Dynamic Spectrum Access Algorithms Based on Survival Analysis

Timothy A. Hall, *Member, IEEE*, Anirudha Sahoo, *Member, IEEE*, Charles Hagwood, and Sarah Streett

***Abstract*—In this study, we design and implement two algorithms for dynamic spectrum access that are based on survival analysis. They use a non-parametric estimate of the cumulative hazard function to predict the remaining idle time available for secondary transmission subject to the constraint of a preset probability of successful completion. In addition to theoretical performance analysis of the algorithms, we evaluate them using data collected from a long term evolution band to model primary user activity to demonstrate their effectiveness in real-world scenarios, even at fine time scales. The algorithms are run in different configurations, i.e., they are trained and run on a few combinations of data sets. Our results show that as long as the cumulative hazard functions are fairly similar across datasets, the algorithms can be trained on one dataset and run on that of another without any significant degradation of performance. The algorithms achieve fairly high white space utilization and have a measured probability of interference that is at or below the preset threshold.**

***Index Terms*—Dynamic spectrum access, spectrum sharing, survival analysis, hazard function.**

I. Introduction

DYNAMIC spectrum access (DSA) seems poised to mitigate the problem of spectrum scarcity. In a typical DSA scenario, a primary user (PU) has priority access to a given band. A secondary user (SU) can transmit during unoccupied (idle) periods opportunistically but must vacate when the PU needs the band again. In order to make efficient use of the spectrum in a DSA environment, an accurate and useful model of spectrum occupancy is needed.

Spectrum occupancy refers to whether or not a particular channel or band is occupied. In this paper, we use the term channel to denote the smallest allocable range of frequencies within a particular communications technology, e.g., 180 kHz for Long Term Evolution (LTE). A band is comprised of multiple channels and represents a single service, e.g., there are 50 channels in a 10 MHz LTE uplink band. We model the occupancy of a given channel as a two-state (binary) random process similar to that used by Spaulding and Hagn [2]:

$$X(t) = \begin{cases} 1 \text{ if } P_R(t) > P_{th} \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

where $P_R(t)$ is the signal power observed at the receiver at time t and P_{th} is a threshold value. $X(t) = 1$ represents the occupied state and $X(t) = 0$ represents the unoccupied state.

A. Previous Work

Various models have been proposed in the literature for spectrum occupancy. A two state Discrete-Time Markov Chain (DTMC) has been used to model spectrum occupancy in [3]. However, stationary DTMC models have been found to be inadequate to represent idle and busy periods. Hence, López-Benítez and Casadevall [3] have proposed a time-inhomogeneous DTMC model. Some authors have also used semi-Markov models for spectrum occupancy [4]. This study assumes a general distribution (rather than exponential) for the idle and busy periods of the spectrum. Further, since there are only two states (ON/OFF), the process is also analyzed as an Alternating Renewal Process [4], [5].

Continuous-Time Markov Chain (CTMC) based models have also been used to represent spectrum idle and busy periods. Since some measurement studies have shown that the ON and OFF periods of spectrum are not exponentially distributed, Geirhofer *et al.* [6], [7] and Stabellini [8] have used semi-Markov models for the purpose. Model occupancy of adjacent channels has been modeled as a two-dimensional Markov chain by Gibson and Arnett [9], [10].

Some studies have shown that busy and idle periods of spectrum exhibit negative correlation, i.e., the idle period following a long busy period is typically short and vice-versa [11]. In this study, the authors have proposed time-correlation models for periodic and non-periodic auto-correlation functions.

There have been few models proposed for predicting spectrum occupancy, which is critical to allocating spectrum to the secondary users. The Partially-Observable Markov Decision Process (POMDP) model has been proposed in [12]. The spectrum sharing scheme proposed in [13] is based on prediction of spectrum occupancy by the primary users in terms of the expected remaining OFF time. A two state semi-Markov model proposed in [4] is used to estimate the distribution parameters of ON/OFF periods. Some methodologies proposed in the literature indirectly predict spectrum occupancy by limiting the duration of transmission of the secondary user (SU) to some constraint. In [14], the transmission duration

Manuscript received May 1, 2017; revised September 15, 2017; accepted October 28, 2017. Date of publication November 9, 2017; date of current version December 22, 2017. Parts of this work were presented at ICC 2017 [1]. The associate editor coordinating the review of this paper and approving it for publication was O. Holland. *(Corresponding author: Timothy A. Hall.)*

T. A. Hall and A. Sahoo are with the Wireless Networks Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 USA (e-mail: tim.hall@nist.gov; anirudha.sahoo@nist.gov).

C. Hagwood is with the Statistical Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 USA (e-mail: charles.hagwood@nist.gov).

S. Streett was with the Statistical Engineering Division, National Institute of Standards and Technology, Boulder, CO 80305 USA.

Digital Object Identifier 10.1109/TCCN.2017.2771761

of an SU is based on the maximum bound on probability of interference to the primary user (PU). Residual idle time of an Alternating Renewal Process is used in [5] to indirectly predict reappearance of the PU. Some researchers have used a Restless Multiarm Bandit formulation for opportunistic channel access [15], [16]. Researchers have also looked at pattern mining of spectrum occupancy data to predict channel availability [17], [18].

B. Motivation for Present Work

The motivation behind the present work is threefold. We want to develop a prediction scheme that is robust, flexible and useful even for very fine time scales. We assume centrally co-ordinated scheduling for the SUs. The scheduler knows when the primary user is no longer active, and when an SU requests a transmission opportunity, the scheduler grants or denies the SU request. Our scheme is not limited to a centralized scheduling architecture, however. It can be used in a carrier sense multiple access (CSMA) system as well. In such a system, the SUs would sense the channel and use our algorithms to predict residual idle time before transmitting as a form of collision avoidance. Analysis and application of prediction schemes presented in this paper to a CSMA based system is beyond the scope of this study.

Most of the stochastic based schemes in the literature either assume a certain distribution (e.g., exponential) of spectrum occupancy data or require that a distribution be fitted to a set of observed data. This study does not have such a requirement. It uses a non-parametric estimate of the cumulative hazard function from historical data to grant dynamic access to the SUs. Hence, our scheme is much simpler to implement in practice.

Finally, most of the DSA schemes in the literature are run over simulated spectrum occupancy data. We ran our algorithms over real spectrum occupancy data to show that they are suitable for implementation on practical systems. We collected occupancy data in the uplink of LTE Band 17, which is centered at 709 MHz with a 10 MHz bandwidth. From these collected datasets we observed that there was a useable amount of white space available in the LTE uplink during peak hours (e.g., 3 PM to 4 PM). During off-peak hours (e.g., 3 AM to 4 AM), the band was very rarely busy. We believe it is possible to exploit white space in the LTE uplink, especially during off-peak hours, for applications that require small and non-delay sensitive data transmissions (e.g., utility meter reading). Hence, we show the effectiveness of our DSA algorithms over this LTE band.

We also envision that our scheme (or some variation thereof) may be used in the Spectrum Sharing architecture proposed in the 3.5 GHz band [19]. In this architecture, there will be three tiers of users in the band. First tier users have the highest priority, but they use the band infrequently. The tier two users, called Priority Access Layer (PAL) users, will be LTE carriers and have medium priority. When tier 1 and tier 2 users are not present in the band, it can be used by tier 3 users called General Authorized Access (GAA) users. It is conceivable that a PAL user can *sell* its white space (idle time) to users who can make use of transmission opportunities on the order of milliseconds as long as the interference to PAL users remains below an agreed threshold. While our scheme is not specific to a particular type or class of SUs, we anticipate increased future traffic from Internet of things (IoT) devices and machine-to-machine (M2M) communications. The result is a large number of entities that each have small, infrequent, non-delay-sensitive data transmissions and can make use of even brief PU idle periods. These opportunistic users can implement our scheme to exploit PAL white spaces.

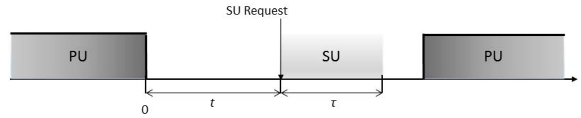

Fig. 1. SU request.

Let us now define the prediction problem upon which our DSA algorithms are based more precisely. We are concerned with how long the channel has been unoccupied by the PU and how much longer the channel will remain unoccupied. Specifically, given that the channel has been unoccupied by the PU for duration t when a request from an SU arrives to transmit for a duration τ, what is the probability that the SU will be able to complete the transmission before the PU appears on the channel? Figure 1 illustrates the relationship between the PU and SU.

