Dynamic Spectrum Access in Open Spectrum Wireless Networks

Yiping Xing, *Student Member, IEEE*, R. Chandramouli, *Member, IEEE*, Stefan Mangold, *Member, IEEE*, and Sai Shankar N

***Abstract*—One of the reasons for the limitation of bandwidth in current generation wireless networks is the spectrum policy of the Federal Communications Commission (FCC). But, with the spectrum policy reform, open spectrum wireless networks, and spectrum agile radios are set to drive next general wireless networks. In this paper, we investigate continuous-time Markov models for dynamic spectrum access in open spectrum wireless networks. Both queueing and no queueing cases are considered. Analytical results are derived based on the Markov models. A random access protocol is proposed that is shown to achieve airtime fairness. A distributed version of this protocol that uses only local information is also proposed based on homo egualis anthropological model. Inequality aversion by the radio systems to achieve fairness is captured by this model. These protocols are then extended to spectrum agile radios. Extensive simulation results are presented to compare the performances of fixed versus agile radios.**

***Index Terms*—Access control, cognitive science, game theory, queueing analysis, spectrum management.**

I. Introduction

THE USAGE OF radio spectrum, and the regulation of radio emissions are coordinated by national regulatory bodies. In the U.S., the main authorities for radio spectrum regulation are the Federal Communications Commission (FCC) for commercial applications, and the National Telecommunications and Information Administration (NTIA) for government use. Historically, FCC and NTIA divided the radio spectrum into many frequency bands, and licenses for the often exclusive usage of these bands are provided to operators, typically for a long time such as one or two decades. Depending on the type of radio service that is then provided by the licensees, frequency bands are often idle in many areas, and inefficiently used. This is not in the interest of the regulatory bodies, because they attempt to achieve high efficiency in the usage of radio resources. The alternative way of radio spectrum regulation is the usage of unlicensed frequency bands, also referred to as open spectrum, that can be used by any radio system, under some restrictions called *spectrum etiquettes*. Open spectrum networking has several advantages. It facilitates mobility and usage efficiency, as a license is not required for this system to operate. It also promotes spectrum sharing (the coexisting of different radio systems in the same spectrum), as one device may transmit, while others in the area are idle. Moreover, radio systems can dynamically use and release spectrum wherever and whenever they are available ("spectrum agile radios"). Dynamic spectrum access by spectral agile radios helps to minimize unused spectral bands ("white spaces" [1]). Open spectrum is expected to be a major driving force behind next-generation wireless networks as more spectrum is opened up for unlicensed use.

Some concerns on the unlicensed spectrum usage are discussed in [2]–[4]. Within unlicensed frequency bands, the radio systems have to coordinate the usage of radio resources among themselves. With this open spectrum approach, there is then of course the problem of how to achieve fair and efficient sharing of radio resources between dissimilar radio systems that cannot communicate with each other. Unlicensed frequency bands are typically used by many dissimilar radio systems to provide a large set of different radio services. However, unlicensed frequency bands may be more efficiently used when the usage of the radio resources is coordinated by means of spectrum etiquette rules. A spectrum etiquette is a set of rules for radio resource management to be followed by radio systems that operate in an unlicensed band. It may help to establish a fair access to the available radio resources, in addition to a more efficient usage of the radio spectrum.

In this paper, we present the following.

- Continuous-time Markov chain modeling of spectrum etiquette for dynamic spectrum access is presented. Systems with and without queueing are investigated.
- A fair, random channel access protocol is derived based on the Markov model.
- A distributed version of the channel access protocol based on an anthropological model that uses only local information is proposed. A homo egualis (HE) society model [5] is used for this purpose.
- The channel access protocol is extended to spectral agile radios.
- Detailed simulation results to evaluate the performance of the proposed protocols.

We also note that our approach based on HE society models the nonrational inequality aversion of decision-making radio systems: This approach bridges between multiple disciplines, social science, and radio communications engineering.

Recently, as more and more communication protocols and commercial wireless devices are being developed to operate in crowded unlicensed spectrum bands, spectrum inefficiency

Manuscript received January 16, 2005; revised June 1, 2005. This work was supported in part by the National Science Foundation (NSF) under Grant NSF CAREER 0133761.

Y. Xing and R. Chandramouli are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: yxing@stevens.edu; mouli@stevens.edu).

S. Mangold is with Swisscom LTD, Innovations, Network Access, CH-3000 Berne, Switzerland (e-mail: stefan.mangold@swisscom.com).

S. S. N is with Qualcomm, Qualcomm Standards Engineering Department, San Diego, CA 92121 USA (e-mail: nsai@qualcomm.com).

Digital Object Identifier 10.1109/JSAC.2005.862415

is becoming a serious problem. Spectral agility is being paid considerable attention for its potential to alleviate the spectrum access inefficiency problem [6]–[8]. If the radio device has the flexibility of switching operating spectral bands, promising improvement of spectral efficiency is expected. Of course, such spectral agility cannot be achieved without developing new hardware/software and changing the current spectrum allocation policies. Fortunately, the advances in software defined radio (SDR) [9], [10] has enabled the development of flexible and powerful radio interfaces for supporting spectral agility. Also, the FCCs ongoing review of the current spectrum regulations is also expediting the adoption of more flexible spectrum allocation policies for spectral agility. We will show how future open spectrum scenarios can be engineered with this approach to improve spectrum efficiency and fairness in spectrum access.

The rest of this paper is organized as follows. In Section II, basic model of the access problem in open spectrum network is presented. Markov models are described in Section III. Then, we formulate and analyze a random access model in Section IV, which provisions desired fairness to open spectrum access. General solutions to *airtime* share and blocking probability are given in Section V. In order to improve the spectrum efficiency, spectral agile radio access scheme is proposed in Section VI. In Section VII, a HE-based practical access scheme is proposed. The simulation results and discussion are given in Section VIII. Finally, the conclusions are drawn in Section IX.

II. Basic Model of the Access Problem in Open Spectrum Network

A. Channel and Traffic Model

We assume a perfect channel. That is, a channel is either busy or idle. It is also assumed that radio systems always detect radio resource allocations of other radio systems, here the radio resource refers to the specified bandwidth.

The offered traffic is modeled with two random processes per radio system. The arrival traffic is modeled as a Poisson random process with rate λ_i for radio system i, so the interarrival time is negative-exponentially distributed with mean time $1/\lambda_i$ ms (1 ms $=$ 1 ms). The radio system access duration is also negative-exponentially distributed with mean time $1/\mu_i$ ms, so the departure of the radio system i is another Poisson random process with rate μ_i. We assume that spectral scanning is performed instantaneously, so there is no scanning delay.

B. Usage Model and Etiquette Definition

The 5 GHz unlicensed frequency band is a candidate for a large set of radio services, and is one of the unlicensed frequency bands that may be efficiently used only with an established spectrum etiquette. We use the same abstract model of an unlicensed frequency band, as in [11], as illustrated in Fig. 1.

Here, two different types of radio systems are assumed to operate in the band, each operating with different frequency channel bandwidths. The radio systems of type A operate on three frequency channels (center frequencies f_2, f_5, f_8), the radio systems of type B operate on nine frequency channels (center frequencies $f_1, \ldots, f_9$). The frequency channels overlap with each other, as indicated in the figure. The number and

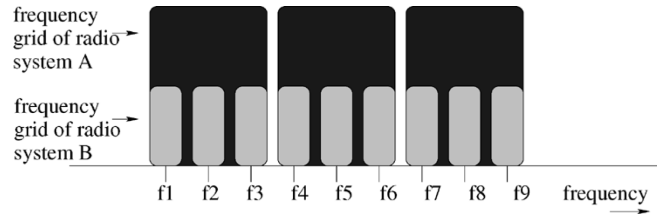


Fig. 1. Frequency channels used by two different types of radio systems (A, B). Each radio system represents a group of communication radio devices.

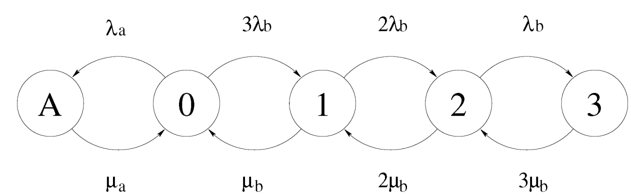


Fig. 2. Continuous time Markov chain with five states to model the unlicensed spectrum access process.

bandwidth of the frequency channels in Fig. 1 do not represent any existing unlicensed band, this usage model serves without loss of generality only as an example model. Here, radio system A can be compared with wireless local area networks (LANs) operating in the 5 GHz band [using orthogonal frequency-division multiplexing (OFDM)]. Radio system B represents narrowband radio systems supporting for example a limited number of voice calls or Bluetooth systems. In our scenario, instead of modeling the detailed protocols, a simplified listen before talk (LBT) is used for all radio systems. A type A radio system requires the respective three frequency channels to be idle before allocating radio resources. Only if the respective channels are idle, a radio system allocates radio resources, otherwise, it will be dropped, i.e., there is no queueing. The radio systems only scan their own frequencies, for example, a radio system B, with center frequency f_2 looks only in its frequency and not any other frequency. Collisions of allocation attempts occur when more than one radio system detects the channel as idle at the same time. In simulations, when collisions happen, one of the radio systems is randomly selected to allocate the radio resource, the other radio systems are dropped.

