QoE-Driven Channel Allocation and Handoff Management for Seamless Multimedia in Cognitive 5G Cellular Networks

Md. Jalil Piran, *Member, IEEE* [1,2], Nguyen H. Tran, *Member, IEEE*[1], Doug Young Suh, *Member, IEEE*[2], Ju Bin Song, *Member, IEEE*[2], Choong Seon Hong, *Member, IEEE*[1], and Zhu Han, *Fellow, IEEE*[3]

***Abstract*—Cognitive radio lies among the promising solutions for overcoming the spectrum scarcity problem in the forthcoming fifth generation (5G) cellular networks, whereas mobile stations are expected to support multi-mode operations to maintain connectivity to various radio access points. However, for multimedia services particularly, because of the time varying channel capacity, the random arrivals of legacy users, and the on-negligible delay caused by spectrum handoff, it is challenging to achieve seamless streaming leading to minimum QoE degradation. The objective of this paper is to manage spectrum handoff delays, by allocating channels based on the user QoE expectations, minimizing the latency, providing seamless multimedia service and improving QoE. First, to minimize the handoff delays, we use channel usage statistical information to compute channel quality. Based on this, the cognitive base station maintains a ranking index of the available channels to facilitate the cognitive mobile stations. Second, to enhance channel utilization, we develop a priority-based channel allocation scheme to assign channels to the mobile stations based on their QoE requirements. Third, to minimize handoff delays, we employ the hidden Markov model to predict the state of the future time slot. However, due to sensing errors, the scheme proactively performs spectrum sensing and reactively acts handoffs. Forth, we propose a handoff management technique to overcome the interruptions caused by the handoff. In such a way that, when a handoff is predicted, we use scalable video coding to extract the base layer and transmit it during a certain interval time before handoff occurrence to be shown during handoff delays, hence providing seamless service. Our simulation results highlight the performance gain of the proposed framework in terms of channel utilization and received video quality.**

***Index Terms*—5G cellular networks, cognitive radio networks, handoff, multimedia communication, quality of experience.**

I. Introduction

Mobile traffic is experiencing its breakneck proliferation and explosive growth by development of diverse applications, e.g. multimedia services, placing a tremendous strain on wireless communications capacity. Vodafone revealed that its customers used 400 petabytes of data on their mobile phones during August to October 2015, which is on average twice as much 4G data as they did on 3G [1]. The popularity of multimedia applications is the main cause of this growth whereas CISCO predicted that 69.1% of mobile traffic would be occupied by video down/uploading [2]. According to the mentioned statistics, the next mobile generation, 5G, will face multifold challenges, e.g. unevenly distribution and diverse increasing of mobile traffic. Therefore, various stringent requirements are solicited, such as near zero delay, scalability to support massive number of devices, high reliability, spectrum and bandwidth flexibility, network and device energy efficiency, etc [3]. Spectrum allocation is considered as one of the most crucial problems among the mentioned requirements. It is due to the inefficient traditional spectrum allocation, in which fixed licensed spectrum bands within ultra-high frequency bands that are allocated to the cellular networks, do not have the capability to satisfy users' expectation of high-speed wireless access for 5G cellular networks. To solve this problem and achieve higher spectrum flexibility in cellular networks, different types of technologies are under consideration such as; carrier aggregation in LTE-U and licensed assisted access (LAA) [4], LTE-WiFi aggregation (LWA) [5], MuLTEFire [6], operation on millimeter wave bands [7], and cognitive radio (CR) [8]. According to CR features such as opportunistic and shared spectrum access, in this paper we consider CR to be utilized in our framework.

[1] Department of Computer Engineering, Kyung Hee University, Yongin-si, 446-701, South Korea, email: ({piran, tran, cshong}@khu.ac.kr). [2] Department of Electrical and Radio Engineering, Kyung Hee University, Yongin-si, 446-701, South Korea, e-mail: ({suh, jsong}@khu.ac.kr). [3]Department of Electrical and Computer Engineering, University of Houston, Houston, TX, 77004, USA, e-mail: (zhan2@uh.edu).

A. Cognitive Radio for 5G Cellular Networks

Due to its unique characteristics such as flexibility, adaptability, and interoperability, CR is a promising candidate technology to address spectrum scarcity issues for 5G cellular networks. Therefore, 5G standardization leaders, such as 5GPPP in [9], ITU in [10], and IEEE in [11], considered CR as a candidate technology for enhancement of spectral efficiency in 5G.

Indeed, there are many essential similarities between 5G and CR: (1) Inter-working with different systems and/or networks. (2) Adaptation based on the access network features (5G case), respective on the primary system network characteristics (CR case). Remark that the 5G integrated access networks and CR primary systems networks have the same meaning. (3) New and flexible protocols. (4) Very advanced PHY and MAC technologies. (5) A very advanced terminal, endowed with the possibility to scan the environment, with intelligence and decision capabilities. (6) Resource management (and end-to-end integrated resource management that should include all the network involved in the data transmission process.)

Thus, whereas CR was proposed to enhance the spectrum efficiency and 5G is going to imply the interconnection of the wireless networks all around the world, integration of these two technologies (*Cognitive 5G Cellular Networks*) is a promising solution to tackle or partly address the next generation mobile communication issues. To enable 5G devices to have cognitive capability, the mobile devices only need to be equipped with Software Defined Radio (SDR) to be reconfigured with the other network protocols, e.g. 802.16, 802.11, 802.22, etc. Thus, with the help of CR technology 5G devices can discover spectrum opportunities with one or two transceivers and there will be no need to have additional devices.

B. Problem Statement

In *cognitive 5G cellular networks*, Cognitive-enabled Users (CUs) are allowed to occupy the unused licensed/unlicensed channels that are not used temporally or spatially by Licensed Users (LUs) and other CUs. The licensed bands are the fixed spectrum bands that are bought by the operators and on the other hand, the unlicensed bands are the bands that have been left free for unlicensed users, e.g. 2.4 GHz ISM, 5 GHz U-NII , and 60 GHz mmWave.

One of the main problems in cognitive 5G cellular networks that need to be considered is dynamic channel allocation, which may seriously degrade service quality and increase transmission failure and delay, particularly in multimedia services. Among different multimedia services, some are considered as more important and critical than the others based on their QoE requirements, delay-sensitive VoIP service as an instance . Therefore, it is necessary that better quality channel should be provided to the traffic classes with higher priority.

Upon successful utilization of a channel by a CU that is not dedicated to CUs, the CU must immediately vacate the primary spectrum bands whenever any LU reclaims the channel (known as spectrum mobility). Moreover, spectrum mobility may happen when a CU moves to another cell or when the quality of the current channel becomes poor. Spectrum mobility presents the spectrum handoff function. During spectrum handoff, the current communication process should be transferred to another available channel.

For channel mobility, it takes significant time called as spectrum handoff delay, which is the time spent to search for another available channel and radio frequency front-end reconfiguration process by CUs. *Thus, another severe issue to be considered in cognitive 5G cellular networks is the spectrum handoff delay, which particularly in multimedia services gives rise to interruption of service that dramatically reduces QoE.*

C. Our Contributions

To overcome with the above-mentioned issues, in this paper, we study a DB-assisted and index-based channel allocation and handoff management framework, based on channel usage statistics, HMM, and scalable video coding (SVC). The main contributions of the paper are described briefly as follows:

1) To predict the arrival time of LUs as well as to select the best available channel, we propose a DB-assisted and index-based scheme to collect channel usage information and rank the available channels in accordance with their quality. The CBS maintains a ranking index of the available channels based on their parameters such as; detection accuracy, false alarm probability, LU idle time, and LU arrivals. The channels are ranked into different classes based on their QoS parameters. The use of ranking index decreases interference with the LUs significantly, maximizes the CU spectrum occupancy, minimizes the number of interruptions, and handoff delay.
2) Whereas, all the traffic sources may not have the same QoE requirements, we develop a priority-based channel assignment paradigm. The scheme classifies CUs' traffics based on their QoE requirements and reserves an appropriate class of the available channels for each traffic class.
3) We formulate channel state estimation in the temporal domain in HMM framework. The transition pattern of the LUs is modeled by a Markov chain. We use CUs' experiences and channel history to predict the next time slot state e.g. handoff occurrence. To take care of spectrum sensing errors as well as prediction errors, our system performs handoff prediction proactively but channel switching is performed reactively upon LU detection.
4) We propose a handoff interruption management to provide seamless multimedia service. If the mechanism of channel state estimation predicts that a handoff will occur in future slots, we use SVC to encode the video sequence into two layers: base layer (BL) and enhancement layer (EL). Due to the lightweight of BL and its higher priority, the server sends only the BL code in a certain interval before handoff occurrence. Then, if arrival of a LU is detected by the CU, while the CU is searching for a substitute available channel in the index (during the handoff delay), the receiver presents the pre-fetched BL code to hide handoff interruption from the user perspective.
5) Our framework can be distinguished from the other related proposed scenarios by its unique merits that can be stated in four aspects.
    a) According to the fact that multimedia services are bandwidth-hungry, and delay- and distortion-sensitive, so they need high quality as well as stable and reliable channels in order to achieve the target QoE. However, identifying such kind of channels needs a longer sensing time that shortens the transmission time and increases delay and distortion. Our proposed framework with its unique features such as QoE-drivenness and channel quality-awareness, interestingly outperforms comparing to the other proposed scenarios.
    b) By employing HMM and reactively handoff performing, the scheme significantly reduces handoff delay, limits interference to LUs, and improves spectral efficiency.
    c) While the other related works ignore QoE, this framework effectively provides seamless multimedia service and improves QoE, by extracting, transmitting, and presenting BL during the handoff delay using SVC.
    d) And finally, the unique feature of the proposed framework is to consider all the three important requirements that an efficient CR-based scheme requires, i.e., channel utilization, QoS provisioning, and QoE-drivenness.
6) We conduct a comprehensive simulation study to evaluate

the effectiveness of our proposed framework in several aspects such as the number of interruptions, average handoff delay, average number of collisions, the quality of the reconstructed video, and end-user perception. In the cases of the number of interruptions, average handoff delay, and average number of collisions, we compare the proposed work with reactive and proactive random channel selection, greedy non-priority channel allocation, and fair proportional channel allocation. The proposed scheme reduces the number of interruptions up to 25%, and for handoff delay and collisions, it outperforms much better than the other four methods. Finally, we have observed how effective the proposed framework improves the quality of the reconstructed video in terms PSNR and MOS.

The rest of this paper is organized as follows. In Section II, we review and discuss the related work. The system model is presented in Section III. Our proposed QoE-driven channel allocation and handoff management framework is described and formulated in Section IV. Section V, validates the derived results and analyzes the performance behaviors. Finally, in Section VI, we draw conclusions.

II. Literature Review

In recent years, plenty of research work has been conducted to manage the issue of spectrum handoff. The authors in [12] investigated the tradeoff between the spectrum handoff delay and channel busy duration time. For initial and target channel selection, a probabilistic approach was introduced in [13]. The authors provided analytical results for the switching-enabled policy according to the connection-based spectrum handoff. The authors in [14] studied spectrum handoff in three aspects: non-spectrum handoff method, the pre-determined channel list handoff, and the spectrum handoff based on radio sensing scheme. Then, they evaluated them in terms of link maintenance probability and the effective data rate of SUs' transmission.

