Dynamic Spectrum Reservation for CR Networks in the Presence of Channel Failures: Channel Allocation and Reliability Analysis

Indika A. M. Balapuwaduge, *Member, IEEE*, Frank Y. Li, *Senior Member, IEEE*, and Vicent Pla

***Abstract*—Providing channel access opportunities for new service requests and guaranteeing continuous connections for ongoing flows until service completion are two challenges for service provisioning in wireless networks. Channel failures, which are typically caused by hardware and software failures or/and by intrinsic instability in radio transmissions, can easily result in network performance degradation. In cognitive radio networks (CRNs), secondary transmissions are inherently vulnerable to connection breaks due to licensed users' arrivals as well as channel failures. To explore the advantages of channel reservation on performance improvement in error-prone channels, we propose and analyze a dynamic channel reservation (DCR) algorithm and a dynamic spectrum access (DSA) scheme with three access privilege variations. The key idea of the DCR algorithm is to reserve a dynamically adjustable number of channels for the interrupted services to maintain service retainability for ongoing users or to enhance channel availability for new users. Furthermore, the DCR algorithm is embedded in the DSA scheme enabling spectrum access of primary and secondary users with different access privileges based on access flexibility for licensed shared access. The performance of such a CRN in the presence of homogeneous and heterogeneous channel failures is investigated considering different channel failure and repair rates.**

***Index Terms*—Cognitive radio networks, dynamic channel reservation, licensed shared access, retainability, CTMC**

I. Introduction

THE current global research efforts on 5th generation (5G) mobile communications have identified the need for large extent improvements of accessibility and reliability of communication services. With limited bandwidth and static spectrum allocations, current cellular networks cannot achieve this target unless more flexible and dynamic spectrum access is enabled. Cognitive radio (CR) allows future wireless networks to dynamically access the licensed spectrum without causing interference to incumbent users [2], [3] and it is considered as a key component in the 5G paradigm which can tackle challenges such as ultra reliable communication (URC). In the context of 5G, URC refers to the provision of certain level of communication services with high degrees of availability and reliability [4]. Therefore performance measures related to dependability attributes such as reliability and availability of CR networks (CRNs) are of major importance to its successful operation in future 5G wireless networks.

When compared with other wireless networks, CRNs are more prone to channel access failures from the secondary network's point of view. It is therefore less predictable with respect to quality of service (QoS) experienced by users. This feature poses challenges to the provisioning of a dependable service that meets the requirements of CR users on reliability and availability. Traditionally, service level performance evaluation of CRNs has been addressed merely by considering resource insufficiency, where the impact of wireless channel failures and recovery has been largely overlooked partially due to the complexity of analysis. In wireless networks, including CRNs, a connection may fail owing to two fundamental reasons. That is, 1) hardware failures or software malfunctioning; and 2) intrinsic features in radio transmissions, such as channel fading and shadowing [5]. Due to those failures, the network capacity or the number of active subscribers that the network can support may decrease and the response time for a service request may degrade. Therefore traditional performance models that ignore failures and recovery of channels generally overestimate network capacity and other performance measures.

To perform a realistic reliability analysis, it is required to consider performance changes that are associated with channel failures. Consequently, the goal of this paper is to study the performance of dynamic spectrum access (DSA) together with dynamic channel reservation (DCR) in CRNs with a *focus on the impact of channel failures and their recovery* based on the concept of licensed shared access (LSA). In LSA [6], the incumbent users and LSA users are authorized to share the spectrum (or part of the spectrum) in accordance with a set of sharing rules agreed in their spectrum access agreements. For instance, when an incumbent user makes a portion of its spectrum available to an LSA user for a given period of time, the incumbent user cannot interfere with the LSA user during that period [6]. Therefore LSA relaxes the concept of opportunistic access which does not provide any protection for secondary network (SN) services against primary network (PN) occupancy.

However, LSA is not the only spectrum sharing model that has attracted attention in recent years although it is gaining its momentum [7]. Dynamic spectrum leasing (DSL) has also

Manuscript received March 21, 2016; revised November 11, 2016, April 6, 2017 and August 8, 2017; accepted November 3, 2017. Date of publication xx xx 2017; date of current version xx xx 2017. This paper was presented in part at the IEEE ICC, May 2016 [1]. The associate editor coordinating the review of this paper and approving it for publication was J.-M. Park. *(Corresponding author: Frank Y. Li.)*

Indika A. M. Balapuwaduge and Frank Y. Li are with the Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway (email: {indika.balapuwaduge; frank.li}@uia.no).

Vicent Pla is with ITACA, Universitat Politècnica de València (UPV), 46022 València, Spain (email: vpla@upv.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier

been in the spotlight as one of the spectrum sharing models. LSA aims at establishing a consensual sharing framework and the sharing rules are set and applied throughout the whole duration of the LSA agreement. In contrast, the negotiation and the transfer of exclusive usage rights in DSL are done at each leasing period. On the other hand, LSA is much more compatible with existing spectrum allocations so that a legacy network can continue to operate with minor effects from the adopted LSA schemes [8]. Following the principles of LSA, the proposed channel reservation algorithm can be applied over a licensed spectrum where incumbent users are willing to share a part of their spectrum. In this way, auctions and bidding are not necessary in the process of channel allocation.

In a legacy CRN, a primary user (PU) always has absolute preemptive priority for channel access while a secondary user (SU) can only opportunistically access the idle channels. When there are no idle channels at a given instant, a new secondary request will simply be blocked. An ongoing SU connection could also be dropped upon PU arrivals or channel failures. Typically, forced termination of PU sessions has not been considered in CRNs since ongoing PU sessions cannot be preempted by new user arrivals. Once channel failures occur, however, an ongoing PU session can also be terminated before service completion [9]. Thus, *channel availability* and *service retainability* which describe the properties of a system from a *dependability perspective* are two important metrics for performance evaluation in a CRN [10]. Note that the *dependability* of a communication system is defined as its ability to deliver services that can justifiably be trusted, when faced with failures of their components [11], [12]. One reason for which a system does not behave as it is specified can be, for instance, any fault in its design, or the failure of some of its components when facing with unpredicted changes in the system's environment. The theory to analyze this type of phenomenon is known as *dependability theory* [13] in the research community. Therein, retainability refers to the ability of finishing a service completely without being terminated before its actual end [14] and it is considered as one of the key performance indicators, for instance in evolved UMTS terrestrial radio access (E-UTRAN) [15].

A. Related Work

The earlier research work on CR has been largely focusing on aspects such as spectrum sensing and dynamic spectrum access [17]. Few studies which take into account channel failure when performing performance modeling exist. A majority of them considered system centric performance analysis rather than an analysis from a dependability perspective. By considering a topology in a CRN, the formation of *blackholes*, i.e., explosive spreading of random failures was investigated in [18]. Even though correlations among failures and cascading effects are studied, the results presented in [18] are limited to the analysis on failure statistics and network percolation. Although heterogeneous failures of network functions were studied in [19] when analyzing service resilience in 5G networks, it targeted only at cloud-based mobile networks.

A cognitive relaying scheme was proposed in [20] to enhance the PU and SU performance in the event of transmission failures. However, the failures of links are assumed to be with a constant probability. Although the secondary throughput maximization was a main target in [20], the forced termination of ongoing connections was not investigated therein. Instead, forced termination and blocking probability analyses were conducted in [21] considering both resource insufficiency and link unreliability. However, the traffic model considered in [21] is a multi-cellular system and the analysis does not directly apply to CRNs with reserved channels. A model for file transfer over an unreliable channel was considered in [22] by taking into account of interruptions during file transfers due to server failures. However, the analysis in [22] was performed for file fragmentation and no analytical or simulation results were presented.

In addition, a few recent studies considered also channel reservation mechanisms which are applicable to CRNs. Channel reservation policies have been proposed for improving spectrum utilization in CRNs by considering both PUs and SUs. In [23], a PU based channel reservation policy is applied to a CRN where a suitable number of channels are initially reserved for PUs. As long as the CRN has available reserved channels, the PUs cannot occupy the unreserved channels. However, the probability of forced SU terminations may not decrease substantially since ongoing SU services in the unreserved band can still be terminated if the whole reserved band is occupied. Another DCR scheme was proposed in [24] to reduce forced termination probability while minimizing the increase of blocking probability. Since the reserved channels can also be occupied, forced terminations of SU services in the non-reserved band may not be decreased adequately [24]. As a solution to this problem, we proposed a static channel reservation (SCR) scheme in [1] which solely targeted on retainability. However, static reservation generally leads to lower channel availability for CRs.

This paper is motivated by the aforementioned interesting previous studies. However, the work presented in this paper is distinct from those related studies with respect to mainly *four aspects*. First, the proposed channel reservation scheme *dynamically adjusts* the number of reserved channels and it targets at minimizing forced terminations for both ongoing PU and SU services. Second, both PUs and SUs can access the reserved band in the proposed scheme. A common assumption in many previous studies on DCR in CRNs is that the reserved spectrum is occupied exclusively by the SUs or the PUs. This type of reservation would lead to considerably reduced spectrum utilization in the network. In contrast, our approach provides higher flexibility for channel access upon a sudden increase of spectrum demand of the PN or the SN. Third, random failures of channels are also considered to obtain more realistic results of performance metrics. We study also the heterogeneity of failures which appears more often in real-life wireless networks. Fourth, the performance of those schemes is investigated from the perspective of dependability theory. This is because the terminology related to QoS of telecommunication services has been defined in the *ITU-T E.800 Recommendation* [25] based on the dependability theory. Correspondingly, the definitions of dependability metrics used in this study conform with the ITU recommendations.

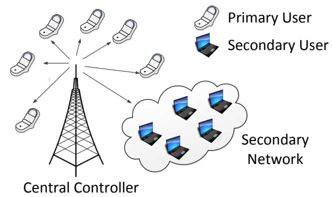

Fig. 1. A CRN with a centralized architecture enabling channel reservation.

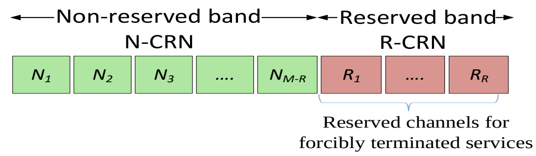

Fig. 2. Reserved (R-CRN) and non-reserved (N-CRN) channel assignment, where N_i, $i \in \{1, 2, \cdots, M - R\}$ represents the channels in the N-CRN and R_j, $j \in \{1, 2, \cdots, R\}$ represents the channels in the R-CRN.

B. PU/SU Spectrum Sharing and Access Privilege Revisited

Traditionally, the spectrum access privilege in an *overlay* CRN has been given exclusively to PUs. With this concept, SUs can only access the idle channels in an opportunistic manner by applying various techniques such as precise spectrum sensing and dynamic channel access. Recently, a novel concept which allows spectrum sharing of PUs and SUs has emerged for the purpose of better QoS provisioning to SUs [26]. For instance, the EU radio spectrum policy group (RSPG) has defined the concept of LSA which allows incumbent license users share their spectrum via supplementary, dynamically, or over long-term spectrum sharing solutions with SUs in order to ensure predictable QoS, as described in [6]. In the US, the FCC is also promoting the *citizens broadband radio service* which allows shared spectrum access among PUs and SUs at the 3.5 GHz radio band through priority access licenses (PAL) and general authorized access (GAA) while still protecting authorized access of PUs [27], [28]. To enable PU and SU spectrum sharing, various techniques including spectrum leasing [29] and channel reservation [30] can be employed. Based on these observations, we design an LSA-oriented DSA scheme with channel reservation which guarantees PU's access privilege in the non-reserved band while protecting ongoing services in the reserved band.

C. Contributions

In brief, the contributions of this paper can be summarized as follows.

