Changepoint Detection for Real-Time Spectrum Sharing Radar

Samuel Haug  
Department of Electrical and Computer Engineering  
Baylor University  
Waco, TX  
Sam_Haug@baylor.edu

Austin Egbert  
Department of Electrical and Computer Engineering  
Baylor University  
Waco, TX  
Austin_Egbert@baylor.edu

Robert J Marks II  
Department of Electrical and Computer Engineering  
Baylor University  
Waco, TX  
Robert_Marks@baylor.edu

Charles Baylis  
Department of Electrical and Computer Engineering  
Baylor University  
Waco, TX  
Charles_Baylis@baylor.edu

Anthony Martone  
DEVCOM Army Research Laboratory  
Adelphi, MD  
Anthony.F.Martone.civ@army.mil

Abstract—Modern radar must adapt to changing environments, and changepoint detection is a method to do so. Many radar systems employ a prediction action cycle to proactively determine transmission modes while spectrum sharing. This method constructs and implements a model of the environment to predict unused frequencies, and subsequently transmits in predicted available bands. For these selection strategies to succeed, the assessment of the underlying environmental models must be robust to change. Changepoint detection increases this robustness. Changepoint detection is the identification of sudden changes, or changepoints, in the distribution from which data is drawn. This information allows the models to discard “garbage” data from a previous distribution, which has no relation to the current state of the environment. For spectrum sharing applications, these changepoints may represent interferers leaving and entering the spectral environment, or changing their spectrum access patterns. In this work, we demonstrate the effectiveness of the addition of changepoint detection to spectrum sharing algorithms in changing environments. Bayesian online changepoint detection (BOCD) is applied to the sense-and-predict algorithm to increase the accuracy of its models and improve its performance. The use of changepoint detection allows for dynamic and robust spectrum sharing even as interference patterns change suddenly and dramatically. BOCD is especially advantageous because it enables online changepoint detection, allowing models to be updated continuously as data are collected. This strategy can also be applied to other predictive algorithms that create and maintain models in changing environments.

Index Terms—Machine learning, prediction methods, radio spectrum management, radiofrequency interference

I. INTRODUCTION

As more wireless users enter the environment, the electromagnetic spectrum becomes more congested. Radar and communication systems both desire greater bandwidth for more efficient operations [1], [2], further limiting availability.

This work has been funded by the Army Research Office (Grant No. W911NF-20-2-0270). The views and opinions expressed do not necessarily represent the opinions of the U.S. Government.

The increased efficiency of spectrum access to accommodate these users is a topic of much research [3]–[9].

Vertical sharing of bandwidth involves secondary users (SUs) using available bandwidth that does not interfere with primary user (PU) operations [10], and can be accomplished with prediction. This prediction requires the creation and maintenance of an environmental model of PU activity. Given a precise and accurate environmental model, a system can predict future availability and transmit accordingly.

Radar secondary users (SUs) that operate in shared frequency bands must be able to adapt to avoid interference. As PUs enter and leave the spectral environment, models can become outdated and poorly represent the current state of the environment. Spectrum sharing systems must overcome this obstacle for optimal performance. Models of rapidly changing environments can be improved with the addition of changepoint detection. Changepoint detection is the identification of sudden changes, or changepoints, in the distribution from which data is drawn. With knowledge of changepoint locations, models can discard data from an unrelated distribution and consider only the most current spectral behavior. This can increase prediction accuracy.

Spectral evaluation and prediction algorithms that employ changepoint detection can be robust with respect to sudden and drastic changes in spectral environment behavior. This can allow the radar system to collect more reliable and undistorted information [1].

In this work we apply changepoint detection to a single spectral sharing algorithm, but changepoint detection can also be applied to other systems that maintain a model of the environment [11] or employ machine learning on observed environmental behavior [12], [13]. Wang et al. [14] give a survey of many such spectrum sharing techniques and algorithms. Our work is unique because it employs changepoint detection for increased robustness to sudden and drastic changes in
environmental behavior.

Changepoint detection has previously been employed to detect interference in cognitive radio systems [15]. In our work large scale changes in interference patterns for the purpose of live prediction is considered. Previous work, on the other hand, considers small scale interference changes for the purpose of detection.

II. METHODS

The effectiveness of the addition of changepoint detection to spectrum sharing algorithm is shown using Bayesian online changepoint detection (BOCD). BOCD enhances existing spectrum sharing algorithms for sense-and-predict.