The remainder of this paper is structured as follows. Section II formulates the prediction problem in terms of survival analysis, resulting in two algorithms for secondary channel requests. Section III presents proofs that the probability of interference threshold is also an upper bound of the overall probability of SUs and the PU experiencing interference. Section IV describes the collected data, simulation environment and metrics we used to evaluate the algorithms. Section V presents our results. Section VI interprets the results and discusses future work.

II. Prediction Algorithms

A. Survival Analysis

Survival analysis has been used to analyze statistical properties of the duration of time until an event, such as failure in a mechanical system, occurs [20]. Our prediction problem can be solved by using survival analysis as presented below.

Let $T_1, S_1, T_2, S_2, \ldots,$ represent the successive idle and busy periods of the spectrum. Thus T_i and S_i represent the i^{th} idle and busy periods respectively. The T_i's can be thought of as survival times. That is, an idle period survives only until the channel becomes busy again. Let random variable T represent an arbitrary survival time and $0 < p < 1$ an adjustable parameter. Assuming the T_i's are independent and identically distributed as T, our prediction problem can be represented by the hypothesis test given by

$$\begin{aligned} \mathcal{H}_0 &: P[T \geq t+\tau \mid T \geq t] > p \quad \textit{versus} \\ \mathcal{H}_1 &: P[T \geq t+\tau \mid T \geq t] \leq p \end{aligned} \tag{2}$$

$\mathcal{H}_0$ holds if the idle period, having lasted t units of time, lasts τ more units of time with probability greater than p. Note that p represents the probability of successful transmission for duration τ, given that the channel has been idle for duration t.

The basic functions of survival analysis are the survival function and the hazard function. The survival function at time t is the probability of surviving at least t units of time and is given by

$$S(t) = P[T \geq t] = 1 - F(t) = \int_t^{\infty} f(s)ds \tag{3}$$

where $f(s)$ and $F(t)$ are the probability density function and cumulative distribution function of T, respectively. The hazard function is the probability of instantaneous failure at time t given survival up to time t and indicates the risk of failure at time t. The hazard function is given by

$$\begin{aligned} h(t) &= \lim_{\delta t \to 0} \frac{P[t \leq T < t + \delta t \mid T \geq t]}{\delta t} \\ &= \lim_{\delta t \to 0} \frac{P[t \leq T < t + \delta t]}{P[T \geq t] \cdot \delta t} \\ &= \frac{1}{P[T \geq t]} \cdot \lim_{\delta t \to 0} \frac{P[t \leq T < t + \delta t]}{\delta t} \\ &= \frac{f(t)}{S(t)} \end{aligned} \tag{4}$$

From (3), it is clear that the derivative of $S(t)$ is $-f(t)$. Hence, (4) can be rewritten as

$$h(t) = -\frac{d}{dt} log S(t) \tag{5}$$

Now integrating both sides of (5) from 0 to t, noting that $S(0) = 1$ and finally taking the exponential on both the sides, we have

$$S(t) = \exp\left(-\int_0^t h(s)ds\right) \tag{6}$$

The function important to us is the cumulative hazard function, defined by $H(t) = \int_0^t h(s)ds, t \geq 0$. Using (6) we have

$$\begin{aligned} P[T \geq t + \tau \mid T \geq t] &= \frac{P[T \geq t + \tau]}{P[T \geq t]} \\ &= \exp\left(-\int_0^{t+\tau} h(s) + \int_0^t h(s)ds\right) \\ &= \exp(-[H(t + \tau) - H(t)]). \end{aligned} \tag{7}$$

Thus, using (7) the hypotheses in (2) can be expressed as

$$\begin{aligned} &\mathcal{H}_0 : \exp(-[H(t + \tau) - H(t)]) > p \quad \textit{versus} \\ &\mathcal{H}_1 : \exp(-[H(t + \tau) - H(t)]) \leq p \end{aligned} \tag{8}$$

Having observed a large sample $T_1, T_2, \ldots, T_n$ of n survival times, a non-parametric estimate of the survival function can be computed using the empirical distribution function, $F_n(t)$ of the data $T_i, i = 1, \ldots, n$, as shown below.

$$S_n(t) = 1 - F_n(t) = 1 - \frac{1}{n} \sum_{i=1}^{n} 1_{T_i < t} \tag{9}$$

where 1_A is the indicator function for event A.

Let $T_{(1)} \leq T_{(2)} \cdots \leq T_{(n)}$ be the ordered $T_i, i = 1, \ldots, n$. Then the survival function at any $T_{(i)}$ can be computed using (9) as follows.

$$\begin{aligned} S_n(T_{(i)}) &= 1 - \frac{1}{n} \sum_{j=1}^{n} 1_{T_j < T_{(i)}} \\ &= 1 - \frac{1}{n} \cdot (i - 1) = \frac{n - i + 1}{n} \end{aligned} \tag{10}$$

In the above derivation, we used the fact that exactly $(i - 1)$ values of T_i are strictly less than $T_{(i)}$. The empirical density function is

$$f_n(T_{(i)}) = \frac{1}{n} \quad \text{and zero elsewhere} \tag{11}$$

because mass $1/n$ is put at each $T_{(i)}, i = 1, \ldots, n$. Since $h(t) = f(t)/S(t)$ from (4), using (10) and (11) we get the following empirical estimate of $h(t)$

$$\begin{aligned} h_n(T_{(i)}) &= \frac{1}{n - i + 1} \text{ for } i = 1, 2, \ldots, n \\ h_n(t) &= 0 \text{ for all other } t \end{aligned}$$

Using the definition of the cumulative hazard function, an estimate of it is

$$H_n(t) = \sum_{i:T_{(i)} \leq t} \frac{1}{n - i + 1} \tag{12}$$

Our test statistic is based on the difference of the cumulative hazard function at two different times. An estimate for the difference of the cumulative hazard function at two different times is given by

$$H_n(t + \tau) - H_n(t) = \sum_{i:t \leq T_{(i)} \leq t+\tau} \frac{1}{n - i + 1} \tag{13}$$

Note that this is a form of the well-known Nelson-Aalen estimator for the cumulative hazard function. We used a more general form of $H_n(t)$ to account for duplicate values of $T_{(i)}$, that is, multiple idle times of the same duration [21]. Therefore, after simple manipulation of $\mathcal{H}_0$ in (8), our prediction algorithms are formulated in terms of an approximate test statistic,

$$\text{Reject } \mathcal{H}_0 \text{ if : } H_n(t + \tau) - H_n(t) \geq (-\ln p). \tag{14}$$

B. Definition of Algorithms

Below are two formulations of the prediction algorithm. Algorithm 1 is a request to transmit on a channel for duration τ. If the channel is occupied at the time of request, the request is denied. If the channel is not occupied, then the algorithm grants the request if it determines that the probability of a successful transmission (i.e., the probability of completing the transmission without colliding with the PU) is above a given threshold.

Algorithm 2 returns the longest estimated duration available for transmission for a request made at a particular time. The time returned is the largest value for which the probability of successful transmission exceeds the given threshold or, equivalently, the largest τ in $[0, T_{(n)}]$ that satisfies $H_n(t_0 + \tau) - H_n(t_0) < \theta$. This τ can be found using a binary search over the n ordered idle times used to compute $H_n(t)$, $\{T_{(i)}\}$. Alternatively, a closed-form expression for τ can be derived as follows.

