantoni-Kazakos, "A random-access algorithm for data networks carrying high-priority traffic," in IEEE Transactions on Communications, vol. 40, no. 1, pp. 84–96.

[20] A. Burrell and P. Papantoni-Kazakos, "Priority users in a multi-channel system," in 2006 40th Annual Conference on Information Sciences and Systems, pp. 1218–1223.

[21] Z. Yuan and G.-M. Muntean, "A prioritized adaptive scheme for multimedia services over IEEE 802.11 WLANs," in IEEE Transactions on Network and Service Management, vol. 10, no. 4, 2013, pp. 340–355.

[22] M. Anedda, G.-M. Muntean, and M. Murroni, "Adaptive real-time multi-user access network selection algorithm for load-balancing over heterogeneous wireless networks," in 2016 IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting (BMSB), Nara, 2016, pp. 1-4.

[23] M. Anedda, M. Murroni, and G.-M. Muntean, "E-ARMANS: Energy-aware device-oriented video delivery in heterogeneous wireless networks," in 2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1–6.

[24] J. Yang, "A semi-markov decision process (SMDP) for active state control of a heterogeneous network," in KSII Transactions on Internet & Information Systems, vol. 10, no. 7, 2016, pp. 3171–3191.

[25] K. Jakimoski and T. Janevski, "Radio access technology selection and vertical handover decision algorithms for heterogeneous mobile and wireless networks," in Adhoc & Sensor Wireless Networks, vol. 27, no. 1, pp. 95–109.

[26] E. Khloussy, X. Gelabert, and Y. Jiang, "Investigation on MDP-based radio access technology selection in heterogeneous wireless networks," in Computer Networks, vol. 91, pp. 57–67.

[27] P. Ameigeiras, J. Navarro-Ortiz, P. Andres-Maldonado, J. M. Lopez-Soler, J. Lorca, Q. Perez-Tarrero, and R. Garcia-Perez, "3gpp QoS-based scheduling framework for LTE," in EURASIP Journal on Wireless Communications and Networking, vol. 2016, no. 1, p. 78.

[28] D. Tsilimantos, T. Karagkioules, and S. Valentin, "Classifying flows and buffer state for youtube's http adaptive streaming service in mobile networks," in Proceedings of the 9th ACM Multimedia Systems Conference, ser. MMSys '18, 2018, pp. 138–149.

[29] H. S. Chang, J. Hu, M. C. Fu, and S. I. Marcus, "Markov decision processes," in Simulation-Based Algorithms for Markov Decision Processes, ser. Communications and Control Engineering. Springer London, 2013, pp. 1–17.

[30] D. Nguyen, T. Nguyen, and X. Yang, "Multimedia wireless transmission with network coding," in Packet Video 2007, pp. 326–335.

[31] D. M. Cifarelli and E. Regazzini, "De finetti's contribution to probability and statistics," in Statistical Science, vol. 11, no. 4, 1996, pp. 253–282.

[32] QoS in LTE. [Online]. Available: http://www.simpletechpost.com/2013/01/quality-of-service-qos-in-lte.html