EFFECT OF SPECTRUM SENSING OVERHEAD ON PERFORMANCE FOR COGNITIVE RADIO NETWORKS WITH CHANNEL BONDING

Haruki Katayama and Hiroyuki Masuyama
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku
Kyoto 606-8501, Japan

Shoji Kasahara
Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma
Nara 630-0192, Japan

Yutaka Takahashi
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku
Kyoto 606-8501, Japan

Abstract. In cognitive radio networks, secondary spectrum users detect available frequency channels by spectrum sensing. In general, the sensing time is communication overhead, and affects system’s performance. In this paper, we theoretically consider the effect of sensing overhead on the system performance for cognitive radio networks with channel bonding. Specifically, we model the system with a multidimensional continuous-time Markov chain whose state is defined by the numbers of primary users, secondary users, and sensing users. The blocking probability, the forced termination probability and the throughput are derived. The analysis is validated by Monte Carlo simulation. Numerical examples show that the forced termination probability is not affected by sensing overhead, while the blocking probability and the throughput degrade with the increase in the sensing time. It is also shown that the optimal number of bonded sub-channels for the throughput performance significantly depends on the offered load from primary users.

1. Introduction. The rapid development of wireless network services imposes a heavy load on the limited radio spectrum resources. In particular, the explosive growth of smartphones and tablet devices causes the shortage of available frequency channels at an unprecedented pace. One of the solutions for this spectrum shortage is the cognitive radio (CR) technology [1].

In CR networks, radio devices recognize the surrounding radio spectrum environment and effectively use frequency channels without interference with other...
systems. This technology enables secondary users (SUs) to use sufficient radio resources without any change of spectrum allocation for various wireless systems.

In December 2002, the Federal Communications Commission (FCC) of US admitted unlicensed secondary use of TV channels which are not used spatially or temporally. Since then, the IEEE 802.22 Working Group (WG) has developed standards of physical (PHY) and medium access control (MAC) layers of the wireless regional area network (WRAN) using TV channels [3]. The IEEE 802.22 WG also considers specifications for coexistence of TV and wireless microphones.

In a CR network, SUs detect available frequency channels by spectrum sensing. It is well known that the sensing accuracy depends on the spectrum channel bandwidth, the signal to noise ratio (SNR) of a signal detector, and the scan time [11]. Among these we focus on the scan time. In general, when the scan time is large, the sensing accuracy is high but the communication overhead is large. When the scan time is small, on the other hand, the communication overhead is small but the sensing accuracy is low. We call the scan time the sensing time hereafter.

It is reported in [3] that one TV channel is insufficient to meet the IEEE 802.22 requirement. One of bandwidth enhancement schemes is channel bonding, by which an SU bonds some spectrum sub-channels into a broadband channel. The channel bonding scheme is classified into contiguous channel bonding and non-contiguous channel bonding. In this paper, we focus on non-contiguous channel bonding. When the number of sub-channels to be bonded is small, the sensing time is small but the capacity of the bonded channel is small. Contrarily, when the number of sub-channels to be bonded is large, the capacity of the bonded channel is large but the sensing time becomes large.

In this paper, we consider the effect of sensing overhead on the system performance, focusing on the CR network where SUs acquire spectrum channels by non-contiguous channel bonding. Specifically, we model the system with a continuous-time Markov chain whose state space consists of numbers of primary users (PUs), SUs, and sensing users. The blocking probability, the forced termination probability, and the throughput are derived. In numerical experiments, the analysis is validated by Monte Carlo simulation. We also investigate the effect of the parameters such as the arrival rate of PUs and SUs, the sensing time, and the number of bonded sub-channels on the performance measures.

The rest of this paper is organized as follows. Section 2 shows some related work. Section 3 describes the details of channel bonding and spectrum sensing. Section 4 shows the analytical model of the CR network. The model is analyzed in section 5. Section 6 shows numerical examples, and finally, section 7 provides some conclusions and future work.

2. Related work. There exists much literature which is related to spectrum handoff schemes for CR networks.

Kannappa and Saquib consider the CR network which dynamically assigns service rates to the SUs depending on the available network spectral resource [5]. They analyze the trade-off between the call completion probability, the blocking probability and the forced termination probability of SUs, using a two-dimensional Markov chain. It is demonstrated that increasing the maximum allowable service rates offers low forced termination probability.
Zhu et al. [15] propose a channel reservation scheme for spectrum handoff, leading to a great decrease in the forced termination probability at a slight increase in the blocking probability.

Zhang [14] proposes dynamic spectrum access schemes without and with buffering for SUs. He analyzes the blocking probability, the interrupted probability, the forced termination probability, the non-completion probability and the waiting time of SUs by a Markov approach. It is indicated that the buffering significantly reduces the blocking probability with very minor increase in the forced termination probability.

Tumuluru et al. [13] develop two different policies to handle the spectrum assignment and handoff for SUs, considering prioritized SU traffic. The blocking probability, the forced termination probability and the throughput for the two priority classes are analyzed with Markov chain. The authors also investigate the effect of sub-channel reservation for the high priority SUs to obtain the optimal reservation.

Lee and So [9] consider the channel aggregation schemes such as the constant channel aggregation scheme, the variable channel aggregation scheme based on the probability distribution, and the variable channel aggregation scheme based on the residual channel. It is shown that the variable channel aggregation based on the residual channel achieves the best throughput performance.

Jiao et al. [4] investigate the channel bonding/aggregation strategies under the scenarios when primary channels are unslotted. They analyze the blocking probability, the forced termination probability and the system capacity for SUs, using continuous time Markov chain models. It is shown that channel bonding/aggregation degrades the system capacity and the blocking probability, while it improves the forced termination probability.

Konishi et al. investigate how the number of sub-channels to be bonded affects the throughput performance for SUs in CR networks [6]. The authors also consider the case where the number of sub-channels to be bonded is variable [7]. They develop a multiserver priority queuing model, deriving the blocking probability, the forced termination probability and the throughput for SUs.