Furthermore, there are some proposed schemes in the literature, which consider two modes: *reactive* or *proactive*. In the reactive mode, the spectrum switching by SUs is performed after detection of the PU arrival, while in the proactive mode SUs forecast the channel status based on PU activity and vacate the channel before reappearance of PUs.

In case of *reactive*, a Markov transition model integrating with the preemptive resume priority (PRP) M/G/1 queuing network was introduced in [15] in order to calculate handoff delay caused by sensing, handshaking, channel switching, and the waiting time. The authors in [16] compared the pros and cons of proactive-sensing and reactive-sensing spectrum handoff techniques. Furthermore, the authors proposed a greedy algorithm that determines suboptimal target available channels. The proposed algorithm can automatically switch between the two techniques. *Although because of on-demand spectrum sensing, reactive schemes may get an accurate spectrum hole, they are involved with longer handoff latency.*

On the other hand, in *proactive* case, a spectrum management scheme called voluntary spectrum handoff (VSH) [17] is introduced for CRNs to minimize SU disruption periods during spectrum handoff. The transition probability selection (TPS) and reliability based selection (RBS) algorithms are employed to estimate the spectrum life time, to determine voluntary spectrum handoff time. *Proactive methods are involved with much smaller latency comparing to reactive methods. However, they are suffering from two common detection errors; i.e.; false alarm and miss detection.* By predicting a handoff and as the SU leaving the channel in the predicted time, there will be no PU arriving then, it will result in waste the precious spectrum resources (false alarm). On the other hand, if the scheme does not detect the PU accurately, then it will result in interference to PUs (miss detection).

Most of the proposed schemes considered reactive handoff or proactive handoff to minimize handoff delay while ignoring QoE and channel quality. This may make them unsuitable for multimedia applications where service interruptions significantly degrade the QoE. To the best of our knowledge, there is no previous research work concentrating on providing seamless streaming and improving QoE during handoff delays for multimedia services.

III. System Model

We adopted a time-slotted cognitive-based cellular system having X primary channels with identical bandwidth. Each primary channel has N time slots and each slot is composed of two parts, i.e., sensing (link setup) and data transmission [37]. At the beginning of a time slot, the CU performs the sensing function via energy-detection technique to discover the status of the channel and sends the collected data to the CBS. The traffic pattern of LUs on the licensed channels is designed as a two state Markov model, the ON and OFF renewal process alternating between busy and idle periods. In the queue theory, LU arrival process follows a Poisson arrival model and exponential service distribution.

Let's suppose that CU arrival also follows a Poisson arrival mode. The Poisson arrival process is employed to inspect the channel usage efficiency of the CUs links according LU arrival rate and the number of channels. The traffic modeled by a Poisson process with the arrival rate of λ is adopted to a Markov chain model by calculating the success probability $P_{success}$ by the chain state estimation.

Although 5G devices owe a licensed spectrum band, with the help of CR they will be able to capture more bandwidth to have a higher data rate in order to improve the end-user QoE, particularly for multimedia services. Hence, in *cognitive 5G cellular networks*, mobile units equipped with CR technology are considered as CUs, whenever they need to access other spectrum bands, whether it is licensed or unlicensed. Compared to normal 5G mobile stations, CUs will be able to adjust their QoE expectations, according to the network fluctuations, e.g. dynamic channel state. According to the coverage range of the various CBSs in 5G networks and also based on the spectrum type that they access, our proposed framework is applicable in the following modes:

- It can help to extend the coverage to rural and indoor areas to provide broadband Internet service by WAN. The broadband Internet can be provided through a CBS that is connected to WWW. To implement this scenario there is a need to deploy one or more fixed CBSs and many available radio channels. The CBS is assumed to adjust its transmission power at or below the threshold that has been defined by the local regulator.

- Another scenario is to use CR as the backhaul, whereas both CBS and relay have CR functionality. So that both of them are able to select the same available channel for the backhaul link dynamically.
- The next scenario that can be considered is to apply CR to small cells. According to the fact that small cells require low transmission power and more spectra and higher bandwidth, CR can give them the capability to access more available channels. An advantage for this scenario is the mitigation of co-channel interference among neighboring small cells and reducing the interference between a small- and a macro-cell. This would result in achieving higher network throughput as well. This scenario can be considered for both indoors and outdoors.
- CR can help 5G to access more spectrum for the capacity enhancement. In this scenario, basically the operators use the licensed bands as the main component carrier, which facilitate a stable link to the mobile devices. In cases where a higher data rate is needed, a CBS discover spectrum is provided as a supplement.
- Moreover, although we considered a CR-based framework, even by eliminating the CR, our proposed handoff management technique is applicable to LTE-U, LAA, LWA, and MuLTEFire as well.

A CU is supposed to utilize two transceivers with SDR capabilities. One of the transceivers is allocated for transmission and the other one is for channel sensing with two functionalities: monitoring the channel usage and collecting channel information to estimate future channel state [38]. The channel status, busy or idle, is detected at the beginning of each slot by the CU with energy detection (ED) spectrum sensing techniques [18]. If the status of the channel in the slot is idle, the CU starts to transmit till the end of the slot.

A CU that occupied a primary channel must leave it upon arrival of incumbents immediately, as recommended in IEEE 802.22-WRAN (Wireless Regional Area Networks) [19]. In the standard it is assumed that the CU needs to vacate the occupied primary channels within two seconds from the arrival of LUs. The CU requires to search and move to a new available channel in three cases; when the legacy users reclaim the channel, when the quality of underused channel become poor, and when CUs move spatially and the CU transmission coverage may overlap with an LU that currently uses the same channel.

The process of switching to another available channel poses a non-negligible delay. Such kind of time delays are called as the handoff delay that is cumulative of the time taken for: channel discovering, spectrum sharing, switching to the new available channel, and link establishment.

Since the CU has to temporarily break the transmission process during spectrum handoff to search and switch to another available channel, and whereas, handoff may occur several times during a transmission session, particularly in the case of multimedia transmission, the delay caused by handoff processes interrupts the transmission for several times that prevents providing seamless transmission and significantly degrades QoE. In order to manage handoff delays in multimedia services, our scheme focuses on three aspects: (1) obtaining channel statistical information and mitigating interference to the LUs, (2) selecting the optimal candidate channels during the handoff process and minimizing handoff delay, and (3) pre-fetching the light weight BL code to present during the handoff process to provide seamless service and improve QoE.

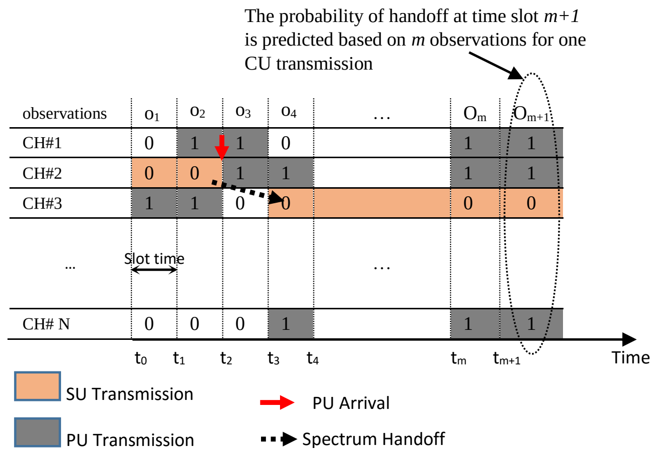


Fig. 1: Spectrum Handoff Process.

Fig. 1, illustrates the operation of a CU in primary channels. We assumed there are X primary channels having synchronized slotted structure. At the beginning, $CH\#2$ is allocated to the CU by the CBS, because it is assumed as the best available and most reliable channel in the ranking list in comparison to the other available channels and it is matched with the traffic QoE requirements, e.g. $CH\#1$. The CU calculates sensing accuracy and idle duration and forward to the CBS to store into the DB. In addition the CU monitors the other channels and maintains the ranking index to choose the backup channel in case of handoff.

At time t_0, a spectrum handoff is predicted to occur at t_2. The handoff delay is predicted to be one slot. The CBS checks the ranking index and allocates $CH\#3$ as the substitute channel to resume CU's transmission over it because it is assumed to be the first channel in the corresponding class in the index. Then, equal to the predicted handoff delay, only the extracted BL code is transmitted in advance to the receiver during the second slot to be viewed during the third slot. At time t_2, an LU is detected by the CU. The sender pauses its transmission, while the receiver starts to show the pre-fetched BL code during the third time slot. As the best matched available channel in the ranking index, $CH\#3$ is facilitated to the CU. The CU moves its transmission to $CH\#3$ and resumes it at t_3 without interference to the LU which starts its transmission over $CH\#2$ at t_2.

IV. The Proposed QoE-Driven Channel Allocation and Handoff Management Framework

Our objective is to allocate the best available channel and to manage the unavoidable spectrum handoff by minimizing the handoff latency and handle the handoff process in such a way to provide seamless multimedia service as well as improving QoE. In doing that, our system consists the three phases, each composed of some functionalities, as illustrated in Fig. 2:

- **Phase I:** *Channel evaluation and allocation*
  - Data collection via ED spectrum sensing technique,
  - Channel quality estimation based on sensing accuracy and channel idle duration,

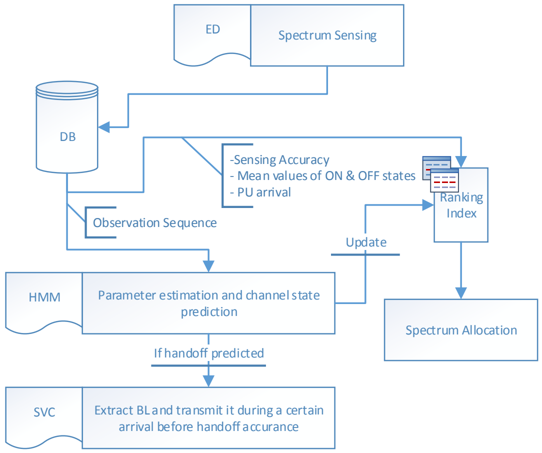


Fig. 2: QoE-Driven Channel Allocation and Handoff Management Scheme.

  - Channel allocation according to the CU's QoE requirements and the available channels' quality,
- **Phase II:** *Channel state estimation*
  - Parameter estimation using the Baum-Welch algorithm (BWA) under the HMM parameters,
  - Channel decoding for channel state estimation
- **Phase III:** *Handoff interruption management*
  - Extraction of the BL code using SVC and sending it during a certain interval before handoff occurrence.

The system functionalities related to each phase are explained and formulated in the next subsections.

A. *Phase I: Channel evaluation and allocation*

In this subsection, we explain and formulate different functionalities of the first phase including data collection, channel quality estimation, and channel allocation. Algorithm 1, illustrates a concise procedure for channel evaluation procedure. The functionalities related to this phase are described and formulated in the following subsections.