We want to find the largest τ such that

$$H(t_0 + \tau) - H(t_0) < \theta.$$

Since $H(t)$ is monotonically non-decreasing, and t_0 and τ are positive, we can find the optimal τ by solving

$$H(t_0 + \tau) - H(t_0) = \theta. \tag{15}$$

Algorithm 1 Request Channel for τ Seconds

input:
τ - the transmit duration requested
parameters:
$H_n(t)$ - the estimated cumulative hazard function
t_0 - the time elapsed since end of last transmission
p - the probability of successful transmission
output: *Grant* or *Deny*

```
if occupied then
    return Deny
end if
θ := −ln p
W_n := H_n(t_0 + τ) − H_n(t_0)
if W_n < θ then
    return Grant
else
    return Deny
end if
```

Algorithm 2 Request Maximum Channel Availability

parameters:
$H_n(t)$ - the estimated cumulative hazard function
$\{T_{(i)}\}$ - the n ordered idle times used to compute $H_n(t)$
t_0 - the time elapsed since end of last transmission
p - the probability of successful transmission
output: τ - the maximum transmit time available now

```
if occupied then
    return 0
end if
θ := −ln p
Find largest τ in [0, T_(n)] such that H_n(t_0 + τ) − H_n(t_0) < θ
return τ
```

Recall that from (7)

$$e^{-(H(t_0+\tau)-H(t_0))} = \frac{P[T > t_0 + \tau]}{P[T > t_0]} = \frac{\bar{F}(t_0+\tau)}{\bar{F}(t_0)}, \quad (16)$$

where $\bar{F}(t) = 1 - F(t)$, and $F(t)$ is the cumulative distribution function of T as defined above. By raising e to the negative power of both sides of (15), we get

$$e^{-(H(t_0+\tau)-H(t_0))} = e^{-\theta}. \quad (17)$$

Then, using the relation in (16),

$$\frac{\bar{F}(t_0+\tau)}{\bar{F}(t_0)} = e^{-\theta}. \quad (18)$$

Recalling that $\theta = -\ln p$, we have

$$\frac{\bar{F}(t_0+\tau)}{\bar{F}(t_0)} = p. \quad (19)$$

Using the quantile function $q_F(\alpha) = F^{-1}(\alpha)$, we solve (19) to find the optimal value of τ,

$$\begin{aligned} \bar{F}(t_o+\tau) &= p\bar{F}(t_0) \\ 1 - F(t_o+\tau) &= p\bar{F}(t_0) \\ F(t_o+\tau) &= 1 - p\bar{F}(t_0) \\ t_0 + \tau &= q_F\big(1 - p\bar{F}(t_0)\big) \\ \tau &= q_F\big(1 - p\bar{F}(t_0)\big) - t_0. \end{aligned} \quad (20)$$

Thus, one can use $\hat{\tau}_n = q_{F_n}\big(1 - p\bar{F}_n(t_0)\big) - t_0$ as an estimate of the maximum transmit time τ in Algorithm 2. For any value $b, 0 \leq b \leq 1$, $q_{F_n}(b)$ is $100b^{th}$ sample quantile of the data $\{T_1, T_2, \ldots, T_n\}$. In the expression for $\hat{\tau}_n$, $b = 1 - p\bar{F}_n(t_0)$.

III. Probability of Interference

Having defined the algorithms, we now prove that $(1-p)$ defines an upper bound on the overall probability of interference to both the PU and the SUs. This distinguishes Algorithms 1 and 2 from other approaches to DSA in that they are tunable by a well-defined, intuitive and critical system parameter. These statistical guarantees are important and useful when defining service contracts between tiers in a spectrum sharing system. The value of p will likely be set by either regulation or by contract.

We first show that both algorithms provide a statistical guarantee on the probability of interference to both the PU and the SUs. Both algorithms have a tunable parameter p that is the requested minimum probability of successful transmission for a given request. Thus, $(1-p)$ is the maximum allowable probability of interference (PoI) for each individual SU request for spectrum.

Note that p and, by extension, $(1-p)$ are run-time parameters that are independent of the estimate of the cumulative hazard function. That is, they are independent of the training data set. In general, p can be different from one request to another, though we anticipate p will be fixed for a given service based on the PU and SU requirements.

Although p and $(1-p)$ are per-request probabilities, we can show that if they remain constant over a period of time, then they serve as lower and upper bounds on the *overall* probabilities of success and interference, respectively, during that time period. Specifically, $(1-p)$ is an upper bound on both the overall probability of an SU transmission experiencing interference, and $(1-p)$ is an upper bound on the overall probability of a PU transmission experiencing interference. We prove both of these assertions below. In our proofs, we assume that the estimated cumulative hazard function of the PU idle time lengths, $H_n(t)$, is equal to the true cumulative hazard function, $H(t)$.

That $(1-p)$ is an upper bound on the probability of SU transmissions is intuitive, and its proof is straightforward. By the definition of both Algorithm 1 and Algorithm 2, $(1-p)$ is always greater than or equal to the probability of any individual granted SU request being interfered with. There are three cases of SU activity during a PU idle period. The first is that no SU requests are granted during the idle period. No interference can occur in this case. The second case is that exactly one SU request is granted in an idle period. The probability that it can be interfered with is less than or equal to $(1-p)$. The third case is that two or more SU requests are granted in an idle period. If this happens, then by the definition of the algorithm only the last SU request granted can be interfered with, with probability less than or equal to $(1-p)$. Thus, the overall probability of interference to the SU is less than or equal to $(1-p)$.

We now show that $1-p$ in Algorithm 1 is an upper bound on the overall probability of a PU being interfered with. A PU transmission starts after an idle period of length T. For any given idle period T preceding a PU transmission, interference to the PU occurs when the following two conditions are true:

1) an SU is granted permission to transmit at time t_0 for a duration τ and
2) $t_0 + \tau > T$

During any given idle period preceding a PU transmission, there are two cases of SU activity. The first is that no SU transmission requests are granted. In this case, the probability of the PU being interfered with is zero. The second case is that one or more SU transmission requests are granted. In this case, by definition, only the last (latest in time) granted SU request can interfere with the PU. The probability of the final granted SU transmission during an idle period interfering with the PU is

$$P[t_0 + \tau > T]. \tag{21}$$

From the definition of Algorithm 1, we know that if this last request was granted, then the following must be true: $W_n < \theta$ and $T > t_0$. Therefore, the proof proceeds as follows:

$$\begin{aligned}
W_n &< \theta \qquad\qquad (22)\\
H(t_0+\tau) - H(t_0) &< \theta \quad \text{// Definition of } W_n\\
e^{-[H(t_0+\tau)-H(t_0)]} &> e^{-\theta}\\
e^{-[H(t_0+\tau)-H(t_0)]} &> p\\
\frac{P[T \geq t_0+\tau]}{P[T \geq t_0]} &> p \quad \text{// From Equation (7)}\\
P[T \geq t_0+\tau] &> p \quad \text{// Since } P[T > t_0] = 1\\
1 - P[t_0+\tau > T] &> p\\
P[t_0+\tau > T] &< 1-p \qquad\qquad (23)
\end{aligned}$$

Thus, in both cases, the probability of interference is less than or equal to $1-p$, making the overall probability of interference to the PU necessarily less than or equal to $1-p$. Though we developed the proof for Algorithm 1, it is also valid for Algorithm 2, as the condition on τ is equivalent to the first statement of our proof, $W_n < \theta$.

IV. Evaluation

Our algorithm is non-parametric; thus, to demonstrate its effectiveness we used a real PU signal, not a modeled one that follows a distribution. We used captured LTE uplink spectrum occupancy data to represent our PU. However, the requirements for SU traffic are different. We need a well-understood and easily interpreted SU traffic model in order to evaluate its performance. Therefore, we assume there is a group of undetermined number of SUs whose aggregate requests for spectrum usage can be modeled as a Poisson arrival process. That is, the inter-arrival times of requests from the SUs are exponentially distributed.

A. Data Collection

Data was collected in Band 17, a 10 MHz uplink (UL) LTE band centered at 709 MHz. A small 10.78 cm rubber duck antenna was connected to an Ettus Universal Sofware Radio Peripheral (USRP)[1] running USRP hardware driver (UHD) version 003.009.001 and GNU Radio version 3.7.9rc1. The sampling rate was 12.5 MHz, resulting in one complex-valued sample, with in-phase and quadrature (I/Q) components every 80 ns. Every 50 μs, 625 consecutive I/Q samples were used to compute a power spectrum over that interval. The power spectra of 20 consecutive 50 μs periods were averaged, and the coefficients were binned to produce a 56 point power spectrum for each 1 ms period. Each power spectrum coefficient is an 8 bit signed integer representing a decibel (dB) value rounded to the nearest integer. Each coefficient corresponds to the power in dB over a 180 kHz range. The middle 50 coefficients correspond to the 50 LTE channels.

An LTE resource block (RB) is 180 kHz wide in frequency and is one slot (0.5 ms) in duration. Thus, our collected dataset consists of an integer power value, in dB, for each pair of RBs. RBs are allocated in pairs, and there are two RBs per LTE subframe (1.0 ms). In [1] we used the same data collection procedure but added an additional step, taking the peak power value over 100 ms. The higher resolution (in time) dataset in this paper lets us evaluate our algorithms on a different time scale that better captures the dynamics of an LTE system.

We applied a noise threshold power value to produce a binary occupancy sequence for each of the 50 channels. The decision threshold was determined as follows. We connected a matched-load terminator to the receiver port of the USRP and collected samples for a one hour period. We found the level at which 1 % of the sample values were above the threshold, i.e., a 1 % probability of false alarm (PFA) [22]. We set this as our internal noise threshold. When we connected the antenna, we also found significant external noise, resulting in an additional 3 dB being added to the threshold.

