Competitive Spectrum Sharing in Cognitive Radio Networks: a queuing theory based analysis

Ilario Filippini, Matteo Cesana
Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
Milano, Italy
Ilario.Filippini@polimi.it, Matteo.Cesana@polimi.it

Ilaria Malanchini
Alcatel-Lucent Bell Labs
Stuttgart, Germany
Ilaria.Malanchini@alcatel-lucent.com

***Abstract*—We consider cognitive radio networks where primary licensed users coexist with unlicensed users, which are allowed to opportunistically access licensed spectrum upon performing spectrum sensing functionality. An analytical framework based on queuing theory and game theory is introduced to assess the performance of the spectrum sharing process. Queuing theory is used to model the achievable throughput of secondary users, whereas game theoretic tools are introduced to capture the competitive dynamics of the spectrum sharing process among multiple secondary users. The proposed framework is used to derive performance measures of spectrum sharing when spectrum access is/is not regulated by a central spectrum management authority.**

***Index Terms*—spectrum sharing, game theory, queuing theory**

I. Introduction

Fixed spectrum assignment policies have been preponderant in the recent past in the management and regulation of wireless systems, where licensed entities are granted the rights for the use of frequency bands on a long term basis over relatively-vast geographical regions. Such assignment policies have shown notable shortcomings in the efficient use of the spectrum with the creation of spectrum portions scarcely utilized by licensed users [1].

Cognitive Radio Networks (CRNs) are commonly recognized as viable architectural solutions to solve the problems of limited spectrum availability and scarce efficiency in the spectrum usage, by letting unlicensed users to opportunistically access under-utilized licensed bandwidth.

CRNs are technically enabled by Cognitive Radio (CR) transceivers which have the capability of dynamically adapting their transmitter parameters (operating spectrum, modulation, transmission power) based on the current "context", thus being able to search for spectrum "holes" in licensed channels to be used for opportunistic transmissions. As an example, recent FCC guidelines allow CRs to opportunistically operate in TV bands white spaces [6].

The most general scenario of CRNs thus features two types of users: licensed users, often referred to as Primary Users (PUs), are priority users in the licensed bands, opportunistic users, or secondary Users (SUs), may access the licensed spectrum in non-intrusive manner.

One of the core functionalities of SUs is *spectrum sensing* to identify available spectrum portions. Spectrum sensing can be distributed and delocalized, with the SUs probing the spectrum searching for spectrum holes, centralized, through the adoption of third-party geo–localized spectrum databases that contain the PUs expected/current channel use, or hybrid. Upon spectrum sensing, SUs need take decision on which channel (spectrum portion) to opportunistically use among the available ones, thus giving raise to a *spectrum sharing* problem among SUs. In this problem, the straightforward choice of accessing a good quality spectrum is undermined by the congestion SUs may experience due to the prevalence of the choice.

This work introduces an analytical framework for assessing the performance of different spectrum sharing scenarios. The framework builds on a queuing model to derive the average transmissions delay experienced by secondary transmissions over licensed channels, which is then used to evaluate the outcome of spectrum access/sharing strategies in regulated and partially unregulated scenarios. In the regulated scenario, a spectrum broker (spectrum database) has full knowledge of the spectrum "context" (channel occupation, load and bandwidth) and can thus orchestrate the spectrum sharing among the secondary users to optimize the average quality perceived by the secondary transmissions (without harming the primary ones); in the partially-unregulated scenario, the spectrum sharing/access is a completely distributed process and the secondary users compete among themselves for the available spectrum each one optimizing the experienced quality, with partial information on the spectrum status. In the former case, the spectrum sharing is formalized as an optimization problem; in the latter case, a non–cooperative game formulation is used to capture the dynamics among competitive users. Numerical results are derived to compare the perceived quality by secondary users under the two spectrum sharing scenarios.

The paper is organized as follows. In Section II we discuss related work and in Section III we introduce the reference models for the regulated and partially-unregulated spectrum sharing scenarios. The quality of the spectrum sharing is numerically evaluated in Section IV. Section V concludes the manuscript.