1) Data Collection: Dynamic spectrum access consists of spectrum sensing, networking, and regulator policy. Spectrum sensing for a fixed spectrum channel can be performed in two domains: *Spatial Domain*, in which the CR-enabled device estimates the location of the primary transmitter and based on that adjusts the transmission power to prevent interference to the transmitted signals, and *Temporal Domain*, in which the CUs try to discover the time intervals that the primary transmitter is idle.

In this paper, we follow the temporal spectrum sensing and data collection is performed through channel sensing at the start of each slot. During sensing time, the CU first senses the LU activity on the target band and decides to transmit if and only if there is no active LU. We use the ED technique for channel sensing due to its efficiency and fast non-coherent features that essentially calculates a running average of the signal power over a window of a given spectrum length.

In our technique, CU compares its received signal with a sensing threshold. The sensing threshold is based on the noise floor that is to specify the transmission state of the primary transmitter. This technique does not need any prior information regarding LU's signal and its performance is poor in the environments with

Algorithm 1: Channel Evaluation

Input: Set of Accessible Channels
Output: Ranking Index of the Available Channels
1 **for** $i = 1$ *to* $i \leq x$ **do**
2 sense channel i at during the sensing part of the first slot
3 **if** $(\sum_{n-1}^{N}(R_s^j[n])^2) > \Theta$ **then**
4 monitor $E_j[T_1]$
5 forward the collected data to the CBS
6 update the index
7 **Break**
8 **else if** $(\sum_{n-1}^{N}(R_s^j[n])^2) < \Theta$ **then**
9 calculate $\omega_i = P_d^i(1 - P_{fa}^i)$
10 estimate $E_i[T_0]$
11 calculate $Q_{CH}^i = (1 + \log_\rho \omega_i)E_i[T_0]$
12 forward the collected data to the CBS
13 CBS assigns the discovered channel to the corresponding class
14 CBS updates the index

low SNR. However, the sensing function is not always precise because of limitation in the observation quantity and noise.

Algorithm 1 (Channel Evaluation)

DURING the sensing part of each slot, the CU senses the target channel to collect the channel's information. If the channel is busy (Line 3), the CU just monitor busy period and forward its collected data to the CBS to update the index (lines 4-6). Otherwise, if the channel is sensed as free (line 8), the sensing accuracy is calculated according to the false-alarm and detection probabilities (lines 9). Then, the duration of idle periods is estimated (line 10). Based on the sensing accuracy and idle duration, the quality of the channel is computed (line 11). The collected data are forwarded to the CBS (line 12). The CBS assigns the discovered channel to its corresponding class (line 12), and update the index with the new changes (line 14). If we consider N as the number of accessible channels, then the the computational complexity that a CU processes n channels is $O(N)$. For the next M time slots, the total computation complexity for which the CU processes N channels is $O(NM)$.

Normally, two types of errors happen in sensing of LU activities, which are known as false-alarm and miss-detection errors. Miss-detection phenomenon occurs when the user fails to sense the presence of the LU, which results to interference to the LUs. And on the hand, the probability of false-alarm occurs when the sensing function declares falsely that there is an active LU, which results in waste of spectrum resources [20]. Assuming Rayleigh fading channels [21], miss-detection probability P_{md} and false-alarm probability P_{fa}, are respectively as follows using complete and incomplete gamma and generalized Marcum Q-

Functions:

$$\begin{cases} H_1 : \text{Presence of Signal} \quad \text{if} \left(\sum_{n-1}^{N} \left(R_s[n] \right)^2 \right) > \Theta, \\ H_0 : \text{Absence of Signal} \quad \text{if} \left(\sum_{n-1}^{N} \left(R_s[n] \right)^2 \right) < \Theta, \end{cases} \tag{1}$$

$$P_{fa} = P\left\{ \left(\sum_{n-1}^{N} \left(R_s[n] \right)^2 \right) > \Theta \mid H_0 \right\} = \frac{\Gamma(m, \frac{\Theta}{2})}{\Gamma(m)}, \tag{2}$$

$$P_{md} = 1 - Q\left(\left(\frac{\Theta}{\delta^2} - 1 - \gamma \right) \sqrt{\frac{\mho N}{1 + 2\gamma}} \right), \tag{3}$$

$$P_d = e^{-\frac{\Theta}{2}} \sum_{k=0}^{m-2} \frac{1}{k!} \left(\frac{\Theta}{2} \right)^2 + \left(\frac{1 + \bar{\gamma}}{\bar{\gamma}} \right)^{m-1} \times \left(e^{\frac{\Theta}{2(1+\gamma)}} - e^{-\frac{\Theta}{2}} \sum_{k=0}^{m-2} \frac{1}{k!} \left(\frac{\Theta \bar{\gamma}}{2(1 + \bar{\gamma})} \right)^k \right), \tag{4}$$

where R_s is the received signal, Θ is the E-D threshold, δ^2 is AWGN noise variance, $\mho$ is the number of CUs, γ is SNR, and $\sum_{n-1}^{N} \left(R_s[n] \right)^2$ is the output of the integrator, and the upper incomplete gamma function [22] is defined as the integral from $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$. The primary signals are modeled as a two-state Markov chain: H_0 and H_1. From (2) it can be seen that the probability of false-alarm and SNR are independent. Thus, when H_1 is satisfied, it means that an LU is active on the channel. The relation between miss-detection and false-alarm probabilities [23] are stated as:

$$P_{fa} = Q\left(\sqrt{N}\gamma + Q^{-1}(1 - P_{md}) \sqrt{1 + 2\gamma} \right). \tag{5}$$

Based on IEEE 802.22 standard [19], the probabilities of miss-detection and false-alarm are assumed to be lower than 0.1: $0.01 \le P_{fa} \le 0.1$ and $0.9 \le P_d \le 0.99$.

2) Channel Quality Estimation: Data collected via spectrum sensing function are stored in a DB at the CBS side. With the help of the collected data, the quality of the available channels is computed based on sensing accuracy and channel idle duration. Using (2) and (4), the accuracy of spectrum sensing is calculated based on false-alarm probability and detection probability:

$$\omega = P_d \left(1 - P_{fa} \right), \tag{6}$$

where $0.81 \le \omega \le 0.99$. In addition, the discussed spectrum sensing function enables the CU to identify LU traffic patterns. The CU may obtain various traffic patterns over different channels over the time. The objective of traffic pattern identification is to predict traffic rate fluctuations based on the evaluation history and the CU experiences.

In this paper, we consider deterministic mode of traffic patterns, and hence we model the channel usage as an ON-OFF source alternating model, where ON means busy period and OFF means idle period. The model is confirmed as a suitable model because it approximates the channel usage pattern at public safety bands [24]. The ON and OFF periods are independent identically distributed (i.i.d.). The alternating renewal process can be designed in form of a two-state birth-death process.

$$f(T_1) = \begin{cases} \nu e^{-\nu T_1}, & T_1 \ge 0, \\ 0, & T_1 < 0, \end{cases} \tag{7}$$

$$f(T_0) = \begin{cases} \upsilon e^{-\upsilon T_0}, & T_0 \ge 0, \\ 0, & T_0 < 0, \end{cases} \tag{8}$$

where ν and υ are the rate parameter of the exponential distribution [25] for busy and idle periods with the average $\frac{1}{\nu}$ and $\frac{1}{\upsilon}$, respectively. If we assume T_0 and T_1 as the duration of OFF and ON states, and $E[T_0]$ and $E[T_1]$ as the vectors related to the mean quantities of the OFF and ON states, the PDF of busy and idle periods will be $f_{T_1}(x)$ and $f_{T_0}(y)$, respectively, the time duration till the transition of the next state is $\frac{f_{T_1}(x)}{f_{T_0}(y)}$, and hence, the secondary opportunity for the CU can be stated as:

$$CU_{opp} = \frac{E[T_0]}{E[T_0] + E[T_1]} = \frac{1/\upsilon}{1/\upsilon + 1/\nu}. \tag{9}$$

The maximum-likelihood estimation method is used to obtain these mean values. To do that, the model uses the CU experiences and the collected information stored in the DB. The mean duration of ON and OFF periods are updated according to the sensing results gradually. The probabilities that an LU on a specific channel remains idle while the CU transmission process is going on and the LU does not arrive while the CU is in sensing mode are calculated based on the scale, ζ^C, and shape, ζ^H, parameters as follows:

$$P_t^T = \left(\frac{\zeta^C}{\zeta^C + \Gamma_T} \right)^{\zeta^H}, \tag{10}$$

$$P_t^S = \left(\frac{\zeta^C}{\zeta^C + \Gamma_T} \right)^{\zeta^C}. \tag{11}$$

Finally, based on sensing accuracy and mean value of OFF state, the channel quality is estimated as [26]:

$$Q_{CH} = \left(1 + \log_\rho \omega \right) E[T_0], \tag{12}$$

where $\rho > 1$ is to taking the preference of the CU into account that is derived using partial derivative of the sensing accuracy and mean duration of idle state:

$$\frac{\partial^2 Q_{CH}}{\partial \omega \partial \rho} = -\frac{\omega}{E[T_0]\rho} \left(\frac{1}{\ln \rho} \right)^2 < 0, \tag{13}$$

$$\frac{\partial^2 Q_{CH}}{\partial E[T_0] \partial \rho} = -\ln \omega \left(\frac{1}{\ln \rho} \right)^2 > 0, \tag{14}$$

where $\rho \in [1.1, 10.0]$. As in (12) the quality of the channels is computed based on sensing accuracy and idle duration. To satisfy some specific QoS parameters, the CU may tend to capture the channels with higher idle duration rather than higher sensing accuracy. So selection of a bigger ρ means that the idle time duration is prior of the sensing accuracy, and vice versa.

3) Channel Allocation: In a CR-based network, the CUs are facilitated with high bandwidth through heterogeneous wireless architecture and dynamic channel access mechanisms. However, such kind of networks face the challenges because of the fluctuating nature of the available channels and various requirements of different applications. Therefore, some special kind of spectrum management schemes are required to address these challenges. In this section, we explain our solution to tackle dynamic channel access issue in *cognitive 5G cellular networks*, which is abstracted in Algorithm 2.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2016.2629507, IEEE Transactions on Vehicular Technology

Generally, the CBS needs to allocate a new channel to a CU in three cases: when a LU reclaims the channel, when the quality of the channel becomes poor, and when the CU moves spatially out of a cell.