Two of the most representative etiquette rules defined in [11] are as follows.

- Rule 4: A radio system of type A or type B should apply LBT when operating.
- Rule 6: In order to protect other radio systems most efficiently, a radio system B that follows Rule 4 should synchronize its LBT process in time across neighboring frequency channels that overlap with the same A channels.

III. Markov Modelling

A. Equal Traffic Load Without Queueing

The unlicensed spectrum access problem can be modeled as a continuous time Markov chain. Without loss of generality, we can model the two radio system access model illustrated in Fig. 1 as a five state Markov chain, as shown in Fig. 2. The five states of the Markov chain are described in Table I. The assumption here is that for each type of the radio system, we have the same

TABLE I
FIVE STATES OF THE MARKOV CHAIN

State	Description
A	Radio system A occupies the reference spectrum range.
0	All the three frequency grids are idle.
1	There is only one type B radio system in the reference range
2	There are two type B radio systems in the reference range
3	There are three type B radio systems in the reference range

traffic load and occupation time. Later, we will relax this assumption and propose a more general model. Since the contentions of the spectrum only take place between radio system A and B, we focus on one of the type A radio systems and the three type B radio systems whose required spectrum is within the type A radio system's spectrum range. Here, we call this spectrum range as reference range. As collisions rarely happen especially with low traffic load, in this Markov model, we omit the collision state.

We define an infinitesimal generator matrix $\mathbf{A}$ to characterize the transition of the states of the Markov chain. The infinitesimal generator matrix with the sum of each row equalling zero is given as follows:

$$\mathbf{A}=\begin{bmatrix} -\mu_a & \mu_a & 0 & 0 & 0 \\ \lambda_a & -\lambda_a-3\lambda_b & 3\lambda_b & 0 & 0 \\ 0 & \mu_b & -\mu_b-2\lambda_b & 2\lambda_b & 0 \\ 0 & 0 & 2\mu_b & -2\mu_b-\lambda_b & \lambda_b \\ 0 & 0 & 0 & 3\mu_b & -3\mu_b \end{bmatrix}. \tag{1}$$

Then, we have

$$\mathbf{\Pi A}=0 \tag{2}$$

where $\mathbf{\Pi} = [\Pi_A, \Pi_0, \Pi_1, \Pi_2, \Pi_3]$ is the steady-state probability vector and Π_i represents the probability of being in state i. Solving recursively, we can get

$$\mathbf{\Pi}=[1, P_0, P_1, P_2, P_3]\, P_A \tag{3}$$

where

$$P_A=\left[\frac{\mu_a}{\lambda_a}+1+\frac{3\lambda_b\mu_a}{\lambda_a\mu_b}+\frac{3\lambda_b^2\mu_a}{\lambda_a\mu_b^2}+\frac{\lambda_b^3\mu_a}{\lambda_a\mu_b^3}\right]^{-1}$$
$$P_0=\frac{\mu_a}{\lambda_a}, P_1=\frac{3\lambda_b\mu_a}{\lambda_a\mu_b}, P_2=\frac{3\lambda_b^2\mu_a}{\lambda_a\mu_b^2}, P_3=\frac{\lambda_b^3\mu_a}{\lambda_a\mu_b^3}. \tag{4}$$

One of the most important metrics in the unlicensed band access is the average airtime per radio system. Airtime refers to the ratio of allocation time per radio system type to the reference time (say one hour) [11]

$$airtime_{type=A,B}=\frac{1}{N_{type}}\sum_{i=1}^{N_{type}}\frac{allocation\ time(i)}{reference\ time} \tag{5}$$

where N_{type} is the number of type i radio systems. Based on the previous Markov model, the airtime can be approximated by

$$\begin{aligned} airtime_{type_A} &= \Pi_A, \\ airtime_{type_B} &= \frac{1}{3}\Pi_1+\frac{2}{3}\Pi_2+\Pi_3. \end{aligned} \tag{6}$$

From (6), we can see that when radio system A and B are given the same high traffic load, $airtime_{type_B} \gg airtime_{type_A}$ which is not fair for radio system A. As the traffic load of radio system B increases, the $airtime_{type_A} \rightarrow 0$ which is unacceptable for the broader band radio system A. Some etiquette rules were proposed in [11] to mitigate this unfairness, but the improvements are limited.

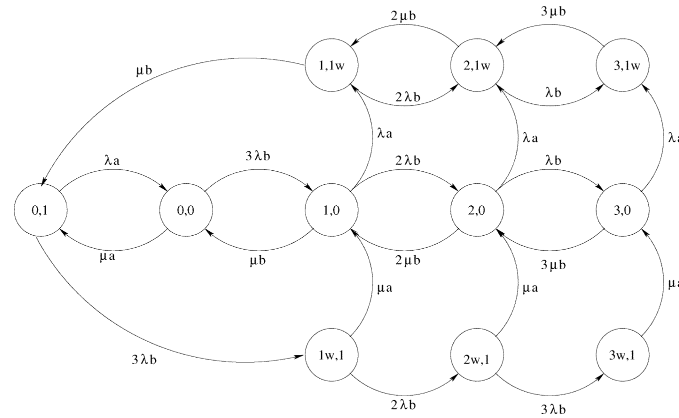


Fig. 3. Markov chains to model the unlicensed spectrum access process with waiting.

B. Equal Traffic Load With Queueing

As queueing can increase the server utilization and, hence, the throughput in classical queueing systems, we also expect increase in terms of *airtime* share of these coexisting radio systems by introducing waiting (queueing). With queueing, instead of dropping the radio system when the channels are busy, they continue scanning until the channel(s) become idle. Of course, the tradeoff here is between the waiting time and the spectrum utilization. When there are more radio systems of the same type than the available spectrum channels, collisions will be pronounced. In order to decrease the probability of collisions, collision avoidance techniques should be used, such as random backoff, or carrier sense multiple access/collision avoidance (CSMA/CA). In this paper, since our focus is on the effect of coexistence of dissimilar radio systems on unlicensed frequency bands, a perfect collision avoidance among resource allocations is assumed.

We again consider the model shown in Fig. 1 with equal traffic load within the same radio system. At this time, only if the respective channels are idle, a radio system allocates radio resources, otherwise, it continues to scan until the channels become idle. The Markov model that characterizes this scheme is illustrated in Fig. 3. Here, each state is (k_1, k_2), where k_i equals the number of radio system i on the spectrum block. While $k_i w$ means the number of waiting radio systems of type i. We can solve for the *airtime* share by the general method described in Section V.

C. Markov Model for General Traffic Load

The Markov chain shown in Fig. 4 models a more general case where within each radio system type, there may exist different traffic loads and occupation times. In Fig. 4, a state is represented by the triplet (n_1, n_2, n_3). In this, n_1, n_2, and n_3 represent the status of the radio system of type B occupying carrier

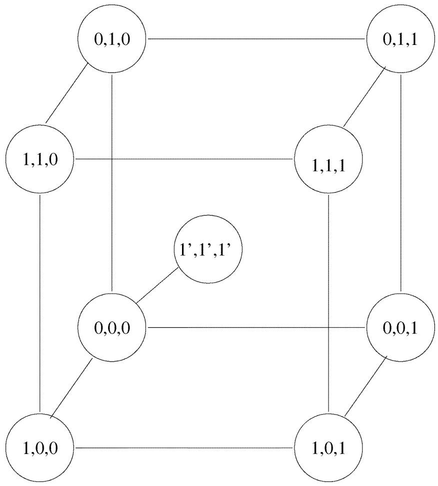


Fig. 4. Markov chain with nine states to model the unlicensed spectrum access process.

frequencies f_1, f_2, and f_3, respectively. Here, $n_i = 0$, $i = 1$, 2, 3 indicates that the radio system of type B requiring carrier frequency f_i is idling, while $n_i = 1$, $i = 1$, 2, 3 indicates that it occupies the carrier frequency f_i. State $(1', 1', 1')$ represents radio system A occupies the three frequency channels.