- A DSA scheme which enables dynamic channel reservations and spectrum access in a multi-channel CRN is proposed based on the concept of LSA. Three variations of this scheme are studied with different PU access priority privileges in the reserved channels.
- A DCR algorithm which can dynamically adjust the number of reserved channels is proposed and it is incorporated with the DSA scheme. Two working modes are proposed in the DCR algorithm targeting at either maintaining service retainability or enhancing channel availability respectively.
- The combined effect of random channel failures and resource insufficiency is investigated while considering both homogeneous and heterogeneous channel failures when evaluating the performance of the proposed scheme.
- A continuous time Markov chain (CTMC) model is developed to evaluate the system performance of *both primary and secondary networks*. Three dependability metrics including channel availability, service retainability and network unserviceable probability are derived based on the CTMC model. Furthermore, simulations are performed to validate the correctness and the preciseness of the derived analytical models.
- Based on different performance requirements and conditions, another algorithm that can identify the optimal upper bound for the number of reserved channels in a CRN is proposed.

The remainder of this paper is organized as follows. An overview of the network scenario and the modeling assumptions are presented in Section II. In Section III, we introduce an algorithm which performs DCR and incorporate them with a DSA scheme for a CRN with error-prone channels. In Section IV, CTMC models are developed and the expressions for performance metrics are deduced. The performance of the PN and the SN is investigated under homogeneous and heterogeneous channel failures in Sections V and VI respectively. An algorithm to find the optimal number of reserved channels is proposed in Section VII. In Section VIII, two additional levels of PN access privilege (full and partial priority) are studied. Finally, the paper is concluded in Section IX.

II. Network Scenario and Assumptions

In this study, an infrastructure-based CRN architecture that has a centralized network entity, such as a base station as illustrated in Fig. 1, is considered. It comprises two types of networks: the PN and the SN with multiple PUs and SUs. The central controller exerts control over all CR users within its transmission range. The CR users perform their observations and analysis on radio environments and feed these facts back to the central controller. It is the central controller which makes decision on spectrum availability and spectrum allocation. Moreover, the CRN may apply a common control channel for channel reservation and allocation. For analysis simplicity, the protocol overhead is not included in our models.

Consider a PN operating on $M \in \mathbb{Z}^+$ equal capacity channels, where $\mathbb{Z}^+$ is a set of positive integers. The SN opportunistically accesses and reuses the licensed spectrum allocated to the PN. As already mentioned, channels are prone to different kinds of failures which could interrupt ongoing communication sessions. In order to avoid forced terminations, a certain, typically small, number of channels are reserved. *Those reserved channels can only be accessed by the SU and PU services that are interrupted due to channel failures or the preempted SU services upon new PU arrivals*. As depicted in Fig. 2, $R \in \mathbb{Z}^+$ out of M channels are reserved. The set of reserved channels is denoted as R-CRN whereas the set of non-reserved channels is denoted as N-CRN.

The number of channels that can be reserved to the R-CRN, i.e., R, is dynamically adjustable depending on ongoing channel occupancy status. However, to enforce a certain degree of fairness between the newly arrived users and ongoing users, we define an upper bound R_{max} on the total number of channels that can be reserved, i.e., $R \leq R_{max}$, where $R_{max} \in \mathbb{Z}^+$. It is worth mentioning that the set of channels allocated to the N-CRN or R-CRN does not necessarily have to be contiguous. For the ease of presentation, R-CRN channels are shown as a contiguous band in Fig. 2. In the next section, we discuss the methodology adopted for adjusting R under different conditions of the network. In our analysis, the following assumptions are made as the basis for developing the analytical model presented in Section IV.

- The arrivals of services are Poisson processes with rates λ_P and λ_S for PU and SU services respectively. Moreover, the service times for PU and SU services are exponentially distributed, with corresponding service rates per channel μ_P and μ_S respectively.
- The on-time duration (during which the channel is working properly until a failure occurs) of a channel is exponentially distributed with a failure rate per channel λ_F in both N-CRN and R-CRN. Moreover, we assume that both occupied and idle channels are prone to failures.
- Targeted at performing a steady-state analysis, the repair time of a failed channel in both N-CRN and R-CRN is considered to be exponentially distributed [31], with a repair rate per channel μ_R. Moreover, multiple failed channels can be repaired simultaneously.
- The sensing and spectrum handoff (if necessary) latency is negligible in comparison with the time between two consecutive service events.

III. The Proposed Dynamic Spectrum Access Scheme and DCR Algorithm

The proposed DSA scheme includes a channel reservation and adjustment procedure and a channel allocation procedure upon six distinct events: PU/SU arrivals, PU/SU departures, channel failures and channel recovery. For channel reservation, a DCR algorithm which determines the required number of reserved channels in the R-CRN is proposed. The proposed DCR algorithm works with two working modes. One of these working modes needs to be selected in order to adjust the number of reserved channels prior to channel allocation to the new users. Herein, we explain the DCR algorithm before presenting the proposed channel allocation scheme.

A. Dynamic Channel Reservation Algorithm

The proposed algorithm is designed with two different working modes considering distinct reliability aspects: service retainability and channel availability. The ongoing traffic load is taken as the basis for determining R in Algorithm 1. In this study, the ongoing traffic load, Φ, is calculated as the ratio between the number of ongoing PU and SU services in the *whole CRN* and the number of operational channels (see Line 1 of Algorithm 1). Note that an *operational channel* indicates a channel that is not in a failed state and each operational channel can be either occupied or idle.

The design principle for resource reservation in the R-CRN with respect to the two working modes in the proposed algorithm is explained as follows.

Algorithm 1: Dynamic channel reservation (DCR) algorithm with two working modes.

Input: M : Total number of channels in the whole CRN
Input: i_n : Total number of channels occupied by PUs in the N-CRN
Input: j_n : Total number of channels occupied by SUs in the N-CRN
Input: i_r : Total number of channels occupied by PUs in the R-CRN
Input: j_r : Total number of channels occupied by SUs in the R-CRN
Input: f : Total number of failed channels in the whole CRN
Input: R_{max} : The upper bound of the number of reserved channels
Input: a_i : $a_{k+1} < a_k < \ldots < a_i < \ldots < a_1 < a_0$; $0 < a_i < 1$, $i = 1, \ldots, k; k \in \mathbb{Z}^+; a_0 = 1, a_{k+1} = 0$
Output: R : The allowable number of reserved channels that will be allocated to the R-CRN

[1] Calculate $\Phi = (i_n + j_n + i_r + j_r)/(M - f)$
[2] Calculate $N_{Avail} = M - (i_n + j_n + i_r + j_r + f)$
[3] Calculate $N_R = i_r + j_r$
[4] **for** $i = 0 : 1 : k$ **do**
[5]   **if** $(a_{i+1} \leq \Phi < a_i)$ **then**
[6]   traffic_load_level$= i$
[7]   **break**;
[8]   **end**
[9] **end**
[10] Select a working mode: *Working mode* $= 1$ intends to enhance the retainability of ongoing services while *Working mode* $= 2$ increases channel availability for new users.
[11] **if** *Working mode* $= 1$ **then**
[12] $R' = R_{max} -$traffic_load_level
[13] **end**
[14] **if** *Working mode* $= 2$ **then**
[15] $R' = R_{max} - (k -$traffic_load_level)
[16] **end**
[17] $R' = \max\{R', N_R\}$: $R' \in \mathbb{Z}^+$ *is a local variable which is created to store the selected value of R where* $0 \leq R' \leq R_{max}$
[18] $R = \min\{N_{Avail} + N_R, R'\}$

1) *Working mode 1* intends to maintain and enhance the *retainability* of ongoing services in the CRN subject to the channel reservation constraint R_{max} and the ongoing traffic load. To improve retainability, *working mode 1* allocates a higher number of channels to the R-CRN when the ongoing traffic load becomes heavier. This is reasonable from the retainability perspective since forced terminations are more likely to occur at a higher traffic load.
2) Conversely, resource reservation in the R-CRN in *working mode 2* is performed for the purpose of reducing blocking probability of new users. In other words, the improvement of channel availability for new users is of most interest in this case. Accordingly when the ongoing traffic load is higher, *the network reserves fewer channels to the R-CRN in working mode 2*.

Hereafter, we refer to *working mode 1* and *working mode 2* in short as *mode 1* and *mode 2* respectively. In both working modes, the ongoing traffic load is characterized by $k+1$ levels, where $k \in \mathbb{Z}^+$. To distinguish these levels, k configurable parameters, $a_1, a_2, \ldots, a_k$ are required as input parameters to Algorithm 1 such that $0 < a_k < a_{k-1} < \cdots < a_2 < a_1 < 1$. For instance, consider a scenario with $k = 2$. The ongoing traffic load is characterized by three levels as high, medium and low traffic levels. Accordingly, we may configure $a_1 = 0.70$ and $a_2 = 0.35$. Therefore, the high, medium and low traffic loads indicate that $\Phi \geq 0.7$, $0.35 \leq \Phi < 0.7$ and $\Phi < 0.35$ respectively.

The basic channel reservation procedure is shown in Algorithm 1. Recall that Algorithm 1 is run prior to channel allocation upon PU and SU service arrivals. As the initial step, the ongoing traffic load, Φ, is calculated on Line 1 and then a corresponding traffic load level is determined. If Φ is higher, *mode 1* tends to allocate a larger value of R, i.e., a value closer to R_{max}. However, *once the traffic load decreases, the parameter, traffic_load_level increases; thus,* mode 1 *assigns a smaller value of* R as indicated at Line 12. By allocating narrower spectrum to the R-CRN at low traffic loads, it is able to avoid spectrum under-utilization of the reserved band. Otherwise, the allocation of additional channels to the R-CRN will not really help much to decrease the rate of forced terminations at low traffic loads.

Unlike in *mode 1*, *mode 2* assigns a lower value of R if the traffic load increases, as indicated in Line 15. This is for the purpose of increasing channel access opportunity for new users. On the other hand, when the current traffic load falls within lower traffic levels, *mode 2* tends to allocate a higher value of R, i.e., closer to R_{max}. With appropriate configurations of parameters a_i, spectrum under-utilization can be minimized.

To avoid spectrum handover of ongoing services from the R-CRN to the N-CRN and to keep the value of R' positive, the algorithm selects the maximum value between R' and N_R as indicated in Line 17. Here N_R is the number of ongoing services in the R-CRN. In this study, spectrum handover from the R-CRN to the N-CRN is not enabled according to the reservation criterion. The last line of the algorithm inspects the availability of idle channels in the CRN in order to *adjust the reserved band* according to the value of R' obtained in Line 17 and to determine an appropriate value for R.

The selection of an appropriate working mode depends on channel failure and recovery rates, the QoS requirements of the PN and the SN, as well as the arrival rates of users. For instance, when most of the SUs demand low forced termination probability while operating in error-prone channels, *mode 1* would be recommended.

B. Dynamic Spectrum Access Scheme

In the following, we provide an overview of the proposed DSA scheme which is operated along with the DCR algorithm proposed above. In this subsection, we focus solely on the first variation in which the reserved channels could be allocated only to the interrupted services to avoid forced terminations. That is, with this variation, a newly arrived service will *not* be assigned to the R-CRN no matter it is a PU or an SU. The reason for applying such a restriction is due to the fact that maintaining a higher retainability level for established connections is one main QoS requirement. The first variation of the proposed DSA channel access scheme is explained below upon the occurrence of six events. The scheme with those six events is partly illustrated in a flow chart in Fig. 3. The step where the operation of the DCR algorithm is required is shown by the process block *X* in the flow chart.