Algorithm 2 (Channel Allocation)

The set of available channels and the set of traffic sources are considered as the input for this algorithm. The output of the algorithm is appropriate channel allocation to the CUs' calls according to the requirements. If there is a call from any CU (line 2), the scheme checks whether there is any available channel or not, if yes (line 3), the call will be accepted, otherwise it will be rejected (line 26). The priority of CUs' traffic source is determined based on a pre-defined pattern (line 5). To improve QoE, channel allocation is based on class priority. So, there is a mapping between channel classes and traffic class. The incoming traffic is mapped to its corresponding class. If the corresponding class is not empty, the CBS selects the best available channel and informs the CU about the channel during link setup (lines 7-9). Then, in order to sense the LU reactively, the CU senses the channel at the start of the first slot, if no active LU detected, the CU begins its transmission over it (lines 10-16). Otherwise, the CU considers the channel as a busy channel for this time slot (lines 17-22), and informs the CBS about the channel state. The CBS decreases the channel ID, and goes to line 7. If we assume N as the number of channel classes in the index and M as the number of channels in each class, then the computational complexity is $O(NM)$

Algorithm 2: Channel Allocation Procedure

Input: Ranking Channel Index and Traffic Classes
Output: Channel allocation to the CU

1 Generate LU arrival by λ
2 **if** *channel requested by any CU* **then**
3   **if** *the index is nonempty* **then**
4   Accept th CU's request
5   determine the CU's QoS traffic class
6   **for** $i = N$ *to* $i = 1$ **do**
7   **if** *the class in nonempty* **then**
8   **for** $j = M$ *to* $j = 1$ **do**
9   select channel y in the class
10   sense the channel at the sensing part of the first slot
11   **if** $(\sum_{n-1}^{N}(R_s^j[n])^2) > \Theta$ **then**
12             $flag_j = 1$; $\lambda++$
13   monitor $E_j[T_1]$
14   update the index
15             $j = j - 1$
16   **Break**
17   **else if** $(\sum_{n-1}^{N}(R_s^j[n])^2) < \Theta$ **then**
18             $flag_j = 0$
19   allocate the channel to the CU
20   estimate $E_j[T_0]$
21   calculate $Q_{CH}^j = (1 + \log_\rho \omega_j) E_j[T_0]$
22   update the index
23   j=j-1
24   i=i-1
25   **else**
26   Reject the CU's request

If an LU reclaims the channel the CU must avoid any harmful interference with the LU. This process, through interference nullification, is done by the listen-before-talk (LBT) policy in which the CU must sense the channel before transmitting any data packet. Hence, it is assumed that the LU arrival is detectable by the CU within an acceptable time duration and collision percentage is negligible.

In any of the mentioned three cases that CUs need to switch to another channel, the CBS looks at the ranking index and selects the best available and the most reliable channel according to the channel quality and the CU's QoE requirement. The CBS informs the CU about the selected channel during the link setup. Then, the CU switches to the allocated channel and resumes its transmission process. With more available channels in the ranking index, the CU may transmit the packets at a higher data transmission rate. The CBS periodically updates the DB of channel usage statistics. Any new discovered idle channel will be added to the index. The process of channel sensing is performed independently on the channels in a mixture of proactive and on-demand modes. Therefore, the CBS is supposed to schedule the concurrent sensing on the channels to avoid duplicate sensing and duplicate data collection on a specific channel.

In the remaining part of this section, we explain the classification of the traffic sources based on their requirements.

In the case of multimedia transmission in *cognitive 5G cellular networks*, it is expected that there will be a lot of different types of bandwidth-hungry and distortion-sensitive traffic sources such as IPTV, video conferencing, video, voice, and image. Each of the traffic sources may require different network QoS parameters. For instance, voice and CBR (constant bit rate) are sensitive to the delay while the others such as data, image, and VBR (variable bit rate) are not. Therefore, for such kind of delay-sensitive multimedia traffic, our framework allocates the channel with smaller handoff probability. Both delay and the quality of the received content at the receiver side are considered as the most important QoE metrics.

Whereas, the users in a CR-based network are not supposed to use the allocated channel permanently, dynamic channel access may seriously degrade service quality and increase transmission failure and delay. Therefore, to tackle this problem, in the proposed scheme we allocate the available channels to the CUs based on their requirements and the quality of the channels. In such a way that the channels with higher sensing accuracy and idle duration are provided to the more delay and failure sensitive traffic sources in order to enhance the QoE.

Traffic mapping to the appropriate channels is done at the CBS for downlink and at the CU for uplink directions. The channels allocated to a CU enable it to transmit over uplink and receive over the downlink [36]. The general scenario is shown in Fig. 3. At the channel scheduler module, the CBS provides an appropriate channel for the traffic coming from a media origin

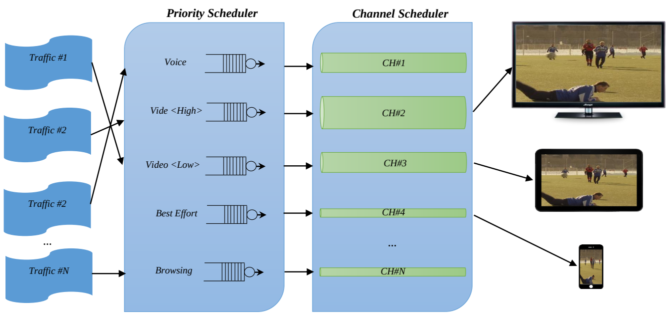


Fig. 3: Traffic and Channel Mapping

server and heading to a CU. And on the other hand, the CU opts a suitable channel based on the traffic requirements. As mentioned already the ranking index is exchanged between the CBS and CU during link setup.

Upon arrival of a channel request from a CU from a specific traffic class, the channel scheduler module checks the type of channel that is required. Then, based on the channel quality that is required by the CU, the scheduler checks the ranking index. If the index is empty, the call will be rejected. Otherwise, it checks the corresponding class in the ranking index. If the corresponding class in non-empty, the call is accepted. Otherwise, it checks the next class with lower quality. A buffer is established for every admitted request.

The scheduler schedules the accepted calls over the allocated channel. To protect the LUs, in the case of capturing licensed bands, upon detection of an LU, the scheduler stops the transmission immediately and allocates another suitable channel from the index if available. If there is no channel in the index the CBS puts the suspended CU into a queue to wait for the next opportunity.

To compare the efficiency of the proposed channel allocation scheme, we study a greedy non-priority channel allocation and a fair proportional scheme as follows.

a) Greedy Non-priority Channel Allocation; In this scheme, the CBS does not classify the incoming traffic based on their QoE requirements and just serves them as First-In, First-Out (FIFO). Upon arrival of a channel request from a CU, the CBS checks whether there is any free channel in the index. If so, the CBS allocate the first available channel to the CU. However, to protect the LUs, the CU senses the channel during the sensing part of the first slot, if there is no active LU on it, the CU starts its transmission. Such kind of channel allocation schemes decreases the channel utilization efficiency.

b) Fair Proportional Channel Allocation; In this scheme, the CBS reserves an equal number of bands and allocates them to different priority classes according to the total number of traffic classes. The channel quality and QoE expectations are not taken into account, which results in wastage of spectrum resources.

B. ***Phase II***: *Channel state estimation*

We use HMM to model the signal characteristics statistically regarded as Markov chains that is observable through the memoryless channels. To form an HMM, we need to incorporate an process that can be observed along with an underlying Markov chain. The observable process is to collect information regarding the underlying Markov chain that is said to be *hidden*. For a underlying Markov chain, the observed process will be independent conditionally. The hidden process is in two modes: discrete or continuous finite-state homogeneous Markov chain. Then, the result of the process that can be observed may have finite-alphabet or general-alphabet. So, it is characterized by a PDF or a PMF appropriately.

The channel state, busy or idle, is hidden due to not being directly observable. Therefore, the results of channel sensing done by the CU are considered as the observation of the channel state. Hence, since the variations of the hidden states are based on LU activities, therefore it is a hidden Markov process while the collected observation information under a certain hidden state is a normal random process.

In this paper, we consider a hidden Markov chain with Guassian densities that is represented by a doubly stochastic process $\{(H_t, O_t)\}$ where $t = 0, 1, \ldots$. The hidden state set is $H_t = \{h_1(t), h_2(t), \ldots, h_N(t)\}$ with state space $h_i \in \{0, 1\}, \forall i \in \{1, \ldots, N\}$ indicating the channel is idle and busy, respectively. $\{H_t\}$ is a discrete-time finite-state homogeneous Markov chain that satisfies $P(H_{t+1} = x \mid H_t = y) = P(H_1 = y \mid H_0 = x) \forall t = 1, 2, \ldots$. The observation state set for M slots is $O_t = \{o_1(t), o_2(t), \ldots, o_M(t)\}$ with state space $o_j \in \{0, 1\}$,$\forall j \in \{1, \ldots, M\}$ indicating idle and busy, respectively. The observation set is the local decisions of the CU regarding M-slot LUs' activities sensing on a specific channel. (H_t, O_t) is the stationary finite state system based on the followings:

- (H_t, O_t) are jointly stationary, $\quad 0 < t \leq N$,
- $P(O_{t+1} = o_{j+1},\ H_{t+1} = h_{i+1} \mid O_1^j = y_1^j,\ H_1^i = h_1^i) = P(O_{t+1} = o_{j+1},\ H_{t+1} = h_{i+1} \mid H_t = h_i)$,

where H_t and O_t are the state and the output of the stationary finite state system, respectively. The random variable H_t is conditionally independent given O_t:

$$p(o_0^K \mid h_0^k) = \prod_{k=0}^{K} p(o_t \mid h_t), \quad (15)$$

where K is a non-negative integer.

At the beginning, the CU can provide the observation set. The observation set is provided to determine the history of the results of channel sensing. We assumed that the channel state are fixed during a time slot (not changed). It means that the status of one slot can be either idle (flag=0) or busy (flag=1). The flag is used to specify the number of OFF and ON slots of a band to conclude the number of LU arrivals.

If we consider $S(t) = \{s_1(t), s_2(t), \ldots, s_N(t)\}$ as the set of channel states, the status of CH_n at time t is $s_n(t)$ based on some corresponding state transitions probabilities, where the state space is $S : \{0, 1\}$. Then, $o_n(t)$ is the corresponding result of spectrum sensing function. We describe our model as a HMM [27] by its parameters $\Lambda = (A, B, \pi)$, where;

- $A = [a_{ij}]_{N \times N}$ $\forall 1 \leq i, j \leq N$ is the state transition matrix that defines transitioning probability from one state to another or to the same state.
- $B = [b_j(k)]_{N \times M}$ is the output symbol probability matrix that computes the probability of providing various output symbols while being in a specific state.
- $\pi = \{P(s_1 = h_i)\}$ is the initial state probability vector.

The parameters i.e., the probabilities of state transition, observation symbol emission, and initial state distributions are calculated as the followings, respectively:

$$a_{ij} = P\Big(h_n(t) = s_j \mid h_n(t-1) = s_i\Big), \sum_{j=1}^{N} a_{ij} = 1, \tag{16}$$

$$b_j(k) = P\Big(o_t = \acute{s_j} \mid s_n(t) = s_j\Big), \qquad \sum_{k=1}^{N} b_j(k) = 1, \tag{17}$$

$$\pi_j = P\Big(s_n(t) = s_j\Big), \qquad \sum_{j=1}^{N} \pi_j = 1, \tag{18}$$

where $0 \leq a_{ij},\ b_j(k) \leq 1;\ i,j \in N,\ k \in M,\ \pi_j \geq 0,\ \pi = \{\pi_1, \pi_2, ..., \pi_N\}$, and $\acute{S} = \{\acute{s_1}, \acute{s_2}, ..., \acute{s_N}\}$ denotes N observation symbols. To estimate the status of the next slot, we need to specify the model parameter of each channel according to the observation set. Therefore, HMM predictor is employed to predict the state of o_{M+1} based on the experienced M observations.