Here, all the states are connected by straight lines that are bidirectional. Let λ_1, λ_2, λ_3 and μ_1, μ_2, μ_3 represent the arrival rates and service rates of the three type *B* systems whose center frequencies are f_1, f_2, f_3, respectively. Let λ_a and μ_a represent the arrival and service rates of type A system, respectively. The basic equation governing the above system is given by

$$\begin{aligned}
\lambda_a P_{0,0,0} &= \mu_a P_{1',1',1'},\\
(\lambda_1+\lambda_2+\lambda_3)P_{0,0,0} &= \mu_1 P_{1,0,0}+\mu_2 P_{0,1,0}+\mu_3 P_{0,0,1},\\
(\mu_1+\lambda_2+\lambda_3)P_{1,0,0} &= \lambda_1 P_{0,0,0}+\mu_2 P_{1,1,0}+\mu_3 P_{1,0,1},\\
(\lambda_1+\mu_2+\lambda_3)P_{0,1,0} &= \mu_1 P_{1,1,0}+\lambda_2 P_{0,0,0}+\mu_3 P_{0,1,1},\\
(\lambda_1+\lambda_2+\mu_3)P_{0,0,1} &= \mu_1 P_{1,0,1}+\mu_2 P_{0,1,1}+\lambda_3 P_{0,0,0},\\
(\mu_1+\mu_2+\lambda_3)P_{1,1,0} &= \lambda_2 P_{1,0,0}+\lambda_1 P_{0,1,0}+\mu_3 P_{1,1,1},\\
(\mu_1+\lambda_2+\mu_3)P_{1,0,1} &= \lambda_3 P_{1,0,0}+\mu_2 P_{1,1,1}+\lambda_1 P_{0,0,1},\\
(\lambda_1+\mu_2+\mu_3)P_{0,1,1} &= \mu_1 P_{1,1,1}+\lambda_2 P_{0,0,1}+\lambda_3 P_{0,1,0},\\
(\mu_1+\mu_2+\mu_3)P_{1,1,1} &= \lambda_1 P_{0,1,1}+\lambda_2 P_{1,0,1}+\lambda_3 P_{1,1,0},\\
P_{0,0,0}+P_{0,0,1}+\cdots+P_{1,1,1}&+P_{1',1',1'}=1 \qquad (7)
\end{aligned}$$

where P_{n_1,n_2,n_3} represents the probability of being in state (n_1, n_2, n_3), $n_1, n_2, n_3 = 0, 1$. The above equation, with the exception of the state $(1', 1', 1')$ represents three independent M/M/1/1 queues. The solution to the above equation is given by

$$P_{n_1,n_2,n_3} = C\left[\frac{\lambda_1}{\mu_1}\right]^{n_1}\left[\frac{\lambda_2}{\mu_2}\right]^{n_2}\left[\frac{\lambda_3}{\mu_3}\right]^{n_3}, n_1, n_2, n_3 = 0, 1 \qquad (8)$$

where C is the normalizing constant. The inclusion of the state $(1', 1', 1')$ also represents a M/M/1/1 queue that is independent of P_{n_1,n_2,n_3}. For simplicity and in order to get more insight into the analysis, we use the five state Markov model (Fig. 2) unless stated, otherwise.

IV. Random Access Model

Efficiency and fairness are obviously the main goals of a spectrum etiquette. As discussed before, if every radio system accesses the unlicensed band in a greedy manner, then the radio system requiring broader band to operate will suffer from an unacceptable low airtime share. So one way to provision more fairness to the etiquette rules would be to require each radio system to work in a cooperative manner. One option would be that each radio system i tries to contend for the spectrum with probability p_i. After the radio system has decided to contend for the spectrum, it accesses the spectrum compliant to etiquette Rule 4 described in Section II-B.

This random access scheme can be approximated by slightly modifying the previous five state Markov model. Here, each radio system will only contend for the spectrum with probability p_i, so the actual traffic load to the system can be approximated by $p_i\lambda_i$. If perfect fairness is achieved, then we have

$$airtime_{type_A}(p_a, p_b) = airtime_{type_B}(p_a, p_b). \qquad (9)$$

Then, from (6), we have

$$p_a(p_b) = \frac{1}{3}p_b P_1 + \frac{2}{3}p_b^2 P_2 + p_b^3 P_3. \qquad (10)$$

So

$$airtime_{type_A}(p_b) = \frac{\frac{1}{3}p_b P_1 + \frac{2}{3}p_b^2 P_2 + p_b^3 P_3}{P_0 + \frac{1}{3}p_b P_1 + \frac{2}{3}p_b^2 P_2 + p_b^3 P_3 + p_b P_1 + p_b^2 P_2 + p_b^3 P_3}. \qquad (11)$$

When the *airtime* for both radio system A and B are the same, we can find the optimal p_b by maximizing the *airtime*. Since

$$\frac{\partial airtime_{type_A}}{\partial p_b} > 0 \qquad (12)$$

$airtime_{type_A}$ is an increasing function of p_b. So, the optimal p_b is the largest possible p_b. Depending on different λ and μ values, we have the following two cases.

- $p_a(p_b = 1) > 1$: We cannot use the maximum value of $p_b = 1$ to maximize the *airtime* function because of $0 < p_a \leq 1$. But as $p_a(p_b)$ is an increasing function of p_b, the maximum possible $p_{b_{\mathrm{opt}}}$ can be calculated when $p_a = 1$ (which is the optimal value for p_a) from (10).
- $p_a(p_b = 1) < 1$: We can get the maximum value of $p_{b_{\mathrm{opt}}} = 1$ and, hence, by (10), we can get $p_{a_{\mathrm{opt}}} = p_a(p_b = 1)$.

Since $\frac{\partial airtime_{type_A}(p_a,p_b)}{\partial p_a} > 0$, $\frac{\partial airtime_{type_A}(p_a,p_b)}{\partial p_b} < 0$, $\frac{\partial airtime_{type_B}(p_a,p_b)}{\partial p_a} < 0$ and $\frac{\partial airtime_{type_B}(p_a,p_b)}{\partial p_b} > 0$, it can be shown that this $(p_{a_{\mathrm{opt}}}, p_{b_{\mathrm{opt}}})$ pair actually corresponds to the strategy that no radio systems can do better in terms of airtime share without harming the other coexisting radio systems. So in this sense, both the efficiency and fairness are obtained by using this optimal probability pair. Obviously, without considering any fairness issues, the most efficient access in the sense of pure spectrum utilization would be that all the users compete for the spectrum greedily. In that way, some type of networks may be totally blocked out of the spectrum. And further, if different types of radio systems have different traffic load, to just equalize the *airtime* may cause low spectrum utilization. Exact fairness does not always apply as far as an operator is concerned, they are likely to be more concerned with revenue. So, users who pay more will get more access, which means different networks may have different priorities. To properly address these problems, we need to define the fairness and efficiency carefully.

Definition 1: Weighted fairness is achieved if the following equation holds:

$$\frac{airtime_{type_i}}{L_i} = \frac{airtime_{type_j}}{L_j} = K, \quad \forall i, j \tag{13}$$

where K is a constant, and the weight $L_i = \theta_i \lambda_i$, here θ_i is the priority parameter and λ_i is the traffic load for type i radio system.

Definition 2: Given the weighted fairness of different types of radio system is satisfied, efficiency is achieved if each of the radio system's *airtime* is maximized.

This efficiency optimal point, actually corresponding to the Pareto efficiency solution in game theory perspective. An strategy profile is Pareto optimal if some players must be hurt in order to improve the payoff of other players [12]. The above analysis with the exact equality fairness condition (9) hold just under a special case with $L_A/L_B = 1$. But the analysis of finding the optimal access probability p_i can be extended to cases with $L_i/L_j \neq 1$.

V. General Solution to *Airtime* Share and Blocking Probability

In order to analyze more complex access process, we need to deal with Markov chains with more states and more complex transition structures. Explicitly solving the balance equations to get the *airtime* share or blocking probability then becomes untractable. Therefore, we introduce a simple numerical technique described below to calculate the *airtime* shares and blocking probabilities.

A. Airtime Share

Given the traffic rates and occupying time, we can define an infinitesimal generator matrix $\mathbf{A}$ as before. Then, by solving the following equation we can get the state probabilities:

$$\mathbf{\Pi A} = 0 \tag{14}$$

where $\mathbf{\Pi} = [\Pi_1, \Pi_2, \ldots, \Pi_K]$, and K is the number of states in the Markov chain. The generator matrix $\mathbf{A}$ is singular so we cannot solve the state probability vector directly. But with the condition that the sum of all the steady-state probabilities should be one, we can put these two conditions into the following compact equation:

$$\begin{bmatrix} \mathbf{A}^T \\ 1_{1\times K} \end{bmatrix} \mathbf{\Pi}^T = \begin{bmatrix} 0_{K\times 1} \\ 1 \end{bmatrix}. \tag{15}$$

Then, by defining $\mathbf{A}' = \begin{bmatrix} \mathbf{A}^T \\ 1_{1\times K} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0_{K\times 1} \\ 1 \end{bmatrix}$, we have

$$\mathbf{A}'\mathbf{\Pi} = \mathbf{b}. \tag{16}$$

By using minimum mean-squared error (MMSE) criterion [13], we obtain the following unique solution:

$$\mathbf{\Pi}^T = (\mathbf{A}'^{\mathrm{T}}\mathbf{A}')^{-1}\mathbf{A}'^{\mathrm{T}}\mathbf{b}. \tag{17}$$

After having the state probabilities, the $airtime_i$ share for radio system i is just the weighed summation of the respective state probabilities.