Note that the above work never considers sensing overhead. It is reported in [8] that sensing overhead becomes critical when spectrum holes are available for very small duration.

In [11], the throughput of SUs is analyzed with a discrete-time Markov model for dynamic spectrum access control systems where sensing errors such as false alarm and misdetection at the PHY layer are taken into consideration. In [10], spectrum handoff and buffering are considered and the throughput performance is analyzed with discrete-time Markov models based on [11]. In the literature, the effect of sensing overhead on system performance is considered, however, the sensing unit is a channel, and the case where some available channels are bonded is not considered.

In [12], the authors consider the CR network where the spectrum is managed with the dedicated control channel and connections are dynamically switched by a software defined radio technology. The authors investigate the effect of two sensing policies on the throughput performance. One is the random sensing policy, where SUs randomly scan the frequency channels, and the other is the negotiation-based sensing policy, where SUs perform sensing based on the frequency usage information by overhearing signals over the control channel. Note that in such networks with cooperative sensing, synchronization causes additional overhead and devices need to be more complicated.
In this paper, we investigate the effect of sensing overhead, considering channel bonding and non-cooperative sensing.

3. **Channel bonding and spectrum sensing.**

3.1. **Channel bonding.** We assume that the spectrum of the system is composed of some primary channels and each primary channel is further divided into sub-channels of equal bandwidth. In the following, we call the users who use primary channels as PUs, and the users who secondarily use sub-channels are referred to as SUs. Each PU uses one primary channel to communicate. On the other hand, each SU requires a certain number of available sub-channels. In order to enhance the channel capacity, we consider channel bonding [3]. In channel bonding, idle sub-channels can be bonded to a broadband spectrum channel. The channel bonding scheme is classified into contiguous channel bonding and non-contiguous channel bonding according to the continuity of bonded frequency channels. In this paper, we focus on non-contiguous channel bonding.

3.2. **Spectrum sensing.** An SU detects available sub-channels in order to avoid interference with PUs. This detection process is called spectrum sensing. The sensing algorithm considered in this paper is as follows. An SU with communication request repeatedly selects and scans a sub-channel until idle sub-channels required by the SU are found. The efficient sensing order was proposed in [2]. However, we assume random order for simplicity. During transmission, each SU periodically senses his/her own sub-channels. When the SU detects a PU arrival, the SU tries to sense other idle sub-channels. If the SU detects idle sub-channels, he/she vacates the sub-channels in use and switches to the detected idle sub-channels. This switching process is called spectrum handoff. If the SU fails in finding idle sub-channels, his/her communication is forced to terminate. We assume that communications of PUs are managed at the base station, and that PUs never sense communication of SUs.

4. **Analytical model.** In this section, we describe the analytical model for the CR network. We assume that the number of primary channels is equal to one in order to avoid state explosion in numerical experiments. Under this assumption, SUs cannot carry out spectrum handoff. We also assume that the primary channel is divided into \( N \) sub-channels.

Each PU arrives at the system according to a Poisson process with rate \( \lambda_p \) and demands one primary channel. When a PU arrives at the system, there exist three possibilities:

(i) if the primary channel is vacant, the request of the arriving PU is accepted;
(ii) if the primary channel is already occupied by a PU, the arriving PU is lost; and
(iii) if any or all of the \( N \) sub-channels are occupied by SUs, the occupied ones are released and then the arriving PU is accepted, i.e., the SUs in transmission are forcibly removed.

The transmission times of accepted PUs are independent and identically distributed (i.i.d.) according to an exponential distribution with rate \( N\mu \), which is independent of the arrivals and transmissions of the other users.

As for SUs, they arrive at the system according to a Poisson process with rate \( \lambda_s \), which is independent of the arrival process of PUs. Each arriving SU demands
r sub-channels and tries to detect idle sub-channels by sensing. For analytical convenience, we assume that the capacity of the “virtual room” for SUs in sensing is equal to \( \lceil N/r \rceil \), which is the maximum number of SUs transmitted simultaneously. Thus, an arriving SU is immediately lost if the number of SUs in sensing is equal to the maximum \( \lceil N/r \rceil \). Otherwise the arriving SU starts the first sensing period to detect an idle sub-channel.

We assume that an SU can find one idle sub-channel, if any, in one sensing period. We also assume that each sensing period of an SU is distributed according to an exponential distribution with rate \( \gamma \), which is independent of those of the other SUs in sensing. Further, we assume that when a PU or SU is accepted, the sensing history of SUs in sensing is completely initialized, i.e., set to the status of “no idle sub-channel detected”.

Under the above assumptions, if an SU detects \( r \) idle sub-channels in some sensing periods (at least \( r \) sensing periods), then the SU immediately occupies the \( r \) sub-channels and requires the transmission time distributed according to an exponential distribution with rate \( r \mu \), which is independent of those of the other users. Recall here that the transmission of an SU may not be completed due to the arrival of a PU. On the other hand, if an SU fails to find an idle sub-channel in some sensing periods, then the SU is lost immediately after the failed sensing period.

5. Analysis. In this section, we first construct a continuous-time Markov chain of the analytical model described in the previous section. We then derive some performance measures for the CR network with multiple primary channels.