To do that, HMM is supposed to produce the observation sequence having maximum likelihood probability. Therefore, the parameters are adapted by maximizing the probability of $P(O|\Lambda)$. Whereas, we supposed to have a DB to store CU experiences, we exploit the collected information to facilitate HMM. The data collected by the CU are used to compute the probabilities of occupancy of the channels using the maximum likelihood approach. To do that, we first calculate the joint distribution $P(o;s)$, of the sensing data and estimated occupancy. Then, the joint distributions for all possible channels busy sequences are calculated, and finally the distribution that results the maximum probability and related channel busy state sequence is obtained according to the estimation of real channel busy status.

$$\begin{aligned} &P\Big((o_1(t), o_2(t), ..., o_M(t))(s_1(t), s_2(t), ..., s_N(t))\Big) = \\ &\Big[P\Big(O_1 = o_1(t)\Big)P\Big(S_1 = s_1(t)\Big)\Big] \times \Big[P\Big(O_2 = o_2(t)\Big) \\ &P\Big(S_2 = s_2(t)\Big)\Big] \times ... \times \Big[P\Big(O_M = o_M(t)\Big)P\Big(S_N = s_N(t)\Big)\Big]. \end{aligned} \tag{19}$$

First, to calculate the parameters and train the HMM model for the future channel state prediction, the observation sequence is used as the training sequence. The observation set regarding the channel state needs to be specified in order to obtain the history of spectrum sensing results. To do this, we used BWA [28] that is derived form of the Expectation-Maximization (EM) algorithm [29] to estimate the HMM parameters. Using BWA, the HMM model parameters, $\Lambda = (A, B, \pi)$ are defined in the following format:

$$\pi = (\pi_0, \pi_1), \tag{20}$$

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}, \tag{21}$$

$$B = \begin{bmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{bmatrix}. \tag{22}$$

Based on a model Λ and an observation set O, the observation evaluation problem $P(O|\Lambda)$ is resolved by employing forward-backward procedure using forward and backward variables:

$$\alpha_t(i) = P(O_t, s_n(t) = i \mid \Lambda), \tag{23}$$

$$\beta_t(i) = P(\{o_{t+1}, o_{t+2}, ..., o_T\}|s_m = i, \Lambda). \tag{24}$$

The recursive relation for both of them is calculated as:

$$\alpha_{t+1}(i) = b_j(t+1)\sum_{i=1}^{N} \alpha_t(i)a_{ij},\ \forall j \in N; 1 \leq t \leq T-1, \tag{25}$$

$$\beta_t(i) = \sum_{i=1}^{N} \beta_{t+1}(j)a_{ij}b(o_{t+1}),\ \forall j \in N; 1 \leq t \leq T-1, \tag{26}$$

where $\alpha_1(i) = \pi_j b_j(o_1), 1 \leq j \leq N$, and $\beta_T(i) = 1, 1 \leq i \leq N$, Then, the model parameters are estimated as:

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{t=1}^{T-1}\sum_{j=1}^{N} \alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}, \tag{27}$$

$$b_j(k) = \frac{\sum_{t=1,o_t=k}^{T}\sum_{j=1}^{N} \alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{t=1}^{T-1}\sum_{j=1}^{N} \alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}, \tag{28}$$

$$\pi_j = \sum_{j=1}^{N} \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{P(O \mid \Lambda)}. \tag{29}$$

Using (25) and (26), $P(O|\Lambda)$can be obtained as:

$$P(O_t \mid \Lambda) = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i). \tag{30}$$

Then, we calculate the joint probability of sensing results followed by either busy or idle slot, o_{M+1}.

$$P(O_t, 1 \mid \Lambda) = \sum_{i=1}^{N}\left(\sum_{j=1}^{N} \alpha_M(j)a_{ij}\right) \cdot b_i(o_{M+1} = 1), \tag{31}$$

$$P(O_t, 0 \mid \Lambda) = \sum_{i=1}^{N}\left(\sum_{j=1}^{N} \alpha_M(j)a_{ij}\right) \cdot b_i(o_{M+1} = 0). \tag{32}$$

By having the estimated parameters and decoded channel state, we predict spectrum handoff occurrence for the coming time slot $(M+1)$ according to:

$$\begin{cases} P(O_t, 1 \mid \Lambda) \geq P(O_t, 0 \mid \Lambda) \Longrightarrow \text{handoff}, \\ P(O_t, 1 \mid \Lambda) < P(O_t, 0 \mid \Lambda) \Longrightarrow \text{no handoff}. \end{cases} \tag{33}$$

To estimate the status of two conductive slots, suppose that the transition rate of idle to busy is $r_{01} = P(X_{t+1} = 1 \mid X_t = 0)$ and the transition rate of busy to idle is $r_{10} = P(X_{t+1} = 0 \mid X_t = 1)$. Then the probability of idle and busy based on transition rate are as follows, respectively;

$$P(0) = \frac{r_{10}}{r_{01} + r_{10}}, \tag{34}$$

$$P(1) = \frac{r_{01}}{r_{01} + r_{10}}, \tag{35}$$

And, using the Chapman-Kolmogorov equation [23], the probability of two consecutive states based on transition rate is calculated by:

$$P\Big(s_{i_+1} = 0 \mid s_i = 0\Big) = \frac{r_{10}}{r_{01} + r_{10}} + \frac{r_{01}}{r_{01} + r_{10}}e^{-(r_{10}+r_{01})t_{slot}}, \tag{36}$$

$$P\Big(s_{i_+1} = 0 \mid s_i = 1\Big) = \frac{r_{10}}{r_{01} + r_{10}} - \frac{r_{10}}{r_{01} + r_{10}}e^{-(r_{10}+r_{01})t_{slot}}, \tag{37}$$

where t_{slot} is time slot duration. Equation (36) calculates the idle probability of two consecutive idle slots whereas (37) computes the probability that slot n is idle and the next slot $n+1$ is busy. Then using (10), (11), and (23)-(28) the probability of availability of at least one channel for CU is formulated as:

$$P_{avail} = 1- \prod_{m=1}^{M}\left[1-\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\alpha_M(j)a_{ij}\right)\cdot b_j(o_{M+1}=0)\right)\cdot \mathbb{P}\left(P_t^T \geq \varphi\right)\right] \quad (38)$$

where φ is collision threshold. And the failure probability, the probability that the CU tries to capture a channel and fails to transmit;

$$P_{fail} = \prod_{m=1}^{M}\left[1-\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\alpha_M(j)a_{ij}\right)\cdot b_j(o_{M+1}=0)\right)\cdot \mathbb{P}\left(P_t^T \geq \varphi\right)\right]. \quad (39)$$

Using (38) and (39) the probability of successful transmission of CU is equal thereby to:

$$P_{succes} = \left(1-\prod_{m=1}^{M}\left[1-\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\alpha_M(j)a_{ij}\right)\cdot b_j(o_{M+1}=0)\right)\cdot \mathbb{P}\left(P_t^T \geq \varphi\right)\right]\right)\left(\prod_{m=1}^{M}\left[1-\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\alpha_M(j)a_{ij}\right)\cdot b_j(o_{M+1}=0)\right)\cdot \mathbb{P}\left(P_t^T \geq \varphi\right)\right]\right). \quad (40)$$

And finally, by having the initial arrival time of the LU that is sensed by the CU and stored in the database, the maximum number of spectrum opportunities [30] captured by the CU is determined as follows:

$$CU_{OFF} = \left[-\frac{1}{t_{slot}}\left(t_{init} + \frac{\log \varphi}{\lambda}\right)\right], \quad (41)$$

where λ is LU arrival rate and t_{init} is initial arrival time of the LU.

C. *Phase III*: *Handoff interruption management*

In the previous subsection, we discussed handoff prediction scenario in details and we stated that to manage handoff interruptions when a handoff is predicted, the server extracts and sends only the lightweight BL code. In this section, we present our proposed handoff management scheme. And subsequently, we discuss the evaluation metrics for the received video quality assessment.

As we stated already, in cognitive 5G networks, the CUs needs to switch to another available channels when an LU reclaims the channel, and/or when the quality of the current channel become poor, and/or even when the CU moves to another cell. The spectrum mobility results to handoff delay, which makes service interruption in multimedia services. In order to overcome with such kind of challanges, in this section, we present our handoff interruption management scheme.

We use SVC to encode the video sequence in a scalable mode. To do that, we use temporal-SNR scalability mode to encode the video sequence into one BL (L0) and one EL (L1) with the same spatial resolution, but with different frame rate and quality. The quality is determined by quantization parameter (QP), the higher QP the lower PSNR and bitrate [32]. We set higher QP and lower

Algorithm 3: Handoff Interruption Management Procedure

Input: λ, the encoded video
Output: transmitted video
1 Generate LU arrival by λ ;
2 **do**
3   **if** $P(O,0 \mid \Lambda) \leq P(O,1 \mid \Lambda)$ **then**
4     $.\backslash encoded.264\ .\backslash L0.264\ \ -sl\ \ 0$;
5   transmit the extracted BL in advance ;
6   sense LU arrival ;
7   **if** $(\sum_{n-1}^{N}(R_s^i[n])^2) > \Theta$ **then**
8   Vacate CH_i;
9       $HO++$;
10       $flag_i = 1$;
11       $\lambda++$;
12   monitor $E_i[T_1]$;
13   update the index ;
14       $i++$;
15   **Break** ;
16   **else**
17   continue transmitting
18   **else**
19   continue transmitting
20 **while** *there is a video seqment to transmit*;

frame rate to the BL to generate lightweight BL. The reason is to overcome with the interruptions caused by spectrum handoff and thereby provide seamless multimedia service. The EL bitstream is to enhance the quality of the received content. However, decoding of the EL depends on the successful decoding of BL. Moreover, the EL must be transmitted, received, and decoded entirely, otherwise it does not imporve the content quality at all.

Algorithm 3 (Handoff Interruption Management Procedure)

If a handoff is predicted by the channel state estimation procedure discussed in the previous section(line 3), the server extracts BL code (L0.264) from the encoded content (encoded.264) and transmits it during a certain arrival before the handoff starts (line 4-6). The CU continues the current transmission till detection of a LU. If an LU arrival is detected (line 7), the CU vacates the channel. The number of handoffs and LU arrival is incremented by 1, the flag is changed to 1, the index is updated with the new changes, and the CU senses the channel to obtain channel ON duration (lines 8-15). Meanwhile, the client shows the pre-fetched BL code during spectrum mobility (from the time that the CU vacates the channel till the time the CU captured a new channel). Otherwise, if no handoff is predicted, or even a handoff is predicted but a real LU arrival is not detected, the CU continues the transmission over the current under-use channel.). The number of repeating time of this algorithm is based on the volume of the video file that is going to be transmitted. Therefore, we just need to compute the complexity of inner operations, those are constant and the complexity is $O(1)$.

In our scheme, when a handoff is predicted, equal to mean duration of handoff delay, only the lightweight BL code is sent in advance. Then, at the time of LU arrival, while the CU is looking at the ranking index and trying to select the best available channel, the receiver views only the pre-fetched BL code. The handoff interruption management procedure is abstracted in Algorithm 3.