II. Related Work

The performance evaluation of spectrum sharing has recently attracted much attention in the research arena. The most common quantitative "tools" used to evaluate the spectrum sharing efficiency include queuing theory, mathematical programming, game theory and decision theory.

The work in this paper is along the same lines of the body of literature which couples game theory with queuing theory with some notable differences and novelties both in the reference scenario and in the modeling/solution approach. As for the reference scenario, most of the contributions in the field consider the case where secondary user can balk access to primary channels, whereas in this paper a secondary transmission does not have this opportunity. Modeling–wise, the work in the field mostly resort to atomic games to represent the competition among secondary users, whereas we consider the case where the population of secondary users is large and thus the competition is better represented by a non–atomic game formulation.

In [2], closed form expressions are derived for the maximum throughput and the average transmission delay of secondary transmissions by modeling the primary/secondary transmissions system as a M/G/1 queue with priority customers (primary users) under preemption repeat policies for the secondary users. Li and Han [3] extend the previous work by addressing the case where competition and collisions may happen among secondary transmissions, by considering a scenario where two queues of secondary users "compete" for accessing a single server (licensed) channel. Habachi and Hayel [4] resort to similar queuing theoretic models to assess the delay performance of secondary transmissions with the goal to design optimal sensing strategies for secondary users. M/G/1 queues are used also in [5] to assess the delay performance of spectrum handover operations upon activation of primary users. Chen and Liu [6] optimize the spectrum access policy of a secondary user to minimize the average transmission delay while imposing constraints on the collision probability onto primary transmissions.

Along the same lines, Pla et al. [7] consider a scenario with multiple available primary channels with heterogeneous data rates where incoming secondary users are assigned to the unoccupied primary channel which has the highest nominal rate (available spectrum bandwidth); preemption repeat discipline is applied if secondary transmissions are interrupted by primary ones. A similar multi–channel scenario is considered in [8], which resorts to M/G/1 queues to characterize the average delay experienced by secondary transmissions on specific licensed channels; an optimization framework is then introduced to minimize the average transmission delay of secondary transmissions over all the available licensed channels. Multi–channel scenarios are addressed also in [9]–[11].

Besides queuing theoretic tools to assess the performance of secondary transmission in presence of primary users, recent work has also focused on coupling performance measures obtained through queuing theory together with decision process analysis. As a matter of fact, the problem of choosing the best available licensed channel for opportunistically use can also be formalized in a decision making framework [12]. [13], [14] introduce a decision making scenario where secondary users accessing a primary channel may decide to actually access the channel or to balk based on the knowledge of the expected delay. A queue model with interruption is considered to derive the average delay. Similar scenarios are considered in [15] and [16] with slight differences in the specific delay model to characterize the queuing delay. Along the same lines, Jagannathan *et al.* propose in [17] and later in [18] a decision making framework for secondary transmissions to choose when to use dedicated spectrum and when to resort to primary spectrum holes. The trade off being between costly but soon available dedicated bandwidth and less expensive but less "reliable" primary user channels. The proposed framework leverages queuing–theoretic tools to assess the expected delay when using the spectrum holes. The competitive case where secondary users selfishly compete for the available channels is modeled through a non–cooperative atomic game whose Nash equilibria are numerically derived and discussed.

It is worth to point out, that a decision making scenario similar to ours is presented in [19] in the field of wireless access networks. The authors consider the case where a flow of users has to choose which access point to get connectivity from. The problem is formalized as a non cooperative non–atomic game whose equilibria are evaluated in the case where two available access points can be used for connectivity.

III. Reference Scenario

We consider a network scenario (see Figure 1) where primary and secondary users potentially share a common set $\mathcal{N}$ of available channels, being $|\mathcal{N}| = N$ the number of channels, and $\mathcal{I} = [1, 2, \ldots, N]$ the set of indices of the available channels. Each channel $i \in \mathcal{I}$ is characterized by a given bandwidth and is primarily used by one flow of licensed user which is characterized by a Poisson distributed traffic process with intensity λ_i^p. Each primary user is further characterized by transmissions with average duration of $1/\mu_i^p$. The secondary transmissions over a given channel i are similarly characterized by Poisson traffic with intensity λ_i^s and an average transmission duration time $1/\mu_i^s$. The superscript *s* is removed from the secondary traffic and service time for the sake of readability throughput the manuscript. The total incoming traffic from secondary users is defined as $\lambda_{tot} = \sum_{i \in \mathcal{I}} \lambda_i$.