1) PU arrival: First, the DCR algorithm presented in Subsection III-A is run to determine the corresponding R and adjust the R-CRN accordingly. If there is an idle channel in the N-CRN, a newly arrived PU will commence transmission on that idle channel. In case where the whole N-CRN is occupied,

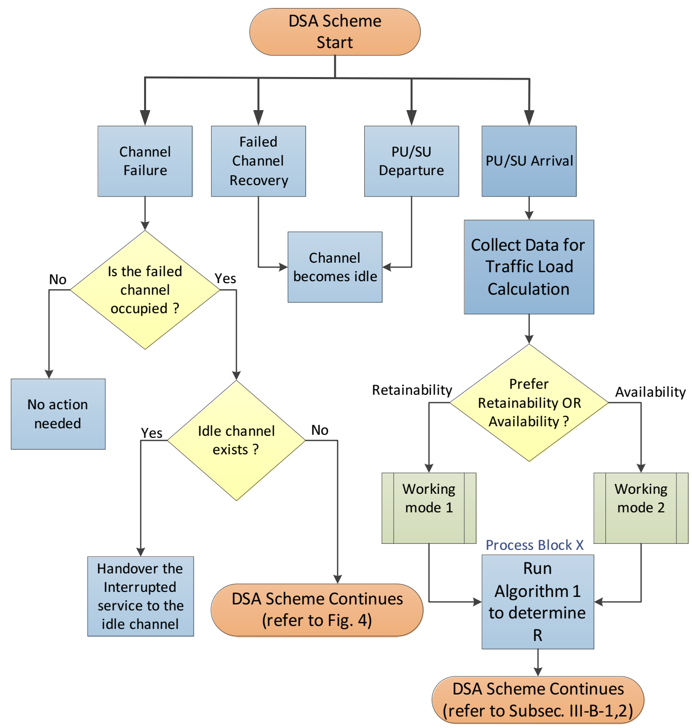

Fig. 3. Flow chart illustration of the proposed DSA scheme in part.

one of the SU services in the N-CRN is interrupted. It performs spectrum handover to a vacant channel in the R-CRN and a channel is released for access by the new PU. This interrupted SU service is allowed to access the R-CRN since it is subject to a forced termination upon the PU arrival. As mentioned earlier, the channels in the R-CRN are exclusively reserved for the ongoing services which are interrupted and cannot find another channel in the N-CRN. If the interrupted SU service cannot find an idle channel in the R-CRN, it is forced to terminate. When all operational channels in the N-CRN are occupied by PUs, a new PU request will be blocked.

2) SU arrival: The same as performed in PU arrivals, the system adjusts the number of channels in the N-CRN and the R-CRN by adopting Algorithm 1. Upon an SU arrival, the system will allocate an idle channel to the new SU from the N-CRN if it exists. If all operational channels in the N-CRN are occupied either by PUs or/and SUs, the new SU request is blocked without considering the channel availability in the R-CRN. Since PUs have priority over the channels in the N-CRN, a newly arrived SU cannot preempt other ongoing services.

3) PU departure: No handover actions are performed after a departure of PU services from the N-CRN or the R-CRN. The reasons are as follows. If an ongoing SU in the R-CRN was handed over to the N-CRN upon a service departure in the N-CRN, the channel occupancy of the N-CRN would increase. As a result, the blocking probability of SUs would increase since newly incoming SUs can only access the N-CRN. Moreover, such a handover process would lead to additional control traffic in the network. Considering those two reasons, plus the additional complexity needed in our CTMC model, we do not consider handovers from the R-CRN to the N-CRN in the design of the proposed DSA scheme.

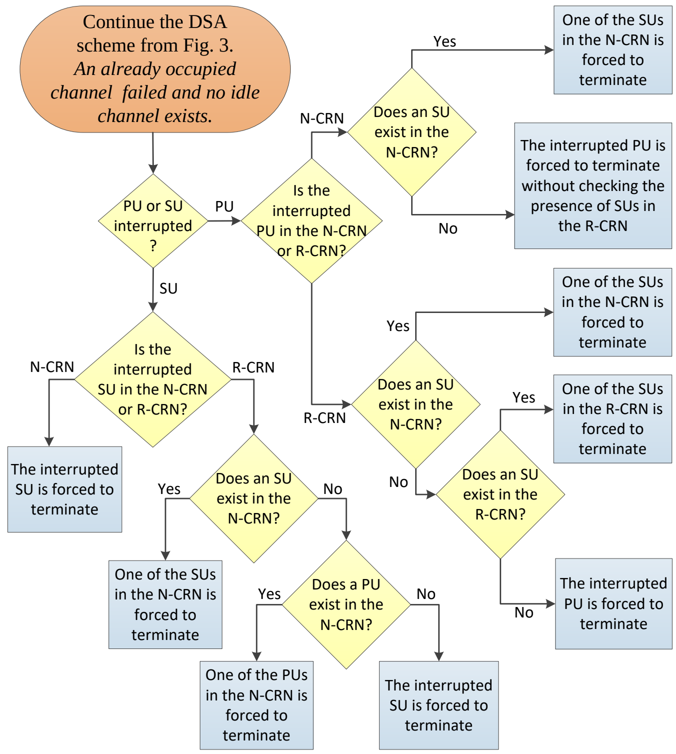

Fig. 4. Flow chart illustration of the proposed DSA scheme *when a currently occupied channel failed: PU priority variation I.*

4) SU departure: Similar to the above mentioned PU departure event, no channel adjustment is required after the departure of an SU service. The network keeps ongoing services in the R-CRN so that they can complete their services in the reserved band without performing spectrum handover back to the N-CRN.

5) Channel failure: If an idle channel in the N-CRN or R-CRN failed, the number of available channels in the network is accordingly decreased by one. However, when *an occupied channel in the N-CRN or R-CRN fails*, a spectrum handover procedure may be required, as illustrated in Fig. 4. Upon a failure of an already occupied channel, the handover process is performed based on the priority level assignment for PU and SU services in the N-CRN and the R-CRN. In this initial variation of the DSA scheme, the ongoing PU services in the R-CRN are considered as the highest priority class while the SU services in the R-CRN are considered as the second highest priority class. The PU and SU services in the N-CRN are regarded as the third and fourth priority classes respectively. Let $PL(s)$ denote the priority level assigned to a service type, s. Accordingly, $PL(PU_{R-CRN}) > PL(SU_{R-CRN}) > PL(PU_{N-CRN}) > PL(SU_{N-CRN})$ where $PL(PU_{R-CRN})$ denotes the priority level of an PU service in the R-CRN and so on. Since the ongoing SU services in the reserved band are considered to have higher priority than the PU services in the N-CRN, the PU services in the N-CRN cannot preempt an SU service in the R-CRN upon a channel failure. Moreover, this priority level assignment is independent of the selected working mode of the DCR algorithm.

However, it is worth mentioning that other PU access priorities may apply when a channel failure occurs. Later on in Section VIII, we propose two other variations of the DSA scheme which give full and partial priority to PUs respectively. In these variations of the DSA scheme, a higher priority is assigned to the PUs in both the R-CRN and the N-CRN.

6) Channel repair: Once a failed channel is re-established to its normal conditions and ready for carrying services, it is said that the channel is repaired or recovered. However, as operational channels, ongoing services do not perform spectrum handover when a failed channel is recovered. Note that R is *not* adjusted upon channel failures or recoveries.

IV. CTMC Modeling and Performance Metrics

We model the considered network and the proposed DSA scheme (variation I) using a Markov chain, with continuous time and discrete states. Let $\mathcal{S}$ be the set of feasible states in the CTMC model. The states of the CTMC model corresponding to the DSA scheme are represented by $\boldsymbol{x} = (i_n, j_n, i_r, j_r, f)$ where i_n and j_n denote the number of PU and SU services in the N-CRN, i_r and j_r denote the number of those services in the R-CRN respectively and f denotes the number of failed channels in the whole CRN. The total numbers of occupied channels in the N-CRN and the R-CRN for a given state $\boldsymbol{x}$ are denoted by $B_n(\boldsymbol{x})$ and $B_r(\boldsymbol{x})$ respectively, i.e., $B_n(\boldsymbol{x}) = i_n + j_n$ and $B_r(\boldsymbol{x}) = i_r + j_r$. Moreover, denote the sum of occupied plus failed channels in the whole CRN as $B(\boldsymbol{x})$, i.e., $B(\boldsymbol{x}) = B_n(\boldsymbol{x}) + B_r(\boldsymbol{x}) + f$. Therefore, the number of idle channels in state $\boldsymbol{x}$ can be obtained as $M - B(\boldsymbol{x})$. Note that the number of reserved channels in state $\boldsymbol{x}$ is denoted as $R(x)$. The state transition rates associated with different events are summarized with different conditions in Table I. The transitions and transition rates mentioned in this table constitute the transition rate matrix $\boldsymbol{Q}$. Let $\pi(\boldsymbol{x})$ denote the steady state probability of being in state $\boldsymbol{x}$. The steady state probabilities of each state can be calculated according to

$$\boldsymbol{\pi Q} = \boldsymbol{0}, \quad \sum_{\boldsymbol{x}\in\mathcal{S}} \pi(\boldsymbol{x}) = 1, \tag{1}$$

where $\boldsymbol{\pi}$ is the steady state probability vector and $\boldsymbol{0}$ is a row vector of all 0's. In the following, we derive mathematical expressions to analyze performance metrics of the CRN.

A. Capacity

In this study, *capacity* is defined as the rate of service completions, i.e., the average number of service completions per time unit. Let ρ_P and ρ_S be the capacity of PU and SU services respectively. Correspondingly, we obtain

$$\rho_P = \sum_{\boldsymbol{x}\in\mathcal{S}} (i_n + i_r)\mu_P \pi(\boldsymbol{x}), \tag{2}$$

$$\rho_S = \sum_{\boldsymbol{x}\in\mathcal{S}} (j_n + j_r)\mu_S \pi(\boldsymbol{x}). \tag{3}$$

B. Channel Availability

Once all channels in a CRN are occupied at a given instant, a new user request will simply be blocked and the CRN is said to be unavailable for new users. Thus, channel availability

TABLE I
TRANSITIONS FROM A GENERIC STATE $\boldsymbol{x} = (i_n, j_n, i_r, j_r, f)$ OF THE DYNAMIC SPECTRUM ACCESS SCHEME UPON PU/SU EVENTS AND CHANNEL FAILURES AND REPAIRS: PU PRIORITY VARIATION I.

Event	Destination State	Tran. rate	Conditions
1. PU AR. A vacant channel exists in the N-CRN.	$(i_n+1, j_n, i_r, j_r, f)$	λ_P	There exists at least one vacant channel in the N-CRN, i.e., $B_n(\boldsymbol{x}) < M - R(x)$.
2. PU AR. No vacant channel exists in the N-CRN. SU_N performs handover to the R-CRN.	$(i_n+1, j_n-1, i_r, j_r+1, f)$	λ_P	$B_n(\boldsymbol{x}) = M - R(x); j_n > 0; B_r(\boldsymbol{x}) < R(x)$.
3. PU AR. An SU_N is forced to terminate.	$(i_n+1, j_n-1, i_r, j_r, f)$	λ_P	$B_n(\boldsymbol{x}) = M - R(x); j_n > 0; B(\boldsymbol{x}) = M$.
4. SU AR. A vacant channel exists in the N-CRN.	$(i_n, j_n+1, i_r, j_r, f)$	λ_S	$B_n(\boldsymbol{x}) < M - R(x); B(\boldsymbol{x}) < M$.
5. PU DP from the N-CRN.	$(i_n-1, j_n, i_r, j_r, f)$	$i_n\mu_P$	There is at least one PU service in the N-CRN, i.e., $i_n > 0$.
6. PU DP from the R-CRN.	$(i_n, j_n, i_r-1, j_r, f)$	$i_r\mu_P$	$i_r > 0$.
7. SU DP from the N-CRN.	$(i_n, j_n-1, i_r, j_r, f)$	$j_n\mu_S$	$j_n > 0$.
8. SU DP from the R-CRN.	$(i_n, j_n, i_r, j_r-1, f)$	$j_r\mu_S$	$j_r > 0$.
9. Idle channel failure.	$(i_n, j_n, i_r, j_r, f+1)$	$(M - B(\boldsymbol{x}))\lambda_F$	There is at least one idle channel in the CRN, i.e., $B(\boldsymbol{x}) < M$.
10. An occupied channel fails. An idle channel exists in the CRN.	$(i_n, j_n, i_r, j_r, f+1)$	$(B(\boldsymbol{x}) - f)\lambda_F$	$f < B(\boldsymbol{x}) < M$.
11. An occupied channel fails. No idle channels exist in the CRN. An SU_N is forced to terminate.	$(i_n, j_n-1, i_r, j_r, f+1)$	$(M-f)\lambda_F$	$B(\boldsymbol{x}) = M; j_n > 0$
12. An occupied channel fails. No idle channels exist in the CRN. A PU_N is forced to terminate†.	$(i_n-1, j_n, i_r, j_r, f+1)$	$(M-f)\lambda_F$	$B(\boldsymbol{x}) = M; j_n = 0; i_n > 0$.
13. An occupied channel fails. No idle channels exist in the CRN. An SU_R is forced to terminate.	$(i_n, j_n, i_r, j_r-1, f+1)$	$(M-f)\lambda_F$	$B(\boldsymbol{x}) = M; B_n(\boldsymbol{x}) = 0; j_r > 0$.
14. An occupied channel fails. No idle channels exist in the CRN. A PU_R is forced to terminate.	$(i_n, j_n, i_r-1, j_r, f+1)$	$(M-f)\lambda_F$	$B(\boldsymbol{x}) = M; B_n(\boldsymbol{x}) = j_r = 0; i_r > 0$.
15. A failed channel is repaired.	$(i_n, j_n, i_r, j_r, f-1)$	$f\mu_R$	$f > 0$.