Data was collected for four different one hour periods (dataset name in parentheses):

- 3:00 PM to 4:00 PM local time, Monday, August 28, 2017 (3pm_day1)
- 3:00 AM to 4:00 AM local time, Tuesday, August 29, 2017 (3am_day1)
- 3:00 PM to 4:00 PM local time, Tuesday, August 29, 2017 (3pm_day2)
- 3:00 AM to 4:00 AM local time, Wednesday, August 30, 2017 (3am_day2)

We chose these four so that we could compare the same time period on two separate days for two levels of expected user activity (i.e., relatively high activity from 3 PM to 4 PM on a weekday and low activity between 3 AM and 4 AM).

B. Simulation

As stated above, an idle period of the spectrum occupancy is a set of one or more consecutive zeros. Each zero represents

[1]The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

an idle period with a duration of one sampling interval (1 ms for our experiments). Thus, the T_i values (in terms of sampling interval) are represented as the number of consecutive zeros. Similarly, a busy period is a set of one or more consecutive ones. In our experiments, we have used the occupancy of LTE uplink channel number 5 as our PU traffic. After building the idle and busy periods, we then compute the cumulative hazard function as per (13). We set the probability of successful transmission (p) and, thus, the interference threshold, which is equal to $(1 - p)$. Note that the PU expects its measured probability of interference (PoI) to be less than this preset interference threshold as well, as we have shown in Section III.

We have evaluated the performance of Algorithm 1 and Algorithm 2 in different configurations as described below. The configurations are denoted as *time_train_run*, where *time* is the start time of both datasets, *train* is the day of the dataset used for training the algorithm (i.e., the cumulative hazard function is built using this data) and *run* represents the day of the dataset which is used to run the algorithm.

As an example, in configuration 3*pm_day*1_*day*1, the algorithms are trained using 3*pm_day*1 data, i.e., the cumulative hazard function is built using 3*pm_day*1 data and then the algorithm is also run on 3*pm_day*1 data. Results from this configuration validate the effectiveness of survival analysis for opportunistic spectrum access.

When using configuration *3pm_day1_day2* the algorithms are trained using 3pm_day1 data but run on 3pm_day2 data. This configuration helps us understand how the algorithms perform when the training and running data are from different week days. Note that in practice, the *3pm_day1_day1* configuration does not correspond to a realistic scenario, since the training has to happen on some historical data and then the algorithm would run on different data. Hence, this configuration is useful for analysis only.

As stated above, SU traffic is modeled as a Poisson arrival process. We test Algorithm 1 using a fixed transmission length request τ. This models SU traffic as a large number of devices each of which has short, infrequent, non-delay-sensitive data transmissions such as Internet of things (IoT) devices and machine-to-machine (M2M) communication. Such devices can transmit opportunistically given even brief PU idle periods. In Algorithm 2, instead of a particular transmission time requested, the maximum available transmission time that meets the PoI threshold is returned. This can be useful for an SU that serves as an aggregator of small messages or is able to transmit large data in multiple chunks.

C. Metrics

We used the following metrics to measure performance of the two algorithms. The first three metrics are common to both the algorithms whereas the remaining four are defined for Algorithm 1 only.

- White Space Utilization (WSU): Given the spectrum occupancy of a channel, White Space Utilization (WSU) of the channel by the secondary users is defined as the fraction of total idle time used by the SUs for their own transmissions. In another words, it is the ratio of the total duration of idle time used by the SUs for their own transmissions to the total PU idle time duration in the channel.
- Probability of Interference (PoI): For a given channel, the PoI of the SUs is defined as the probability that a transmission of an SU collides with that of the PU. Thus, it is the ratio of the number of times an SU transmission collides (or runs into a busy period) with a PU transmission to the total number of SU transmissions over a statistically long observation period.
- Percentage Overlap of SU Transmission (POST): This is the duration of the overlap of SU transmissions with PU transmissions expressed as a percentage of total SU transmission time.
- Desirable Accept Ratio (DAR): This is defined as the fraction of requests that were accepted, and the corresponding transmissions were successful. In these cases the algorithm correctly predicted the remaining idle time.
- Undesirable Accept Ratio (UAR): This is defined as the fraction of requests that were accepted, and the corresponding transmissions were not successful, i.e., these transmissions resulted in a collision with a PU transmission. In these cases the algorithm incorrectly predicted the remaining idle time.
- Desirable Reject Ratio (DRR): This is defined as the fraction of requests that were rejected and would have resulted in collision with the PU if they were accepted. So, in these cases the algorithm correctly predicted the remaining idle time and rejected the requests.
- Undesirable Reject Ratio (URR): This is defined as the fraction of requests that were rejected and would have resulted in successful transmission if they were granted. In these cases the algorithm incorrectly predicted the remaining idle time and rejected the requests. This metric represents lost opportunities for the SU.

D. Baseline WSU for Algorithm 1

Since the SU requests are modeled as a Poisson arrival process with a fixed τ, our WSU will depend in part on the offered load defined by these parameters. We need a benchmark WSU that accounts for this in order to evaluate the performance of Algorithm 1. We can do this by computing the WSU for a simpler system. First, we treat the segments of PU idle time as one long continuous stretch of time. Second, we assume that all SU requests, that arrive when no other SU is transmitting, are granted. In other words, we do not test that W_n is below θ before granting the request. Of course, SU requests that arrive when another SU is transmitting will be denied (dropped). This gives us a bound for a maximum achievable WSU for a given inter-arrival time and transmission time τ.

We can model this simpler system using an $M/D/1/1$ queue, which is a single server Erlang loss queue (i.e., zero length buffer) with deterministic service time (i.e., fixed $\mu = \frac{1}{\tau}$). If the SU requests follow a Poisson arrival process of rate λ, where λ is the inverse of the average inter-arrival

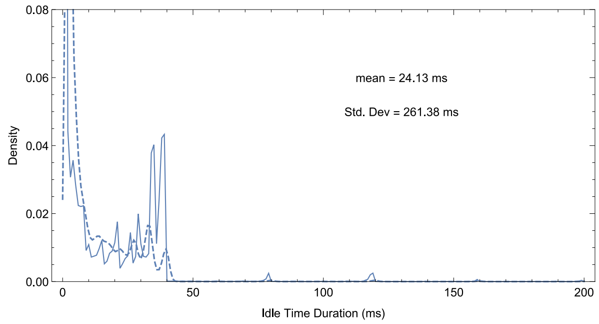

Fig. 2. Idle time duration density of 3pm_day1 dataset discretized to 1 millisecond. Dashed line is a smooth kernel density approximation. The large standard deviation is a result of the density having a heavy tail. Tail part > 200 ms not shown.

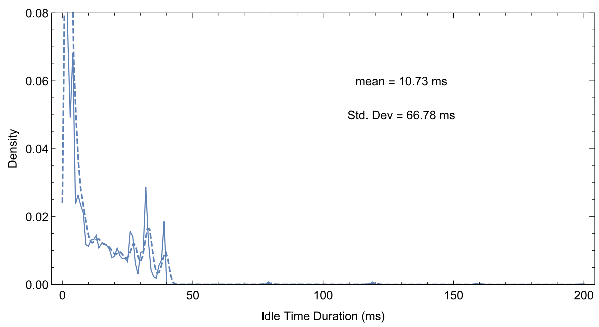

Fig. 3. Idle time duration density of 3pm_day2 dataset discretized to 1 millisecond. Dashed line is a smooth kernel density approximation. The large standard deviation is a result of the density having a heavy tail. Tail part > 200 ms not shown.

times, we can compute the following quantities [23]:

$$\begin{aligned} \rho &= \lambda\tau && \text{(offered load)} \\ P_B &= \frac{\rho}{1+\rho} && \text{(blocking probability)} \\ U &= \rho(1-P_B) && \text{(utilization)} \end{aligned} \tag{24}$$

The last quantity gives us an upper bound for our Algorithm 1 simulation WSU results.

Algorithm 2 returns the largest τ value that still meeets the probability of interference threshold. Thus, it depends on the distribution of PU idle times. It is not as straightforward to determine analytically the distribution of the SU transmission times (i.e., τ values) that would be needed to model it as an M/G/1/1 queue, which is left as future work.