A secondary transmission which is "active" on a given channel has to back off as soon as the corresponding primary transmission kicks in and repeat the whole interrupted transmission as the channel frees up. We assume ideal collision detection, which means that secondary ongoing transmissions leave the channel as soon as a primary one is detected without causing any interference to the primary incoming transmission.

We consider as performance measure the transmission delay, which is defined as the time that it takes for a secondary transmission to go through a channel. From the perspective of

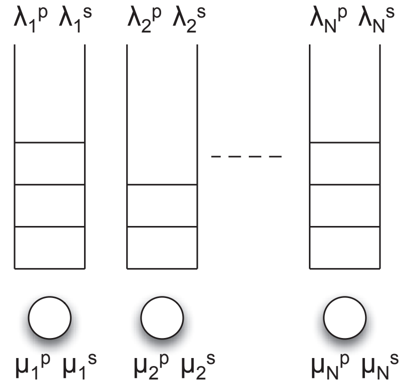

Fig. 1. Reference scenario with N primary channels.

secondary transmissions, the reference channel can be modeled as a queue with preemption by primary users. We leverage the results in [2] which models each one of the channels as a M/G/1 queue where secondary users can be preempted by primary transmissions. Under this scenario, the average secondary transmission delay, $d_i(\lambda_i)$, defined as the time that it takes for a secondary transmission to be served after entering channel i, can be expressed using the well-known Pollaczek-Khintchine result [20]:

$$d_i(\lambda_i) = \frac{\frac{\lambda_i}{\mu_i} E[Z_i^s]}{1 - \frac{\lambda_i}{\mu_i}} + E[C_i^s] \tag{1}$$

where $E[C_i^s]$ is the extended service time considering PU interruptions and $E[Z_i^s]$ is the residual extended service time seen by secondary packet entering channel i. Closed form expressions for the average transmission delay of secondary transmissions are derived in [2] under different statistical forms for primary/secondary users service time. Note that Eq. (1) is a function of the incoming primary/secondary traffic as well as of the primary/secondary service time.

In this work, we are interested in studying the secondary user performance under different channel sharing scenarios. Namely, we consider two distinct scenarios: in the first one, a spectrum broker orchestrates the access to the licensed resources by optimally assigning secondary traffic to the available channels in order to minimize the average transmission delays of secondary users; in the second scenario, secondary users compete for the available channels by selfishly minimizing their perceived transmission delay.

IV. Numerical Analysis

In this section, we introduce two different mathematical models in order to reflect the different sharing scenarios described in the previous section. Namely, when a spectrum broker is available, the secondary incoming traffic can be optimally subdivided among the available channels. In this case, the problem can be cast as a mathematical formulation that minimizes the transmission delay for secondary transmissions in order to find the optimal secondary traffic assignment $\boldsymbol{\lambda_{opt}} = [\lambda_1, \lambda_2, \ldots, \lambda_N]$. Formally, we can write:

$$\begin{aligned} \text{minimize} \quad & S(\boldsymbol{\lambda}) = \sum_{i=1}^{N} \lambda_i d_i(\lambda_i) \\ \text{s. t.} \quad & \sum_{i=1}^{N} \lambda_i = \lambda_{tot}, \\ & \lambda_i \geq 0 \; i \in \mathcal{I}, \end{aligned} \tag{2}$$

where $S(\boldsymbol{\lambda})$ is the social welfare.

In contrast, when a spectrum broker is not available, the secondary users tend to behave selfishly choosing the "best" channel to access to. Therefore, we model this process as a non-cooperative game. Furthermore, we assume that the number of secondary users is large enough, so that each users demand is infinitesimal with respect to the overall demand (non-atomic game). In this game setting, we leverage the Wardrop equilibrium to represent a stable outcome of the traffic repartition among primary channels.

