A Novel Data-Driven Method for Nonstationarity Detection in Radar Target Detection

Murat Akcakaya, *Member, IEEE,* Satyabrata Sen, *Senior Member, IEEE,* and Arye Nehorai, *Fellow, IEEE*

***Abstract*—Most existing radar algorithms are developed under the assumption that the environment (clutter) is stationary. However, in practice, the characteristics of the clutter can vary enormously depending on the radar operational scenarios. If unaccounted for, these nonstationary variabilities may drastically hinder the radar performance. Therefore, to overcome such shortcomings, we develop a data-driven method for target detection in nonstationary environments. In this method, the radar dynamically detects changes in the environment, and adapts to these changes by learning the new statistical characteristics of the environment and by intelligibly updating its statistical detection algorithm. Specifically, we employ drift detection algorithms to detect changes in the environment; and incremental learning, particularly learning under concept drift algorithms, to learn the new statistical characteristics of the environment from the new radar data that become available in batches over a period of time. The newly learned environment characteristics are then integrated in the detection algorithm. We use Monte Carlo simulations to demonstrate that the developed method provides a significant improvement in the detection performance compared with detection techniques that are not aware of the environmental changes.**

Keywords*—*Data-driven adaptive radar, cognitive radar, nonstationary environment, incremental learning, active drift learning

I. Introduction

Increasing the accuracy of target detection and tracking is of great importance in military and coastal security operations, navigation, and maritime rescue operations. To guarantee the accuracy of both detection and estimation, clutter, interference, and noise must be suppressed to make the target signal distinguished. However, in practical scenarios, the characteristics of clutter backscattering can vary enormously depending on the region where the radar focuses its beam at a particular instant (or, maybe during a few successive instants). Even when the radar operation is restricted to a particular region, the change in weather conditions (temperature, wind speed, humidity, etc.) can drastically change the clutter statistics in maritime environments. Additionally, a highly-dynamic target, for example, moving through an urban canyon, can show high nonstationarity due to frequent appearance and disappearance in the radar coverage area.

A considerable amount of research work has been done in order to design radar systems that improve the performance of target detection and tracking in the presence of these various forms of nonstationarity. Particularly, in recent years, the desire for a radar system that can effectively *sense* its scenario, *learn* from its experience, and *adapt* to the changes in environment has been emphasized, and a conceptual framework for a cognitive fully adaptive radar that includes these three components has been laid out [1]-[3]. To implement this cognitive radar framework, several intelligent methods have been developed (mostly considering each component separately):

- Practical methods to identify the distribution of the radar measurements using a pre-learned dictionary of possible probability density functions representing the clutter characteristics [4]-[6], and to develop auto-regressive modeling of clutter distribution with a knowledge-aided Bayesian covariance estimation [7]-[10] for target detection;
- Adaptive weighted sum of clutter covariance estimates with exponentially decaying weights and predetermined delay constants [11], [12], waveform design that depends on the clutter characteristics [13], and nonhomogeneity detector in training data [14] for spatio-temporal adaptive processing applications,
- Target-state-dependent radar measurement models integrated in a fully adaptive cognitive radar framework [1], [15], [16] for target tracking.

These existing methods address important practical issues in the realization of the cognitive radar framework. However, these methods suffer from three shortcomings: (1) they have been developed under the assumption of parametric distribution for the clutter and noise processes; (2) a dictionary of possible probability density functions is assumed to exist; and (3) the problem of estimating the time instant when the statistical characteristics of the data changes is not answered. To overcome these shortcomings, in this paper, we develop machine learning based target detection algorithms that enable us to take a holistic approach to fully realize the cognitive radar framework. Our proposed radar system provides a unifying approach by autonomously learning and estimating the change-point in the statistical characteristics of the scenario, and by accordingly adapting to the environment with the incorporation of newly learned environment statistics into the target detection

M. Akcakaya is with the Department of Electrical and Computer Engineering, University of Pittsburgh, PA 15261, USA, e-mail: Akcakaya@pitt.edu

S. Sen is with the Computer Science and Mathematics Division, Oak Ridge National Laboratory, TN 37831, USA, e-mail: sens@ornl.gov. Dr. Sen's work has been supported in part by the U.S. Missile Defense Agency (MDA) at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy, under contract DE-AC05-00OR22725.

A. Nehorai is with the Department of Electrical and Systems Engineering, Washington University in St. Louis, MO 63130, USA, e-mail: nehorai@ese.wustl.edu. Dr. Nehorai's work has been supported in part by AFOSR Grant No. FA9550-11-1-0210.

algorithm.

Specifically, we utilize *incremental learning* and *drift detection* algorithms for building a detection method that incrementally learns the environment and updates the system parameters on the fly. This becomes possible as the radar always provides labels or labeled data (e.g., labels for target presence or absence, etc.) along with different confidence measures from the received measurements in a sequential manner. Thus, we also bypass the requirement of obtaining training data for infinitely many representations of the nonstationary characteristics to devise an efficient supervised learning algorithm.