A. Sense-and-Predict

The sense-and-predict algorithm [16] employs the prediction action cycle to inform intelligent radar frequency selection. The spectrum is divided into frequency sub-bands, and each sub-band is considered independently as an alternating renewal process. Busy and idle times for each sub-band are modeled as a log-normal distribution [17], [18]. During initialization, a period of passive spectral evaluation (training) is required before prediction and transmission. When training is completed, model statistics used to inform predictions are generated, and transmission begins. Transmission is sub-band specific; lack of transmission in one sub-band does not mean that the radar as a whole ceases to function, but rather that the specific sub-band is avoided. The time duration allocated for the training phase is denoted as a spectrum evaluation interval (SEI). The data gathered in this SEI are used to create the model for the next SEI, and training data gathered in this second SEI are used to create the model for the successive SEI. In this way, the system uses training data from a given SEI to create the working model for the next SEI. The length of this SEI is chosen by the user. A block diagram of the sense-and-predict system is shown in Figure 1. The system characterizes the observed busy $B$ and idle $I$ time durations. The timeline is therefore divided up at each instant as either being busy or idle. The time durations that a sub-band spends being both busy and idle are recorded, and two distributions are developed for both busy and idle time durations. An illustration of sub-band behavior and notation is shown in Figure 2. The present state of the environment is shown at time $t = 0$. The units of time are omitted because they vary widely by application, but can often be measured in units of the system pulse repetition interval. The length of time that a sub-band remains busy is referred to as a busy time duration or busy interval. The length of time that a sub-band remains idle is referred to as an idle time duration or idle interval. Note that busy and idle intervals that intersect with the present (at the bottom of the illustration) are unlabeled, because the system has not observed the end of those intervals and does not know how long each sub-band will remain in its current state. The system’s model of the environment therefore includes two parametric probability distributions for each sub-band, one for busy durations and one for idle durations. These distributions are used to predict future sub-band behavior. To determine distribution parameters, a sample mean estimate $\mu_B$ and variance estimate $\sigma_B^2$ are first computed for busy times in all $M$ sub-bands

$$
\mu_{B_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} B_{t_{ij}}
$$

$$
\sigma_{B_i}^2 = \frac{1}{n_i} (B_{t_{ij}} - \mu_{B_i})^2.
$$

The index of each frequency sub-band is denoted by $i \in \{1, 2, 3, ..., M\}$. The number of observed busy intervals in the $i$-th band is $n_i$. The number of observed idle intervals is also $n_i$. The duration of the $j$-th busy interval in channel $i$ is $B_{t_{ij}}$. The statistics of the time process are assumed to be log-normal with a mean of $\mu_B$ and a standard deviation of $\sigma_B$. From this assumption the computed sample statistics are log-normal parameters. The logarithm of a log-normal distribution is a Gaussian distribution with mean $\mu$ and variance $\sigma^2$. For the $i$-th sub-band busy time durations the corresponding parameters
The log-normal mean and standard deviation can be computed as a function of the Gaussian mean and standard deviation. Specifically,

$$\hat{\mu}_{B_i} = \ln \left( \frac{\mu_{B_i}^2}{\sigma_{B_i}^2 + \mu_{B_i}^2} \right)$$ \hspace{1cm} (1)$$

$$\hat{\sigma}_{B_i} = \sqrt{\ln \left( \frac{\sigma_{B_i}^2}{\mu_{B_i}^2} \right) + 1}.$$ \hspace{1cm} (2)$$

The idle mean and variance are calculated in the same way except B subscripts are replaced by I subscripts.

The system uses these calculated parameters to compute availability probabilities for each sub-band at every time-step in the next spectrum evaluation interval. These availability probabilities are calculated by referencing the distribution’s cumulative distribution function (CDF). The CDF for a log-normal distribution is

$$\text{cdf}_L(\hat{\mu}, \hat{\sigma}, t) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^t \frac{1}{\tau} \exp \left( -\frac{(\ln \tau - \hat{\mu})^2}{2\hat{\sigma}^2} \right) d\tau.$$ \hspace{1cm} (3)

This CDF is used to determine the likelihood of a state change in a busy interval

$$p_{B_i}(t_B + \Delta t) = \text{cdf}_L(\hat{\mu}_{B_i}, \hat{\sigma}_{B_i}, t_B + \Delta t)$$ \hspace{1cm} (4)$$

$$p_{I_i}(t_I + \Delta t) = 1 - \text{cdf}_L(\hat{\mu}_{I_i}, \hat{\sigma}_{I_i}, t_I + \Delta t)$$ \hspace{1cm} (5)$$

where \(p_{I_i}(t + \Delta t)\) and \(p_{B_i}(t + \Delta t)\) are the probabilities that the sub-band will be available after \(\Delta t\) time-steps if the sub-band has already been idle or busy, respectively, for \(t\) time-steps. When required, this equation allows for a system latency of multiple time-steps between observation and transmission. If there is no system latency, \(\Delta t\) is set to a single pulse repetition interval.

At each time-step after availability probabilities are calculated for each sub-band, a threshold \(\theta\) is applied to predict the set of available sub-bands \(A = \{A_1, A_2, A_3, \ldots, A_M\}\). Because busy and idle intervals are modeled and evaluated separately, separate thresholds are used to evaluate availability probabilities of busy and idle sub-bands

$$A_i = \begin{cases} p_{B_i} \geq \theta_B, & S_i = 1 \vspace{0.5em} \end{cases} p_{I_i} \geq \theta_I, & S_i = 0.$$ \hspace{1cm} (6)

where \(\theta_B\) represents the probability threshold for busy statistics, \(\theta_I\) represents the probability threshold for idle statistics \(\theta_I, \theta_B \in [0, 1]\), and \(S_i\) represents the current status of the \(i\)-th sub-band, with a one representing busy and a zero representing idle.