In the proposed network, the CBS is assumed to allocate the channels to the CUs based on both the quality of the channels and the requirements of the traffic sources. In order to overcome with the issue of signaling overhead, in addition to the bitrate required by the traffic source, an approximate amount of bitrate is considered for the extra BL bitrate, which may be required to be sent additionally in advance in order to manage the handoff delay. One of the salient advantages of the proposed framework is that the CBS is aware of both channel quality and CUs' requirements. Thereby, our system enhances both channel utilization and service quality. For the cases that the CBS predicts that maybe the current channel does not have the capacity to carry the extra BL code and there is the probability of overhead, for the current slot also transmits only BL code due to the higher priority of BL code. It means, in case of overhead, only BL code will be transmitted for both the current slot and the channel switching (handoff delay) slots. So, there will not be service interruption at the receiver side.

Although, comparing to the other related proposed schemes, the distinctive superiority of the proposed framework can be seen in this algorithm. Whereas, in multimedia services the degree of delight (seamless service) or annoyance (interrupted service) of the end users is the main efficiency determinative factor, which is ignored by the other proposed spectrum handoff schemes, our framework interestingly is able to improve the end user satisfaction by providing seamless multimedia services.

In the remaining part of this section we explain the evaluation metrics for the reconstructed video quality.

To evaluate the efficiency of video transmission schemes, traditionally the delivered video quality was measured in terms of PSNR or distortion as a metric scale of QoS. For a video sequence, PSNR or distortion is defined as the average of the corresponding assessments over all of the frames. However, visual masking phenomenon is not considered in PSNR evluation mode. Therefore, we consider PSNR as an objective evluation metric and MOS as a subjective evaluation mode. We describe and formulate both PSNR and MOS as follows:

1) PSNR: Modeling of PSNR or distortion is done in different ways such as: modeling of distortion as a continuous function of the content rate [31] and discreet values based on the number of received layers [33], and modeling PSNR as a linear/piece-wise linear non-decreasing utility function [34]. We model PSNR as a linear function of the bit rate:

$$\psi_{org} = \theta \cdot (r_{tot} - r_{L0}) + \psi_{L0} = \theta r_{L1} + \psi_{L0}, \tag{42}$$

where r_{tot} is the total bitrate that is the sum of BL bitrate (r_{L0}) and EL bitrate (r_{L1}), and the $R-D$ model parameter (θ) is chosen according to the spatial-temporal characteristics of the video and codec. If we consider r_{L0} as ratio of packet loss in BL, then PSNR of the reconstructed video at the receive side will be:

$$\psi_{rec} = \theta(r_{L0} - r_{L0} r_{L1}) + \psi_{L0}. \tag{43}$$

TABLE I: Simulation Parameters

Parameter	Value
φ	0.2
x	5
δ^2	-87dB
P_d	[0.9,0.99]
P_{fa}	[0.01,0.1]
r_{L0}	1Mbps
t_{slot}	100ms
A	$\begin{bmatrix} 0.2 & 0.8 \\ 0.3 & 0.7 \end{bmatrix}$

TABLE II: Channel Ranking Parameters

	CH#1	CH#2	CH#3	CH#4	CH#5
P_{fa}	0.091	0.092	0.089	0.091	0.091
P_d	0.921	0.942	0.971	0.991	0.901
ν	0.114	0.378	0.367	0.075	0.170
υ	0.674	0.515	0.515	0.575	0.749
λ	0.632	0.532	0.436	0.321	0.632

2) MOS: In CR-based networks the condition of the channels varies over time. Therefore, CUs may experience packet loss because of the low quality of the captured channels or channel switching as we discussed in the previous sections. Thus, based on frame rate r_f, transmission rate r_t, packet error rate r_e, modulation η, and coding scheme σ, the MOS is calculated as:

$$\Upsilon = \frac{a_1 + a_2 r_f + a_3(\ln r_t)}{1 + a_4 r_e + a_5 (r_e)^2}, \tag{44}$$

where $r_e = \frac{1}{1+e^{\eta(SINR-\sigma)}}$, and the coefficients $a_1 - a_5$ are derived as in [35]. Generally packet delivery fails due to spectrum handoff and poor channel quality.

V. Performance Evaluation

In this section, the effectiveness of our proposed framework is evaluated through a simulation study. According to our objectives, i.e., (1) handoff delay minimization, (2) seamless transmission, and (3) QoE improvement, we assess the effectiveness of our framework in terms of (1) handoff latency, (2) average number of interruptions, and (3) the quality of reconstructed video at the receiver side, in terms of PSNR in dB and MOS.

A. Simulation Setup

Evaluation parameters used to assess the performance of our framework are listed in Table I. Moreover, sensing period, switching delay, and transmission duration supposed to be 10ms, 10ms, and 90ms, respectively. CU arrival is assumed to follow a Poisson arrival mode. We set 10 and 20 HMM states and assumed that the states of channels follow an exponential distribution. The CU estimates the parameters through the discussed techniques in the previous sections and then based on the estimated metrics the quality of the channels is predicted. The parameters are: spectrum sensing accuracy, mean duration of ON and OFF states, and LU arrival rate.

B. Simulation Results and Discussions

We categorize the results are in three parts according to our objectives. In each subsection, we compare our result with the other related schems as well.

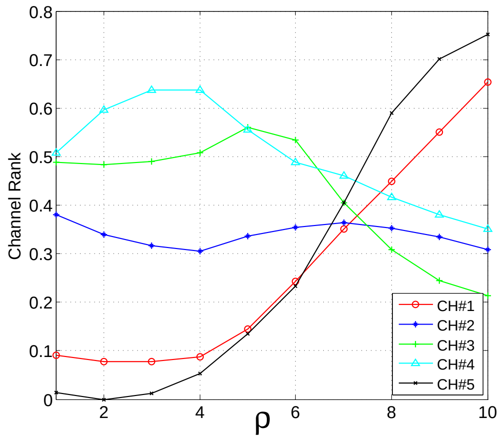


Fig. 4: Channel Ranking based on the Channel Quality Estimation.

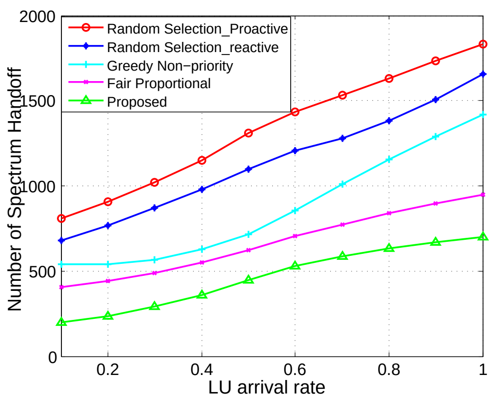


Fig. 5: Number of Handoff vs. LU Arrival Rate.

1) Channel Evaluation: In our framework, we considered channel ranking index to facilitate the best and most available channels to the CU based on its requirements to decrease the handoff delay and the number of interruptions. For the ranking, we consider various random parameters that are listed in Table II for five primary channels as instances. However, we conduct simulation on an example scenario with 20 primary channels. The obtained results can be readily generalized to any number of primary channels. Based on the parameters, the CU estimates the channel quality and ranks in a descending order.

Fig. 4, shows the result of channel ranking. From the figure it is clear to see that the channel rank changes with ρ in such a way that channel sensing accuracy is the main metric with smaller ρ, and with a bigger ρ the channel idle duration is in priority. For example, because $CH\#4$ has the highest sensing accuracy when ρ is low, it is the best available channel. While on increasing ρ, $CH\#5$ demonstrates to have higher idle duration time, the rank of this channel increases as well. This kind of ranking procedure fits well for multimedia services due to the intelligent channel allocation that greatly satisfies the user requirements. Moreover, it reduces the number of spectrum handoff as well, since one of the reason for handoff was poor channel quality. Such kind of channel allocation prevent to allocate poor quality channel to high bandwidth demand applications.

Fig. 5, compares the number of interruptions of our proposed method with random channel selection in two proactive and reactive modes, greedy non-priority, and fair proportional channel allocation algorithms. The proposed method decreases the number of interruptions compared with random selection-proactive method, because in the random selection-proactive method, the CU needs to leave the channel before LU arrival where the arrival of LU is predicted based on channel statistical information. However, due to common channel sensing errors (e.g. false-alarm) in which, the CU predicts a wrong arrival of LUs, the CU vacates the band where there is no claim from any LU. This results into wasting of precious spectrum resources. However, although our method predicts proactively, it performs handoff action reactively, which resolves the issue of false-alarm. In Addition, our proposed framework performs better than the random selection-reactive mode, because it selects already evaluated and stable channels, while random selection blindly selects the available channels. It results into selecting even poor quality channels, and hence the CU needs to perform handoff again when the quality of the selected channel becomes unacceptable.

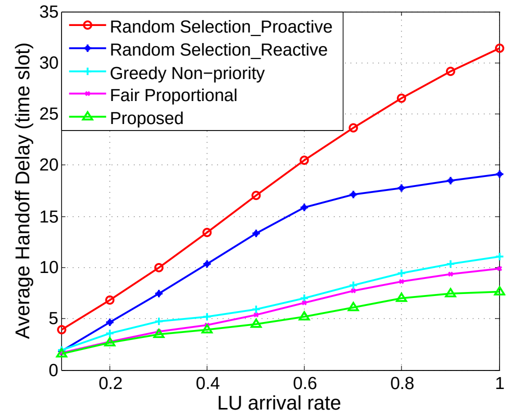


Fig. 6: Average Handoff Delay.

Moreover, comparing with greedy non-priority channel allocation, although this scheme allocate already evaluated channels to the CU, the channel allocation is done in a FIFO mode. Thus, because of non-consideration of the traffic requirements of the traffic, it is possible that after some time the channel capacity does not satisfy the requirements and the CU needs to change the channel. Hence, increasing the number of interruptions. And finaly compared with fair proportional channel allocation scheme, because of static channel reservation in this scheme for each priority class, the quality of channels and CU requirements are not taken into account. However, our scheme performs intelligently. Where with the help of the ranking index and the DB, the best available and the most stable channels are facilitated to the CU based on its requirements and the number of interruptions decreases significantly. As a result that is shown in Fig. 5, the proposed framework, comparing the average number of interruptions for the other schemes, reduces the number of interruptions up to 25%.

Fig. 6, compares the spectrum handoff delay of our framework with the random selection scenario in reactive and proactive methods, greedy non-priority, and fair proportional schemes. As discussed for the last simulation result, the other schemes are involved in a higher number of interruptions and thereby longer handoff delay. Our scheme has the lowest latency due to the assistance of the ranking index. The reason is that the CU does

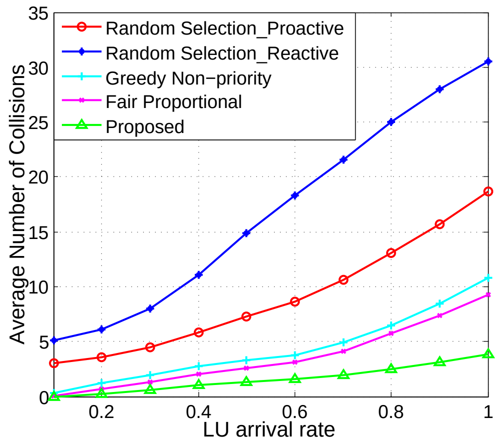


Fig. 7: Average Number of Collisions.

not need to search and compare the quality of available channel at the time of LU arrival. The channel searching and quality evaluation are already done and in the shortest time possible, hence the CBS allocates the best available channel to the CU in minimum duration of time and the CU switches to it in order to continue its transmission.