In our case, a Wardrop equilibrium can be defined as a traffic repartition at which all the used channels (that is, with non–null incoming secondary traffic) feature a transmission delay which is equal or less than the transmission delay of any other used channel. More formally, a traffic repartition among primary channels $\boldsymbol{\lambda_w} = [\lambda_1, \lambda_2, \ldots, \lambda_N]$ is a Wardrop equilibrium if the following holds:

$$\lambda_k > 0 \quad \text{iff} \quad d_k(\lambda_k) \leq d_i(\lambda_i) \quad \forall i, k \in \mathcal{I}, \quad i \neq k \tag{3}$$

As the delay function in (1) is continuous and non-decreasing in the flow variables, the corresponding non–cooperative game admits a unique Wardrop equilibrium [21]. To practically derive the equilibrium, we observe that, by definition, the average delay at the equilibrium is the same for all the primary channels used by secondary users. That is, a non-negative flow repartition (such that $\lambda_i \geq 0$, $\forall i \in \mathcal{I}$) is a Wardrop equilibrium if it satisfies the following system of equalities:

$$\begin{cases} d_i(\lambda_i) = d_k(\lambda_k) & \forall i, k \in \mathcal{I} : \lambda_i > 0, \lambda_k > 0 \\ \sum_{i \in \mathcal{I}} \lambda_i = \lambda_{tot} \end{cases} \tag{4}$$

Algorithm 1 can consequently be used to calculate the Wardrop equilibrium of a given network scenario. In a nutshell, the algorithm proceeds by iteratively solving (4) and eliminating at each iteration the primary channels (and the corresponding equation in (4)) with negative assigned secondary flows. The algorithm stops when the systems provides at some iteration a non-negative flow repartition. The Wardrop equilibrium is finally obtained by setting to 0 the flows for primary channels eliminated during the iterations and adding them up to the solution of the system. In a real system, this reflects the fact that channels that are overused by primary users will not be used by secondary users.

It is easy to show that the following holds:

Proposition 1. $\boldsymbol{\lambda_w}$ *obtained through Algorithm (1) is a Wardrop equilibrium.*

Algorithm 1 Wardrop equilibrium calculation

procedure WARDROP($\mathcal{C}, \boldsymbol{\lambda_p}, \boldsymbol{\mu_p}, \boldsymbol{\mu}$)
  $\boldsymbol{\lambda}^* \leftarrow Solve$ (4)
  while $\boldsymbol{\lambda}^* \not\geq \mathbf{0}$ **do** $\triangleright$ Some variable is negative
    Select $i \in \mathcal{I} | i = \arg\min \lambda_i^*$
    $\lambda_{WE}(i) = 0$
    $\mathcal{I} \leftarrow \mathcal{I} \setminus \{i\}$ $\triangleright$ Remove index i from set $\mathcal{I}$
    $S \leftarrow S(\mathcal{I})$ $\triangleright$ Set up (4) without channel i
    $\boldsymbol{\lambda}^* \leftarrow Solve$ (4)
  end while
  $\lambda_{WE}(i) \leftarrow \lambda^*(f(i)) \quad \forall i \in \mathcal{I}$
    $\triangleright$ Equilibrium found, and $f := \mathcal{I} \rightarrow \{1, 2, \ldots, |\mathcal{I}|\}$
  return $\boldsymbol{\lambda_{WE}}$
end procedure

Proof: The proof ends up by observing that the $\boldsymbol{\lambda_{WE}}$ derived through Algorithm (1) verifies the definition of Wardrop equilibrium given in (3). ■