If the availability probability is above the threshold, the system will predict that sub-band to be available and attempt to transmit in that sub-band. If the availability probability is below the threshold, the system will predict it to be unavailable and will not attempt to transmit in that sub-band. These thresholds control a direct trade-off between false alarms and missed detections with regard to predicted state change. A false alarm in this context signifies that the system believes a sub-band to be available when it is not, which results in a collision with interference. These collisions degrade signal accuracy [1]. Missed detections in this context signify that the system predicted a sub-band to be unavailable when it is available, which results in a missed opportunity for transmission. Higher values for both \(\theta_B\) and \(\theta_I\) mean that the system requires more evidence to believe that a sub-band will be available, and will therefore cause more missed opportunities and fewer collisions. Conversely, lower values will have the opposite effect, causing more collisions and fewer missed opportunities.

Availability probability thresholds are optimized as sub-band models are created with a grid search, exhaustively testing 100 threshold values for both \(\theta_I\) and \(\theta_B\) between 0.05 and 0.95. The error rates for each threshold combination are then compared and the threshold combination that minimizes the total weighted error \(\rho\) is selected. This total weighted error is calculated with the rate of collisions \(C\) and missed opportunities \(D\), implementing a convex weighting factor \(\alpha\) between them

$$\rho = \alpha C + (1 - \alpha)D$$ \hspace{1cm} (7)

where \(0 \leq \alpha \leq 1\).

### B. Bayesian Online Changepoint Detection

While several changepoint detection algorithms take place offline after all data has been collected [19]–[22], Bayesian online changepoint detection (BOCD) can be implemented...
is denoted by $t$ in one dimension. \( P(\tau_1|x_{1:t}) \) is a probabilistic model of the environment for run length $\tau$. \( \pi_0 \) represents the time since the last changepoint. Large values of $\tau$ indicate that the current partition is long and a changepoint has not occurred recently. As data are introduced to the system, all possible run lengths are considered, and each of these possible run lengths is assigned a probability $P(\tau_1|x_{1:t})$. Upon initialization in a new environment the first datum must be the first in its run length, and so $P(\tau_1 = 0 | x_1) = 1$. As time increases a given run length may only exhibit one of two behaviors: either the run length increases by one (no changepoint) or the run length remains at zero (changepoint). The probability that the run length will increase is called the growth probability, and the probability that the run length will drop to zero is called the changepoint probability. As more data are collected, probabilities for every combination of changepoints are calculated, creating a stochastic process called the discrete posterior. The discrete posterior models discrete time $t$ in one dimension and the probability mass function (PMF) for $\tau$ in the other dimension. Online evaluation is performed recursively

$$P(\tau_1, x_{1:t}) = \sum_{\tau_{1-1}} P(\tau_1 | \tau_{1-1}) P(x_t | \tau_{1-1}, x_{1:t-1}) P(\tau_{1-1}, x_{1:t-1})$$

where

$$P(\tau_1 | \tau_{1-1}) = \begin{cases} H(\tau_{1-1} + 1), & \tau_1 = 0 \\ 1 - H(\tau_{1-1} + 1), & \tau_1 = \tau_{1-1} + 1 \\ 0, & \text{otherwise} \end{cases}$$

and $H(t)$ is the hazard function. In the case of a memoryless hazard function such as an exponential decay, $H(t)$ is a constant $h$. This recursive operation does not produce a normalized PMF, and so the resulting mass function must be normalized to have a total value of 1. $P(x_t | \tau_{1-1}, x_{1:t-1})$ represents the probability of generating the newest datum using a model of the environment for run length $\tau$, represented below by $\pi_0^{(\tau)}$. An example of the discrete posterior for a time series in the $x$ domain is shown in Figure 3. Note that the algorithm correctly identifies changepoints by dropping the run length $\tau$ to zero at each vertical line.