The average number of collisions for the proposed scheme compared with the other schemes is presented in Fig. 7. Collision may occur in two cases; (1) when the CU cannot detect the arrival of LUs and does not leave the channel upon arrival of LUs (miss-detection), and (2) when a channel is detected as free, but it is occupied by an LU (false alarm).

As shown in the figure, the error probability for random selection-proactive is more than random selection-reactive, because in addition to the two cases, proactive method is involved with the prediction error, as well. It means that if the prediction output is not accurate, the CU does not even try to sense the arrival of the LU. The reactive method is involved with the two cases, but the involvement is less than the proactive method. However, our proposed scheme works better, because it employs two techniques for LU detection, namely prediction and detection. And for the channel selection, it selects only the evaluated channels in the ranking index. In greedy and fair proportional schemes since the handoff occurrence is higher than the proposed scheme, the chance of collision is higher as well.

2) Evaluation of Video Quality: Finally, the effectiveness of our proposed framework in terms of quality of the reconstructed video is evaluated.

We used Soccer standard SVC test video sequences in 704×576 pixels resolution (600 frames). JSVM 9.19.7 reference software was used to encode the video sequences with a GOP length of 30 frames at 30 frames per second and the rate of 200Kbps. The video sequence is encoded into one BL and one EL. We consider a delay deadline for each traffic class based on their requirements. For instance, 0.25, 0.5 second, no delay deadline for conversational video class, streaming video class, and interactive class, respectively. Whereas, the conversational video class, video conferencing for example, has the lowest delay deadline, it has the highest priority. An instance for interactive traffic class is non-real-time data application with no delay restriction.

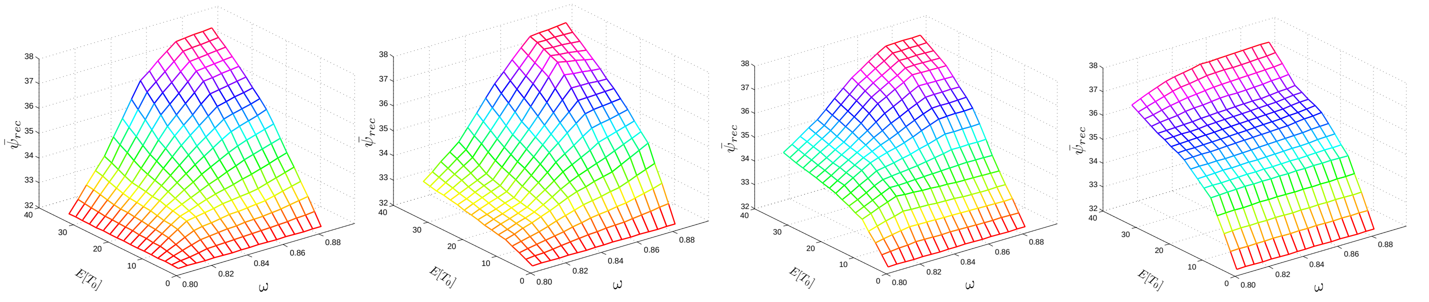
The end user would decode the received ELs if a significant amount of parity packet of the layer were received otherwise it is discarded. Using JSVM reference software, the received data are decoded and PSNR is measured for the reconstructed video to be compared with the original video PSNR. In the normal cases when the CU is transmitting over a captured channel, both BL and EL can be received by the receiver. Thus, the EL is decoded to improve the quality of video. However, when a handoff occurs, normally there should be a service interruption. But in our scheme, before handoff occurrence, the receiver receives BL code in advance for the estimated handoff delay duration. Hence, during handoff delay the receiver shows the pre-fetched BL code. Despite the lower quality, the end user does not realize the handoff interruption.

Fig. 8, shows the impact of the estimated accuracy of channel sensing and the idle duration on the quality of the reconstructed sequence. Clearly, the quality of the reconstructed video is correlated with both of the metrics, i.e., sensing accuracy and idle duration. It can be concluded from the figures that the highest video quality is obtained from the channels that can facilitate more opportunities to the CU, and the CU can detect the channels with less sensing false alarm or miss detection. We adjusted the preference parameter [1.2, 1.8, 2.7, 5]. Increasing ρ means that the idle time duration is preferred by the CU, so sensing accuracy has less impact on the video quality.

Fig. 9, compares MOS vs. LU arrival rate for different traffic class based on the proposed scheme, random channel selection, greedy non-priority channel allocation, and fair proportional schemes. For conciseness, we present the simulation results for three traffic classes i.e., conversational video class, streaming video class, and interactive classes to see the performance of the proposed scheme. And for the other schemes, we just show their performance for streaming class. The proposed handoff interruption management scheme achieves obvious improvement in average MOS compared with the other schemes. The conversational video class has lower MOS compared with the streaming class because of its larger delay deadline, while the interactive class with no delay deadline has minimal changes in its MOS.

The performance of the proposed scheme compared to the other schemes is shown in streaming traffic class. When the rate of LU arrival increases, the number of channel switching increases as well. It happens because of several reasons as explained in the previous sections as well: sensing errors, prediction errors, channel allocation regardless of channel quality and CU requirements. Anyhow, by increasing the channel switching, the probability of packet loss increases as well, which results to deterioration of the end user satisfaction. In this case, the behavior of the proposed handoff management framework is interesting. As shown in Figures 5 and 6, the other schemes are involved with higher number of interruptions and longer handoff delay. In those schemes, when an LU reclaims the channel, the CU has to stop transmission, find a new available channel, and resume its transmission over it. This process causes interruptions, and hence decreases QoE. However, our scheme performs intelligently. By increasing the number of LU arrival, although the number of spectrum handoff increases, at least the end user can see the video sequence (BL only) even with a lower quality. Thereby, the MOS of the proposed scheme is higher than the other schemes.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2016.2629507, IEEE Transactions on Vehicular Technology



(a) $\rho = 1.2$ (b) $\rho = 1.8$ (c) $\rho = 2.7$ (d) $\rho = 5$

Fig. 8: PSNR of the Reconstructed Video based on Channel Quality

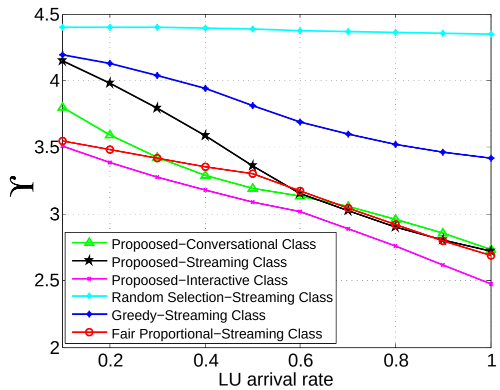


Fig. 9: MOS of Different Traffic Classes vs. the Arrival Rate of LUs.

VI. Application Scenarios and Requirements

The proposed framework is a promising solution to be applied in many distinct application areas that are envisioned for 5G where current wireless networks will struggle to deliver. Specifically, CR will facilitate various versions of radio technology to share the same spectrum band in a more efficient way. The followings are just some instances of the target applications that do not only need higher data rate but need an improved QoE:

- ***Internet of Media Things (IoMT)***, which is the collection of interfaces, protocols and associated media-related information representations that enable advanced services and applications based on human-to-machine and M2M interaction in physical and virtual environments. Media Things refers to the Things with at least one of audio and/or visual sensing and/or actuating capabilities.
- ***Mobile Health Care Applications***, like smartphone-based applications for the monitoring and treatment of long-term medical conditions from mental health problems to diabetes. And even more sophisticated application can be envisioned such as far-distance robotic surgery using haptic technology.
- ***Virtual Reality (VR) and Augmented Reality (AR) applications***, which is for military application as an instance (superimposing a digital view on a physical view). Such kind of applications need large volumes of data for the end-user devices e.g. headsets and displays with very low delay and high replibility.
- And many other applications like smart cities, moving networks, industrial tele-control, e-transportation, etc.

New applications and key design principles of the system lead to many stringent requirements that are needed to meet the coming mobile broadband system. Some of the non-negligible requirements are: (1) The system should be prepared to service a tremendous number mobile users. (2) The radio latency should be lower than one millisecond in order to achieve fast procedure response time and high data rates with low cost. (3) New evolution of battery technology is expected to increase batter life time. (5) Very low cost of devices is needed to be ensured. (6) A peak data rates of 10 Gbps is expected and even more data rate for cell-edges.

VII. Conclusion

In this paper, we addressed the issue of efficient channel allocation based on QoE requirements as well as service interruptions caused by spectrum handoff, specifically for multimedia applications. The proposed framework efficiently evaluates the quality of available channels based on sensing accuracy, idle duration and the arrival rate of licensed users. In our priority-based channel allocation scenario, according to the quality of the available channels, the CBS maintains a channel ranking index to provide the CUs the most reliable channel based on the quality of the available channels and QoE requirements of the various traffic classes. Hence, improving channel utilization and decreasing the number of interruptions. Then, we developed a scheme by employing HMM to estimate the status of the future time slot, handoff occurrence for instance. Handoff prediction significantly reduces interference to the LUs and enhance spectral efficiency as well. After that, by taking into account the unavoidable spectrum handoff and non-negligible delay caused by handoff, we proposed a handoff interruption management scheme. Our framework interestingly is able to provide seamless multimedia service over *cognitive 5G cellular networks*.

Acknowledgment

This research was supported by *Ministry of Science, ICT and Future Planning, Korea, in the ICT R&D Program, 2015* and *Basic Science Research Program through National Research Foundation of Korea (NRF) funded by the Ministry of Education(NRF-2014R1A2A2A01005900)*.

We would like to thank to US NSF CPS-1646607, ECCS-1547201, CCF-1456921, CNS-1443917, ECCS-1405121, and NSFC61428101.

References

[1] L. Lin, "Unlicensed spectrum will improve customer 4G mobile experience," *Vodafone Webpage, http://www.vodafone.com/*, Nov. 2015.

[2] Cisco VNI Forecast, "CISCO visual networking index: global mobile data traffic forecast update 2015-2020," *CISCO VNI 2016*, Feb. 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2016.2629507, IEEE Transactions on Vehicular Technology

[3] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K. K. Wong, R. Schober, and L. Hanzo, "User Association in 5G Networks: A Survey and an Outlook," *IEEE Communications Surveys and Tutorials* , no.99, pp.1-1, 2015

[4] N. Rupasinghe, and I. Guvenc. "Licensed-assisted access for WiFi-LTE co-existence in the unlicensed spectrum." *Globecom Workshops (GC Wkshps)*, Austin, USA, 8-12 Dec. 2014.

[5] D. Marcin and M. Szydelko. "A unified traffic steering framework for LTE radio access network coordination." *IEEE Communications Magazine*, vol. 54, no. 7, pp. 84-92, 2016.