B. Blocking Probability

Despite the fairness concerns, another important metric wireless users care about is the instant access probability, or the blocking probability. Our model can be considered as a finite population queueing model, where the time blocking which is the proportion of time the system spends in the blocking states is not equal to the call blocking which is the probability that an arriving call is blocked. So PASTA property does not apply here. Instead, we can compute the state probability seen by an arriving call (radio system) as follows:

$$\pi_j^* = \frac{\lambda_j \pi_j}{\sum\limits_{k=0}^{s} \lambda_k \pi_k} \tag{18}$$

where $s+1$ is the total number of states. Considering a long period of time T, on the average the system spends in state j the time $\pi_j T$. During this time, there are on the average $\lambda_j \pi_j T$ call arrivals which find the system in the state j. The total number of calls arriving in time T is on the average $T\sum_{k=0}^{s} \lambda_k \pi_k$. Then, the proportion of calls which find the system in the state j is as given by the above expression. When the random access probability is p_i the blocking probability experienced by the radio system i is

$$P_{\mathrm{BLK}_i} = 1 - p_i \left[1 - \sum_{j:\text{blocked states}} \pi_j^*\right]. \tag{19}$$

VI. Spectral Agile Radio Access

The proposed random access scheme makes it possible to achieve the desired fairness in open spectrum access with different types of radio systems. But with the increasing number of such devices, reducing the blocking probability and increasing the *airtime* share become a critical issue. Spectrum agility-based channel access helps this cause.

A. Physical-Layer Aspects of Agile Radio

Traditionally, frequency-division has been adopted to divide the electromagnetic spectrum between different wireless technologies. In frequency-division, portions of the spectrum are statically assigned to particular wireless technology to support their transmissions. This assignment has some inherent disadvantages with respect to spectral efficiency. For example, consider two operators operating on their licensed bands (i.e., exclusive use model). It is possible that the one wireless network operating in a particular frequency band is fully loaded, while the second may have unused resources (or *vice versa*). It would

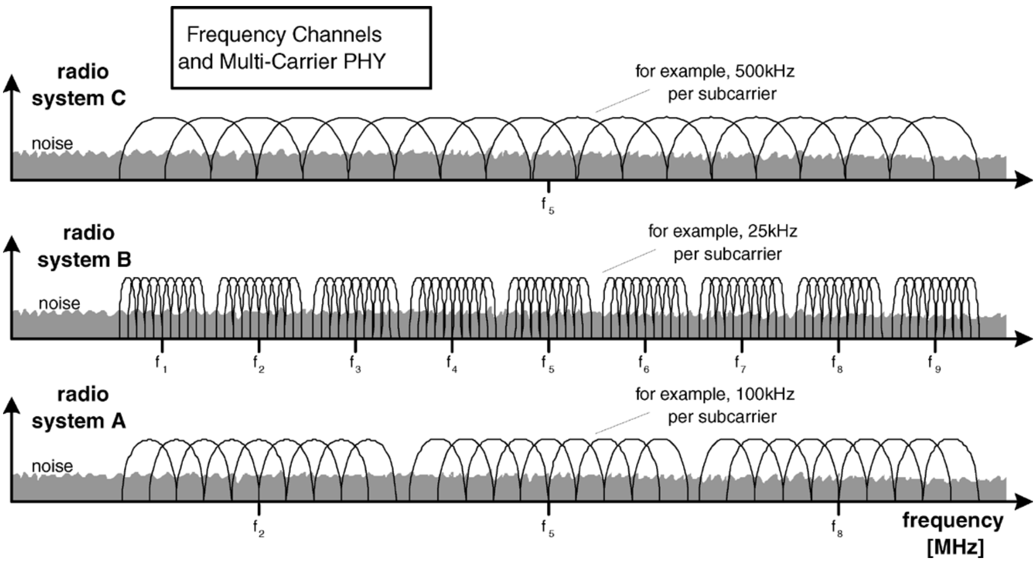


Fig. 5. Frequency channels and multicarrier PHY.

be profitable for both wireless network technologies, if these unused resources were shared to allow more capacity for the fully loaded operator. This form of spectral sharing is much easier to accomplish if the underlying access technique in the two operators is multicarrier-based. In case of a single operator, the operator can access the spectrum opportunistically based on the load of the system.

In order to allow different wireless technologies to borrow/lend spectral bands from/to different wireless networks residing in the same or different spectral region, we could conceive of a simple ad hoc strategy for selecting Δf (the frequency separation among carriers in a OFDM-based system design). For example, two different strategies to enable spectral sharing are as follows.

1) Different Δf for different regions.
2) Common Δf for all regions

In the first strategy, different values of Δf are chosen for different spectral regions, based on the following practical considerations: 1) Δf is selected much smaller than the coherence bandwidth of the channel (to ensure that each carrier undergoes a flat fade) and 2) Δf is selected big enough to prevent the use of large number of carriers (to allow practical implementation of the system via fast Fourier transform/inverse (FFT/IFFT). Therefore, we could choose values of Δf that are proportional to the carrier frequency. This choice allows the system to maintain the transceiver complexity within reasonable levels. Specifically, we could have the following subcarrier spacing.

1) $\Delta f = 25$ kHz for the low band region.
2) $\Delta f = 100$ kHz for the medium band region.
3) $\Delta f = 500$ kHz for the high band region.

These choices were made based on coherence bandwidth measurements for a typical indoor small office channels.

In the second strategy, we suggest the use of the same carrier separation (e.g., $\Delta f = 25$ kHz) across all regions. Even though this strategy facilitates orthogonal sharing across the entire bandwidth, it results in prohibitive complexity in IFFT/FFT implementation. For example, choosing $\Delta f = 25$ kHz for a 1 GHz service in region 3 requires a 40 000 -point IFFT. This can be chosen based on the current UWB standards. Fig. 5 shows how the multicarrier OFDM is used for agile radios.

B. MAC-Layer Aspect of Agile Radio

With the advances in SDR, spectral agile networks become more and more tractable. Such radio devices can dynamically utilize idle spectrum bands. One of the interesting concerns here is, given the additional freedom of carrier frequency switching, what is the gain in efficiency for radio systems with different bandwidth requirements. There are lots of ways to take advantage of this switching. To achieve the upper bound for the agile spectrum efficiency, one way is to "pack" all the radio systems tightly together in the spectral domain. Such a packing would ensure that there is no spectral hole/white space.

We describe the "packing" behavior as follows. If one radio system releases the spectrum, the other radio systems will switch their operating frequency band so that the vacant band is occupied (if it meets the demands of at least one of the active radio systems). Then, a new accessing radio system scans to find spectrum opportunities from the beginning of the spectral band, and occupies the first idle spectrum opportunity it finds. In this way, all the spectrum fragments can be saved for future accessing radio systems. A simple example given in Fig. 6 explains this process clearly, where when radio system B_2 releases the spectrum, radio systems B_4 (also could be B_3) then switches to frequency f_5. When a new accessing radio system A_2 scans the spectrum, it will occupy frequency f_7, f_8, and f_9. Without this spectral agility, it will have to queue or be dropped.

All the switching and scanning may introduce considerable signaling overhead and delay. To alleviate this, in practical schemes, we can divide the spectrum into blocks, and implement this packing in each blocks. This can considerably save the switching and scanning. Besides, software radios work very well with short scanning range.

Without loss of generality, we still use the basic scenario discussed before in Fig. 1 to illustrate this scheme, and compare it

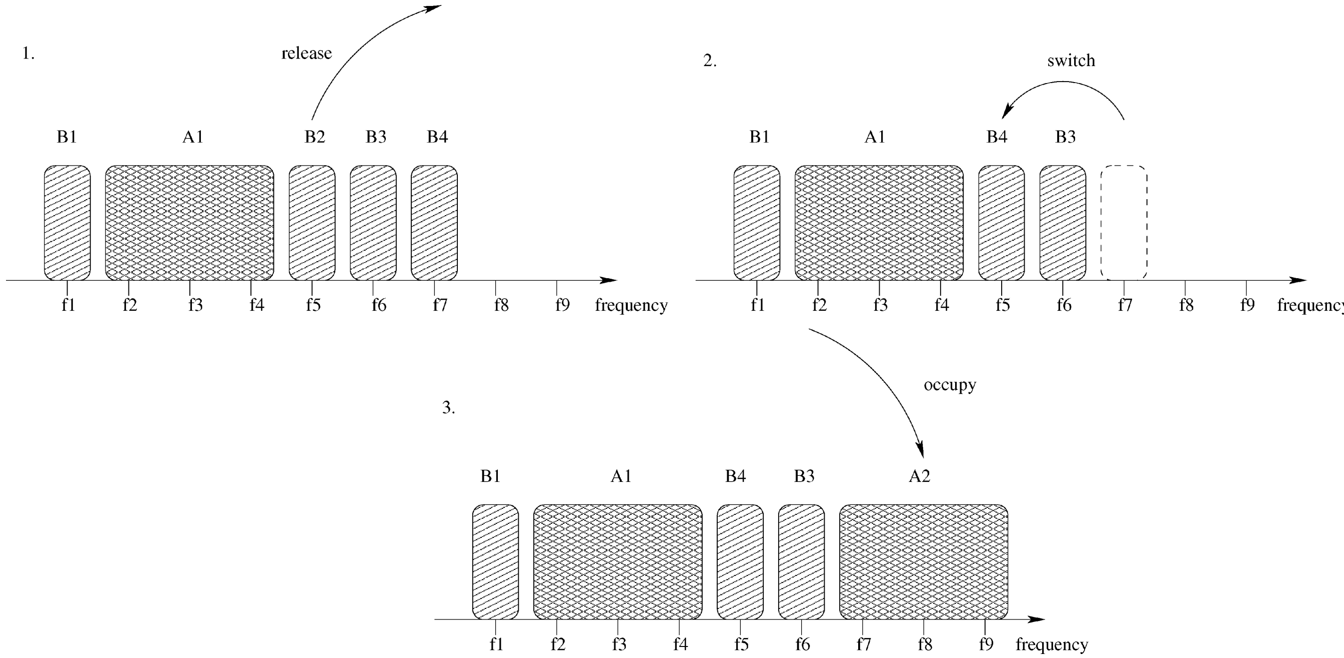


Fig. 6. An example for "packing" behavior.