BOCD uses a Bayesian prior to help inform its model $\pi^{(\tau)}_t$, which requires a priori knowledge of the upcoming data series, including probability distribution hyperparameters. BOCD originally required an estimate of the changepoint prior, which is the changepoint frequency $h$ of future data. Later versions of this technique allow for online estimation of this changepoint prior with unknown changepoint locations, which reduces the amount of a priori information that is required [24]. This system uses a joint probability mass function of $\tau_1$ and $\alpha_t$, where $\alpha_t$ is the number of changepoints that have occurred since $t = 0$. In this joint distribution a similar method is used to determine changepoint and growth probabilities, though any instances of hypothesized changepoints not only reset $\tau$, but $\alpha_t$ also increases by one. A time-step with no changepoint increases $\tau$ by one and $\alpha_t$ remains the same. The joint probability mass function is calculated recursively as

$$P(\tau_1, \alpha_t | x_{1:t}) = \sum_{\tau_{1-1}} \sum_{\alpha_{t-1}} P(\tau_1, \alpha_t | \tau_{1-1}, \alpha_{t-1})$$

where

$$P(\tau_t, \alpha_t | \tau_{t-1}, \alpha_{t-1}) = \begin{cases} \frac{b_{t-1} + 1}{\alpha_{t-1} + b_{t-1} + 2}, & \tau_t = \tau_{t-1} + 1 \& \alpha_t = \alpha_{t-1} \\ \frac{a_{t-1} + 1}{\alpha_{t-1} + b_{t-1} + 2}, & \alpha_t = \alpha_{t-1} + 1 \& \tau_t = 0 \\ 0, & \text{otherwise} \end{cases}$$

and $b_t$ is a count of how many nonchangepoints have occurred since time $t = 0$, given by $b_t = t - \alpha_t$. This joint probability mass function is used to estimate the changepoint prior.

III. BOCD ALTERATIONS

In order for the BOCD technique to be compatible with the sense-and-predict system, several alterations are required. First, the system will be required to analyze an environment, several alterations are required. In an effort to reduce the risk of an incorrect prior in a completely unknown environment, these priors are not employed, and model parameters are drawn entirely from collected data. The underlying probabilistic model $\pi^{(\tau)}_t$ for a run length of $\tau$ is generically

\[ P(\tau_t, x_{1:t}) = \sum_{\tau_{1-1}} \sum_{\alpha_{t-1}} P(\tau_t, \alpha_t | \tau_{1-1}, \alpha_{1-1}) \]

\[ \times P(x_t | \tau_{1-1}, x_{1:t-1}) \]

Fig. 3: Discrete posterior generated by BOCD on test data, ground truth changepoints are represented by vertical lines.
chosen to be a normal distribution with parameters $\mu^{(r)}$ and $\sigma^{(r)}$. Because $r$ is the run length, it is also the number of data in consideration.

The mean of a run length is determined by averaging the observed values in a given partition

$$\mu^{(r)} = \frac{1}{r} \sum_{i=0}^{r-1} x_i$$

where $x_0$ is the most recent observed data point. Similarly, the variance of the data is calculated using the first and second moments of the data in the partition

$$(\sigma^{(r)})^2 = \frac{1}{r} \sum_{i=0}^{r-1} x_i^2 - (\mu^{(r)})^2.$$ 

As hypothesized run lengths of a single observation do not yield a useful standard deviation, all run lengths are assumed to contain at least two data

$$P(r_t|x_{1:t}) = \left\{ \begin{array}{ll}
\sum_{r_{t-1} \neq 0} h(\pi_t^{(r)}) P(r_{t-1}|x_{1:t-1}), & r_t = 0 \\
0, & r_t = r_{t-1} + 1 \\
(1 - h)\pi_t^{(r)} P(r_{t-1}|x_{1:t-1}), & r_t = r_{t-1} + 1 \neq 1
\end{array} \right.$$ 

Thus, the standard deviation of a new partition is initially scaled to the first two data in that partition. As more data are collected, this hypothesis is either supported or rejected.

As a result of this lack of a priori information, a tuning parameter $\gamma \in (0, \infty)$ has been added to control the sensitivity of the algorithm

$$P(r_t|x_{1:t}) = \left\{ \begin{array}{ll}
\sum_{r_{t-1} \neq 0} h(\pi_t^{(r)}) P(r_{t-1}|x_{1:t-1}), & r_t = 0 \\
0, & r_t = r_{t-1} + 1 \\
(1 - h)\pi_t^{(r)} P(r_{t-1}|x_{1:t-1}) \gamma, & r_t = r_{t-1} + 1 \neq 1
\end{array} \right.$$ 

Higher values of $\gamma$ favor growth probabilities and will less often detect changepoints. This single dimensional tuning parameter can be more easily tuned in a live system than the previous multi-dimensional parameters.

In addition to removing these priors, online estimation of the hazard rate is employed [24], as summarized in Section II-B.

B. Computational Cost Improvements

The original changepoint detection algorithm’s computational complexity increases quadratically with time, which can quickly becomes infeasible. To combat this, run length probabilities in the tail of the distribution with probabilities less than a given threshold $\theta$ are discarded [23]. While this does reduce the computational cost slightly, the overall quadratic trajectory of the computational cost is unchanged. To limit the worst-case cost of the algorithm to a constant level, a run length maximum $L$ is imposed. Probability mass that would move above this limit in the discrete posterior is instead accumulated at $r_t = L$

$$P'(r_t = L|x_{1:t}) = P(r_t = L|x_{1:t}) + \sum_{k=L+1}^{\infty} P(r_t = k|x_{1:t})$$ 

This alteration to the algorithm does not affect its ability to detect changepoints, though it does limit the history window of the algorithm to $L$ data. For run length estimations less than $L$, estimations are unaffected. For run length estimations greater than $L$, the effect of this data truncation is mitigated because all truncated data are assumed i.i.d. This truncation does, however, limit the ability of the algorithm to identify small changes to a given environment over a long period of time. The value of $L$ directly controls a trade-off between decreasing computational cost and increasing accuracy of the changepoint detection.