The notations AR and DP indicate an arrival event and a departure event respectively. An SU service and a PU service in the N-CRN are denoted as SU_N and PU_N respectively. An SU service and a PU service in the R-CRN are denoted as SU_R and PU_R respectively.

†This event happens only when there are no SUs in the N-CRN (condition $j_n = 0$). If there is at least one SU service in the N-CRN, the PU is not terminated. This follows from the priority rule $PL(PU_{N-CRN}) > PL(SU_{N-CRN})$. Moreover, the network does not need to check the existence of any ongoing SU services in the R-CRN before terminating the PU in the N-CRN. This is because of the priority rule $PL(SU_{R-CRN}) > PL(PU_{N-CRN})$ according to PU priority variation I of the DSA scheme.

based measurements are of importance for both PUs and SUs when performing channel access in a CRN [32]. In this paper, channel availability for PU or SU services is defined as *the probability that the CRN will allocate a channel to a new PU or SU arrival* without blocking the request. Blocking of a new PU service occurs when all the operational channels in the N-CRN are occupied by PUs. Let A_P denote the channel availability of PU services. We obtain

$$A_P = 1 - \sum_{\substack{\boldsymbol{x} \in \mathcal{S} \\ (B(\boldsymbol{x})=M \text{ or } B_n(\boldsymbol{x})=M-R(x)); \quad j_n=0}} \pi(\boldsymbol{x}). \tag{4}$$

Moreover, an SU service request will be blocked if all the operational channels in the N-CRN are occupied by PUs or/and SUs. Therefore, the channel availability of SU services can be obtained by

$$A_S = 1 - \sum_{\substack{\boldsymbol{x} \in \mathcal{S} \\ B(\boldsymbol{x})=M \text{ or } B_n(\boldsymbol{x})=M-R(x)}} \pi(\boldsymbol{x}). \tag{5}$$

Accordingly, the blocking probabilities of PU and SU services, denoted as P_P^B and P_S^B respectively, are obtained as

$$P_P^B = 1 - A_P \quad \text{and} \tag{6}$$

$$P_S^B = 1 - A_S. \tag{7}$$

C. Retainability

A session-oriented service often explicitly measures the probability that a session delivers service with an acceptable level of quality until the session terminates [33].

1) Retainability definition: Retainability is defined as the probability that a connection, once established, will operate within the specified transmission quality without interruption for a given time interval [25]. Mathematically, the retainability of a service, θ, is expressed as

$$\theta = 1 - P_F \tag{8}$$

where P_F is the forced termination probability of that service.

2) Retainability of the SN: The probability of forced termination represents the probability that an ongoing SU service in the network is forced to terminate before its communication is finished [34]. Note that when considering random channel failures, forced terminations could happen to both SU and PU services. The forced termination probability of SUs, P_F^S, can be expressed as the ratio between the mean forced termination rate of SU services, P_R^{SE}, and the effective rate in which a new SU service is assigned a channel, Λ_S [34]. Denote the rate of forced terminations of SUs due to PU arrivals as R_S. Then we have

$$R_S = \lambda_P \sum_{\substack{\boldsymbol{x}\in\mathcal{S}\\ B(\boldsymbol{x})=M;\ j_n>0}} \pi(\boldsymbol{x}). \tag{9}$$

In addition, ongoing SU services can also be terminated upon a channel failure when all other channels in the CRN are busy. Denote the rate of forced terminations of SUs due to channel failures as $R_S^{'}$. It is obtained by

$$R_S^{'} = \lambda_F \sum_{\substack{\boldsymbol{x}\in\mathcal{S}\\ B(\boldsymbol{x})=M;\\ ((j_n>0)\text{ or }(B_n(\boldsymbol{x})=0;\ j_r>0))}} (M-f)\pi(\boldsymbol{x}). \tag{10}$$

Since the effective rate in which a new SU service is assigned a channel is $\Lambda_S = A_S\lambda_S$, we have $P_S^F = \frac{R_S+R_S^{'}}{\Lambda_S}$. Correspondingly, the retainability of SU services, θ_S, can be expressed as

$$\theta_S = 1-\left(\frac{R_S+R_S^{'}}{\Lambda_S}\right). \tag{11}$$

3) Retainability of the PN: Similarly, the forced termination probability of PU services due to channel failures, P_P^F, can be expressed as

$$P_P^F = (R_P + R_P^{'})/\Lambda_P, \tag{12}$$

where $R_P^{'}$ and Λ_P are given by

$$R_P^{'} = \lambda_F \sum_{\substack{\boldsymbol{x}\in\mathcal{S}\\ B(\boldsymbol{x})=M;\\ ((j_n=0;\ i_n>0)\text{ or }(B_n(\boldsymbol{x})=0;\ j_r=0;\ i_r>0))}} (M-f)\pi(\boldsymbol{x}) \tag{13}$$

and $\Lambda_P = A_P\lambda_P$ respectively. Note that R_P, which denotes the forced termination rate of PUs due to new user arrivals, always equals zero since none of the ongoing PUs can be terminated due to the arrivals of new users. Therefore, the retainability of PU services, θ_P, is given by

$$\theta_P = 1-\frac{R_P^{'}}{\Lambda_P}. \tag{14}$$

D. Network Unserviceable Probability (NUP)

While an individual analysis of the blocking probability helps us to estimate the channel occupancy status of the network with respect to network resources, an individual analysis of the forced termination probability determines the retainability of the services. However, none of them alone is able to evaluate the overall satisfaction of the network from a user's point of view. More specifically, *to complete a requested service successfully, the service should neither be blocked upon its arrival nor be terminated before its completion*. Therefore, the proportion of services that can be completed successfully is an important metric to be evaluated.

Network unserviceable probability, which reflects this metric, is an uncharted item in CRN research. From the viewpoint of users in the network, the service can be accomplished only if the network provides a connection *when it is required and as long as it is required* [35]. Thus the considered CRN is unserviceable if either of the following conditions happens:

1) No channel is available upon a service arrival due to failed channels and/or traffic congestion;
2) The channel allocated for the connection failed before the session is finished due to channel failures (for both PUs and SUs) or PU arrivals (for SUs only).

Accordingly, the NUP for SU services, Q_S, can be defined as the probability that an SU service cannot be completed successfully. It is obtained by calculating the ratio between the rate of service completions and the rate of arrivals as follows:

$$\begin{aligned} Q_S &= 1-(\text{Prob. of successfully finishing an SU service}),\\ &= 1-\frac{\lambda_S(1-P_S^B)(1-P_S^F)}{\lambda_S},\\ &= P_S^B + P_S^F - P_S^B P_S^F. \end{aligned} \tag{15}$$

Note that $1-P_S^B$ represents the probability that a new SU request is admitted to the network and $1-P_S^F$ represents the probability that an ongoing SU flow could finish its service successfully. These two probabilities together decide the probability of successfully initiating and completing a service. From (15), it is clear that both blocking probability and forced termination probability are reflected when evaluating the NUP performance. Similarly, the NUP for PU services, denoted as Q_P, can be derived as follows.

$$Q_P = P_P^B + P_P^F - P_P^B P_P^F. \tag{16}$$

Generally speaking, the lower the NUP, the better the network performance. Moreover, there is a tradeoff between channel availability and service retainability when performing the proposed DCR algorithm. That is, a larger R reduces channel availability for newly arrived users but leads to higher retainability for ongoing services. Therefore, deriving such a metric, i.e., the NUP which can evaluate the joint effect of those performance measures, will help us better understand the overall performance of the proposed DCR algorithm.

V. Network Performance under Homogeneous Channel Failures

In this section, we provide numerical results to investigate the impact of channel failures and traffic conditions on the defined performance metrics considering that there is only one type of channel failure. The default values of the network configuration are listed in Table II unless otherwise stated. Four traffic load levels are considered when calculating R. Correspondingly, we configure $a_1 = 0.75, a_2 = 0.50$ and $a_3 = 0.25$ in Algorithm 1. Note that the number of channels allocated to the R-CRN in the SCR scheme is denoted as R in the presented numerical results. The curves representing *Without channel reservation* in this section are obtained by configuring $R = 0$.

For homogeneous channels, we configure the channel failure rate as 0.05 failure per time unit by default for most results

TABLE II
SUMMARY OF SYMBOLS IN THE CTMC AND VALUE CONFIGURATION

Symbol	Description	Default Value
M	Total number of channels in the CRN	8
R_{max}	Upper bound of number of reserved channels in DCR	2 or 4
R	Number of reserved channels	2 (for SCR); variable (for DCR and decided by Alg. I)
λ_P, λ_S	PU and SU arrival rates	5 services per unit of time
μ_P, μ_S	PU and SU service rates per channel	2 services per unit of time
λ_F	Failure rate of channels	0.05 failures per unit of time
μ_R	Repair rate per channel	1.0 repairs per unit of time

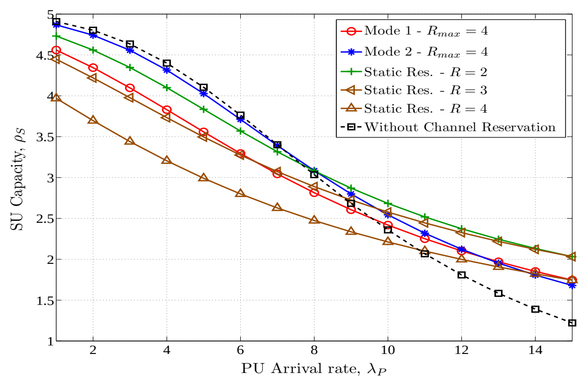

Fig. 5. Capacity achieved by SUs as a function of the PU arrival rate.

illustrated in this section. Other values of λ_F are also studied in a few figures. With $\lambda_F = 0.05$, a low failure rate is targeted in comparison with the arrival rates of PU and SU services. For instance, when the PU arrival rate is 10 flows per time unit, on average a channel failure occurs every 200 PU arrivals. This is a reasonable configuration, meaning that the termination of a service upon a channel failure does not occur frequently. On the other hand, the channel should be repaired at a much higher rate in order to maintain a sufficiently high level of user satisfaction once channel failure occurs. Therefore, a repair rate of $\mu_R = 1.0$ per time unit is configured for our performance evaluation.

A. Capacity of the PN and the SN

In Fig. 5 and Fig. 6, we plot the achieved capacity for the SN and the PN respectively as a function of the PU arrival rate considering both working modes in the DCR algorithm with different configurations of R_{max}, given that $\lambda_F = 0.05$. From Fig. 5, we notice that, if the PU arrival rate, λ_P, is low, the capacity of the SN decreases when channel reservation is adopted, no matter a static or a dynamic reservation scheme is adopted. Generally speaking, when a certain amount of spectrum is reserved, fewer channels are available for new users, leading to lower SU capacity. The situation changes when more PUs become active, i.e., with a higher λ_P.