V. Results

A. Hazard and Cumulative Hazard Function

Before we present the results of our algorithms, we want to present the hazard and cumulative hazard functions of idle time durations of the collected data, which will help in understanding the results better. The cumulative hazard function is used directly in Algorithms 1 and 2; however, it is also useful to look at the corresponding instantaneous hazard function, since it is more widely known and often more easily interpreted. For example, the characteristic "bathtub" shape often found in many other applications is visible in Figure 8,

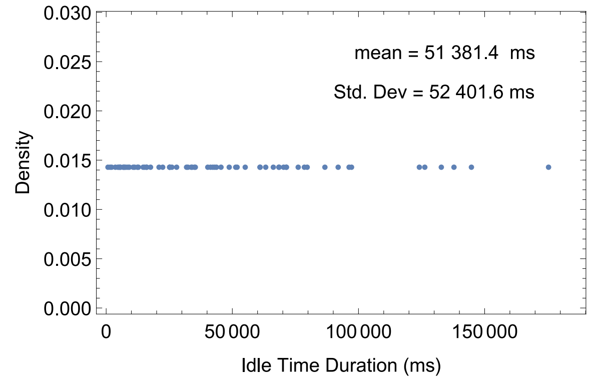

Fig. 4. Idle time histogram of 3am_day1 dataset.

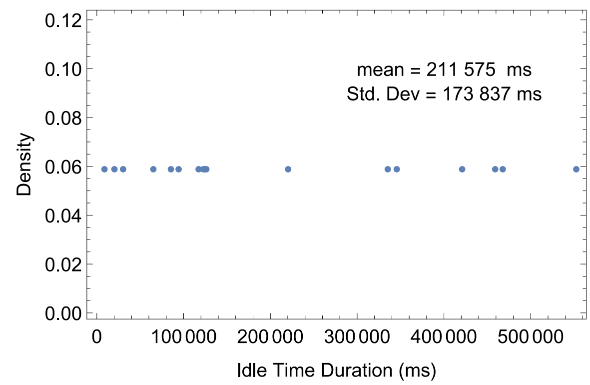

Fig. 5. Idle time histogram of 3am_day2 dataset.

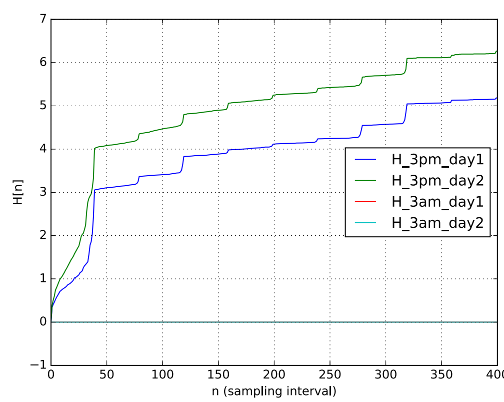

Fig. 6. Cumulative Hazard Function of Idle time Lengths (up to 400 sampling intervals).

most prominently between the values of 1 ms and 40 ms, but also repeated over successive 40 ms intervals.

Figure 2 and 3 show the idle time duration density of 3pm_day1 and 3pm_day2 dataset respectively. The 3pm_day1 dataset has more number of long idle durations below 50 ms. Hence, the hazard function of 3pm_day1 has values lower than 3pm_day2 (Figure 8) in this time range. Idle time duration density presented Figure 2 and 3 show that on both the days most of the idle times are of very long duration at 3am. Hence, the corresponding hazard funtions are zero for the initial part for both the days. Since the range of idle periods is very large, we show the $H(\cdot)$ function up to 400 sampling intervals (400 ms) in Figure 6, whereas Figure 7 covers the

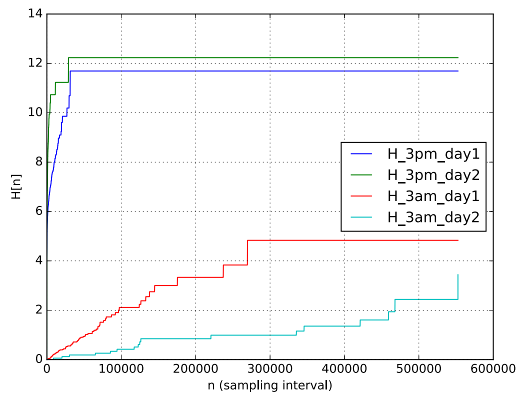

Fig. 7. Cumulative Hazard Function of Idle Time Lengths (for the entire range of idle period).

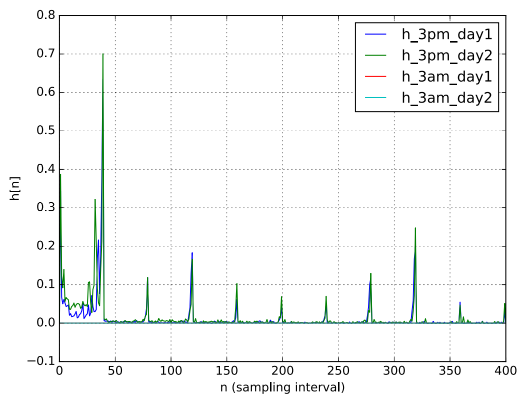

Fig. 8. Hazard Function of Idle Time Lengths (up to 400 sampling intervals).

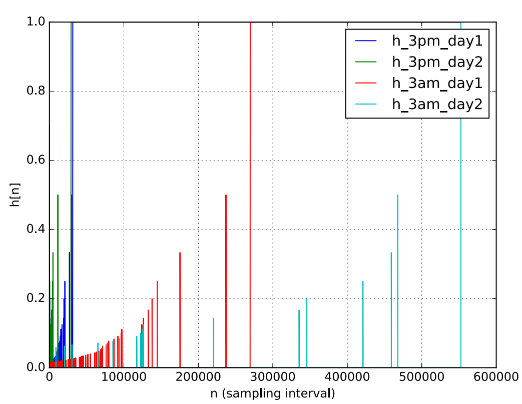

Fig. 9. Hazard Function of Idle Time Lengths ((for the entire range of idle period)).

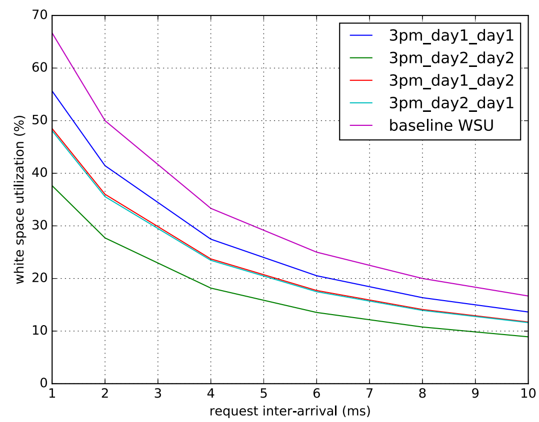

Fig. 10. WSU vs inter-arrival time for 3pm dataset for Algorithm 1.

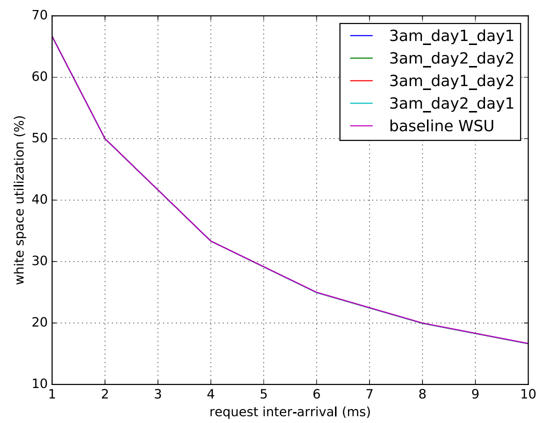

Fig. 11. WSU vs inter-arrival time for 3am dataset for Algorithm 1.

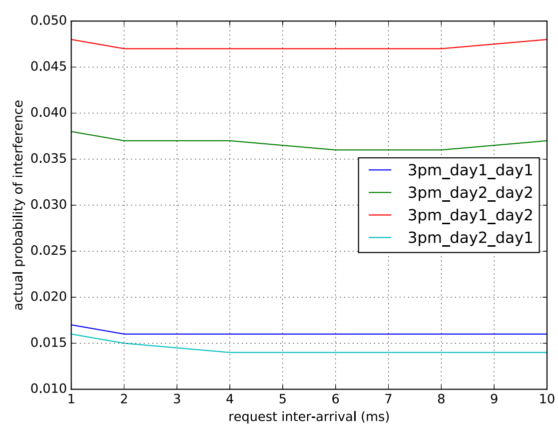

Fig. 12. Probability of Interference vs inter-arrival time for 3pm dataset for Algorithm 1.

entire range of idle periods. The slope of $H(\cdot)$ of 3pm_day1 is less steep than 3pm_day2 for about the first 40 sampling periods. Above 40 ms, the slopes for the two days are more or less equal. This is also clear from the instantaneous hazard function $h(\cdot)$ shown in Figure 8 and 9. The hazard functions of 3pm_day1 and 3pm_day2 have periodic spikes at about 40 ms intervals. These periodic spikes are due to the periodic *Sounding Reference Signal (SRS)* sent by the UEs in the uplink. The first spike at about 40 ms is higher for 3pm_day2 and then the next spikes are almost equal to 3pm_day1. The cumulative hazard functions of 3am_day1 and 3am_day2 are initially both zero and then aftwerwards the slope of 3am_day1 becomes steeper than 3am_day2. This is also evident from the instantaneous hazard function shown in Figure 8 and 9, where the hazard value is zero for both the days in the beginning.