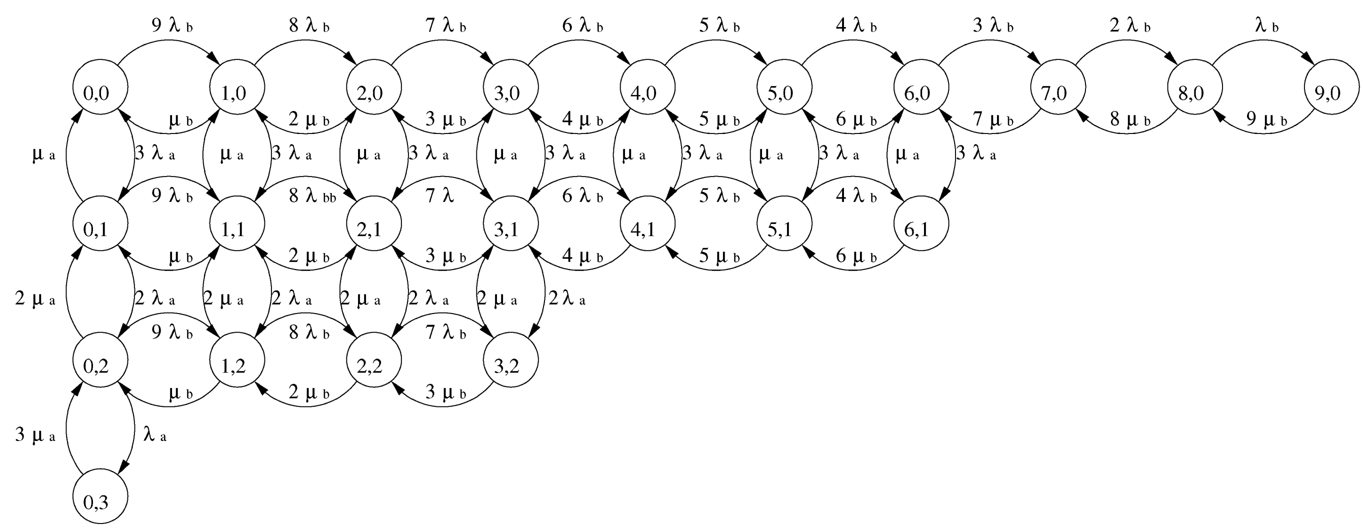


Fig. 7. Markov chain to model the unlicensed spectrum access process with spectral agility.

with the other schemes. The most significant difference for this scheme is that each radio system here is an agile radio system. It behaves in the "packing" way. The corresponding Markov model is shown in Fig. 7. Here, each state $\Pi_j = (k_1, k_2)$, where k_i equals the number of radio system i on the spectrum block. We can solve for the *airtime* share and the blocking probability by the general method described previously.

VII. Homo Egualis (HE) Society Model-Based Access Scheme

To obtain $(p_{a_{\rm opt}}, p_{b_{\rm opt}})$, we need the information of all the λ's and μ's, which is impractical in a real access scenario. A more realistic scheme would be to allow the radio systems to learn these p_a and p_b themselves with only local information or measurement. We present such a technique next.

A. Agent Egualis Society

In many decision-making and strategy-setting people do not behave like the self-interested "rational" actor depicted in neoclassical economics and classical game theory [5]. In an HE society, individuals have an inequality aversion. As a result altruists appear in ultimatum and public good games. As Gintis states in [5], support for HE comes from the anthropological literature, describing how Homo sapiens evolved in small hunter-gather groups. Such societies had no centralized structure of governance, so the enforcement of norms depends on the voluntary participation of peers. An HE society [5] can be modeled as follows, where the utility function of player i, u_i in an n-player game is

$$u_i = x_i - \frac{\alpha_i}{n-1} \sum_{x_j > x_i} (x_j - x_i) - \frac{\beta_i}{n-1} \sum_{x_j < x_i} (x_i - x_j) \quad (20)$$

where $x = (x_i, \ldots, x_j)$ are the payoffs for each player and $0 \leq \beta_i < \alpha_i \leq 1$. $\beta_i < \alpha_i$ reflects the fact that HE exhibits a weak urge to inequality when doing better than the others and a strong urge to reduce inequality when doing worse than the others. In [5], it is shown that in this model the salient behaviors in ultimatum and public goods games, where fairness does matter, can be reproduced.

B. Proposed HE-Based Access Scheme

The inequality aversion property of the HE agents can be utilized to achieve fairness in the spectrum access problem. In this scheme, each radio system learns the access probability p_i by itself. Here, we define $Onlinetime_i$ as the averaged cumulative "on" spectrum time per radio system of type i. Then, we define $x_i = Onlinetime_i/L_i$, where $L_i = \theta_i\lambda_i$ is the same as used before in (13). The cumulative $Onlinetime_i$ is normalized by the radio system's traffic load and priority, which makes this spectrum access scheme able to adapt to different traffic loads and priority, hence, achieve more efficiency and maintain our defined weighted fairness. With initial $p_i = 1$, each time the probability p_i is updated, as follows:

$$p_i = \max\left(0, \min\left(1, p_i + \frac{\alpha_i}{n-1}\sum_{x_j \geq x_i}\left(\frac{x_j - x_i}{x_j}\right) - \frac{\beta_i}{n-1}\sum_{x_j < x_i}\left(\frac{x_i - x_j}{x_i}\right)\right)\right) \quad \text{for all } j \neq i \tag{21}$$

where n is the number of different radio system types, $0 < \beta_i < \alpha_i$ reflects the fact that radio system exhibits a weak urge to inequality when doing better than others and a strong urge to reduce inequality when doing worse than the others. This forces each radio system to make an effort to efficiently use the idle spectrum while taking fairness into consideration. Here the only local information needed is the radio system's own history of the *Onlinetime* and the *Onlinetime* of the other radio systems whose spectrum is within the same spectrum block. This can be obtained by keeping a record of the busy time of the required spectrum, which can be obtained by periodical spectrum scanning. When there are more than two different radio systems trying to coexist in the same spectrum, different radio systems can be identified by some smart technologies (e.g., we can detect the different transmitting power levels to distinguish from different radio systems), so each radio system can access the spectrum based only on its own recorded history and the local measurements performed by itself. While λ_i can be estimated by historical usage records of radio system type i. Only the priority parameter α_i needs to be announced by the radio system or be broadcasted by some operator. If each radio system can only achieve imperfect knowledge of the spectrum usage time of other radio systems, there will be some degradation in the achieved weighted fairness.

In our analysis, the arrival traffic is modeled as a Poisson process, in conventional wired systems it has been demonstrated that the traffic can be more heavy tailed than Poisson-based models. As we move to more and more packet wireless systems we may see similar distributions. As the weighted fairness

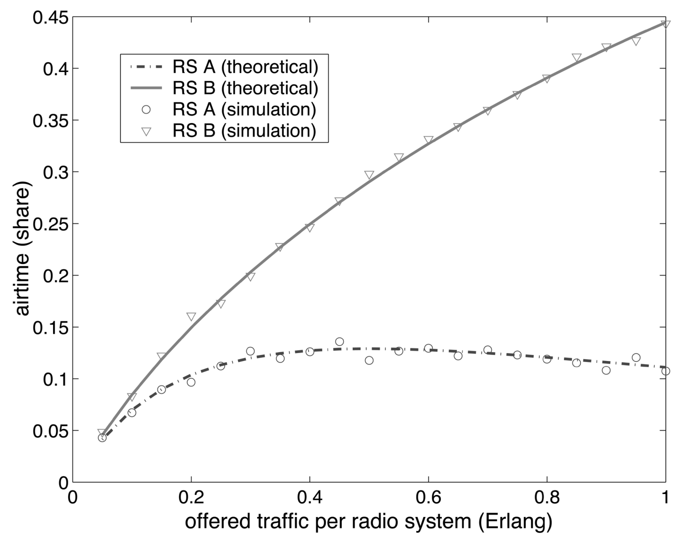


Fig. 8. Simulated and the Markov modeled spectrum access *airtime* under Rule 4. No queueing case.

is achieved through random access scheme, when the arrival traffic has heavy tailed distribution, we can still use our proposed HE algorithm to balance different types of radio systems. But the theoretical analysis based on continuous-time Markov chain will be invalid.