An example of BOCD with an imposed run length maximum is shown in Figure 4. Note the accumulation of probability mass at $L$ for large run lengths. In order to reduce the number of calculations required to maintain a joint probability mass function of both $r_t$ and $\alpha_t$, probability masses in $\alpha_t$ lower than a given threshold $\theta_{\alpha}$ at either tail of the distribution are also discarded.

IV. SENSE-AND-PREDICT ALTERATIONS

A. Conditional Failure Probability

While the original sense-and-predict algorithm equates a probability distributions CDF with its failure probability, we
employ a distribution’s conditional failure probability. Failure in this context represents a sub-band state change. The conditional failure probability of a probability distribution is the probability that a failure will occur in an interval given no previous failure. For any generic probability distribution $X$ this probability is given by

$$P(a \leq X \leq b | X \geq a) = \frac{\int_{a}^{b} f_X(\tau) d\tau}{\int_{a}^{\infty} f_X(\tau) d\tau}$$

where $f_X(x)$ is the distribution’s probability density function (PDF). This quantity is also given by

$$P(a \leq X \leq b | X \geq a) = \frac{F_X(b) - F_X(a)}{1 - F_X(a)}$$

where $F_X(x)$ is the distribution’s CDF.

This method is used to determine the availability probabilities for the $i$-th sub-band

$$p_{B_i}(t_{B_i}, \Delta t) = \frac{F_{B_i}(t_{B_i} + \Delta t) - F_{B_i}(t_{B_i})}{1 - F_{B_i}(t_{B_i})}$$

$$p_{I_i}(t_{I_i}, \Delta t) = 1 - \frac{F_{I_i}(t_{I_i} + \Delta t) - F_{I_i}(t_{I_i})}{1 - F_{I_i}(t_{I_i})}$$

where $t_{B_i}$ and $t_{I_i}$ is the current length of the busy or idle interval respectively, $F_{B_i}$ and $F_{I_i}$ are the sub-band’s computed lognormal busy and idle interval distribution CDF’s respectively

$$F_{B_i}(t) = \text{cdf}_{E}(\hat{\mu}_{B_i}, \hat{\sigma}_{B_i}, t)$$

$$F_{I_i}(t) = \text{cdf}_{E}(\hat{\mu}_{I_i}, \hat{\sigma}_{I_i}, t),$$

and $\Delta t$ is the duration of the action latency period, in units of radar pulse repetition intervals.

The set of available sub-bands $\mathbf{A}$ is then predicted by applying (6).

B. Nonparametric Model of Interval Data

Though alternating renewal processes are traditionally modeled using a log-normal distribution [17], [18], nonparametric modeling of sub-band busy and idle distributions may be appropriate. Nonparametric modeling creates an empirical and discrete probability mass function (PMF) of a distribution using a normalized histogram of observed occurrences. A cumulative sum of this PMF provides the distribution’s CDF

$$\text{cdf}_{E}(t) = \sum_{i=1}^{t} \text{pmf}_{E}(i)$$

where $\text{cdf}_{E}(t)$ is the empirical CDF and $\text{pmf}_{E}(t)$ is the empirical PMF. The conditional failure probabilities are then calculated with (8) and (9) using each sub-band’s empirical CDF.

A significant advantage of the empirical data model is large-scale pattern recognition. A parametric model attempts to fit data into a given shape, but a unique underlying data pattern may be lost. A nonparametric model however, may be able to identify and adapt to such patterns. An example of this nonparametric pattern recognition is shown in Figure 5.

To emphasize pattern recognition, very predictable alternating data is chosen. This unique data pattern does not conform to the lognormal model and is therefore ignored by the parametric model. The nonparametric model, however, can distinguish the pattern.

This pattern recognition may conversely lead to poor predictions in certain circumstances. Shortly after a changepoint in a data series, for example, the model may interpret probabilistic variation as a persistent pattern. In this situation, the model will incorrectly perceive a pattern where none exists. A parametric model can provide better predictions in this case as it is able to overlook these phenomena.

The success of either the lognormal (parametric) or nonparametric methods of modeling is environmentally dependent and either is available to the system user.