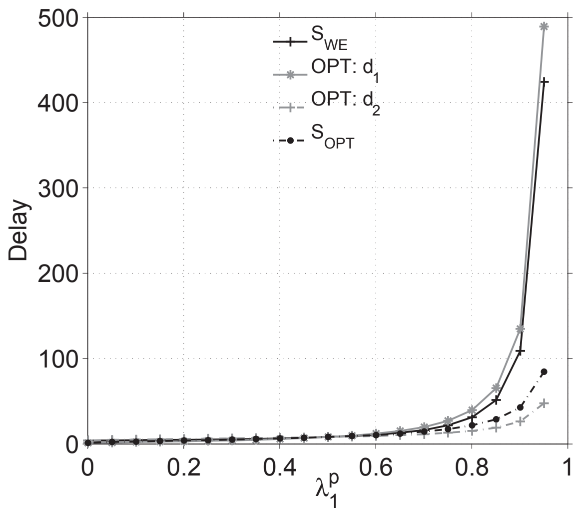

Fig. 2. Transmission delays (normalized to μ_p) in case of two available primary channels ($N = 2$) and different combination of primary traffic λ_p^1, λ_p^2: comparison between the optimal case (spectrum broker) and the distributed one.

In the following, we provide a set of numerical results, if not differently specified we assume $\lambda^p = 0.5$ and $\mu = \mu^p = 1$. Figure 2 compares the transmission delay for secondary users in case of optimal flow repartition (spectrum broker) and the transmission delay at the Wardrop equilibrium, when $N = 2$ and for different combinations of the primary user traffic $(\lambda_1^p, \lambda_2^p)$. The figure reports the average delay at the optimum as well as the average delay experienced in each one of the two channels. As expected, the delay curves increase as the primary traffic intensity increases. The average delay perceived in the two channels is different at the optimum flow repartition, whereas the two channels have the same delay at the Wardrop equilibrium (by definition).

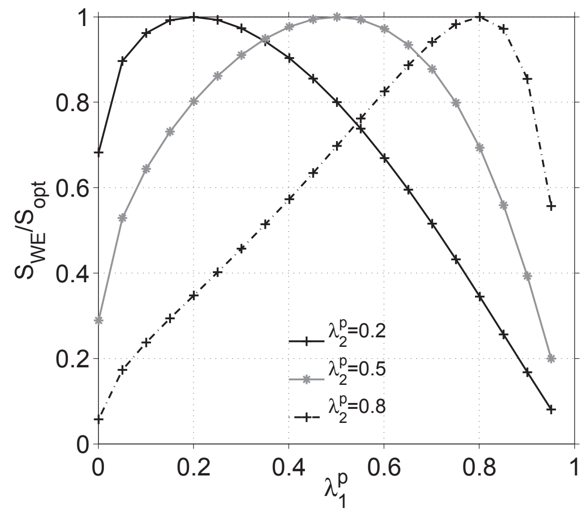

Fig. 3. Ratio between the optimal social welfare and the one at the Wardrop equilibrium in case of two available primary channels ($N = 2$) and different combination of primary traffic λ_1^p, λ_2^p.

The ratio between the social welfare at the Wardrop equilibrium ($S_{WE} = S(\boldsymbol{\lambda_{WE}})$) and the social welfare at the optimum ($S_{OPT} = S(\boldsymbol{\lambda_{OPT}})$) is reported in Figure 3 for the very same settings as 2. As clear from the figure, the flow repartition at the Wardrop equilibrium is optimal if the two primary channels are homogeneous in terms of primary traffic intensity. On the other side, under heterogeneous primary channel traffic, the "quality" of the Wardrop equilibrium considerably deteriorates.

Figure 4 extends the delay analysis when varying the number of available channels N in homogeneous primary traffic conditions. In this case, the Wardrop equilibrium is always optimal, and adding up channels leads to lower average transmission delay for secondary transmissions.