B. Performance of Algorithm 1

Figure 10 shows the performance of Algorithm 1 in terms of WSU as average request inter-arrival time varies for the 3pm dataset. As request inter-arrival time increases, WSU decreases

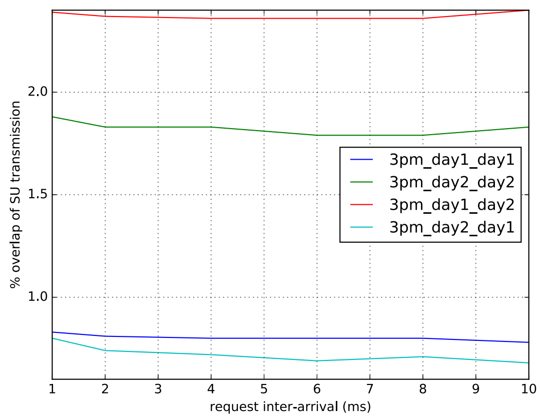

Fig. 13. Percent overlap vs inter-arrival time for 3pm dataset for Algorithm 1.

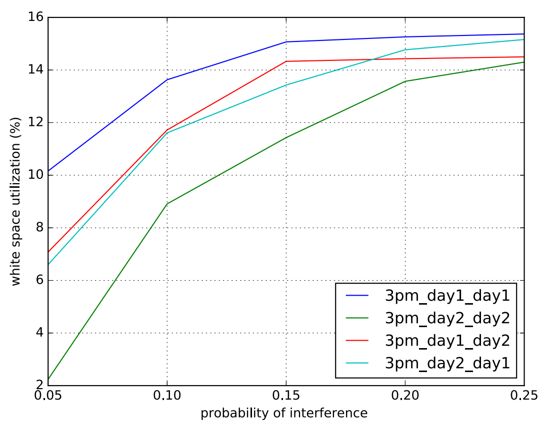

Fig. 14. WSU vs PoI threshold for 3pm dataset for Algorithm 1 for request inter-arrival time of 10 ms.

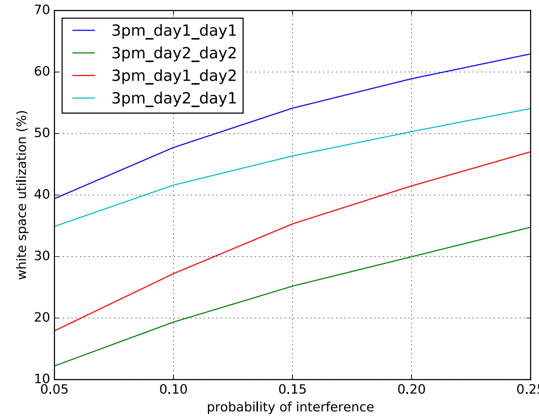

Fig. 15. WSU vs PoI threshold for 3pm dataset for Algorithm 2 for request inter-arrival time of 10 ms.

since the offered load from the SU requests decreases. Since slope of $H(\cdot)$ function for 3pm_day2 is steeper than that of 3pm_day1, WSU of 3pm_day2_day2 is always lower than 3pm_day1_day1. For 3pm_day2_day1, SU transmission grants are computed using 3pm_day2 dataset which has steeper slope in its $H(\cdot)$ function. So the WSU of 3pm_day2_day1 is less than 3pm_day1_day1. Likewise, WSU of 3pm_day1_day2 is more than 3pm_day2_day2, since 3pm_day1 dataset has lower slope in its $H(\cdot)$ function. For comparison, the baseline WSU computed using Equation (24) is also plotted in the figure. WSU of all configurations always remains below this theoretical upper bound.

From Figure 12, we observe that the PoI is always well below the set threshold of 0.1 for all configurations of 3pm dataset. Since 3pm_day2 dataset has fewer idle times (see Figure 2 and 3) and 3pm_day1 data has a less steep $H(\cdot)$ function, when the algorithm is run on 3pm_day2 data using 3pm_day1 data for learning, (i.e., Configuration *3pm_day1_day2*), the algorithm grants relatively longer transmission durations (based on the training dataset). This leads to more interference because of the relatively shorter idle periods of 3pm_day2. Thus the PoI for this configuration is higher than the other three.

From Figure 13 it can be seen that the POST for 3pm datasets are low for all configurations, and for a given configuration it is almost constant across inter-arrival times. These properties can be attributed to the short duration of SU transmissions used (2 ms) in the experiments.

For the 3am datasets Figure 11 shows the WSU versus request inter-arrival time. WSUs of all configurations are almost identical to each other. From the idle time histogram of 3am data on both days (Fig. 4 and 5), it can be seen that the channel mostly has very long idle periods. From these datasets we also observed that the busy periods were very short and very sparse. Each SU transmission is of constant duration of 2 ms which is very short compared to the length of idle periods. The $H(\cdot)$ functions of 3am dataset on both the days are at zero for a long duration. Thus, almost equal number of requests are granted in each configuration, which leads to almost identical WSU for all configurations. It is interesting to see that the baseline WSU computed using Equation (24) very closely matches with the WSU computed by Algorithm 1. The baseline WSU computation assumes that the PU idle time is a long continuous stretch of time and that all SU requests, which arrive when no other SU is transmitting, are granted. The low PU activity in the 3am datasets and DAR value of 100 % (See Table I) very closely matches those assumptions.

The PoI and POST values for 3am dataset in all configurations are always very close to zero for all configurations and for all values of request inter-arrival times. Hence, we have not provided graphs for these cases. This behavior can be explained as follows. Since the idles periods are very long and the busy periods are very short and extremely sparse in 3am dataset, the number of interference events and the duration of any overlap of SU transmissions with PU activity are both extremely small.

Figure 14 shows the variation of WSU with respect to the PoI threshold for an average request inter-arrival time of 10 ms for 3pm dataset. As the PoI threshold increases, WSU also increases, because with a higher PoI threshold, the SUs are granted more transmission opportunities as they are allowed to interfere with the PU with higher probability.

We show the DAR, UAR, DRR and URR for Algorithm 1 when the SU request inter-arrival time is 4 ms. The URR is zero for all configurations. Thus Algorithm 1 results in no lost opportunities across all the configurations. The algorithm also

TABLE I
VARIOUS ACCEPT AND REJECT RATIOS FOR REQUEST INTER-ARRIVAL TIME 4 ms FOR ALGORITHM 1

Configuration	DAR (%)	UAR (%)	DRR (%)	URR (%)
3pm_day1_day1	74.9	1.2	23.9	0
3pm_day1_day2	60.3	3.0	36.7	0
3pm_day2_day2	43.4	1.6	54.9	0
3pm_day2_day1	60.7	0.9	38.4	0
3am_day1_day1	100.0	0	0	0
3am_day1_day2	100.0	0	0	0
3am_day2_day2	100.0	0	0	0
3am_day2_day1	100.0	0	0	0

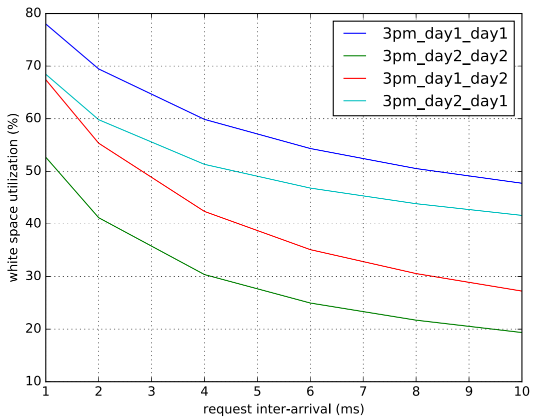

Fig. 16. WSU vs inter-arrival time for Algorithm 2 for 3pm dataset.

has very low UAR, which is good, since this metric shows how well the algorithm avoids making bad decisions while accepting a request. Although not shown, results for other request inter-arrival times are equally good.