V. ALGORITHM INTEGRATION

In the combined sense-and-predict system with changepoint detection, the BOCD algorithm is employed to provide accurate and current estimates of busy and idle interval times $\mu_{B_i}$, $\sigma_{B_i}$, $\mu_{I_i}$, and $\sigma_{I_i}$ for each sub-band. Two changepoint detection models are allocated for each of the $M$ sub-bands, one for idle interval statistics, and one for busy interval statistics. As the end of a sub-band’s idle or busy interval is observed, the length of that interval is passed to the appropriate BOCD model as an $x_t$ sample, and the joint probability mass function is updated. Subsequently, the run length $r_t$ with the highest probability is selected and the corresponding mean $\mu^{(r)}$ and variance $(\sigma^{(r)})^2$ supply parameters $\mu_i$, and $\sigma_i$ for

![Figure 5: Pattern recognition with nonparametric model, $\Delta t = 1$.](image-url)

In the case of a rigid pattern, the empirical model is able to successfully predict change while the lognormal model is not.
idle interval models and parameters $\mu_{B_i}$ and $\sigma^2_{B_i}$ for busy interval models.

If the system user chooses to employ the lognormal model, these statistics are then used to compute log-normal distribution parameters $\mu_{B_i}$, $\sigma_{B_i}$, $\tilde{\mu}_{B_i}$, and $\tilde{\sigma}_{B_i}$ using (1) and (2). These log-normal parameters are then used to calculate availability probabilities for each sub-band using (8) and (9).

If the system user chooses instead to employ the nonparametric method of modeling, these statistics instead generate an empirical PMF and an empirical CDF. These are then used to calculate availability probabilities (8) and (9) using the empirical CDF.

In this system employing changepoint detection, an initial passive spectrum evaluation interval may be much shorter. Initial distribution parameters are determined after two idle and busy intervals are observed for each sub-band. In addition, subsequent training intervals are not required because learning is continuous. Early models of the environment are based on very few data and may yield poor predictions. As more intervals are observed, the system models the environment with an increasing history of behavior. A system diagram of the sense-and-predict system with changepoint detection is shown in Figure 6.

A. Offline Threshold Determination

Due to frequent model parameter change, exhaustive online threshold optimization with a grid-search is not feasible. As a result, thresholds are determined offline. Optimal thresholds vary based on various environmental parameters, resulting in generic initial threshold selection

$$\theta_I = \theta_B = 1 - \alpha$$

where $\alpha$ is the error weight factor from (7). These thresholds will directly control the trade-off between transmitted collisions and missed opportunities.

VI. RESULTS

To compare the effectiveness of the sense-and-predict algorithm with changepoint detection to the original algorithm, each algorithm is applied to a simulated spectral environment. There are a total of three versions of the sense-and-predict system that are tested: the system without changepoint detection, the system employing changepoint detection with a lognormal model of sub-band intervals, and the system employing changepoint detection with an empirical model of sub-band intervals.

Each version of the sense-and-predict algorithm models and predicts the behavior of all sub-bands independently. As a result, presented simulation results consider only a single sub-band for clarity without loss of generality. The simulation environment used for testing draws busy and idle intervals from a probability distribution to present to the different versions of the sense-and-predict algorithm. At each time-step a Bernoulli trial with probability $h$ determines changepoint occurrence. This probability $h$ is called the changepoint frequency or hazard rate. If a changepoint does occur, the means of both busy and idle distributions will change. The magnitude of this change, $|\Delta|$, is drawn from its own probability distribution.

Specific algorithm performance is dependent on a large variety of parameters. Parameters for each sub-band include: parameters for ground truth busy and idle time interval probability distributions, changepoint frequency $h$, and parameters for the changepoint magnitude distribution $|\Delta|$. The latter distribution controls the magnitude of changepoints when they occur. Additional design parameters (where applicable) include collision error weighting $\alpha$, the length of the spectrum evaluation interval in units of system PRI, the length of the system’s action-latency period in units of system PRI $\Delta t$, the maximum run length parameter $L$, and the changepoint detection sensitivity parameter $\gamma$.

Sample simulation results are shown in Table I which highlights the effect of several key parameters. The performance metric is the weighted prediction error rate $\rho$, evaluated by a weighted combination of the missed opportunity rate $\tilde{D}$ and the collision rate $C$, using (7). The performance of the
original sense-and-predict algorithm is denoted by $\rho_O$, while that of the altered algorithm employing the lognormal model is denoted by $\rho_L$, and that of the altered algorithm employing the nonparametric model is denoted by $\rho_N$. These results show that the introduction of changepoint detection provides a significant error reduction to the algorithms performance when the underlying environmental model is changing ($h = 0.03$) while also maintaining a comparable level of performance in the absence of changepoints ($h = 0$).

Test 1 demonstrates algorithm performance in an environment with small variance in busy and idle distributions. Such an environment is very consistent within a single changepoint partition. Test 2 demonstrates the effect of large variations in busy and idle distributions even in the absence of changepoints. While there are no changepoints, the behavior of each sub-band is consistently varied. The sub-band behavior in Test 3 is more consistent, similar to that of Test 1, but no changepoints are present. The comparison of algorithm performance between Test 1 and Test 3 most directly demonstrates the effect of changepoints on algorithm performance. Test 4 primarily demonstrates the effect of the reduction of the maximum run length parameter $L$ on the accuracy of the altered algorithms. Though there is a trade-off between algorithm accuracy and computational cost as explained in Section III-B, the results of this test show that computational cost can be greatly reduced with slight reduction in accuracy.