VIII. Numerical Results

We describe the simulation results in this section and compare it with the theoretical analysis. Equal loads on the radio systems are assumed unless otherwise stated. Fig. 8 shows the simulated average *airtime* per radio system and the theoretical results obtained from the Markov model in Fig. 2. We see that the proposed Markov model fits the simulation results very well. Here, radio systems are blocked when there are no idle frequency bands, and each radio system's carrier frequency is fixed. Radio systems only scan their own fixed frequency for spectrum opportunities. As one can imagine, it is seen here that with only Rule 4 (LBT), the narrow bandwidth radio system B will dominate the airtime share over the broadband radio system A. This dominance emphasizes the issue of unfairness, which is a prominent problem in coexistence of different types of radio systems.

The proposed HE-based random access scheme can mitigate this dramatically, as illustrated in Fig. 9. The theoretical *airtime* share for radio systems A and B are the same when optimal access probability pair $(p_{a_{\rm opt}}, p_{b_{\rm opt}})$ is used, and both of them increase with the increase of the traffic load. The HE access scheme using only local information is observed to produce a close to optimal solution. The HE access scheme is observed to produce a performance gain even higher than the optimal achievable solution using $(p_{a_{\rm opt}}, p_{b_{\rm opt}})$, as shown in Fig. 9 for some traffic load. This is not surprising, because in the HE access, the p_a and p_b values change during the access, while the optimal probability solution pair $(p_{a_{\rm opt}}, p_{b_{\rm opt}})$ is obtained with the assumption that both of them will remain unchanged throughout the access. So when using the HE access in a real system, it may sometimes perform better than the predicted optimal results from the proposed Markov model solution of the

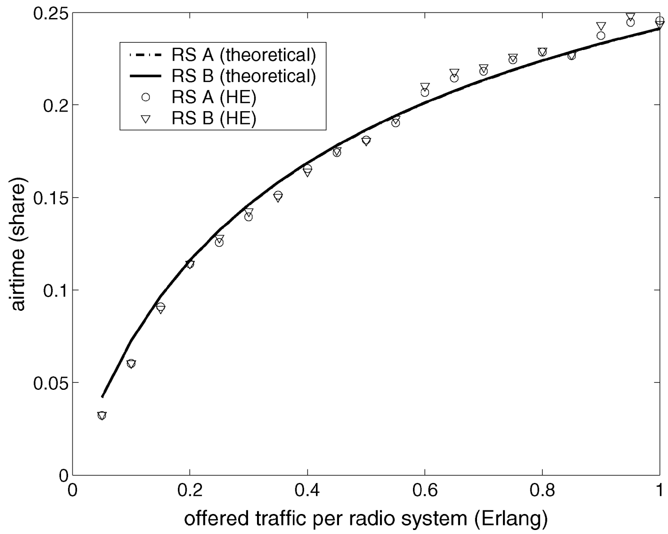


Fig. 9. Proposed HE access scheme compared with theoretical optimal solution in terms of *airtime*. No queueing case.

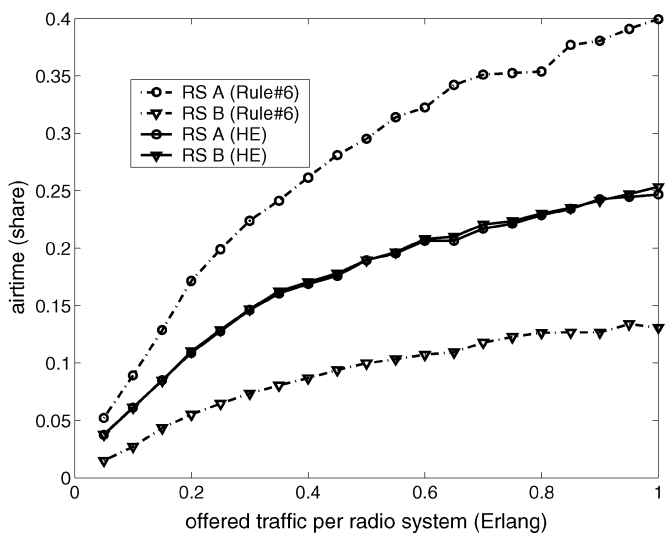


Fig. 10. Proposed HE access scheme compared with Rule ♯6 in terms of *airtime*. No queueing case.

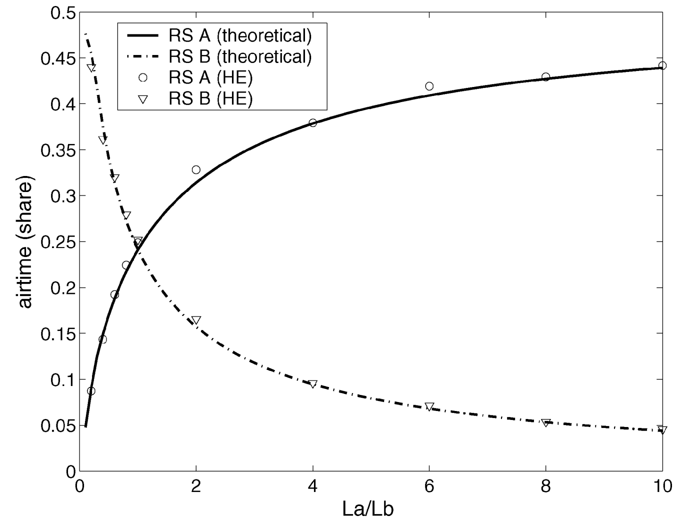


Fig. 11. *airtime* achieved with different weight ratios L_A/L_B. No queueing case.

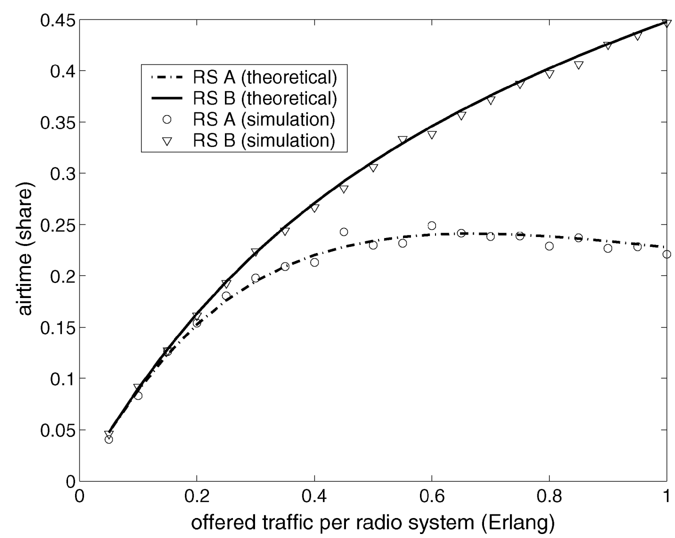


Fig. 12. Simulated and the Markov modeled spectrum access *airtime* under Rule 4. Queueing of radio systems case.

random access scheme. Of course, for some instances, it may produce lower performance than the optimal solution as seen later.

The etiquette Rule 6 can protect the broadband radio system A by requiring a radio system of type B that follows Rule 4 to synchronize its LBT process in time across neighboring frequency channels that overlap with the same reference channel [11]. But as can be seen in Fig. 10, although the *airtime* share for radio system A increases, the cost is a significant decrease of *airtime* share for radio system B. The HE-based access is seen to be much better in terms of fairness compared to the etiquette Rule 6.

With different traffic load for different types of radio systems and different priorities, the ratio of L_i/L_j will no longer equal one. In Fig. 11, we illustrate the *airtime* achieved by radio system A and B versus different L_A/L_B ratios. As can be seen, $airtime_{type_A}/airtime_{type_B} = L_A/L_B$ is achieved, which means the weighted fairness is achieved.

When each radio system is allowed to wait (queue) instead of being dropped if the desired frequency band is busy, the efficiency of the spectrum usage is expected to increase as illustrated in Fig. 12, where only LBT etiquette is used. The theoretical result is derived from the Markov model in Fig. 3. With waiting, as can be seen, the *airtime* share for the broadband radio system A almost doubled while it increased a little bit for radio system B. Therefore, the total utilization of the spectrum increases. The tradeoff here is the delay due to waiting, which is shown in Fig. 13. The total waiting time is normalized to the *reference time*. Furthermore, when there are more radio systems, collisions will be prominent and a sophisticated collision avoidance scheme is required. Fig. 14 shows that the HE access scheme works well for this case as well. Comparing Figs. 14 and 9, we can see that when using HE with waiting, the radio systems' *airtime* share increases significantly.