[6] B. Fernando, S. Kumar Ray, and J. A. Gutiérrez. "Understanding the Current Operation and Future Roles of Wireless Networks: Co-Existence, Competition and Co-Operation in the Unlicensed Spectrum Bands." IEEE Journal on Selected Areas in Communications, vol. 34, no. 11, pp. 2829-2837, 2016.

[7] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, Feb. 2014.

[8] X. Hong, J. Wang, C. Wang, and J. Shi, "Cognitive radio in 5G: a perspective on energy-spectral efficiency trade-off," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 46-53, Jul. 2014.

[9] 5G Vision. The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services. available onlin: https://5g-ppp.eu/.

[10] Colin Langtry, "ITU-R activities on 5G", IEEE World Forum on the Internet of Things 14-16 December 2016, Milan, Italy.

[11] W. Jiang and H. Cao, "SIG on Cognitive Radio for 5G", available online at http://cn.committees.comsoc.org/

[12] A. Lertsinsrubtavee, N. Malouch, and S. Fdida, "Spectrum handoff strategy using cumulative probability in cognitive radio networks," *International Congress on Ultra-Modern Telecommunication and Control Systems*, Budapest, Hungary, Oct. 2011.

[13] F. Sheikholeslami, M. Nasiri-Kenari, and F. Ashtiani, "Optimal probabilistic initial and target channel selection for spectrum handoff in cognitive radio networks," *IEEE Transaction on Wireless Communication*, vol. 14, pp. 570-548, Jan. 2015.

[14] L. C. Wang and A. Chen, "On the performance of spectrum handoff for link maintenance in cognitive radio," *IEEE 3^{rd} International Symposium on Wireless Pervasive Computing*, Santorini, Greece, May 2008.

[15] C. W. Wang and L. C. Wang, "Analysis of reactive spectrum handoff in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no.10, pp. 2016-2028, Nov. 2012.

[16] L. C. Wang and C. W. Wang, "Spectrum handoff for cognitive radio networks: reactive-sensing or proactive-sensing," 27^{th} *IEEE International Performance Computing and Communication Conference*, Austin, USA, Dec. 2008.

[17] S. Yoon and E. Ekici, "Voluntary spectrum handoff: a novel approach to spectrum management in CRNs," *IEEE International Conference on Communication*, Cape Town, South Africa, May 2010.

[18] W. Zhang and C. K. Yeo, "Sequential sensing based spectrum handoff in cognitive radio networks with multiple users," *Computer Networks, Elsevier*, vol. 58, pp. 87-98, Jan. 2014.

[19] IEEE Standard 802.22, "Wireless regional area network (WRAN) specific requirements part 22: cognitive wireless RAN medium access control (MAC) and physical layer (PHY) specification: policies and procedures for operation in the TV bands," Jul. 2011.

[20] Md. Jalil Piran, M. E. Ahmed, A. Ali, J. B. Song, and D. Y Suh, "Channel Allocation Based on Content Characteristics for Video Transmission in Time-Domain-Based Multichannel Cognitive Radio Networks." Mobile Information Systems 2015 (2015).

[21] M. Jin, Q. Guo, J. Xi, Y. Li, Y. Yu, and D. Huang, "Spectrum sensing using weighted covariance matrix in Rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5137-5148, Nov. 2015.

[22] S. Atapattu, C. Tellambura, and H. Jiang. "Relay based cooperative spectrum sensing in cognitive radio networks," *IEEE Global Telecommunications Conference*, Hawaii, USA, Dec. 2009.

[23] D. J. Lee and W. Y. Yeo, "Channel availability analysis of spectrum handoff in cognitive radio networks," *IEEE Communications Letters*, vol. 19, no. 3, pp. 435-438, Jan. 2015.

[24] B. Vujicic, N. Cackov, S. VujicCic, and L. Trajkovic, "Modeling and characterization of traffic in public safety wireless networks," *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, Philadelphia, USA, Jul. 2005.

[25] G. Yuan, R. C. Grammenos, and Y. Yang, "Performance analysis of selective opportunistic spectrum access with traffic prediction," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4. pp. 1949-1959, Dec. 2009.

[26] X. Xing, T. Jing, Y. Huo, H. Li, and X. Cheng, "Channel quality prediction based on Bayesian inference in cognitive radio networks," *IEEE INFOCOM*, Turin, Italy, Apr. 2013.

[27] S. R. Eddy, "Hidden Markov models." *Current opinion in structural biology*, vol. 6, no.3, pp. 361-365, Jun. 1996.

[28] L. R. Welch, "Hidden Markov models and the Baum-Welch algorithm," *IEEE Information Theory Society Newsletter*, vol. 53, no. 4, pp. 10-13, Dec. 2003.

[29] A. P. Dempster, N. M. Laid, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," Royal Statistical Society, vol. 39, no. 1, Apr. 1977.

[30] J. Lee and H. Park, "Channel prediction-based channel allocation scheme for multichannel cognitive radio networks," *Journal of Communications and Networks*, vol. 16, no. 2, pp. 209-216, Apr. 2014.

[31] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H. 264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103-1120, Sep. 2007.

[32] S. Ullah, K. Thar, and C. S. Hong. "Management of scalable video streaming in information centric networking," Multimedia Tools and Applications, pp. 1-28, 2016.

[33] T. Schierl, S. Johansen, A. Perkis, and T. Wiegand, "Rateless scalable video coding for overlay multisource streaming in MANETs," *Journal of Visual Communication and Image Representation*, vol. 19, no. 8, pp. 500–507, Dec. 2008.

[34] M. Schaar, S. Krishnamachari, S. Choi, and X. Xu, "Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs," *IEEE Journal on Selected Areas in Communication*, vol. 21, no. 10, pp. 1752-1763, Dec. 2003.

[35] A. Khan, L. Sun, E. Jammeh, and E. Ifeachor, "Quality of experience-driven adaptation scheme for video applications over wireless networks," *IET Communications*, vol. 4, pp. 1337–1347, Jul. 2010.

[36] L. Chen, W. Chen, B. Wang, X. Zhang, H. Chen, and D. Yang, "System-Level Simulation Methodology and Platform for Mobile Cellular Systems," *IEEE Communications Magazine*, vol. 49, no. 7, pp. 148-155, Jul. 2011.

[37] C. Wang, H.Chen, Q. Yin, A. Feng, and A. F. Molisch,"Multi-user Two-Way Relay Networks with Distributed Beamforming," *IEEE Transactions on Wireless Communications*, vol. 10. no. 10, pp. 3460-3471, Aug. 2011.

[38] Y. Li, Q. Yin, L. Sun, H. Chen, and H. Wang, "A Channel Quality Metric in Opportunistic Selection with Outdated CSI over Nakagami-m Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 3, pp. 1427-1432, Jan. 2012.

Md. Jalil Piran (S'10–M'16) received his MSc degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, India, and the Ph.D. degree in Electronics and Radio Engineering from Kyung Hee University, South Korea, in 2011 and 2016, respectively. Currently he is a Postdoctoral Researcher with the Department of Computer Engineering, Kyung Hee University. Dr. Jalil Piran was Research Assistant with Media Lab. at Kyung Hee University as well as Communication Research Center at International Institute of Information Technology (IIIT) Hyderabad, India. His research interests include: Mobile and Wireless Communication, Cognitive Radio Networks, and Multimedia communication. He is an IEEE Member since 2011 and has been working as a delegate from South Korea for ISO/IEC MPEG Forum since 2013.

Nguyen H. Tran (S'10-M'11) received the BS degree from Hochiminh City University of Technology and Ph.D degree from Kyung Hee University, in electrical and computer engineering, in 2005 and 2011, respectively. Since 2012, he has been an Assistant Professor with Department of Computer Science and Engineering, Kyung Hee University. His research interest is to applying analytic techniques of optimization, game theory, and stochastic modelling to cutting-edge applications such as cloud and mobile-edge computing, datacenters, heterogeneous wireless networks, and big data for networks. He received the best KHU thesis award in engineering in 2011 and best paper award at IEEE ICC 2016. He is the Editor of IEEE Transactions on Green Communications and Networking.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2016.2629507, IEEE Transactions on Vehicular Technology

Doug Young Suh (S'89-M'90) received the BSc degree in Department of Nuclear Engineering from Seoul University, South Korea in 1980, MSc and Ph.D. degrees from Department of Electrical Engineering in Georgia Institute of Technology, Atlanta, Georgia, USA, in 1986 and 1990 respectively. In September 1990, he joined Korea Academy of Industry and Technology and conducted research on HDTV until 1992. Since February 1992, he is a professor in College of Electronics and Information Engineering in Kyung Hee University, South Korea. His research interests include networked video and video compression. He has been working as a Korean delegate for ISO/IEC MPEG since 1996.

Ju Bin Song Ju Bin Song (S'99-M'02) received the B.Sc. and M.Sc. degrees from the Department of Electronics, Kyung Hee University, Seoul, Korea, in 1987vand 1989, respectively, and the Ph.D. degree from the Department of Electronic and Electrical Engineering, University College London (UCL), London, U.K., in 2001. From 1992 to 1997, he was a Senior Researcher at the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. He was a Research Fellow with the Department of Electronic and Electrical Engineering, UCL, in 2001. From 2002 to 2003, he was an Assistant Professor with the School of Information and Computer Engineering, Hanbat National University, Daejeon. He has been a Professor and the Department Head with the Department of Electronics and Radio Engineering, Kyung Hee University, since 2003. He has been the Vice Dean with the School of Electronics and Information, Kyung Hee University, since 2015. From 2009 to 2010, he was a Visiting Professor with the Department of Electrical Engineering and Computer Science, University of Houston, Houston, TX, USA. His current research interests include resource allocation in communication systems and networks, cooperative communications, game theory, optimization, cognitive radio networks, and smart grid. He has served as a technical program member for international conferences and as the workshop chair for 7th EAI International Conference on Game Theory for Networks, 2017. Dr. Song was the recipient of the Kyung Hee University Best Teaching Award in 2004 and 2012.

Choong Seon Hong (S'95–M'97–SM'11) received the B.S. and M.S. degrees in electronic engineering from Kyung Hee University, Seoul, South Korea, in 1983 and 1985, respectively, and the Ph.D. degree from Keio University, Minato, Japan, in 1997. In 1988, he joined KT, where he worked on broadband networks as a Member of Technical Staff. In September 1993, he joined Keio University. He worked for the Telecommunications Network Laboratory, KT, as a Senior Member of Technical Staff and the Director of the Networking Research Team until August 1999. Since September 1999, he has been a Professor with the Department of Computer Science and Engineering, Kyung Hee University. His research interests include future Internet, ad hoc networks, network management, and network security. He is a member of ACM, IEICE, IPSJ, KIISE, KICS, KIPS, and OSIA. He has served as the General Chair, a TPC Chair/Member, or an Organizing Committee Member for international conferences such as NOMS, IM, APNOMS, E2EMON, CCNC, ADSN, ICPP, DIM, WISA, BcN, TINA, SAINT, and ICOIN. In addition, he is currently an Associate Editor of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, International Journal of Network Management, and Journal of Communications and Networks and an Associate Technical Editor of the IEEE Communications Magazine.

Zhu Han (S'01–M'04-SM'09-F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Currently, Dr. Han is an IEEE Communications Society Distinguished Lecturer.