Even in the absence of changepoints, the altered algorithm performs comparably to the original algorithm, demonstrated in Tests 2 and 3.

Algorithm performance is largely dependent on environmental conditions, and the BOCD sensitivity parameter $\gamma$ can aid the algorithm in adapting to various changepoint frequencies. A sample sweep of this tuning parameter and its effect on algorithm error rate is shown in Fig. 7.

VII. DISCUSSION

The sense-and-predict algorithm with changepoint detection performs significantly better than the original algorithm in the presence of changepoints. When each changepoint occurs, the modified algorithm is able to quickly adjust its models to accommodate the change in data, while the original is unable to do so until the next spectrum evaluation interval. Though the static model periodically retrains its models, it often includes unrelated data from a previous partition. This leads to uncharacteristic sample estimates. The altered algorithm is also able to begin transmission on initialization much more quickly than the original, as it does not require a passive spectrum evaluation interval. An example of the differences in model update behavior is shown in Figure 8. Solid model lines denote estimated $\mu_I$ and dashed lines denote estimated $\mu_I \pm \sigma_I$. Note that spectrum evaluation intervals are evaluated by pulse number not interval number, so vertical lines are not evenly dispersed in the interval domain.

The unaltered algorithm is significantly affected both by the presence of changepoints and the variation of busy and idle distributions, while the altered versions of the algorithm are

### TABLE I: Simulation results comparison. Altered algorithm error rates are significantly reduced in the presence of changepoints and comparable in their absence.

| Test # | 1       | 2       | 3       | 4       |
|--------|---------|---------|---------|---------|
| Busy Interval Distribution | $\sim N(150, 4)$ | $\sim N(50, 10)$ | $\sim N(150, 4)$ | $\sim N(150, 4)$ |
| Idle Interval Distribution  | $\sim N(150, 4)$ | $\sim N(50, 10)$ | $\sim N(150, 4)$ | $\sim N(150, 4)$ |
| Changepoint Frequency $h$  | 0.03     | 0.03    | 0.03    | 0.03    |
| Changepoint Magnitude $\Delta_i$ | $\sim N(40, 10)$ | -       | -       | $\sim N(40, 10)$ |
| Spectrum Evaluation Interval Length | 5000    | 5000    | 5000    | 5000    |
| Maximum Run Length $L$     | 60       | 60      | 60      | 30      |
| Algorithm Tuning Parameter $\gamma$ | 1100    | 1100    | 1100    | 1100    |
| System Latency $\Delta t$ | 5        | 5       | 5       | 5       |
| Collision Weighting $\alpha$ | 0.5     | 0.5     | 0.5     | 0.5     |

Original Algorithm Error Rate $\rho_O$ | 0.2644 | 0.0993 | 0.0209 | 0.2594 |
Lognormal Model Error Rate $\rho_L$ | 0.0943 | 0.0998 | 0.0197 | 0.0913 |
Nonparametric Model Error Rate $\rho_N$ | 0.1221 | 0.1010 | 0.0200 | 0.1153 |

Fig. 7: BOCD sensitivity parameter sweep. The parameter is chosen to minimize system error rate, $\gamma = 1100$. 
significantly affected by the variation of busy and idle distributions but affected much less significantly by the presence of changepoints. This increase in robustness of the spectrum sharing system is greatly desired to operate in many unique environments, including those with and without large numbers of changepoints.

VIII. CONCLUSION

In this work Bayesian online changepoint detection is applied to the sense-and-predict algorithm to increase model accuracy in a changing environment. Changepoint detection is the process of identifying sudden changes in the distribution from which data is drawn. This detection allows only relevant data while maintaining dynamic models of the environment, rather than including irrelevant data from a previous distribution.

As expected, the addition of changepoint detection to the sense-and-predict algorithm greatly increases its performance in the presence of changepoints. Informed knowledge of changepoint locations allows the system to more quickly and efficiently diminish the importance of old and unrelated data. In the absence of changepoints, the modified algorithms maintain their performance, while the original algorithm’s performance is degraded significantly in the presence of changepoints. The performance of all versions of the algorithm are dependent on busy and idle duration partition statistics.

The additional accuracy in spectral model-based predictions provided by changepoint detection may enhance vertical sharing of bandwidth by allowing secondary users to avoid collisions with primary users with unknown spectrum access patterns.

Though we specifically apply changepoint detection to this spectrum sharing algorithm, the addition of changepoint detection can similarly be implemented into other systems that maintain a model of the environment.