At a lower λ_P, an SU service is unlikely to be terminated forcibly and consequently the reserved channels are rarely accessed. On the other hand, the incoming SU services cannot access the R-CRN although there are vacant channels in the reserved band, regardless of the working mode adopted. Therefore, most of the reserved channels are under-utilized and

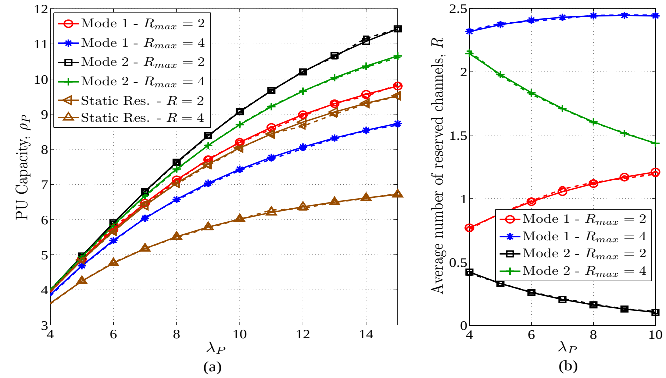

Fig. 6. Capacity achieved by PUs as a function of the PU arrival rate. The dashed lines represent the results obtained by simulations.

this fact leads to capacity degradation compared with the case without channel reservation. Furthermore, at a higher λ_P, the forced termination probability of SUs increases and some of the interrupted services will get channel access opportunities in the reserved band to continue their services until completion. Consequently the capacity of the SN with DCR is increased compared with the case without channel reservation. Therefore, it is interesting to investigate under which conditions a network needs channel reservation as well as under which conditions this is not needed. Accordingly, we can conclude that the advantage of DCR over the scheme without channel reservation in terms of the achieved SU capacity is evident when $\lambda_P > 9.5$ for mode 1 and when $\lambda_P > 7$ for mode 2, as shown in Fig. 5.

Fig. 6(a) shows the capacity of the PN as λ_P varies. The performance diversity of the proposed working modes is shown more clearly in this figure. It is observed that the capacity achieved in the PN becomes higher in mode 2 than in mode 1 due to the following reasons. The number of reserved channels, R, in mode 1 is increased when the ongoing traffic load in the network rises as illustrated in Fig. 6(b). In contrast, mode 2 reserves a lower value of R with an increasing traffic load. This is because that, for a given high value of λ_P, the network allocates a comparatively larger number of channels to the R-CRN in mode 1 than in mode 2. Consequently, new PUs lose channel access opportunities in mode 1, leading to a capacity reduction when compared with mode 2. However, the PU capacity loss is less than 3% for an SU retainability increase of 20% when $\lambda_P = 8$. Therefore, PUs' performance degradation is not that significant in comparison with the retainability gain achieved in the SN. To explore how this PU capacity loss can be minimized, we propose in Section VIII two other variations of the DSA scheme which give PU higher access privilege in the R-CRN.

Let us further analyze the SU capacity in Fig. 5. It is shown that mode 2 can outperform mode 1 only if $\lambda_P < 13$. Again, this is due to the priority consideration for channel access opportunities between new SU requests and ongoing SU services. In other words, when the traffic arrival rate is low, mode 2 serves better since it opens more opportunities for new arrivals. At higher arrival rates, mode 1 is preferred since it protects ongoing services. Moreover, at higher arrival

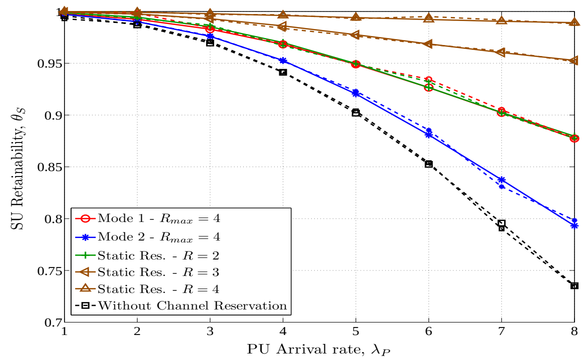

Fig. 7. Retainability of SU services as a function of the PU arrival rate. The dashed lines represent the results obtained by simulations.

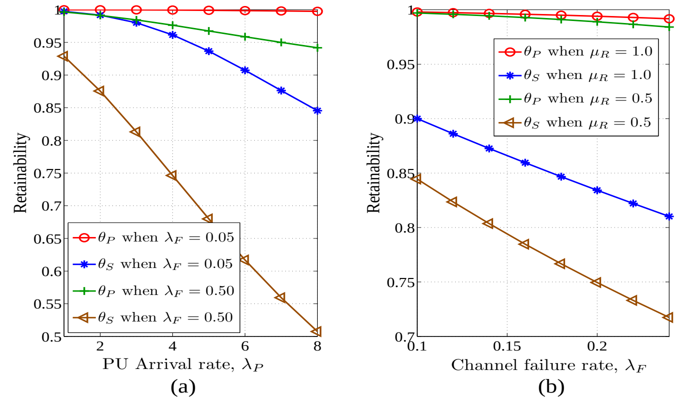

Fig. 8. Retainability of SU and PU services as a function of the PU arrival rate and channel failure rate when mode 1 is selected in Algorithm 1.

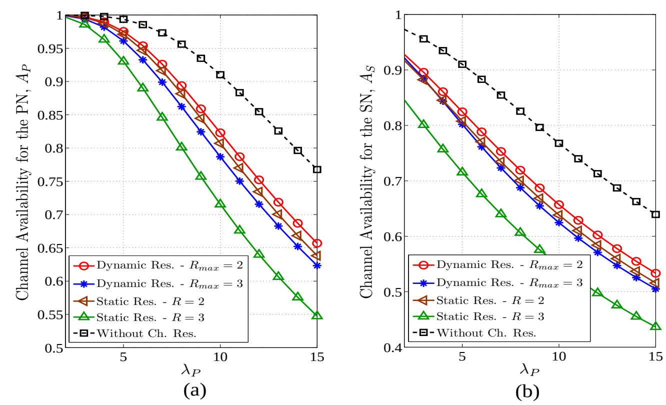

Fig. 9. Channel availability for PN and SN as a function of the PU arrival rate. DCR indicates mode 1 in Algorithm 1.

rates, both static and dynamic channel reservations show better performance compared with the scheme without channel reservation in terms of SU capacity. As already mentioned, the provision of channel access opportunities for interrupted SU services in both DCR and SCR at higher λ_P leads to increased capacity. On the other hand, without channel reservation, those channel access opportunities would be obtained by PUs, leading to lower SU capacity.

Furthermore, extensive simulations have been conducted by utilizing MATLAB to validate the correctness and the preciseness of the obtained analytical results. For the sake of illustration clarity, we represent simulation results by dashed lines only in Fig. 6 and Fig. 7. The simulations were performed in a similar way as presented in our previous study [34]. The results clearly demonstrate that the results obtained from the theoretical derivations match closely the simulation results.

B. Retainability

Let us now observe the achieved retainability by SU services as the PU arrival rate varies for both working modes, as plotted in Fig. 7. As shown in this figure, the retainability of services can be increased significantly when one of the working modes is employed in the DCR algorithm, compared with a CRN without channel reservation. For instance, when $\lambda_P = 8$, mode 1 or mode 2 based DCR improves retainability by approximately 20% and 10% respectively in comparison with the CRN without channel reservation. This result confirms that better retainability is achieved by selecting mode 2.

To achieve higher retainability, more channels should be reserved in the R-CRN, according to the number of ongoing sessions in the CRN. An ongoing session in the N-CRN needs channel access in the R-CRN when it is interrupted due to PU arrivals or channel failures if all other operational channels are busy. Mode 1 follows this principle. This is most likely to happen when the traffic load is high. Conversely, mode 2 may allocate channels to the R-CRN even when there is a low demand for channel reservation. At a higher traffic load, the number of channels in the R-CRN is reduced in mode 2 and therefore the possibility of improving retainability diminishes. For this reason, the retainability exhibits always a higher value with mode 1 than with mode 2. Moreover, via SCR with a higher number of reserved channels, it is able to provide higher service retainability than in DCR since the number of channels in the R-CRN is kept fixed and they are exclusively allocated to ongoing services [1]. However, SCR is not flexible and it shows comparatively poorer performance in terms of channel availability, as to be pointed out in the next subsection.

In Fig. 8, we investigate the retainability level of PU and SU services with respect to different channel failure rates and repair rates. As observed in the figure, higher channel failure rates and lower channel repair rates lead to a decreased retainability level. With a large λ_F, channels are more likely to fail. With a lower μ_R, the channels have to stay for a longer period of time as an unavailable (or a failed) channel. Consequently, the number of idle channels in the network diminishes and thus the retainability degrades. However, the retainability level of PU services is always higher than that of SU services owing to its channel access priority privileges.

C. Channel Availability

In Fig. 9, we evaluate the channel availability of PU and SU services respectively considering different configurations. In the presence of random channel failures, channel availability reveals access opportunities upon PU and SU arrivals at different traffic load levels. From this figure, it is clear that the achieved channel availability decreases due to channel reservation and the degree of degradation is proportional to the number of reserved channels. If channel reservation is not adopted, i.e., when $R = 0$, all channels would be available for new services. Therefore, the scheme without channel

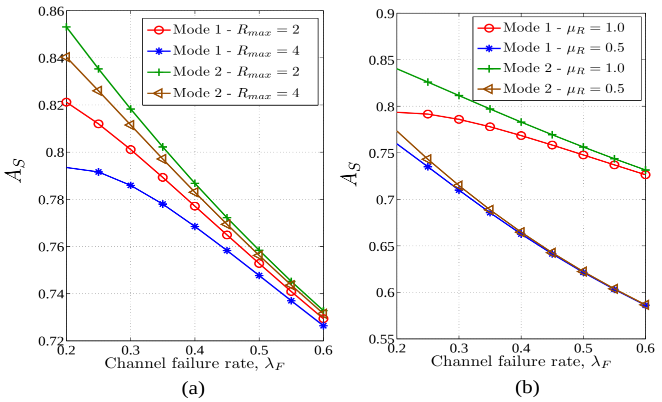

Fig. 10. Channel availability for SUs as a function of the channel failure rate.

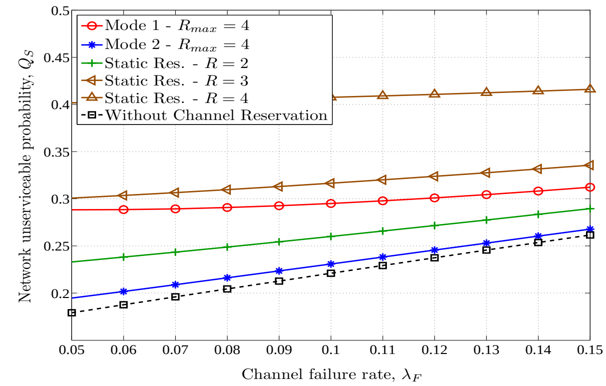

Fig. 11. Network unserviceable probability as the channel failure rate varies.

reservation always exhibits the highest availability for both networks.

Fig. 9 shows also that the channel availability is significantly influenced by the number of reserved channels in static reservation more than in the DCR scheme. For instance, consider SCR at $\lambda_P = 15$. The percentage of channel availability reduction for SUs with $R = 3$ is 15.4% with respect to the case of $R = 2$. On the other hand, the DCR scheme with mode 1 achieves a percentage of channel availability reduction for SUs as 5.3% for the case of $R_{max} = 2$ with the same configuration. The reason is as follows. Mode 1 intends to assign a larger value for R when the ongoing traffic load in the network becomes higher than a certain threshold, resulting in better spectrum utilization. Fig. 9 illustrates further that the PUs' channel availability is higher than that of the SN. This is due to the priority access of the PN, i.e., the ongoing SU services could be preempted by an incoming PU, but not the other way round.

Furthermore, the achievable channel availability for the SN is reported in Fig. 10 considering different channel failure and recovery rates. As expected, with a higher channel failure rate, channel availability tends to monotonically decrease due to lack of idle channels. In this study, channel availability for SU services, A_S, is given by $A_S = 1 - P_S^B$ where P_S^B denotes the blocking probability. When a few number of channels is reserved, i.e., at smaller values of R, more channel access opportunities can be given to the new users. Consequently, the channel availability becomes higher. On the other hand, with a higher channel repair rate, fewer PU and SU services are blocked. This is because that, when μ_R is higher, the failed channels are recovered shortly and become idle again, leading to increased channel access opportunities for new users.

As expected, the achieved channel availability for new SU services becomes higher if mode 2 is selected. On the other hand, selecting mode 2 leads to a lower retainability level in the DCR algorithm. This result indicates a need for an assessment of the tradeoff between the new users' channel availability and the ongoing services' retainability. The NUP-based channel reservation algorithm presented in Section VII provides an insight regarding this tradeoff.