Where possible, the τ value for Algorithm 1 should be chosen carefully as per the $H(\cdot)$ function. If the $H(\cdot)$ function has a very low slope, then a large τ value can be chosen to obtain high WSU. However, if τ is large but the slope of $H(\cdot)$ is very steep, then many requests may be rejected, resulting in low WSU. Therefore, keeping τ relatively low is a better approach to achieving high WSU. Although WSU per request will be lower, it is more than compensated for by a higher number of granted SU requests.

C. Performance of Algorithm 2

Figure 16 and Figure 17 show the performance of Algorithm 2 in terms of WSU for the *3pm* and *3am* datasets respectively. As the average request inter-arrival time increases, the SUs exploit less white space for transmission. Hence, the WSU decreases for both cases. In Figure 16, WSU for 3pm_day1_day1 is higher than 3pm_day2_day2. The slope of the $H(\cdot)$ function for 3pm_day1 dataset is less steep than that of the 3pm_day2 dataset in the initial part of the curve. Also, 3pm_day1 has longer idle periods than 3pm_day2. So for 3pm_day1_day1, the algorithm grants longer transmission times for many SU requests. For 3pm_day2_day1, the algorithm determines SU transmission using the 3pm_day2-trained $H(\cdot)$ function, which has a steeper slope than 3pm_day1. Hence, the WSU in this case is lower

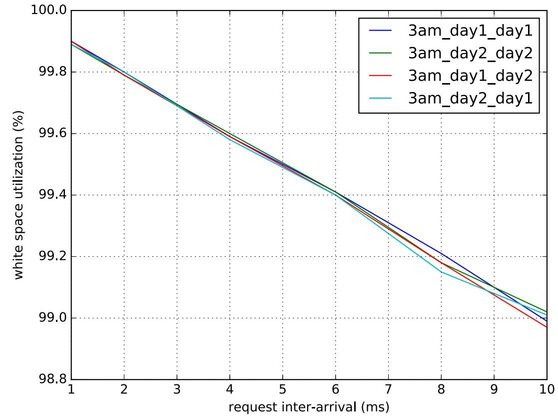

Fig. 17. WSU vs inter-arrival time for Algorithm 2 for 3am dataset.

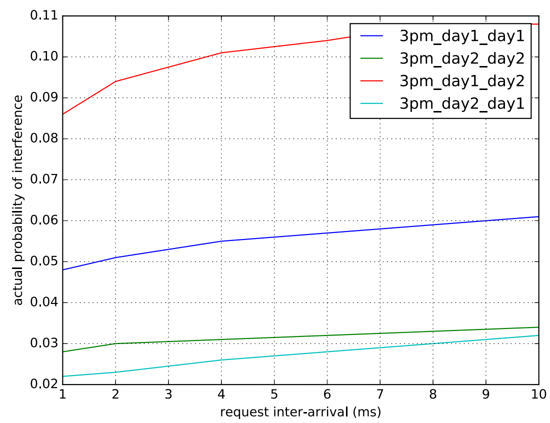

Fig. 18. PoI vs inter-arrival time for Algorithm 2 for 3pm dataset.

than for 3pm_day1_day1. Likewise, since the $H(\cdot)$ function of 3pm_day1 has a lower slope than that of 3pm_day2, the WSU of 3pm_day1_day2 is higher than that of 3pm_day2_day2. For the 3am datasets, WSU is very high for all configurations. The channel was very rarely busy during that time. Hence, the idle time durations are very long, and the total idle time is also very high. Since the slope of $H(\cdot)$ remains at zero in the initial part for both 3am_day1 and 3am_day2, SU transmission grants are long. Since both the days have very sparse busy periods, the WSU for all configurations of 3am dataset are almost the same.

When we compare the WSU performance of Algorithm 1 with Algorithm 2 for a given request inter-arrival time and a given configuration, we notice that the WSU of Algorithm 1 is much lower than that of Algorithm 2. The fundamental design of the two algorithms gives rise to this behavior. For a given request, Algorithm 2 maximizes the granted SU transmission duration, whereas Algorithm 1 only checks to see if it can grant a request for a constant transmission duration (2 ms in our experiment). Thus, Algorithm 2 is more aggressive than Algorithm 1 in terms of duration of SU transmission grants and is able to achieve higher WSU.

Figure 18 and 19 show the variation of PoI as the inter-arrival time of the requests increases. For 3pm dataset, in all configurations, the PoI is mostly below the set threshold (0.1). For 3pm_day1_day2, for higher values of

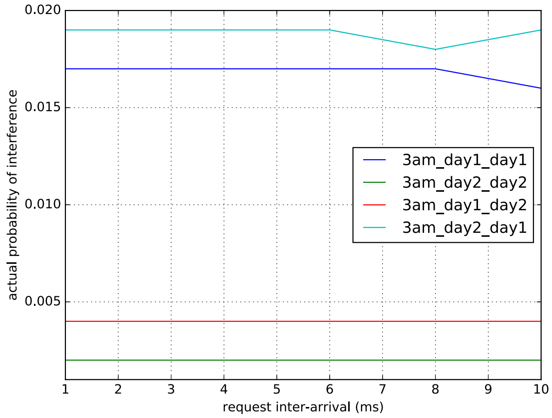

Fig. 19. PoI vs inter-arrival time for Algorithm 2 for 3am dataset.

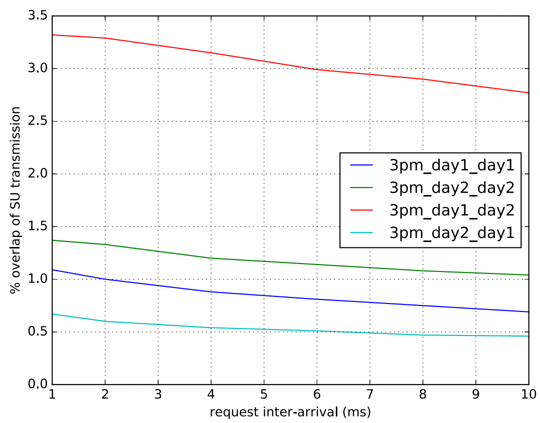

Fig. 20. POST vs inter-arrival time for Algorithm 2 for 3pm dataset.

request interarrival, the observed probability of interference is negligibly higher than the set threshold. Since $H(\cdot)$ function of 3pm_day1 is less steep, SU transmission duration is more aggressively granted than 3pm_day2. Hence, when 3pm_day2 data is run using 3pm_day1 data for learning, it tends to interfere more. Hence, PoI for 3pm_day1_day2 is the highest. For 3am data, the actual probability of interference is always much below the set threshold. Since the channel is rarely busy in these datasets, the PoI is almost constant across all configurations.

Figure 20 and 21 present POST versus request inter-arrival time. Since the channel was mostly idle during 3am, POST is very close to zero for all the configurations of this dataset. In Figure 21, the POST values of 3am_day1_day2, 3am_day2_day1 and 3am_day2_day2 are almost identical to each other (0.03 %). Hence, the curves are indistinguishable from each other. For the 3pm dataset POST values are small but higher than those of the 3am datasets. This is attributed to very sparse busy periods in the 3am datasets compared to the 3pm datasets.

Finally, Figure 15 shows the variation of WSU as the PoI threshold is varied for an average request inter-arrival time of 10 ms for the 3pm datasets. As the PoI threshold increases, WSU increases. With a higher PoI threshold, Algorithm 2 grants longer SU transmissions, resulting in higher WSU. Since Algorithm 2 returns a maximum allowable transmission

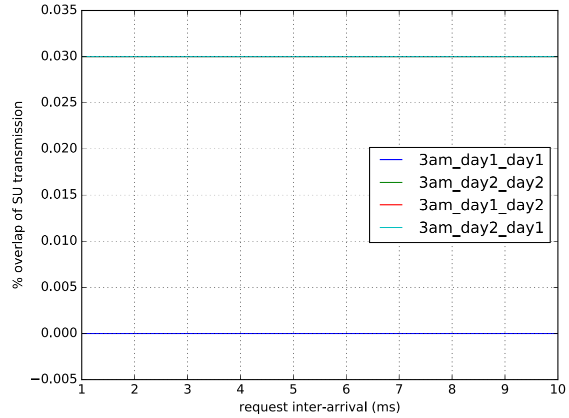

Fig. 21. POST vs inter-arrival time for Algorithm 2 for 3am dataset.

time for a given PoI threshold instead of a simple *Grant* or *Deny*, the DAR, UAR, DRR and URR metrics do not apply to it.