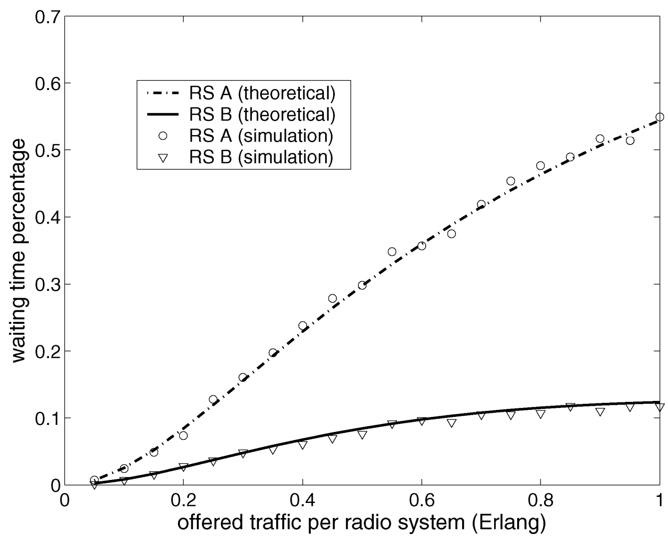


Fig. 13. Simulated and the Markov modeled spectrum access waiting time. Queueing of radio systems case.

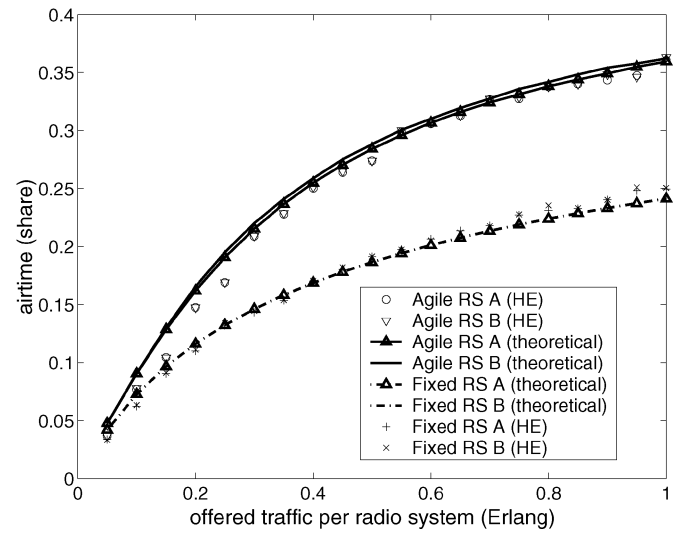


Fig. 15. Spectral agile access scheme compared with original fixed access scheme in terms of *airtime*. No queueing case.

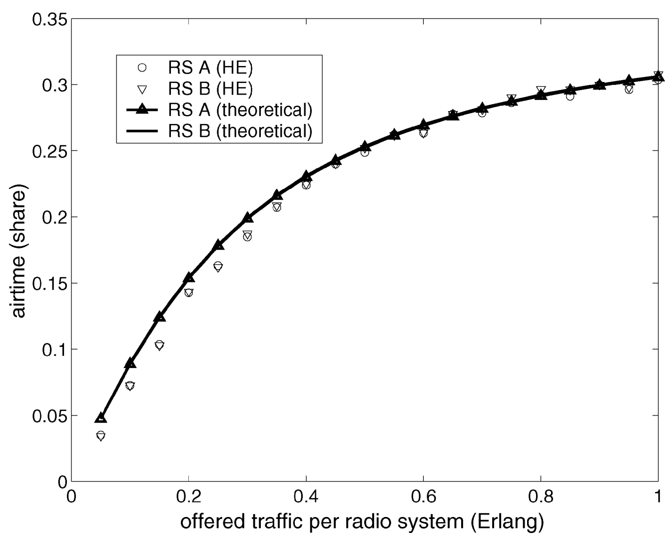


Fig. 14. Proposed HE access scheme compared with theoretical optimal solution in terms of *airtime*. Queueing of radio systems case.

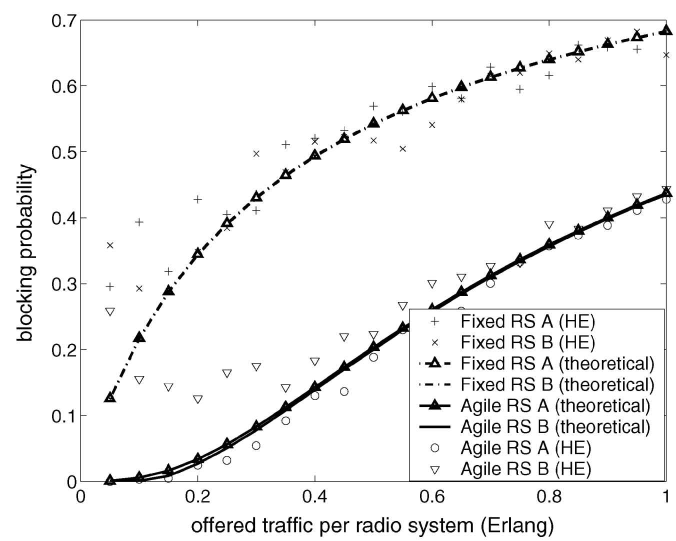


Fig. 16. Spectral agile access scheme compared with original fixed access scheme in terms of blocking probability. No queueing case.

Agile radio systems can potentially increase the spectrum efficiency significantly, as illustrated in Fig. 15. When radio systems have the ability to switch their carrier frequency, they can pack themselves together in the spectral domain and hence will increase the overall *airtime* share. This agility can as well decease the blocking probability experienced by the radio systems, as shown in Fig. 16. In Fig. 16, the HE access scheme will experience higher blocking probability than the expected theoretical prediction especially at low traffic load. This phenomena can be explained by the same reason as mentioned before, which is because the p_a and p_b change during the access when using HE. So, when the traffic load is low, the optimal access probabilities are 1, but this may not be so when using HE. They may be lower than 1 for some time, which will result in a higher blocking probability as can be calculated through (19). Here, the theoretical solution is derived using the Markov model given in Fig. 7.

Figs. 17 and 18 show that the HE scheme can also work well when there are multiple different types of radio systems. Here, we considered three different types of agile radio systems. Radio systems A and B are the same as before, while radio system C represents radio system that uses broadband transmission schemes such as UWB or spread spectrum. It requires even the whole spectrum (all nine spectrum bands) to be idle before allocating radio resources in the scenario discussed in Fig. 1. It can be seen from Fig. 17 that all these different radio systems almost have the same airtime share, which is what is desired for fairness. The theoretical solution is obtained from the Markov model described in Fig. 7 by adding one more state for radio system C. Due to the same reason of the changing of access probabilities p_a, p_b, and p_c in HE, the HE access scheme is observed to produce a performance gain higher than the optimal theoretical solution both in airtime share and the blocking probability in this case.

α and β are two important control parameters in the HE access scheme. Therefore, it is important to study the influence of α and β to the HE scheme. We find that when one of the radio

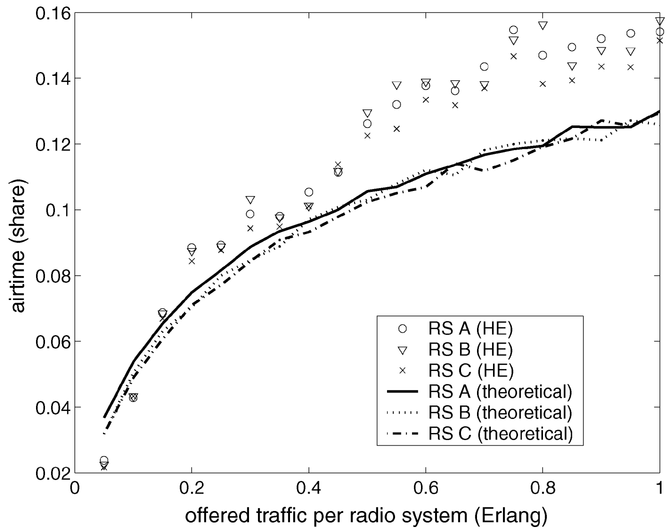


Fig. 17. Proposed HE access scheme compared with theoretical optimal solution in terms of *airtime*. Three different types of spectral agile radio systems. No queueing case.