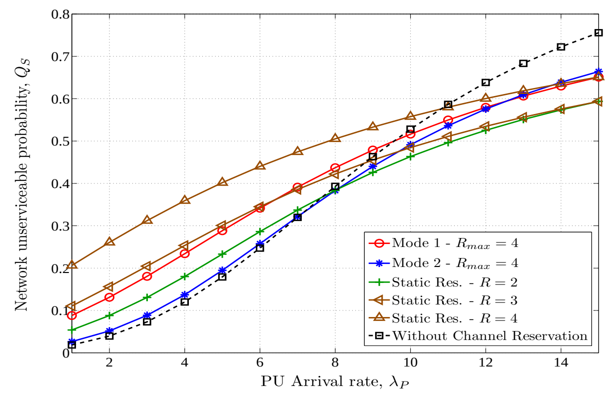

Fig. 12. NUP as a function of the PU arrival rate.

D. Network Unserviceable Probability

Fig. 11 confirms that NUP increases with a higher channel failure rate. Moreover, DCR schemes generally outperform the SCR scheme in terms of the NUP. To illustrate this, let us consider the scheme "Mode 2 with $R_{max} = 4$". At $\lambda_F = 0.1$, the network's NUP is only 0.23. Under the same configuration, the network experiences a higher NUP with static reservation when $R \geq 2$. Compared with mode 2, the NUP of SUs increases by 13% if the static scheme with $R = 2$ is employed. This percentage becomes 39% and 73% for static channel reservation with $R = 3$ and $R = 4$ respectively. Even though DCR reduces NUP considerably, the channel access schemes without channel reservations cannot be outperformed in terms of NUP for the arrival rates of PU and SU services used in Fig. 11. However, by employing mode 2, SUs experience lower NUP compared with DCR with mode 2.

Another important observation regarding the NUP performance of the SN can be found in Fig. 12. This figure depicts the NUP of SU services, as a function of the PU arrival rate. In the DCR algorithm presented in Subsection III-A, channel allocation to the R-CRN is controlled based on the ongoing traffic load. Therefore, the NUP performance of SUs is sensitive to the traffic arrival rate. The results presented in Fig. 12 indicate that, under certain traffic loads, channel reservation provides better performance than the CRN without channel reservations. For instance, when $\lambda_P < 7$, the best NUP performance is observed under the DCR scheme with

TABLE III
TRANSITIONS FROM A GENERIC STATE $\boldsymbol{x} = (i_n, j_n, i_r, j_r, f_1, f_2)$ OF THE CHANNEL ACCESS SCHEME UPON *Type 1* CHANNEL FAILURES AND REPAIRS: PU PRIORITY VARIATION I.

Event	Destination State	Tran. rate	Conditions
9. Idle channel fails due to a first type failure.	$(i_n, j_n, i_r, j_r, f_1+1, f_2)$	$(M-B(\boldsymbol{x}))\lambda_{F_1}$	$B(\boldsymbol{x}) < M$.
10. An occupied channel fails due to a *Type 1* failure. An idle channel exists in the CRN.	$(i_n, j_n, i_r, j_r, f_1+1, f_2)$	$(B(\boldsymbol{x})-F)\lambda_{F_1}$	$F < B(\boldsymbol{x}) < M$.
11. An occupied channel fails due to a *Type 1* failure. No idle channels exist in the CRN. An SU_N is forced to terminate.	$(i_n, j_n-1, i_r, j_r, f_1+1, f_2)$	$(M-F)\lambda_{F_1}$	$B(\boldsymbol{x}) = M; F < M; j_n > 0$
12. An occupied channel fails due to a *Type 1* failure. No idle channels exist in the CRN. A PU_N is forced to terminate.	$(i_n-1, j_n, i_r, j_r, f_1+1, f_2)$	$(M-F)\lambda_{F_1}$	$B(\boldsymbol{x}) = M; F < M; j_n = 0; i_n > 0$.
13. An occupied channel fails due to a *Type 1* failure. No idle channels exist in the CRN. An SU_R is forced to terminate.	$(i_n, j_n, i_r, j_r-1, f_1+1, f_2)$	$(M-F)\lambda_{F_1}$	$B(\boldsymbol{x}) = M; F < M; B_n(\boldsymbol{x}) = 0; j_r > 0$.
14. An occupied channel fails due to a *Type 1* failure. No idle channels exist in the CRN. A PU_R is forced to terminate.	$(i_n, j_n, i_r-1, j_r, f_1+1, f_2)$	$(M-F)\lambda_{F_1}$	$B(\boldsymbol{x}) = M; F < M; B_n(\boldsymbol{x}) = j_r = 0; i_r > 0$.
15. A failed channel due to a *Type 1* failure is repaired.	$(i_n, j_n, i_r, j_r, f_1-1, f_2)$	$f_1\mu_{R_1}$	$f_1 > 0$.

For the activities corresponding to user arrivals and departures, refer to the events numbered **1** to **8** in Table I. The only modification required in those events is the replacement of the final element, f, in each state by f_1, f_2.

The notations AR and DP indicate an arrival event and a departure event respectively. An SU service and a PU service in the N-CRN are denoted as SU_N and PU_N respectively. An SU service and a PU service in the R-CRN are denoted as SU_R and PU_R respectively. Denote $B_n(\boldsymbol{x}) = i_n + j_n$, $B_r(\boldsymbol{x}) = i_r + j_r$, $B(\boldsymbol{x}) = B_n(\boldsymbol{x}) + B_r(\boldsymbol{x}) + F$, where $F = f_1 + f_2$.

$R = 0$, i.e., without any channel reservation. However, once $\lambda_P > 11$, the network experiences the highest NUP without channel reservation. The reason for this result is explained below considering the joint effect of channel availability and service retainability.

When the PU traffic load is low, forced terminations of SU services do not happen with higher probabilities. However, once some channels are reserved, a new user could be blocked even though there are vacant channels available in the R-CRN. This leads to a significant increase of the NUP in the channel reservation scheme because of high blocking probabilities. Contrarily, at heavier PU traffic loads, the ongoing SU services are interrupted by PU arrivals and thus the forced termination probability of SUs increases. With reserved channels, those forced terminations can be alleviated and therefore the network could avoid potential capacity losses, leading to a low NUP. Unlike in the channel reservation enabled networks, the CRNs without channel reservation cannot protect ongoing SU services at higher PU traffic loads. Consequently, the NUP for those systems exhibits a large value. In brief, it can be concluded that the proposed DCR helps to reduce the NUP in a CRN when λ_P is higher. However, it is recommended to limit the number of reserved channels a low arrival rates.

VI. MODELING AND PERFORMANCE ANALYSIS UNDER HETEROGENEOUS CHANNEL FAILURES

As wireless networks may interconnect with each other under diverse environments, different types of failures may happen. Consequently, some failures may occur more frequently than others and some failed channels may be recovered faster than others. By considering these factors, we develop in this section a failure model which consists of two types of channel failures in order to investigate the effect of heterogeneous failures on the performance of the studied multi-channel error-prone CRN.

A. CTMC Model for Heterogeneous Failures in CRNs with Reserved Spectrum

The CTMC model presented in Section IV is extended to model the heterogeneous failure types. Consider two types of failures which have distinct failure rates and repair rates. Let the failure rate and the repair rate of a channel for *Type 1 failure* be λ_{F_1} and μ_{R_1} respectively. Similarly those values for *Type 2 failure* will be λ_{F_2} and μ_{R_2}. Let $\mathcal{S}$ be the set of feasible states in the CTMC model. The states of the CTMC model corresponding to the proposed DCR based spectrum access scheme are then represented by $\boldsymbol{x} = (i_n, j_n, i_r, j_r, f_1, f_2)$ where terms f_1 and f_2 represent the number of failed channels in the CRN due to the first and second failure types respectively. Term $B(\boldsymbol{x})$ which represents the sum of the occupied and failed channels in the whole CRN becomes $B(\boldsymbol{x}) = B_n(\boldsymbol{x}) + B_r(\boldsymbol{x}) + F$ where $F = f_1 + f_2$, $B_n(\boldsymbol{x}) = i_n + j_n$ and $B_r(\boldsymbol{x}) = i_r + j_r$ as mentioned in the homogeneous channel failure case.

The corresponding state transition table is obtained by updating Table I with necessary modifications. In Table III, we list the updated state transitions numbered **9** to **15** for *Type 1* failures. Note that the first 8 events are the same as

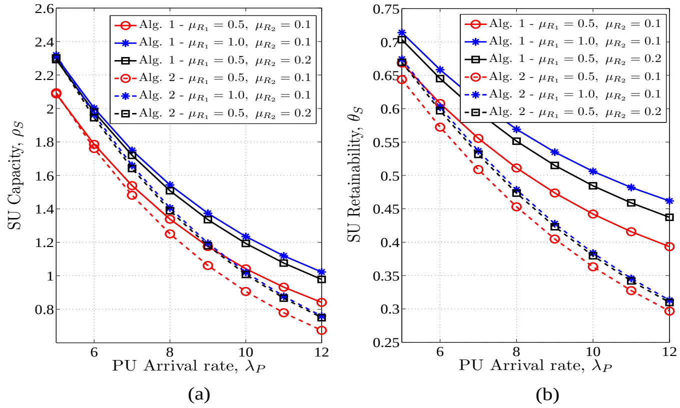

Fig. 13. Capacity and service retainability of the SN in the presence of heterogeneous channel failures when $\lambda_{F_1}=0.1$ and $\lambda_{F_2}=0.01$. The solid lines and the dashed lines represent the numerical results corresponding to mode 1 and mode 2 respectively.

listed in Table I, but we need to replace the final element, f, of each state by f_1, f_2. Similarly, the transitions for *Type 2* failures can be obtained, however, they are not listed here. In order to accommodate heterogeneous failure in Algorithm 1, the only update that needs to be done is located in *Line 1* and *Line 2* of Algorithm 1. Accordingly, the expressions in *Line 1* and *Line 2* should be changed as

Line 1: $\Phi=(i_n+j_n+i_r+j_r)/(M-(f_1+f_2))$ and

Line 2: $N_{Avail}=M-(i_n+j_n+i_r+j_r+(f_1+f_2))$

respectively, in Algorithm 1. Furthermore, the corresponding equations which represent the forced termination rates of SU and PU services due to channel failures in the heterogeneous case are obtained as

$$R_S^{'}=(\lambda_{F_1}+\lambda_{F_2})\sum_{\substack{\boldsymbol{x}\in\mathcal{S}\\ B(\boldsymbol{x})=M;\ ((j_n>0)\ \text{or}\\ (B_n(\boldsymbol{x})=0;\ j_r>0))}}(M-F)\pi(\boldsymbol{x}) \quad (17)$$

and

$$R_P^{'}=(\lambda_{F_1}+\lambda_{F_2})\sum_{\substack{\boldsymbol{x}\in\mathcal{S}\\ B(\boldsymbol{x})=M;\ ((j_n=0;\ i_n>0)\ \text{or}\\ (B_n(\boldsymbol{x})=0;\ j_r=0;\ i_r>0))}}(M-F)\pi(\boldsymbol{x}). \quad (18)$$

The other expressions which represent the PU/SU capacity, blocking probability, NUP and retainability are the same as in the homogeneous failure case.

B. Numerical Results with Heterogeneous Failure Rates

The numerical results presented in this subsection are obtained for a CRN with $M=6$ channels and $R_{max}=2$.