5.1. Construction of Markov chain. Let \( N_p(t) \) and \( N_s(t) \) denote the numbers of PUs and SUs, respectively, in transmission at time of \( t \) \( (t \geq 0) \). Let \( \vec{N}_s(t) = (N_s^{(1)}(t), \ldots, N_s^{(r)}(t)) \), where \( N_s^{(k)}(t) (k = 1, 2, \ldots, r) \) denotes the number of SUs who already find \( k-1 \) idle sub-channels and are detecting the \( k \)th idle sub-channel. It follows from the model assumptions (described in the previous section) that

\[
N_p(t) \in \{0, 1\}, \quad \tag{1}
\]

\[
0 \leq N_s(t) \leq \left\lfloor \frac{(1 - N_p(t))N}{r} \right\rfloor, \quad \tag{2}
\]

\[
\sum_{k=1}^{r} N_s^{(k)}(t) \leq \left\lfloor \frac{N}{r} \right\rfloor. \quad \tag{3}
\]

Note here that the total number of idle sub-channels is equal to \( (1 - N_p(t))N - N_s(t)r \). Thus no SU in sensing has found more than \( (1 - N_p(t))N - N_s(t)r \) idle sub-channels, i.e.,

\[
N_s^{(k)}(t) = 0, \quad \text{for all } k \geq (1 - N_p(t))N - N_s(t)r + 2. \quad \tag{4}
\]

It follows from (1)–(4) that the state space of the stochastic process \( \{(N_p(t), N_s(t), \vec{N}_s(t)); t \geq 0\} \), denoted by \( \mathbb{F} \), is given by

\[
\mathbb{F} = \{(i, j, \mathbf{n}); \mathbf{n} \in T_{i,j}, \ i \in \{0,1\}, \ j \in \{0,1,\ldots,\lfloor(1-i)N/r\rfloor\}\},
\]

where

\[
T_{i,j} = \{\mathbf{n} \in \mathbb{Z}_+^r; \sum_{k=1}^r n_k \leq \lceil N/r \rceil, \ n_k = 0 \ (k \geq (1-i)N - jr + 2)\},
\]

\[
\mathbf{n} = (n_1, \ldots, n_r), \ \mathbb{Z}_+ = \{0, 1, 2, \ldots\}.
\]
Clearly, the stochastic process \( \{(N_p(t), N_s(t), \mathcal{N}_s(t))\} \) is an irreducible continuous-time Markov chain with state space \( \mathbb{F} \) due to the model assumptions, such as Poisson arrivals of both PUs and SUs; and exponential transmission times and sensing periods. Thus \( \{(N_p(t), N_s(t), \mathcal{N}_s(t))\} \) has a unique stationary distribution, denoted by \( \pi = (\pi_{i,j,n})_{(i,j,n) \in \mathbb{F}} \).

Let \( Q \) denote the transition rate matrix of \( \{(N_p(t), N_s(t), \mathcal{N}_s(t))\} \). We can then obtain the stationary distribution \( \pi \) by solving the system of equations \( \pi Q = 0 \) and \( \pi e = 1 \), where \( e \) denotes the \( r \times 1 \) vector of ones.

In what follows, we construct the transition rate matrix \( Q \). For this purpose, we first partition \( Q \) as

\[
Q = \begin{pmatrix}
F_0 & F_1 \\
F_0 & F_1
\end{pmatrix}
\begin{pmatrix}
Q^{(0)} & Q^{(1)} \\
Q^{(2)} & Q^{(3)}
\end{pmatrix},
\]

where \( F_\nu = \{(i, j, n) \in \mathbb{F}; i = \nu\} \) (\( \nu = 0, 1 \)). Note here that

\[
(N_p(t), N_s(t)) \in \{0\} \times \{0, 1, \ldots, \lfloor N/r \rfloor \} \cup \{1\} \times \{0\}.
\]

Note also that while \( \{N_p(t)\} \) is taking a value of zero, \( \{N_s(t)\} \) increases or decreases by at most one at any time, i.e., \( |N_s(t) - N_s(t^-)| \leq 1 \). In addition, \( N_s(t) \leq \lfloor N/r \rfloor =: s_0 \) for all \( t \geq 0 \). Therefore \( Q^{(0)} \) has a tridiagonal block structure as follows:

\[
Q^{(0)} = \begin{pmatrix}
F_{0,0} & F_{0,1} & F_{0,2} & \cdots & \cdots & \cdots & F_{0,s_0} \\
F_{0,1} & A^{(0)} & B^{(0)} & O & \cdots & \cdots & O \\
F_{0,2} & O & C^{(1)} & A^{(1)} & B^{(1)} & \cdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
F_{0,s_0} & O & \cdots & \cdots & O & C^{(s_0-1)} & A^{(s_0)}
\end{pmatrix},
\]

where

\[
F_{0,j} = \{(0, j, n): n \in \mathcal{T}_{0,j}\}, \quad j = 0, 1, \ldots, s_0,
\]

and where

\[
A^{(j)} = (A^{(j)}_{n,n'})_{(n,n') \in \mathcal{T}_{0,j} \times \mathcal{T}_{0,j}}, \quad j = 0, 1, \ldots, s_0,
\]

\[
B^{(j)} = (B^{(j)}_{n,n'})_{(n,n') \in \mathcal{T}_{0,j} \times \mathcal{T}_{0,j+1}}, \quad j = 0, 1, \ldots, s_0 - 1,
\]

\[
C^{(j)} = (C^{(j)}_{n,n'})_{(n,n') \in \mathcal{T}_{0,j} \times \mathcal{T}_{0,j-1}}, \quad j = 1, 2, \ldots, s_0.
\]

Further \( Q^{(1)} \) can be partitioned as

\[
Q^{(1)} = \begin{pmatrix}
F_0 & \cdots & \cdots & \cdots & \cdots \\
F_0 & D^{(0)} & \cdots & \cdots & \cdots \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
F_{0,s_0} & \cdots & \cdots & \cdots & D^{(s_0)}
\end{pmatrix},
\]
where \( \mathbf{D}^{(j)} = (\mathbf{D}_{n,n'}^{(j)})_{(n,n') \in \mathbb{T}_{0,j} \times \mathbb{T}_{0,j}} \) for \( j = 0, 1, \ldots, s_0 \). Note here that when \( \{N_p(t)\} \) changes its value from one to zero, the only possible state transition of \( \{(N_p(t), N_s(t))\} \) is the one from \( (1, 0) \) to \( (0, 0) \). Thus \( \mathbf{Q}^{(2)} \) has the following structure:

\[
\mathbf{Q}^{(2)} = F_1 \begin{pmatrix}
F_{0,0} & F_{0,1} & F_{0,2} & \cdots & F_{0,s_0} \\
F_{1,0} & O & O & \cdots & O \\
O & C^{(1)} & A^{(1)} & B^{(1)} & \cdots & D^{(1)} \\
O & O & C^{(2)} & O & \cdots & O \\
\vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\
O & O & \cdots & \cdots & C^{(s_0-1)} & B^{(s_0-1)} \\
E^{(0)} & O & \cdots & O & C^{(s_0)} & A^{(s_0)} & D^{(s_0)} \\
F_{0,0} & F_{0,1} & F_{0,2} & \cdots & F_{0,s_0} & F_1
\end{pmatrix},
\]

where \( \mathbf{E}^{(0)} = (\mathbf{E}_{n,n'}^{(0)})_{(n,n') \in \mathbb{T}_{1,0} \times \mathbb{T}_{0,0}} \). As for \( \mathbf{Q}^{(3)} = (\mathbf{Q}_{n,n'}^{(3)})_{(n,n') \in \mathbb{T}_{1,0} \times \mathbb{T}_{1,0}} \), the matrix does not have a block structure like \( \mathbf{Q}^{(2)} \) \( (j = 0, 1, 2) \) because it corresponds to the transition of \( \{(N_p(t), N_s(t))\} \) from state \( (1, 0) \) to itself. As a result, \( \mathbf{Q} \) can be partitioned as

\[
\mathbf{Q} = \begin{pmatrix}
\mathbf{F}_{0,0} & \mathbf{F}_{0,1} & \mathbf{F}_{0,2} & \cdots & \mathbf{F}_{0,s_0} & \mathbf{F}_{1} \\
\mathbf{F}_{0,0} & \mathbf{F}_{0,1} & \mathbf{F}_{0,2} & \cdots & \mathbf{F}_{0,s_0} & \mathbf{F}_{1} \\
\mathbf{A}^{(0)} & \mathbf{B}^{(0)} & \mathbf{O} & \cdots & \mathbf{O} & \mathbf{D}^{(0)} \\
\mathbf{O} & \mathbf{C}^{(1)} & \mathbf{A}^{(1)} & \mathbf{B}^{(1)} & \cdots & \mathbf{D}^{(1)} \\
\mathbf{O} & \mathbf{O} & \mathbf{C}^{(2)} & \mathbf{O} & \cdots & \mathbf{O} \\
\mathbf{O} & \mathbf{O} & \cdots & \cdots & \mathbf{C}^{(s_0-1)} & \mathbf{B}^{(s_0-1)} \\
\mathbf{E}^{(0)} & \mathbf{O} & \cdots & \mathbf{O} & \mathbf{C}^{(s_0)} & \mathbf{A}^{(s_0)} & \mathbf{D}^{(s_0)} \\
\mathbf{F}_{0,0} & \mathbf{F}_{0,1} & \mathbf{F}_{0,2} & \cdots & \mathbf{F}_{0,s_0} & \mathbf{F}_{1}
\end{pmatrix}.
\]

In the rest of this subsection, we describe the details of the following block matrices of \( \mathbf{Q} \):

\[
\begin{align*}
\mathbf{A}^{(j)} &= (A_{n,n'}^{(j)})_{(n,n') \in \mathbb{T}_{0,j} \times \mathbb{T}_{0,j}}, & j = 0, 1, \ldots, s_0, \\
\mathbf{B}^{(j)} &= (B_{n,n'}^{(j)})_{(n,n') \in \mathbb{T}_{0,j} \times \mathbb{T}_{0,j+1}}, & j = 0, 1, \ldots, s_0 - 1, \\
\mathbf{C}^{(j)} &= (C_{n,n'}^{(j)})_{(n,n') \in \mathbb{T}_{0,j} \times \mathbb{T}_{0,j-1}}, & j = 1, 2, \ldots, s_0, \\
\mathbf{D}^{(j)} &= (D_{n,n'}^{(j)})_{(n,n') \in \mathbb{T}_{0,j} \times \mathbb{T}_{1,0}}, & j = 0, 1, \ldots, s_0, \\
\mathbf{E}^{(0)} &= (E_{n,n'}^{(0)})_{(n,n') \in \mathbb{T}_{1,0} \times \mathbb{T}_{0,0}}, \\
\mathbf{Q}^{(3)} &= (Q_{n,n'}^{(3)})_{(n,n') \in \mathbb{T}_{1,0} \times \mathbb{T}_{1,0}}.
\end{align*}
\]

To avoid repeating the same phrases, we denote by \( (i, j, n) \) and \( (i', j', n') \), the states of the Markov chain before and after respective transitions associated with the above block matrices.

5.1.1. Matrix \( \mathbf{A}^{(j)} \). Matrix \( \mathbf{A}^{(j)} = (A_{n,n'}^{(j)})_{(n,n') \in \mathbb{T}_{0,j} \times \mathbb{T}_{0,j}} \) \( (j = 0, 1, \ldots, s_0) \) corresponds to the event where the process \( \{(N_p(t), N_s(t))\} \) remains in state \( (0, j) \) (i.e., \( i = i' = 0 \) and \( j = j' \)), which is classified into the following three cases:

(i) an arriving SU starts the first sensing period;
(ii) an SU finishes the \( l \)th \( (1 \leq l \leq r - 1) \) sensing period and starts the \( l + 1 \)st sensing period; and
(iii) one of the SU's in the \( l \)th \( (1 \leq l \leq r) \) sensing period fails to find an idle sub-channel and thus is lost.