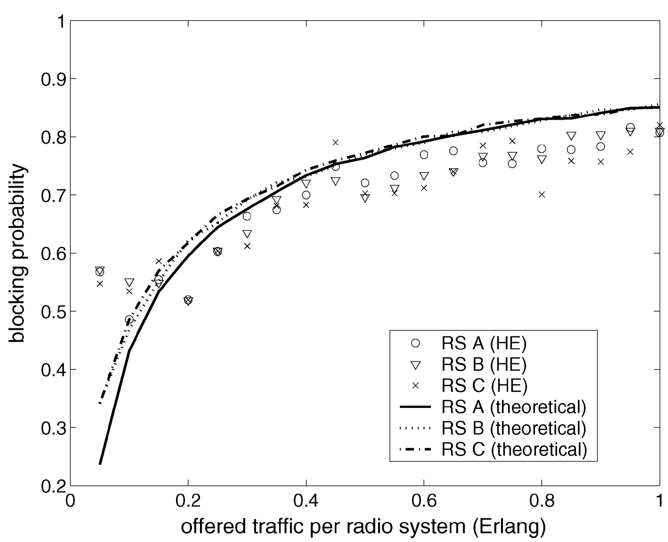


Fig. 18. Proposed HE access scheme compared with theoretical optimal solution in terms of blocking probability. Three different types of spectral agile radio systems. No queueing case.

systems uses $\alpha < \beta$, both of them will experience much lower airtime share, while as long as all of them use $\alpha > \beta$, they will almost get the same airtime performance no matter what specific α and β value they are using. Fig. 19 illustrates one case when both of the radio systems use $\alpha = 0.01 < \beta = 0.1$ compared with both of them use $\alpha = 0.1 > \beta = 0.01$. It can be seen that when α is smaller than β, the performance for both of the two radio systems degrades significantly.

IX. Conclusion

It is shown that continuous time Markov chain models are accurate in predicting the behavior of open spectrum access under the assumption that the arrival traffic has Poisson distribution. The proposed random access protocol achieves theoretical weighted airtime fairness. A distributed version of this access protocol that uses only local information based on a HE society

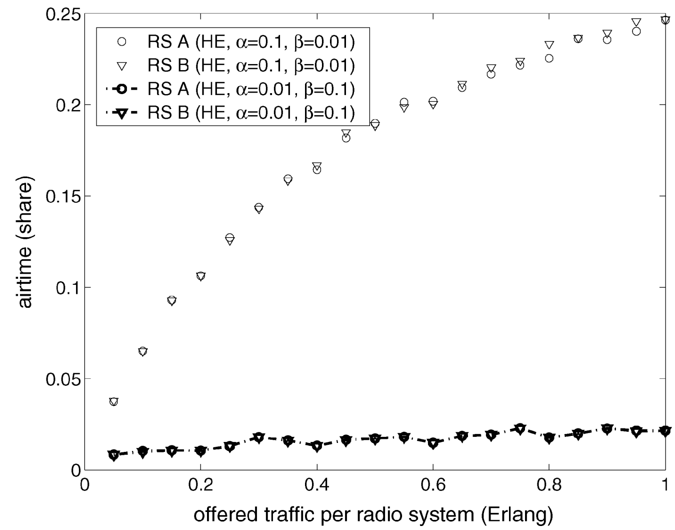


Fig. 19. Two original fixed radio systems use proposed HE access scheme. No queueing case. The comparison is between different parameter pairs ($\alpha = 0.1, \beta = 0.01$) and ($\alpha = 0.01, \beta = 0.1$).

model is also observed to work well. It is seen to produce results close to the theoretically optimal performance. Dynamic spectrum access in wireless networks is observed to outperform the fixed spectrum access counterpart. The Markov model investigated in this paper matches the simulated performance of an agile radio accurately. The HE version of spectrum access protocol produces near-optimal results. It is seen that spectrum agile radios produce superior airtime performance and blocking probabilities making them an attractive option for next general open spectrum wireless networks.

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive comments which led to a significant improvement of the manuscript.

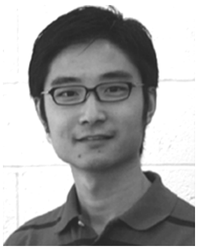
References

[1] XG Working Group. (Jul. 2003) The XG Vision. Request for Comments, BBN Technologies, Cambridge, MA. [Online]. Available: http://www.darpa.mil/ato/programs/XG/rfcs.htm
[2] D. P. Satapathy and J. M. Peha, "Spectrum sharing without licenses: Opportunities and dangers," in *Proc. Telecommun. Policy Res. Conf. Section 3*, 1996, pp. 15–29.
[3] ——, "Performance of unlicensed devices with spectrum etiquette," *Proc. IEEE GLOBECOM*, pp. 414–418, Nov. 1997.
[4] ——, "Etiquette modifications for unlicensed spectrum: Approach and impact," in *Proc. 48th Annu. Int. IEEE Veh. Technol. Conf.*, vol. 1, May 1998, pp. 272–276.
[5] H. Gintis, *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Behavior*. Princeton, NJ: Princeton Univ. Press, 2000.
[6] C. T. Chou, H. Kim, S. S. N, and K. G. Shin, "What and how much to gain from spectrum agility," Real-Time Comput. Lab., Univ. Michigan, Ann Arbor, MI, Tech. Rep., 2005.
[7] S. Mangold, Z. Zhong, K. Challapali, and C. T. Chou, "Spectrum agile radio: Radio resource measurements for opportunistic spectrum usage," in *Proc. IEEE GLOBECOM*, Dallas, TX, Nov. 2004, pp. 3467–3471.
[8] K. Challapali, D. Birru, and S. Mangold, "Spectrum agile radio for broadband applications," *EETimes in Focus Article*, Aug. 2004.
[9] J. Mitola, "The software radio architecture," *IEEE Commun* , vol. 33, no. 5, pp. 26–38, 1995.
[10] "Regulatory aspects of software defined radio," SDR forum document number SDRF-00-R-0050-v0.0, White Paper.

[11] S. Mangold and K. Challapali, "Coexistence of wireless networks in unlicensed frequency bands," in *Wireless World Research Forum 9*, Zurich, Switzerland, Jul. 2003.

[12] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.

[13] G. Strang, *Introduction to Linear Algebra*. Cambridge, MA: Wellesley-Cambridge Press, 1998.



Yiping Xing (S'03) received the B.S. degree in electrical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2001 and the M.E. degree in electrical engineering from the Stevens Institute of Technology, Hoboken, NJ, in 2004. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, Stevens Institute of Technology.

His current research interests include radio resource management for cellular and ad hoc networks, access control for spectrum agile networks, and game theory for wireless networks.

R. Chandramouli (M'00) is an Associate Professor in the Department of Electrical and Computer Engineering (ECE), Stevens Institute of Technology, Hoboken, NJ. Prior to jointing Stevens Institute of Technology, he was on the faculty of the Department of Electrical and Computer Engineering, Iowa State University, Ames. His research interests include steganography, steganalysis, encryption, wireless networking, and applied probability theory. His research in these areas is sponsored by the National Science Foundation (NSF), the Air Force Research Laboratory, and industry.

Dr. Chandramouli is a recipient of the National Science Foundation (NSF) CAREER Award. He has been serving as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY since 2000. He is a Cofounder and Co-Program Chair for the IEEE International Workshop on Adaptive Wireless Networks (2004 and 2005). He is also involved with several conference organization committees as a Technical Program Committee member.

Stefan Mangold (M'01) received the Dipl.-Ing. degree in electrical engineering and the Dr.-Ing. degree (*summa cum laude*) from ComNets RWTH Aachen University, Aachen, Germany, in 1997 and 2003, respectively.

He is a Project Leader at Swisscom Innovations Ltd., Berne Switzerland, where he is working on cognitive radio, spectrum management, and IEEE 802 standards. His current research focuses on operator assisted cognitive radios for spectrum sharing, and spectrum etiquette. He authored technical papers in various IEEE ComSoc conferences and journals, and contributes to IEEE 802 standardization. Before joining Swisscom in April 2005, he was with Philips Research, Briarcliff Manor, NY (2003–2005), and with the Chair of Communication Networks (ComNets) of RWTH Aachen University (1998–2003).

Dr. Mangold is a member of the Association of Computing Machinery (ACM), German Verband Deutscher Elektrotechniker (VDE), and Swiss ElectroSuisse. He is involved with several conference organization committees as a Technical Program Committee member.

Sai Shankar N received the Ph.D. degree from the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, in the area of ATM networks.

In 1999, he joined Philips Research, Eindhoven, The Netherlands, where he served as a Research Scientist working on various problems involving hybrid, fiber, coaxial cable (IEEE 802.14) networks and IP protocols. In the 2001, he joined Philips Research, Briarcliff Manor, NY, and worked in the area of wireless LANs/UWB, cognitive radios, and cooperative communications. He was an active contributor of the wireless LAN standard and has submitted more than 15 proposals in shaping QoS related issues in the IEEE 802.11e. He is also an active participant in the Ultra-Wideband (UWB) Working Group of WiMEdia Alliance and has contributed in shaping the new MAC at the Multiband OFDM Alliance (MBOA) Forum. He has been the reviewer of almost all important journals in the area of networking and has chaired lots of conferences. Currently, he is with Qualcomm, Inc., San Diego, CA, researching on UWB, cognitive radios, and cooperative communications. He has authored more than 45 conference and journal papers and holds more than 40 patents.

Dr. Shankar N was awarded the German Fellowship, DAAD, in the Department of Mathematics, University of Kaiserslautern, Germany, to work on queueing approaches in manufacturing in 1998. He was nominated as one of the five finalists in the Innovator of the Year category by *EETimes*.