1) Capacity and retainability: In Fig. 13, the capacity and the retainabilty of SUs are studied when two types of channel failures exist, with the failure rates, $\lambda_{F_1}=0.1$ and $\lambda_{F_2}=0.01$ respectively. Here we assume that a *Type 2 failure* occurs 10 times slower than a *Type 1 failure*. In other words, channel failures occur at a rate of 0.11 failure per time unit and that 0.1/0.11= 10/11 is the fraction of all failures being Type 1 (i.e., 10 out of 11). Relevant to this assumption, a *Type 1 failure* could happen due to path loss, multi-path fading or shadowing,

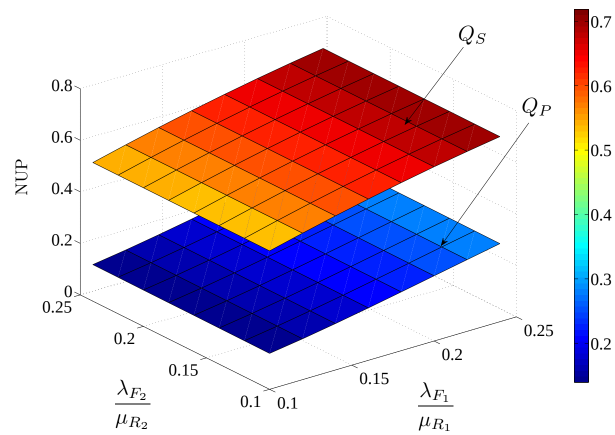

Fig. 14. The NUP as a function of λ_{F_1}/μ_{R_1} and λ_{F_2}/μ_{R_2}. The numerical results are corresponding to mode 1 in Algorithm 1 when $\lambda_P=\lambda_S=2$ and $\mu_P=\mu_S=1$, given that $\mu_{R_1}=0.5$ and $\lambda_{F_1}/\mu_{R_1}=\lambda_{F_2}/\mu_{R_2}$ at each point on the mesh grid. Moreover, the ranges of parameters are $0.05\le\lambda_{F_1}\le0.12$, $0.01\le\lambda_{F_2}\le0.045$ and $0.04\le\mu_{R_2}\le0.45$.

while hardware or power failures, which are unlikely to happen often, could be represented by *Type 2 failures*.

The performance of SUs is shown in Fig. 13 under different configurations of channel repair rates. Consider a reference configuration when $\mu_{R_1}=0.5, \mu_{R_2}=0.1$. As illustrated in this figure, the achieved capacity is increased with a higher repair rate in both working modes. The percentage of increments is approximately 18% and 12% for mode 1 and mode 2 respectively, given that $\lambda_P=12$ when the repair rate of a failed channel due to *Type 1* failure is doubled. If we now double the repair rate of *Type 2* failed channels instead of *Type 1*, the above mentioned percentage increments become 14% and 10% for mode 1 and mode 2 respectively. Therefore, it is observed that the effect of an increased repair rate is more evident for failures with higher failure rates. From a retainability point of view, this observation appears to be valid.

As we observed with the homogeneous channel failures, the retainability of SU services shows higher values in mode 1. Furthermore, the figure depicts that the retainability of services can be improved with higher channel repair rates. When μ_{R_1} or μ_{R_2} is raised, the frequency of recovering failed channels becomes higher and consequently there exists a comparatively larger number of available channels in the network. Then the interrupted ongoing SU services can find channel access opportunities more easily, leading to higher retainability.

2) Network unserviceable probability: In order to investigate the effect of heterogeneous channel failures on the NUP performance of CRNs, we illustrate the numerical results for the NUP for both PU and SU services in Fig. 14. As expected, the NUP of the PN is always lower than that of the SN, owing to the channel access priority of the PN. We observe also that a higher channel failure rate leads to a raise of the NUP for both services. However the rate of the increase of the NUP highly depends on the failure type characteristics. The two failure types considered in Fig. 14 have different failure and repair rates. Those values are configured to satisfy, $\alpha=\lambda_{F_1}/\mu_{R_1}=\lambda_{F_2}/\mu_{R_2}$ where α denotes the ratio between the failure rate to the repair rate of a channel. Note that, at

Algorithm 2: NUP-based algorithm for identifying an optimal R_{max}.

Input: M : Total number of channels in the CRN
Input: $Q_S|_{R_{max}=k}, k = 0, 1, 2, \cdots$: NUP of SU services when $R_{max} = k$ is configured in Algorithm 1.
Input: A : Parameter which constraints the amount of spectrum to be reserved where $A > 1$.
Input: $(Q_S)_{max}$: Maximum tolerable NUP for SUs
Output: R^{OPT}_{max} : The optimal value of R_{max}

Calculate $R_a = \lfloor \frac{M}{A} \rfloor$
if $(Q_S)_{max} \le Q_S|_{R_{max}=0}$ **then**
  $R^{OPT}_{max} = 0$
else if $Q_S|_{R_{max}=0} < (Q_S)_{max} \le Q_S|_{R_{max}=1}$ **then**
  $R^{OPT}_{max} = 1$
else if $Q_S|_{R_{max}=1} < (Q_S)_{max} \le Q_S|_{R_{max}=2}$ **then**
  $R^{OPT}_{max} = 2$
$\cdots$
else if $Q_S|_{R_{max}=k-1} < (Q_S)_{max} \le Q_S|_{R_{max}=k}$ **then**
  $R^{OPT}_{max} = k$
$\cdots$
else if $Q_S|_{R_{max}=R_a-1} < (Q_S)_{max} \le Q_S|_{R_{max}=R_a}$ **then**
  $R^{OPT}_{max} = R_a$
else
  $R^{OPT}_{max} = R_a$
end

each point on the mesh grid, $\lambda_{F_1} > \lambda_{F_2}$ and $\mu_{R_1} > \mu_{R_2}$. In generally, a network that has a higher channel repair rate would be available for users for a longer period than a network with a lower channel repair rate.

Moreover, the advantage of higher repair rates would be dominated by higher failure rates, as can be observed in Fig. 14. Interestingly, although the repair rate of *Type 1* failures is higher than that of *Type 2* failures, the rate of the increase of NUP is more significant in *Type 1* failures when the ratio, α, is increased. This fact can be clearly observed from the gradient of the NUP mesh plot with respect to each axis representing the ratio, α. The reason is explained as follows. The rate of occurring *Type 1* failures is twice (or more) as higher as that of the second type. Therefore, the time between failures tends to be shorter in *Type 1* failures. That is, *the network with Type 1 failures cannot be serviceable for a longer period of time even though the failed channels are recovered shortly*. Accordingly, we conclude that, the failure type with a higher repair rate does not necessarily provide better performance in terms of the NUP since the failure rates would also influence the duration of channel available period.

VII. Identifying the Optimal R_{max} for DCR

Identifying an optimized upper bound for the number of reserved channels in the DCR algorithm is also important since there is a tradeoff between channel availability and service retainability as pointed out earlier. In this section, we propose an NUP-based algorithm to identify the optimal R_{max} which guarantees a bounded NUP for the network.

Algorithm 2 is proposed to obtain the optimal upper bound for the number of reserved channels, denoted as R^{OPT}_{max} in the R-CRN for a given NUP constraint. Note that, the algorithm can be extended to calculate R^{OPT}_{max} also for any other performance metric mentioned in Section IV. The NUP constraint in this algorithm means that the system should guarantee the NUP of SU services always being below a certain threshold level given by $(Q_S)_{max}$ while *achieving the highest possible retainability for SU services*. In this algorithm, a CRN consisting of M channels is considered where no more than $\lfloor \frac{M}{A} \rfloor$ channels could be reserved. Here $A > 1$ is a parameter which constraints the amount of spectrum to be reserved. Herein, we determine R^{OPT}_{max} while investigating the impact of λ_P on the optimal solution. This optimization problem may also be formulated as

$$\begin{aligned} &\text{maximize}_{R_{max}=0,1,\cdots,\lfloor \frac{M}{A} \rfloor} && \theta_S \\ &\text{subject to:} && Q_S \le (Q_S)_{max}. \end{aligned} \tag{19}$$

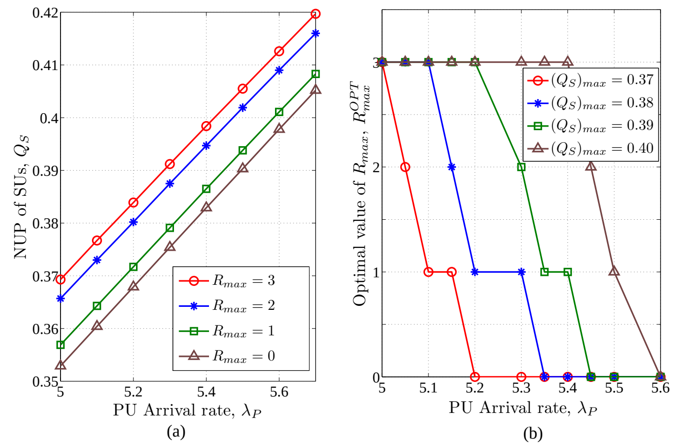

Fig. 15. The optimal value of R_{max} that can be allocated in Algorithm 1 when mode 1 is selected in order to have a guaranteed NUP for the SN given that $M = 8$ and $A = 2.5$. The results shown in sub-figure (b) are obtained based on the NUP performance shown in sub-figure (a).

Configure respectively the maximum tolerable NUP for SUs as $(Q_S)_{max} = 0.40$, $\lambda_{F_1} = 0.01, \lambda_{F_2} = 0.005$, $\mu_{R_1} = 0.5, \mu_{R_2} = 0.1$, $M = 8$ and $A = 2.5$. As observed in Fig. 15(a), Q_{SU} can be maintained below 40% even when $R_{max} = 3$ if $\lambda_P \le 5.4$. Therefore, in order to fulfill the requirement of maintaining a higher retainability level, R^{OPT}_{max} should be 3 when $\lambda_P \le 5.4$. However, if the PU arrival rate appears in the range of $5.4 < \lambda_P < 5.9$, the system should not allocate $R_{max} = 3$ since SUs experience an NUP greater than 40%. Instead $R_{max} = 2$ can guarantee the required NUP in that range. In general, when λ_P becomes higher, the NUP of SU services rises owing to the access priority of PUs. To keep the NUP below the threshold level, the network has to reduce R_{max}. Similarly, through the analysis of the curves obtained from Fig. 15(a), the value of R^{OPT}_{max} which satisfies the given conditions can be obtained for a given range of λ_P. The obtained numerical results are illustrated in Fig. 15(b) for different values of $(Q_S)_{max}$.

VIII. DCR with Higher PN Access Privilege

In this section, we further develop two additional variations of the proposed channel access scheme which give higher access privilege to PUs in the considered CRN. These two variations are slightly modified from the DSA scheme (variation I) presented earlier, and they are referred to as PU's partial priority (PPP) and PU's full priority (PFP) respectively.

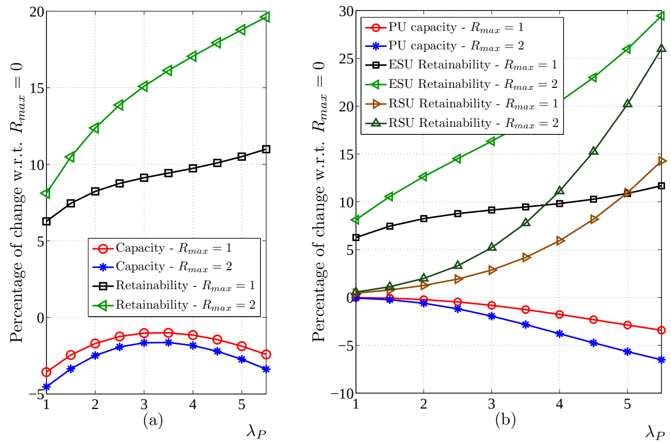

Fig. 16. Percentage of the capacity and retainability increment or decrement w.r.t. no reservation as λ_P varies. The CRN is configured as $M = 6, \lambda_P = 5, \lambda_{ES} = \lambda_{RS} = 2.5, \lambda_F = 0.05, \mu_P = 2, \mu_{ES} = \mu_{RS} = 1, \mu_R = 1$.

A. PPP and PFP

Unlike the initial DSA scheme, both PPP and PFP grant a newly arriving PU access privilege also in the R-CRN. The main features of these two new flavors of the access scheme are listed as follows, while the other principles are the same as in PU priority variation I of the DSA scheme presented in Section III:

- In both PPP and PFP, a newly arriving PU is allowed to access channels in the R-CRN and an ongoing PU service has higher priority than all other SU services if a channel failure happens;
- In PPP, an ongoing SU service in the R-CRN will *not* be forced to terminate upon new PU arrivals;
- In PFP, however, a PU arrival can preempt ongoing SU services from both the N-CRN and the R-CRN.