We first consider case (i). Recall here that an SU arrives according to a Poisson process with rate \( \lambda_s \), and starts the first sensing period (i.e., \( \{N_s^{(1)}(t)\} \) increases by
one) if the total number of SUs in sensing is less than or equal to $s_0 - 1$. Thus,
\[ A^{(j)}_{n, n'} = \lambda_s, \quad n' = n + e_1, \quad |n| \leq s_0 - 1, \]
where $|n| = \sum_{k=1}^{r} n_k$ and $e_k$ ($k = 1, 2, \ldots, r$) denotes the $1 \times r$ unit vector whose $k$th element is equal to one.

We move on to case (ii). Note that one of the SUs in the $l$th ($1 \leq l \leq r - 1$) sensing period finishes its sensing period at rate $n_t \gamma$. The SU finds an idle sub-channel in the $l$th sensing period and starts the next sensing period if there are not less than $l$ idle sub-channels, i.e., $N - jr \geq l$. In this case,
\[ n' = n - e_l + e_{l+1}. \]

Therefore,
\[ A^{(j)}_{n, n'} = n_t \gamma, \quad \begin{cases} n' = n - e_l + e_{l+1}, \\ n_t \geq 1, \quad 1 \leq l \leq \min(r - 1, N - jr). \end{cases} \]

Finally, we consider case (iii). Suppose there are less than $l$ ($1 \leq l \leq r$) idle sub-channels, i.e., $N - jr < l$. In this case, when an SU finishes the $l$th sensing period, the SU fails to find an idle sub-channel in the sensing period and thus is lost. This event occurs at rate $n_t \gamma$ and results in
\[ n' = n - e_l. \]

As a result, we have
\[ A^{(j)}_{n, n'} = n_t \gamma, \quad \begin{cases} n' = n - e_l, \\ n_t \geq 1, \quad N - jr < l \leq r. \end{cases} \]

From the above discussion, the off-diagonal elements of $A^{(j)}$ are given as follows: for $n' \neq n$,
\[ A^{(j)}_{n, n'} = \begin{cases} \lambda_s, \quad n' = n + e_1, \quad |n| \leq s_0 - 1, \\ n_t \gamma, \quad \begin{cases} n' = n - e_l + e_{l+1}, \\ n_t \geq 1, \quad 1 \leq l \leq \min(r - 1, N - jr), \end{cases} \\ n_t \gamma, \quad \begin{cases} n' = n - e_l, \\ n_t \geq 1, \quad N - jr < l \leq r, \end{cases} \\ 0, \quad \text{otherwise}. \end{cases} \]  

(6)

For completeness, we mention the diagonal elements of $A^{(j)}$. Clearly, each of them is negative and its absolute value is equal to the sum of the off-diagonal elements in the corresponding row of the transition matrix $Q$. It thus follows from (5) that
\[ A^{(0)}_{n, n} = - \sum_{n' \neq n, n' \in T_{0,0}} A^{(0)}_{n, n'} - \sum_{n' \in T_{0,1}} B^{(0)}_{n, n'} - \sum_{n' \in T_{1,0}} D^{(0)}_{n, n'}, \]
\[ A^{(s_0)}_{n, n} = - \sum_{n' \neq n, n' \in T_{0,s_0}} A^{(s_0)}_{n, n'} - \sum_{n' \in T_{0,s_0-1}} C^{(s_0)}_{n, n'} - \sum_{n' \in T_{1,s_0}} D^{(s_0)}_{n, n'}, \]
and for $j = 1, 2, \ldots, s_0 - 1$,
\[ A^{(j)}_{n, n} = - \sum_{n' \neq n, n' \in T_{0,j}} A^{(j)}_{n, n'} - \sum_{n' \in T_{0,j+1}} B^{(j)}_{n, n'} - \sum_{n' \in T_{0,j-1}} C^{(j)}_{n, n'} - \sum_{n' \in T_{1,0}} D^{(j)}_{n, n'}. \]
where the elements of $B^{(j)}$, $C^{(j)}$ and $D^{(j)}$ are given in the rest of this subsection. In fact, substituting (6)–(9) into the above equations yields

$$A_{n,n}^{(j)} = \begin{cases} -\lambda_p + \lambda_n + jr\mu + |n|\gamma, & |n| \leq s_0 - 1, \\ -\lambda_p + jr\mu + s_0\gamma, & |n| = s_0. \end{cases}$$

5.1.2. Matrix $B^{(j)}$. Matrix $B^{(j)} = (B_{n,n'}^{(j)})_{(n,n')\in \mathcal{T}_{0,j}\times \mathcal{T}_{0,j+1}} (j = 0, 1, \ldots, s_0 - 1)$ corresponds to the event where one of the $n_r$ SUs in the $r$th sensing period finishes its sensing and the SU occupies $r$ idle sub-channels (i.e., starts its transmission). This event occurs at rate $n_r\gamma$ and results in the initialization of the history of all the SUs in sensing, i.e.,

$$n' = (|n| - 1)e_1.$$

Therefore we have

$$B_{n,n'}^{(j)} = \begin{cases} n_r\gamma, & n' = (|n| - 1)e_1, n_r \geq 1, \\ 0, & \text{otherwise}. \end{cases}$$

(7)

5.1.3. Matrix $C^{(j)}$. Matrix $C^{(j)} = (C_{n,n'}^{(j)})_{(n,n')\in \mathcal{T}_{0,j}\times \mathcal{T}_{0,j-1}} (j = 1, 2, \ldots, s_0)$ corresponds to the event where one of the $j$ SUs in transmission departs from the system due to the completion of its transmission. Since this event occurs at rate $jr\mu$, it follows that

$$C_{n,n'}^{(j)} = \begin{cases} jr\mu, & n' = n, \\ 0, & \text{otherwise}. \end{cases}$$

(8)

5.1.4. Matrix $D^{(j)}$. Matrix $D^{(j)} = (D_{n,n'}^{(j)})_{(n,n')\in \mathcal{T}_{0,j}\times \mathcal{T}_{1,0}} (j = 0, 1, \ldots, s_0)$ corresponds to the event where an arriving PU starts its transmission. This event occurs at rate $\lambda_p$ and results in

$$n' = |n|e_1,$$

which is due to the initialization of the sensing history. As a result,

$$D_{n,n'}^{(j)} = \begin{cases} \lambda_p, & n' = |n|e_1, \\ 0, & \text{otherwise}. \end{cases}$$