Our results indicate that the algorithms can be trained using data from one day at a given time and run during the same time on another day to get reasonable WSU while satisfying the PoI threshold and with low overlap transmission with the PU.

All our experiment runs were for a very long duration (3.6 million sampling intervals). Hence, the number of SU requests were very large, and the computed performance metrics of the two algorithms had very little variation across different runs.

VI. Conclusion

We introduced DSA algorithms based on survival analysis that make efficient use of white space in an LTE band, even at very fine time scales. They are stochastic but non-parametric and therefore do not require the assumption of a particular distribution. This makes the implementation simple. The tuning parameter for the algorithm is the probability of successfully completing a transmission or, equivalently, the PoI threshold. Thus, it is easy to interpret and directly reflect desired system performance metrics. We used real LTE band occupancy data for the PU activity in our simulations. Our results show that if the cumulative hazard functions are fairly similar (in terms of slope) across different datasets, the algorithms can be trained on one dataset and run on another dataset without significant degradation of performance. This is a very important property of the algorithms, since in practice, the algorithms will be trained on historical data and then run in real-time. We expect that in actual spectrum sharing systems the PUs will be wary of sharing their spectrum with SUs for fear of too much interference. This is addressed in our algorithms by showing that the PoI is around or below the preset threshold in all configurations.

This paper provides an initial performance analysis of the algorithms in an LTE band. Evaluation using datasets collected in different bands at varying locations with other traffic characteristics needs to be done. Other time scales need to be investigated to show the range over which the algorithms are effective. A theoretical performance analysis and comparison with other prediction schemes are needed as well.

Depending on the SU application, alternative forms of the algorithms presented in this paper can easily be developed using the same fundamental approach. One can imagine a form of spectrum requests that includes a maximum or desired transmit time and a minimum acceptable time. The algorithm would then either deny the request or return a grant duration in the requested range. Another form could have the user requesting a minimum initial grant and then the scheduler can add additional follow-on transmission time, if available, once the initial request has elapsed. An adaptive version of the algorithm may be more attractive for implementation on practical systems. It would update the estimated cumulative hazard function as new idle periods appear in the spectrum.

In summary, we have introduced and analyzed two DSA algorithms that are tunable by a well-defined, intuitive and critical system parameter, namely, the probability of successful transmission, p. Furthermore, p is independent of the training process and data. Training is fast and computationally inexpensive, and the algorithms' run time performance is not overly sensitive to the training data.

REFERENCES

[1] T. A. Hall, A. Sahoo, C. Hagwood, and S. Streett, "Exploiting LTE white space using dynamic spectrum access algorithms based on survival analysis," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–7.

[2] A. D. Spaulding and G. H. Hagn, "On the definition and estimation of spectrum occupancy," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, no. 3, pp. 269–280, Aug. 1977.

[3] M. López-Benítez and F. Casadevall, "Discrete-time spectrum occupancy model based on Markov chain and duty cycle models," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, May 2011, pp. 90–99.

[4] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 533–545, May 2008.

[5] M. Sharma and A. Sahoo, "Stochastic model based opportunistic channel access in dynamic spectrum access networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 7, pp. 1625–1639, Jul. 2014.

[6] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in WLAN channels: Empirical model and its stochastic analysis," in *Proc. 1st Int. Workshop Technol. Policy Accessing Spectr. (TAPAS)*, Boston, MA, USA, Aug. 2006, Art. no. 14.

[7] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 66–72, pp. 66–72, May 2007.

[8] L. Stabellini, "Quantifying and modeling spectrum opportunities in a real wireless environment," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Sydney, NSW, Australia, Apr. 2010, pp. 1–6.

[9] A. J. Gibson and L. Arnett, "Statistical modelling of spectrum occupancy," *Electron. Lett.*, vol. 29, no. 25, pp. 2175–2176, Dec. 1993.

[10] A. Gibson and L. Arnett, "Measurements and statistical modelling of spectrum occupancy," in *Proc. 6th Int. Conf. HF Radio Syst. Tech.*, Jul. 1994, pp. 150–154.

[11] M. L. López-Benítez and F. Casadevall, "Modeling and simulation of time-correlation properties of spectrum use in cognitive radio," in *Proc. Int. Conf. Cogn. Radio Oriented Wireless Netw. (CROWNCom)*, Osaka, Japan, Jun. 2011, pp. 326–330.

[12] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in Ad Hoc networks: A POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[13] K. W. Sung, S.-L. Kim, and J. Zander, "Temporal spectrum sharing based on primary user activity prediction," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3848–3855, Dec. 2010.

[14] A. Plummer, M. Taghizadeh, and S. Biswas, "Measurement-based bandwidth scavenging in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 1, pp. 19–32, Jan. 2012.

[15] C. Tekin and M. Liu, "Online learning of rested and restless bandits," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5588–5611, Aug. 2012.

[16] Y. Gai and B. Krishnamachari, "Decentralized online learning algorithms for opportunistic spectrum access," in *Proc. IEEE Glob. Telecommun. Conf. (GLOBECOM)*, Houston, TX, USA, Dec. 2011, pp. 1–6.

[17] S. Yin, D. Chen, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: A large-scale spectrum measurement study," *IEEE Trans. Mobile Comput.*, vol. 11, no. 6, pp. 1033–1046, Jun. 2012.

[18] P. Huang, C.-J. Liu, X. Yang, L. Xiao, and J. Chen, "Wireless spectrum occupancy prediction based on partial periodic pattern mining," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 7, pp. 1925–1934, Jul. 2014.

[19] (Apr.2014). *Further Notice of Proposed Rule Making*. Accessed: 25th Mar. 25, 2015. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/FCC-13-100A1.pdf

[20] R. G. Miller, Jr., *Survival Analysis*, vol. 66. New York, NY, USA: Wiley, 2011.

[21] O. Aalen, O. Borgan, and H. Gjessing, *Survival and Event History Analysis: A Process Point of View*. New York, NY, USA: Springer, 2008.

[22] M. López-Benítez and F. Casadevall, "Methodological aspects of spectrum occupancy evaluation in the context of cognitive radio," in *Proc. Eur. Wireless Conf.*, May 2009, pp. 199–204.

[23] R. B. Cooper, *Introduction to Queueing Theory, Second Edition*. New York, NY, USA: North-Holland, 1981.

Timothy A. Hall (M'88) received the B.E.E. and M.E.E. degrees, and the Ph.D. degree in electrical engineering from the University of Delaware, in 1994. He joined NIST in 1999, where his past work included network modeling and simulation, voice quality in Internet telephony, wireless networks, and conformance testing of cryptographic algorithms. He is currently working on aspects of spectrum sharing, including dynamic spectrum access, operational security considerations, and interference protection.

Anirudha Sahoo (M'95) received the B.Sc. (Engg.) degree in electrical engineering from the National Institute of Technology, Rourkela, India, and the M.S. degree in computer science from the University of Louisiana at Lafayette, Lafayette, LA, USA, and the Ph.D. degree in computer science from Texas A&M University, College Station, TX, USA. He worked as a Software Engineer with Intergraph Corporation, Huntsville, AL, USA, and then as a Senior Software Engineer with Cisco Systems, San Jose, CA, USA. He was an Associate Professor with IIT Bombay, India. He is currently a Computer Scientist with the National Institute of Standards and Technology. He is also an Adjunct Professor with the Department of Computer Science, University of Maryland, College Park, MD, USA.

Charles Hagwood received the B.S. degree in mathematics from NC A&T State University, Greensboro, NC, USA, and the Ph.D. degree in mathematics from the University of Michigan, Ann Arbor, MI, USA. He taught in the Department of Mathematics, Dartmouth College, Hanover, NH, USA, and with the Department of Mathematics, University of Virginia, Charlottesville, VA, USA. He is with the National Institute of Standards and Technology (NIST). As a Ford Foundation Fellow, he spent one year with the Department of Statistics, Stanford University, Stanford, CA, USA. He currently with the Statistical Engineering Division, NIST. His current research and consulting involve stochastic differential equations, shape analysis, dynamic spectrum access, standard reference materials, and point process modeling.

Sarah Streett received the M.S. degree in mathematics from Northern Arizona University, in 1994, and the Ph.D. degree in statistics from Colorado State University, in 2000. She has worked as a Post-Doctoral Researcher with the National Center for Atmospheric Sciences, as a Staff Scientist with the National Ecological Observatory Network and as a Mathematical Statistician with the National Institute of Standards and Technology. She is currently self employed as a Contract Statistician.