B. PPP and PFP Performance under Hybrid Traffic Types

Consider now two types of SU flows, i.e., elastic traffic and real-time SU traffic, denoted by *ESU* and *RSU* respectively. Once a channel failure occurs, an already commenced PU services has higher priority than all other SU services in PPP and PFP. With respect to SU services, an already commenced RSU service would have higher priority than an already commenced ESU. Correspondingly, the order of priority defined in Subsection III-B5 is updated as $PL(PU_{R-CRN}) > PL(PU_{N-CRN}) > PL(RSU_{R-CRN}) > PL(RSU_{N-CRN}) > PL(ESU_{R-CRN}) > PL(ESU_{N-CRN})$ for both variations.

As shown in Fig. 16(a), with PFP, the network achieves approximately 20% improvement of the SU service retainability at a cost of 3% SU capacity reduction, given that $R_{max} = 2$ and $\lambda_P = 5.5$. Note that in the PFP variation the capacity of the PN is the same as the one obtained without reservation. Therefore a tradeoff exists between the SU retainability and the SU capacity in this case. On the other hand when the PPP variation is employed, the SN achieves approximately 30% improvement of the service retainability at a cost of 6% PU capacity loss with the same configuration, as illustrated in Fig. 16(b). In this case, the PUs have to sacrifice their capacity in order to maintain a high service completion rate for SUs. From Fig. 16(b), it is also evident that the retainability of ESUs increases faster than that of RSUs' thanks to dynamic channel reservation and spectrum access. In other words, more ESUs may survive even though PUs are active and RSUs have higher priority upon the occurrence of channel failures.

Finally, based on the performance of the DSA scheme with various flavors, an operator may decide which variation to adopt. The initial variation is recommended when the rate of PU arrivals is not high. If a PN demands a higher QoS level at high PU arrival rates, the PFP variation is preferable. On the other hand, the PPP variation may apply to scenarios where both PNs and SNs have flexible QoS requirements.

IX. Conclusions

In this paper we proposed a dynamic spectrum access scheme with channel reservation which enhances the SU performance in error-prone CRNs, in terms of service retainability and network unserviceable probability. The proposed DSA scheme enables flexible channel sharing among PUs and SUs in the context of licensed shared access. An embedded dynamic channel reservation algorithm associated with the DSA scheme can dynamically configure the number of reserved channels depending on the channel occupancy and channel failure status. Through CTMC modeling and simulations, we demonstrate the benefits brought by the DSA scheme and study the tradeoffs among various parameters at diverse traffic loads and under homogeneous and heterogeneous channel failures. Although channel reservation may not be needed with low PU activities and at very low channel failure rates, a DSA scheme with an optimally allocated number of reserved channels is recommended for QoS provisioning. Furthermore, an operator may select an appropriate access privilege level for their PUs when employing such a DSA scheme which facilitates multiple PU access priorities.

References

[1] I. A. M. Balapuwaduge, F. Y. Li, and V. Pla, "Significance of channel failures on network performance in CRNs with reserved spectrum," in *Proc. IEEE ICC*, May 2016, pp. 1-6.

[2] Y. Xiao, Z. Han, C. Yuen, and L. A. DaSilva, "Carrier aggregation between operators in next generation cellular networks: A stable roommate market," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 633-650, Jan. 2016.

[3] A. Kumar, A. Sengupta, R. Tandon, and T. C. Clancy, "Dynamic resource allocation for cooperative spectrum sharing in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5232-5245, Nov. 2015.

[4] H. Shariatmadari, S. Iraji, and R. Jäntti, "Analysis of transmission methods for ultra-reliable communications," in *Proc. IEEE PIMRC*, Aug./Sep. 2015, pp. 2303-2308.

[5] A. Azarfar, J.-F. Frigon, and B. Sansò, "Improving the reliability of wireless networks using cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 338-354, Second Quarter, 2012.

[6] DIGITALEUROPE, "Position paper on licensed shared access (LSA) common understanding, status and next steps," Feb. 2013. [Online]. Available: http://www.digitaleurope.org/.

[7] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moessner, "Licensed spectrum sharing schemes for mobile operators: A survey and outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2591-2623, Fourth Quarter, 2016.

[8] A. Nicita and M. A. Rossi "Spectrum crunch vs. spectrum sharing: Exploring the 'authorized shared access' model," *Communications & Strategies*, pp. 17-40, no. 90, 2nd Quarter 2013. [Online]. Available: https://ssrn.com/abstract=2421169.

[9] K. Y. Wong, Y. Keung, C. Li, and B. Li, "Dynamic call admission control with connection level service guarantee in multi-service wireless cellular networks," in *Proc. IEEE ICC*, Jun. 2004, pp. 4248-4252.

[10] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14-76, Jan. 2015.

[11] S. Ni, Y. Zhuang, Z. Cao, and X. Kong, "Modeling dependability features for real-time embedded systems," *IEEE Trans. Dependable Secure Comput.*, vol. 12, no. 2, pp. 190-203, Mar./Apr. 2015.

[12] A. Banchs, D. Qiao, and D. Westhoff, "Dependability issues with ubiquitous wireless access [Guest Editorial]," *IEEE Wireless Commun. Mag.*, vol. 15, no. 5. pp. 8-9, Oct. 2008.

[13] G. Rubino and B. Sericola, *Markov Chains and Dependability Theory*, Cambridge, UK.: Cambridge Univ. Press, 2014.

[14] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26-35, May 2014.

[15] 3GPP, "Key performance indicators (KPI) for evolved universal terrestrial radio access network (E-UTRAN): Definitions," *3GPP TS 32.450*, pp. 6-10, Jun. 2010. [Online]. Available: http://www.3gpp.org/DynaReport/32450.htm.

[16] A. G.-Andrades, P. Muñoz, E. J. Khatib, I. d.-l.-Bandera, I. Serrano, and R. Barco, "Methodology for the design and evaluation of self-healing LTE networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6468-6486, Aug. 2016.

[17] Z. He, S. Mao, and S. Kompella, "A decomposition approach to quality-driven multiuser video streaming in cellular cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 728-739, Jan. 2015.

[18] L. Sun, W. Wang, and Z. Lu, "On topology and resilience of large-scale cognitive radio networks under generic failures," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3390-3401, Jun. 2015.

[19] T. Taleb, A. Ksentini, and B. Sericola, "On service resilience in cloud-native 5G mobile systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 483-496, Mar. 2016.

[20] D. Hamza and S. Aïssa, "Enhanced primary and secondary performance through cognitive relaying and leveraging primary feedback," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2236-2247, Jun. 2014.

[21] C. B. R.-Estrello, G. H.-Valdez, and F. A. C.-Pérez, "System-level analysis of mobile cellular networks considering link unreliability," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 926-940, Feb. 2009.

[22] J. Nair, M. Andreasson, L. L. Andrew, S. H. Low, and J. C. Doyle, "On channel failures, file fragmentation policies, and heavy-tailed completion times," *IEEE/ACM Trans. Networking*, vol. 24, no. 1, pp. 529-541, Feb. 2016.

[23] T. Chakraborty and I. S. Misra, "Design and analysis of channel reservation scheme in cognitive radio networks," *Computers & Electrical Engineering*, vol. 42, pp. 148-167, Elsevier, Dec. 2014.

[24] S. Deng, B. Wang, Y. Zheng, and K. K. Chai, "Dynamic channel reservation based on forecast in cognitive radio," in *Proc. IEEE VTC Spring*, Jun. 2013, pp. 1-6.

[25] Telecommunication Standardization Sector of ITU, "Quality of Service and Dependability Vocabulary," *Recommendation ITU-T E.800*, 2007. [Online]. Available: https://www.itu.int/rec/T-REC-E.800-198811-S/en.

[26] F. Beltran, S. K. 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 J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2829-2837, Nov. 2016.

[27] FCC, "Report to the president: Realizing the full potential of government- held spectrum to spur economic growth," President's Council of Advisors on Science and Technology (PCAST), Jul. 2012.

[28] M. M. Sohul, M. Yao, T. Yang, and J. H. Reed, "Spectrum access system for the citizen broadband radio service," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 18-25, Jul. 2015.

[29] I. A. M. Balapuwaduge, F. Y. Li, A. Rajanna, and M. Kaveh, "Channel occupancy-based dynamic spectrum leasing in multi-channel CRNs: Strategies and performance evaluation," *IEEE Trans. Commun.*, vol. 64, no.3, pp. 1313-1328, Mar. 2016.

[30] J. J. Alcaraz, M. López-Martínez, and J. Vales-Alonso, "Bandwidth reservation as a coexistence strategy in opportunistic spectrum access environments," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 478-488, Mar. 2014.

[31] M. M. A. Ananada, "Estimation and testing of availability of a parallel system with exponential failure and repair times," *J. of Statistical Planning and Inference*, vol. 77, no. 2, pp. 237-246, Mar. 1999.

[32] A. Amich, M. A. Imran, R. Tafazolli, and P. Cheraghi, "Accurate and efficient algorithms for cognitive radio modeling applications under the i.n.i.d. paradigm," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1750-1765, May 2015.

[33] E. Bauer, R. Adams, and D. Eustace, *Beyond Redundancy: How Geographic Redundancy can Improve Service Availability and Reliability of Computer-based Systems*, Hoboken, NJ, USA: Wiley, 2011.

[34] L. Jiao, I. A. M. Balapuwaduge, F. Y. Li, and V. Pla, "On the performance of channel assembling and fragmentation in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5661-5675, Oct. 2014.

[35] L. Lee, "The unserviceable probability of a class of telecommunications networks," *IEEE Trans. Rel.*, vol. R-20, no. 3, pp. 132-135, Aug. 1971.

Indika A. M. Balapuwaduge (S'12-M'17) received the B.Sc. Engineering (First Class Honors) degree from University of Ruhuna, Sri Lanka, in 2008, and the M.Sc. and the Ph.D. degrees in Information and Communication Technology (ICT) from University of Agder (UiA), Norway in 2012 and 2016, respectively. His Master thesis was awarded as the Best Master's thesis in ICT at UiA in 2012. Currently he is a Post-Doctoral Research Fellow at the Department of ICT, UiA. He spent one year as an engineer at Huawei technologies, Sri Lanka, from Oct. 2008 to Aug. 2009 and he worked as a lecturer at the Department of Electrical and Information Engineering, University of Ruhuna from Aug. 2009 to Aug. 2010. Dr. Indika's current research interest covers various areas of mobile and wireless communications, including ultra-reliable communication, spectrum management, modeling and performance evaluation of communication networks, in particular, cognitive radio networks.

Frank Y. Li (S'99-M'03-SM'09) received the Ph.D. degree from the Department of Telematics, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. He worked as a Senior Researcher at UniK-University Graduate Center (now Department of Technology Systems), University of Oslo, Norway before joining the Department of Information and Communication Technology, University of Agder (UiA), Agder, Norway, in August 2007 as an Associate Professor and then a Full Professor. During the past few years, he has been an active participant in several Norwegian and EU research projects. He is listed as a Lead Scientist by the European Commission DG RTD Unit A.03-Evaluation and Monitoring of Programmes in Nov. 2007. Dr. Li's research interests include MAC mechanisms and routing protocols in 5G mobile systems and wireless networks, mesh and ad hoc networks; wireless sensor network; D2D communication; cooperative communication; cognitive radio networks; green wireless communications; Internet of Things; reliability in wireless networks, QoS, resource management and traffic engineering in wired and wireless IP-based networks; analysis, simulation and performance evaluation of communication protocols and networks.

Vicent Pla received the M.E. and Ph.D. degrees in Telecommunication Engineering from the Universitat Politècnica de València (UPV), Spain, and the B.Sc. in Mathematics from the Universidad Nacional de Educación a Distancia (UNED), Spain. In December 1999, he joined the Department of Communications at the UPV, where he is currently an Associate Professor. His research interests lie primarily in the area of modeling and performance analysis of communication networks. During the past few years, most of his research activity has focused on resource management in wireless networks. In these areas he has published numerous papers in refereed journals and conference proceedings, and has been an active participant in several national and European research projects.