REFERENCES

[1] H. Griffiths, L. Cohen, S. Watts, E. Mokole, C. Baker, M. Wicks, and S. Blunt, “Radar spectrum engineering and management: Technical and regulatory issues,” Proceedings of the IEEE, vol. 103, no. 1, pp. 85–102, 2015.
[2] L. Zheng, M. Lops, Y. C. Eldar, and X. Wang, “Radar and communication coexistence: An overview: A review of recent methods,” IEEE Signal Processing Magazine, vol. 36, no. 5, pp. 85–99, 2019.
[3] H. Griffiths, S. Blunt, L. Cohen, and L. Savy, “Challenge problems in spectrum engineering and waveform diversity,” in 2013 IEEE Radar Conference (RadarConf), pp. 1–5, IEEE, 2013.
[4] H. Deng and B. Himed, “Interference mitigation processing for spectrum-sharing between radar and wireless communications systems,” IEEE Transactions on Aerospace and Electronic Systems, vol. 49, no. 3, pp. 1911–1919, 2013.
[5] D. W. Bliss, “Cooperative radar and communications signaling: The estimation and information theory odd couple,” in 2014 IEEE Radar Conference, pp. 0050–0055, IEEE, 2014.
[6] H. Hayvaci and B. Tavli, “Spectrum sharing in radar and wireless communication systems: A review,” in 2014 International Conference on Electromagnetics in Advanced Applications (ICEAA), pp. 810–813, IEEE, 2014.
[7] C. Baylis, M. Fellows, L. Cohen, and R. J. Marks II, “Solving the spectrum crisis: Intelligent, reconfigurable microwave transmitter amplifiers for cognitive radar,” IEEE Microwave Magazine, vol. 15, no. 5, pp. 94–107, 2014.
[8] F. Paisana, N. Marchetti, and L. A. DaSilva, “Radar, tv and cellular bands: Which spectrum access techniques for which bands?,” IEEE Communications Surveys & Tutorials, vol. 16, no. 3, pp. 1193–1220, 2014.
[9] K.-W. Huang, M. Bicó, U. Mitra, and V. Koivunen, “Radar waveform design in spectrum sharing environment: Coexistence and cognition,” in 2015 IEEE Radar Conference (RadarConf), pp. 1698–1703, IEEE, 2015.
[10] N. Devroye, P. Mitran, and V. Tarokh, “Cognitive decomposition of wireless networks,” in 2006 1st International Conference on Cognitive Radio Oriented Wireless Networks and Communications, pp. 1–5, IEEE, 2006.
[11] Z. Zhang, K. Zhang, F. Gao, and S. Zhang, “Spectrum prediction and channel selection for sensing-based spectrum sharing scheme using online learning techniques,” in 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 355–359, IEEE, 2015.
[12] N. Rastegardoost and B. Jabbari, “A machine learning algorithm for unlicensed lte and wifi spectrum sharing,” in 2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), pp. 1–6, IEEE, 2018.
[13] P. Zhu, J. Li, D. Wang, and X. You, “Machine-learning-based opportunistic spectrum access in cognitive radio networks,” IEEE Wireless Communications, vol. 27, no. 1, pp. 38–44, 2020.
[14] Q. Wang, H. Sun, R. Q. Hu, and A. Bhuyan, “When machine learning meets spectrum sharing security: Methodologies and challenges,” IEEE Open Journal of the Communications Society, 2022.
[15] A. Chêde, F. Rottenberg, M. Dossou, and J. Louveaux, “Spectrum sensing in mobile cognitive radio: A bayesian changepoint detection approach,” in 41st WIC Symposium on Information Theory and Signal Processing in the Benelux (SITB 2021), 2021.
[16] J. Kovarskiy, Real-time Spectral Prediction and Metacognition for Spectrum Sharing Radar, PhD thesis, Pennsylvania State University, 2021.
[17] W. R. Blischke and D. P. Murthy, Reliability: modeling, prediction, and optimization. John Wiley & Sons, 2011.
[18] G. M. Scherrer, *A simulation of a transient alternating renewal process.* PhD thesis, Citeseer, 1971.

[19] D. Barry and J. A. Hartigan, “A bayesian analysis for change point problems,” *Journal of the American Statistical Association*, vol. 88, no. 421, pp. 309–319, 1993.

[20] S. Chib, “Estimation and comparison of multiple change-point models,” *Journal of econometrics*, vol. 86, no. 2, pp. 221–241, 1998.

[21] P. J. Green, “Reversible jump markov chain monte carlo computation and bayesian model determination,” *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[22] L. Schwaller and S. Robin, “Exact bayesian inference for off-line change-point detection in tree-structured graphical models,” *Statistics and Computing*, vol. 27, no. 5, pp. 1331–1345, 2017.

[23] R. P. Adams and D. J. MacKay, “Bayesian online changepoint detection,” *arXiv preprint arXiv:0710.3742*, 2007.

[24] R. C. Wilson, M. R. Nassar, and J. I. Gold, “Bayesian online learning of the hazard rate in change-point problems,” *Neural computation*, vol. 22, no. 9, pp. 2452–2476, 2010.