(9)

5.1.5. Matrix $E^{(0)}$. Matrix $E^{(0)} = (E_{n,n'}^{(0)})_{(n,n')\in \mathcal{T}_{1,0}\times \mathcal{T}_{0,0}}$ corresponds to the event where the PU in transmission departs from the system due to the completion of its transmission. Since this event occurs at rate $N\mu$, we have

$$E_{n,n'}^{(0)} = \begin{cases} N\mu, & n' = n, \\ 0, & \text{otherwise}. \end{cases}$$

(10)

5.1.6. Matrix $Q^{(3)}$. Matrix $Q^{(3)} = (Q_{n,n'}^{(3)})_{(n,n')\in \mathcal{T}_{1,0}\times \mathcal{T}_{1,0}}$ corresponds to the event where the process $\{(N_p(t), N_s(t))\}$ remains in state $(1, 0)$, i.e., $i = i' = 1$ and $j = j' = 0$. When this event occurs, all the SUs in sensing (if any) are in the first sensing period, i.e., $n = |n|e_1$. The event considered here is classified into the following two cases:

(i) an arriving SU starts its first sensing period; and
(ii) one of the SUs in the first sensing period finishes its sensing period, and is lost because all the sub-channels are busy.
The event of case (i) occurs at rate $\lambda s$ if the total number of SUs in sensing is less than $s_0$ (i.e., $|n| \leq s_0 - 1$). In this case,

$$n' = n + e_1.$$ 

Since $n = |n|e_1$, $|n| \leq s_0 - 1$ is equivalent to $n_1 \leq s_0 - 1$. Therefore,

$$Q^{(3)}_{n,n'} = \lambda s, \quad n' = n + e_1, \quad n_1 \leq s_0 - 1.$$  \hspace{1cm} (11)

On the other hand, the event of case (ii) occurs at rate $n_1 \gamma$ and results in

$$n' = n - e_1, \quad n_1 \geq 1.$$ 

It thus follows that

$$Q^{(3)}_{n,n'} = n_1 \gamma, \quad n' = n - e_1, \quad n_1 \geq 1.$$ \hspace{1cm} (12)

From (11) and (12), the off-diagonal elements of $Q^{(3)} = (Q^{(3)}_{n,n'})_{n,n' \in T_1 \times T_1}$ are given as follows: for $n' \neq n$,

$$Q^{(3)}_{n,n'} = \begin{cases} 
\lambda s, & n' = n + e_1, \quad n_1 \leq s_0 - 1, \\
n_1 \gamma, & n' = n - e_1, \quad n_1 \geq 1, \\
0, & \text{otherwise}.
\end{cases} \hspace{1cm} (13)$$

As for the diagonal elements of $Q^{(3)}$, each of them is negative and its absolute value is equal to the sum of the off-diagonal elements in the corresponding row of the transition matrix $Q$. Thus from (5), we have

$$Q^{(3)}_{n,n} = - \sum_{n' \neq n, n' \in T_1} Q^{(3)}_{n,n'} - \sum_{n' \in T_0} E^{(0)}_{n,n'}.$$ 

Substituting (10) and (13) into the above equation yields

$$Q^{(3)}_{n,n} = \begin{cases} 
-(\lambda s + N\mu + n_1 \gamma), & n_1 \leq s_0 - 1, \\
-(N\mu + s_0 \gamma), & n_1 = s_0.
\end{cases} \hspace{1cm} (14)$$

5.2. Performance measures for SUs. In this subsection, we derive three performance measures for SUs: the throughput, the blocking probability, and the forced termination probability.

We first consider the throughput of SUs, denoted by $T$, which is defined as the mean number of SUs who finish their transmissions and leave the system per unit time in steady state. When there exist $j$ SUs in transmission, one of them finishes its transmission at rate $jr\mu$. Further, the steady-state probability of $j$ ($j \geq 1$) SUs being in transmission is equal to $\sum_{n \in T_0,j} \pi_{0,j,n}$. Therefore,

$$T = \sum_{j=1}^{s_0} rj\mu \sum_{n \in T_0,j} \pi_{0,j,n} = \sum_{(0,j,n) \in T_0} rj\mu \cdot \pi_{0,j,n}. \hspace{1cm} (14)$$

Next, we derive the blocking probability $P_B$, which is defined as the steady-state probability that an SU is not accepted, i.e., fails to occupy $r$ idle sub-channels. To derive the blocking probability $P_B$, we consider the mean number of SUs in transmission who are forcibly removed due to the arrivals of PUs during the unit time in steady state. We denote this mean number by $T_F$.

It follows from the model assumptions that if an PU arrives at the system with $j$ SUs in transmission, then all the $j$ SUs are forcibly removed. Further, the
arrival rate of PUs is equal to $\lambda_p$, and the steady-state probability of $j$ ($j \geq 1$) SUs being in transmission is equal to $\sum_{n \in T_{0,j}} \pi_{0,j,n}$. Thus we have

$$T_F = \sum_{j=1}^{\infty} j \lambda_p \sum_{n \in T_{0,j}} \pi_{0,j,n} = \sum_{(0,j,n) \in \mathcal{F}_0} j \lambda_p \cdot \pi_{0,j,n}. \quad (15)$$

Note here that $\lambda_s(1 - P_B)$ is equal to the mean number of SUs who are accepted during the unit time in steady state, and that $T + T_F$ is equal to the mean number of SUs who depart from the system during the unit time in steady state due to the completion or the forcible termination of their transmissions. Note also that in steady state, the input rate to the system is equal to the output rate from the system. Therefore,

$$\lambda_s(1 - P_B) = T + T_F.$$ 

Combining this with (14) and (15) yields

$$P_B = 1 - \frac{T + T_F}{\lambda_s} = 1 - \frac{1}{\lambda_s} \sum_{(0,j,n) \in \mathcal{F}_0} j(\rho \mu + \lambda_p) \cdot \pi_{0,j,n}. $$