Furthermore, it is worth analyzing heterogeneous traffic scenarios where the primary traffic intensity is different for the different available primary channels. Namely, we have considered two sample cases where one channel out of 10 is slightly less/more congested by "primary" traffic. The two cases, referred to as "1 Best" and "1 Worst", are characterized by homogeneous primary traffic for 9 channels, $\lambda_i^p = 0.5$ $\forall i = 2 \ldots 10$, whereas the 1-st channel has $\lambda_1^p = 0.4$ and $\lambda_1^p = 0.6$ in the "1 Best" and "1 Worst" case respectively. Figure 5 shows the flow repartition in the 10 channels for the two different primary traffic scenarios. At the Wardrop equilibrium, the secondary users tend to concentrate to the best resources, that is, under the "1 Best" setting, most of the secondary traffic chooses the first (and faster) channel, whereas in the "1 Worst" case, the "slowest" primary channels is never used. The ratio between the social welfare at the Wardrop equilibrium and the social welfare at the optimum (S_{WE}/S_{OPT}) is 0.96 and 0.98 in the "1 Best" and "1 Worst" cases respectively.

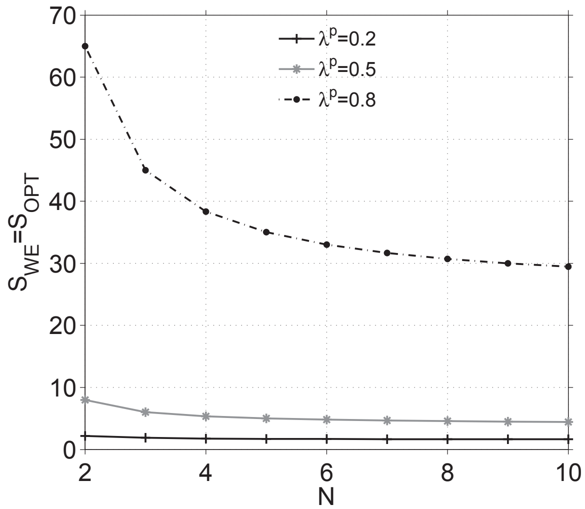

Fig. 4. Average transmission delay (normalized to μ_p) when varying the number of available channels and the primary traffic λ^p.

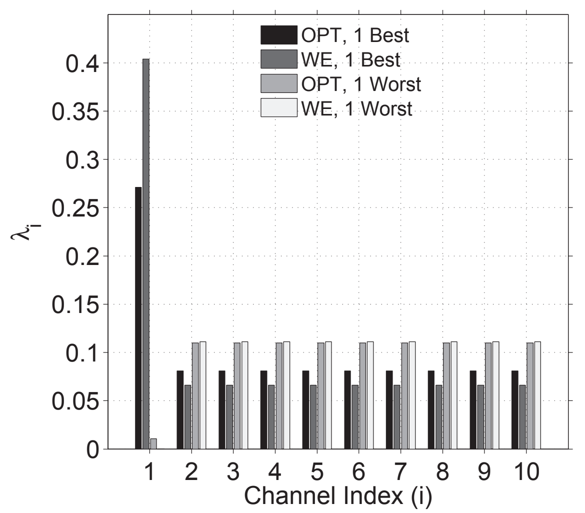

Fig. 5. Optimal and equilibrium subdivision of secondary traffic for two different conditions of primary traffic.

V. Concluding Remarks

A framework has been proposed to evaluate the performance of spectrum sharing policies in Cognitive Radio Networks. The framework leverages queuing theoretic models to evaluate the average transmission delay of secondary transmissions over primary channels with possible preemption by primary licensed users. Two different scenarios have been considered; in the centralized one, a spectrum broker can optimally assign secondary transmissions to primary channels, whereas in the unregulated one secondary users compete for the available primary resources in an uncoordinated way. The unregulated scenario has been modeled as a non-atomic non-cooperative game which has been solved to obtain the corresponding equilibria. The quality of the optimal solution (centralized scenario) and the equilibria of the game have been compared in reference network scenarios.

References

[1] FCC, *ET Docket No 03-322 Notice of Proposed Rule Making and Order*, Dec 2003.

[2] F. Borgonovo, M. Cesana, and L. Fratta, "Throughput and delay bounds for cognitive transmissions," in *Advances in Ad Hoc Networking*, ser. IFIP International Federation for Information Processing, P. Cuenca, C. Guerrero, R. Puigjaner, and B. Serra, Eds. Springer US, 2008, vol. 265, pp. 179–190.

[3] H. Li and Z. Han, "Queuing analysis of dynamic spectrum access subject to interruptions from primary users," in *Cognitive Radio Oriented Wireless Networks Communications (CROWNCOM), 2010 Proceedings of the Fifth International Conference on*, 2010, pp. 1–5.

[4] O. Habachi and Y. Hayel, "Optimal opportunistic sensing in cognitive radio networks," *Communications, IET*, vol. 6, no. 8, pp. 797–804, 2012.

[5] L.-C. Wang, C.-W. Wang, and C.-J. Chang, "Modeling and analysis for spectrum handoffs in cognitive radio networks," *Mobile Computing, IEEE Transactions on*, vol. 11, no. 9, pp. 1499–1513, 2012.

[6] R.-R. Chen and X. Liu, "Delay performance of threshold policies for dynamic spectrum access," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 7, pp. 2283–2293, 2011.

[7] V. Pla, S. De Vuyst, K. De Turck, E. Bernal-Mor, J. Martinez-Bauset, and S. Wittevrongel, "Saturation throughput in a heterogeneous multichannel cognitive radio network," in *Communications (ICC), 2011 IEEE International Conference on*, 2011, pp. 1–5.

[8] X. Li, J. Wang, H. Li, and S. Li, "Delay analysis and optimal access strategy in multichannel dynamic spectrum access system," in *Computing, Networking and Communications (ICNC), 2012 International Conference on*, 2012, pp. 376–380.

[9] P. Smith, A. Firag, P. Dmochowski, and S. M., "Analysis of the m/m/n/n queue with two types of arrival process: Applications to future mobile radio systems," *Journal of Applied Mathematics*, vol. 2012, 2012.

[10] J. Wang, A. Huang, L. Cai, and W. Wang, "On the queue dynamics of multiuser multichannel cognitive radio networks," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 3, pp. 1314–1328, 2013.

[11] Y. Zhao, M. Song, and C. Xin, "Delay analysis for cognitive radio networks supporting heterogeneous traffic," in *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*, 2011, pp. 215–223.

[12] R. Hassin, M. Haviv, and S. Hassin, *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*. Kluver, 2003.

[13] H. Li and Z. Han, "Socially optimal queuing control in cognitive radio networks subject to service interruptions: To queue or not to queue?" *Wireless Communications, IEEE Transactions on*, vol. 10, no. 5, pp. 1656–1666, 2011.

[14] H. Li, "Socially optimal queuing control in cognitive radio systems: Pricing and learning," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, 2010, pp. 1–6.

[15] C. Do, N. Tran, and C. S. Hong, "Optimal queueing control in hybrid overlay/underlay spectrum access in cognitive radio networks," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, 2012, pp. 1–5.

[16] C. Do, N. Tran, M. V. Nguyen, C. seon Hong, and S. Lee, "Social optimization strategy in unobserved queueing systems in cognitive radio networks," *Communications Letters, IEEE*, vol. 16, no. 12, pp. 1944–1947, 2012.

[17] K. Jagannathan, I. Menache, E. Modiano, and G. Zussman, "Non-cooperative spectrum access: The dedicated vs. free spectrum choice," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 11, pp. 2251–2261, 2012.

[18] K. Jagannathan, I. Menache, G. Zussman, and E. Modiano, "Non-cooperative spectrum access: the dedicated vs. free spectrum choice," in *Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '11. New York, NY, USA: ACM, 2011, pp. 10:1–10:11. [Online]. Available: http://doi.acm.org/10.1145/2107502.2107515

[19] N. Kaci, P. Maille, and J. Bonnin, "Performance of wireless heterogeneous networks with always-best-connected users," in *Next Generation Internet Networks, 2009. NGI '09*, 2009, pp. 1–8.

[20] L. Kleinrock, *Queueing Systems: Computer Applications - Vol. 2*. John Wiley & Sons New York, 1976, vol. 82.

[21] I. Milchtaich, "Topological conditions for uniqueness of equilibrium in networks," *Mathematics of Operations Research*, vol. 30, no. 1, pp. 225–244, 2005.