Finally, we consider the forced termination probability $P_F$, which is defined as the steady-state probability that an accepted SU fails to complete its transmission due to the arrival of a PU. Recall here that $\lambda_s(1 - P_B)$ is equal to the mean number of SUs accepted during the unit time in steady state. We then have

$$T_F = \lambda_s(1 - P_B) P_F,$$

from which and (15) it follows that

$$P_F = \frac{T_F}{\lambda_s(1 - P_B)} = \frac{1}{\lambda_s(1 - P_B)} \sum_{(0,j,n) \in \mathcal{F}_0} j \lambda_p \cdot \pi_{0,j,n}. $$

5.3. Approximation to CR network with multiple primary channels. In this subsection, we provide an approximation to performance measures for the CR network with multiple primary channels. For this purpose, we consider a simple combination of $M$ independent copies of our analytical model described in section 4, and approximately evaluate the system performance of the CR network with $M$ primary channels. Figure 1 shows the combination model.

Let $P_B^{(M)}$, $P_F^{(M)}$ and $T^{(M)}$ denote the blocking probability, the forced termination probability and the throughput, respectively, of SUs for the $M$-combination model. It is easy to see that

$$P_B^{(M)} = P_B, \quad P_F^{(M)} = P_F, \quad T^{(M)} = MT.$$ 

The validity of this approximation is discussed in subsection 6.6.

6. Numerical examples. In this section, we show some numerical examples of the performance measures.

6.1. Parameter setting. Table 1 shows the basic parameter setting used in the numerical experiments. These values follow [10]. Here, $\rho_p$ and $\rho_s$ are given by

$$\rho_p = \frac{\lambda_p}{N \mu}, \quad \rho_s = \frac{\lambda_s}{N \mu},$$

$\lambda_s/N$ means the amount of packets brought by SUs per unit time for a sub-channel, and $\mu$ means the service rate of a sub channel. Therefore $\rho_s$ means the offered load on a sub-channel by SUs. According to these parameters, we set $\mu = 25$ [packet/s],
$\lambda_p = 6 \ [\text{packet/s}], \ 1/\gamma = 10^{-4} \ [\text{s}], \ and \ \lambda_s = 60 \ [\text{packet/s}] \ for \ \rho_s = 0.1, \ and \ \lambda_s = 240 \ [\text{packet/s}] \ for \ \rho_s = 0.4, \ and \ \lambda_s = 540 \ [\text{packet/s}] \ for \ \rho_s = 0.9.$

**Table 1.** Basic parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of primary channels ($M$)</td>
<td>1</td>
</tr>
<tr>
<td>Total number of sub-channels ($N$)</td>
<td>24</td>
</tr>
<tr>
<td>Number of sub-channels required by an SU ($r$)</td>
<td>4</td>
</tr>
<tr>
<td>Offered load by PUs ($\rho_p$)</td>
<td>0.01</td>
</tr>
<tr>
<td>Offered load on a sub-channel by SUs ($\rho_s$)</td>
<td>0.1, 0.4, 0.9</td>
</tr>
<tr>
<td>Average size of a packet</td>
<td>5 KB</td>
</tr>
<tr>
<td>Data transmission rate for a sub-channel</td>
<td>1 Mb/s</td>
</tr>
<tr>
<td>Mean sensing time ($1/\gamma$)</td>
<td>100 $\mu$s</td>
</tr>
</tbody>
</table>

### 6.2. Simulation setting
We compare the analysis with Monte Carlo simulation in order to validate the analytical model. There are two differences between the analytical model and the simulation model.

One is the limit of the number of SUs who perform sensing at the same time. In the analytical model, we limit the number of sensing users for the feasibility of numerical calculation. On the other hand, in the simulation, the number of SUs in sensing is not limited.

The other is the behavior of SUs being in sensing. In the analytical model, all the SUs in sensing restart spectrum sensing when a communication request by a PU or an SU is accepted. In the simulation, on the other hand, SUs in sensing never recognizes the beginning of transmission by other users. Instead, an SU who just completes the $r$th sensing checks again whether the $r$ detected sub-channels are available or not. We assume that the processing time of this check is very small and negligible. If there exist sub-channels used by a PU or other SUs, the SU restarts spectrum sensing from the beginning. If all the detected sub-channels are still available, the SU acquires these sub-channels and begins transmission.
Under the setting, the system state is initialized with no users in system, and we generate \(2 \times 10^5\) arrivals of SUs per sample path. Then we observe the behavior of the system after \(1 \times 10^5\) arrivals of SUs, and we calculate the 95% confidence interval of each performance measure.

6.3. **Effect of sensing time on performance.** Figure 2 represents the log-log plot of the blocking probability of SUs against \(\gamma\). It is observed from the figure that the analytical results are larger than the simulation ones when \(\gamma\) is small. However, the discrepancy between the analysis and the simulation decreases with the increase in \(\gamma\). In particular, the analytical results are almost the same as the simulation ones for \(\gamma > 2 \times 10^3\). Remind that in the analytical model, the number of SUs in sensing is limited. When \(\gamma\) is small, the mean sensing time is large and this makes the number of SUs in sensing large, resulting in a large blocking probability. This blocking event doesn’t occur in the simulation. When \(\gamma\) is large, on the other hand, the sensing time is small and the blocking event due to the limitation of the number of SUs in sensing rarely occurs. Therefore the analytical results agree well with the simulation ones.

Figure 3 represents the log-log plot of the forced termination probability of SUs against \(\gamma\). It is observed from the figure that the forced termination probability is constant and insensitive to \(\gamma\) and \(\rho_s\). This is because \(\gamma\) and \(\rho_s\) are not related to the events which affect the SUs in communication, i.e., PU arrivals and the finish of transmission by the SUs.

Figure 4 represents the throughput of SUs against \(\gamma\). In this figure, the horizontal axis is plotted in log scale. It is observed from the figure that the analysis shows the smaller throughput than that of the simulation when \(\gamma\) is small. This is due to the same reason as that in the case of the blocking probability. We also observe that the throughput increases with the increase in \(\gamma\). This is because the forced termination probability is constant while the blocking probability decreases as \(\gamma\) grows.

We can claim from the above results that the forced termination probability is not affected by the sensing time while the blocking probability and the throughput degrade with the increase in the sensing time. We also confirm that this analytical model is appropriate for evaluation of the system performance when \(\gamma\) is greater than \(10^3\).

6.4. **Effect of number of required sub-channels on performance.** Figure 5 represents the blocking probability of SUs against \(r\). It is observed from the figure that the blocking probability exhibits an increasing tendency when \(r\) grows. This result implies that a large \(r\) is not effective for improving the blocking probability. It is also observed from the figure that the blocking probability is slightly improved when \(r\) is the divisor of the total number of sub-channels.

Figure 6 represents the forced termination probability of SUs against \(r\). It is observed from the figure that the forced termination probability decreases with the increase in \(r\). This is because the transmission time for each SU becomes small as the number of required sub-channels becomes large. We also observe that the forced termination probability is insensitive to \(\rho_s\).

Figure 7 represents the throughput of SUs against \(r\). In the figure, the throughput for SUs remains constant when \(\rho_s\) is small, while that exhibits a decreasing tendency with the increase in \(r\) for a large \(\rho_s\). We also observe that for each \(\rho_s\), a large throughput is achieved when \(r\) is in the range from two to four.
6.5. Effect of offered load by PUs on performance. Figure 8 represents the throughput of SUs against $\rho_p$, in cases of $r = 2$ and $8$. It is observed from the figure that the throughput for SUs decreases with the increase in $\rho_p$, as expected. We also observe that the throughput for $r = 2$ is greater than that for $r = 8$ when $\rho_p = 0.01$. 

Figure 2. Blocking probability vs. sensing rate.

Figure 3. Forced termination probability vs. sensing rate.
When $\rho_p$ is large, however, the throughput for $r = 8$ exceeds that for $r = 2$. This is because for small $r$, the transmission time for each SU is large, and most of SUs cannot complete their transmission due to the forcible termination when $\rho_p$ is large.
This result suggests that the optimal value of $r$ to achieve the highest throughput significantly depends on $\rho_p$. 

\textbf{Figure 6}. Forced termination probability vs. number of required sub-channels.

\textbf{Figure 7}. Throughput vs. number of required sub-channels.
6.6. Approximation to CR network with multiple primary channels. In this subsection, we validate the approximation to the CR network with multiple primary channels in subsection 5.3 by Monte Carlo simulation.

In this simulation, each SU senses and acquires $r$ idle sub-channels from $MN$ sub-channels. When a PU arrives and occupies some sub-channels used by an SU, the SU vacates all of the sub-channels in use and restarts sensing. If the SU succeeds in detecting $r$ idle sub-channels again, spectrum handoff occurs and the SU resumes their transmission. If the SU fails in finding $r$ idle sub-channels, on the other hand, the SU is forcibly removed. The rest of the setting is same as that shown in subsection 6.2.

Figure 9 represents the blocking probability of SUs against $r$ for $M = 3$. It is observed from the figure that the analytical results are larger than the simulation ones, and the discrepancy between the analysis and the simulation decreases with the increase in $\rho_s$. Figure 10 illustrates the forced termination probability of SUs against $r$ for $M = 3$. In Figure 10, the analytical results for three $\rho_s$ cases are the same because the forced termination probability is insensitive to $\rho_s$, as explained in Figure 6. We observe from the figure the same tendency as Figure 9. These results imply that the approximation is effective to predict the blocking probability and the forced termination probability when $\rho_s$ is large.

Figure 11 represents the throughput of SUs against $r$ for $M = 3$. It is observed from the figure that the analytical results and the simulation ones exhibit similar tendencies with the increase in $r$. We also observe that the analytical results are in the range from 70% to 100% of the simulation ones for each $r$ and $\rho_s$.

7. Conclusion. In this paper, we focused on the CR network where each SU independently and autonomously carries out spectrum sensing and acquires spectrum channels by non-contiguous channel bonding. We theoretically considered the effect
of sensing overhead on the system performance. Specifically, we modeled the system with a multidimensional continuous-time Markov chain whose state consists of the
numbers of PUs, SUs, and sensing users. The blocking probability, the forced termination probability and the throughput were derived. In numerical experiments, the analysis was validated by Monte Carlo simulation and we investigated the effect of the parameters such as the arrival rate of PUs and SUs, the sensing time, and the number of bonded sub-channels on the performance measures.

Numerical examples showed that the forced termination probability is not affected by sensing overhead, while the blocking probability and the throughput degrade with the increase in the sensing time. It was also shown that the optimal number of bonded sub-channels for the throughput performance significantly depends on the offered load by primary users.

In the multiple primary channel model, we considered a model consisting of independent and identical single-primary-channel models. It is important to extend the single primary channel model to a multiple primary channel model, in which spectrum handoff for secondary users is taken into consideration. However, this extension based on the same Markovian framework is difficult due to state explosion. Developing an analytical model which can scale with the number of primary channels is our future work. It is also important to develop our model to consider the effective mechanism of selecting frequency channels in spectrum sensing.

REFERENCES


Received September 2012; 1st revision December 2012; 2nd revision June 2013.

E-mail address: katayama@sys.i.kyoto-u.ac.jp
E-mail address: masuyama@sys.i.kyoto-u.ac.jp
E-mail address: kasahara@is.naist.jp
E-mail address: takahashi@i.kyoto-u.ac.jp