While incrementally learning in a nonstationary environment, one major problem that the proposed intelligent radar system encounters is *concept drift*. Concept drift refers to changes in the distributions of the measured data (correspondingly the environment model) that are used for incremental learning [17]-[30]. Therefore, in order to learn in the presence of a concept drift, the radar system needs to employ incremental learning together with the active drift detection algorithms for detecting the nonstationarities in the environment. This can be implemented in two different ways: (1) active and (2) passive concept drift learning [22], [23].

We consider an *active drift learning* setup in this paper. This implies that the radar system first estimates the time when a change occurs in the environment distribution from which the measurement data are drawn, and then modifies its detection algorithm with the new distribution to continue the sequential learning. Consequently, this *data-driven* approach complements the existing techniques as it is not limited to any specific clutter/environment type and to any parametric modeling approach.

Mathematical Notations: In the rest of the paper, we assume that C_c for $c \in \{0,1\}$ denote the two classes that represent the absence and presence of the target, respectively. In other words, C_0 and C_1 are equivalent to $\mathcal{H}_0$ and $\mathcal{H}_1$ of radar target detection hypotheses, respectively. For every $c \in \{0,1\}$, $\mathcal{H}_0^c$ and $\mathcal{H}_1^c$ are the hypotheses corresponding to 'no-change' and 'change' in the distribution of data corresponding to class C_c. We denote the calibration data set collected for both classes as $\mathcal{Y}_0 = \{\boldsymbol{y}_t, r_t\}$ for $t = 1, \ldots, T_0$, with $\boldsymbol{y}_t$ as the radar measurement, T_0 as the number of measurements in the calibration data, and $r_t = 0$ if $\boldsymbol{y}_t \in C_0$ and $r_t = 1$ if $\boldsymbol{y}_t \in C_1$.

II. Proposed Detection Method

In order to analyze the nonstationarity in the environment returns, we assume that some initial distribution models have already been estimated from the pre-collected data. In this paper, we refer to the pre-collected data as *calibration data*. We mention here that the calibration data is more general than the well-known radar secondary data; this is because the calibration data include measurements of both the target and non-target components, whereas the conventional secondary data represents only the non-target (clutter) data.

After modeling the initial distributions, our next task is to develop learning algorithms that can effectively model the nonstationarities caused by the changes in the environment. An important step in that endeavor is to detect whether or not there is any change in the statistical properties of the environmental measurements. Assuming the change point to be deterministic but unknown, we calculate it by minimizing the supremum of the average detection delay conditioned on the observed radar measurements with a constraint on the mean time between false alarms.

Mathematically, this means that using the batches of radar data observed up until the measurement time n, $\mathcal{Y}_0^c, \ldots, \mathcal{Y}_n^c$, $c \in \{0,1\}$, we estimate the time when the change in the statistical characteristics of the data from class c is occurring by employing the following constrained optimization problem [31], [32]

$$\begin{aligned} \operatorname{argmin}_\tau \ \sup_{n\geq 1} \ \text{ess sup} \quad & E_n\left[(\tau-n)^+ | \mathcal{Y}_0^c, \ldots, \mathcal{Y}_{n-1}^c\right] \\ \text{such that} \quad & E_\infty[\tau] \geq \alpha \end{aligned} \tag{1}$$

where τ is the stopping time such that when it takes a value k (it means that there is a change point at or prior to time k); sup stands for supremum; ess sup is the essential supremum of a set of random variables that we define below in more detail; $x^+ = \max(0,x)$ for any variable x; $E_n[\cdot]$ is the expectation taken with respect to a distribution p_n^c (such that under p_n^c, $\{\mathcal{Y}_0^c, \ldots, \mathcal{Y}_{n-1}^c\}$ are independent and identically distributed (i.i.d.) with a fixed marginal distribution for $c \in \{0,1\}$); $E_\infty[\tau]$ represents the mean time between false alarms assuming that change never happens in the data stream; and α is a predefined threshold. Essential supremum of a set of random variables $\mathcal{X}$ is a random variable Z with the following properties: (i) $P(Z \geq X) = 1 \ \forall X \in \mathcal{X}$; and (ii) $\{P(Y \geq X) = 1, \forall X \in \mathcal{X}\} \rightarrow P(Y \geq Z) = 1, \forall X \in \mathcal{X}$, where $P(\cdot)$ represents the probability [31]. In light of the definitions above, the solution to the constraint optimization problem in (1) minimizes the supremum of the average delay conditioned on the worst case realization of $\{\mathcal{Y}_0^c, \ldots, \mathcal{Y}_{n-1}^c\}$ over all p_n, $n \geq 1$ [31], [32].

Now, depending on the characterization of p_n^c, the solution of the problem in (1) can be further sub-categorized into two different cases: (i) known parametric distribution of p_n^c, i.e., when we have prior knowledge on the calibration data such that the radar scattering from the environment follows a specific parametric distribution family; and (ii) unknown distribution of p_n^c, i.e., when we have no prior knowledge about the family of the distribution of the radar data. We consider the first case in this paper.

We obtain an optimal solution of (1) under the assumption that $\{\mathcal{Y}_0^c, \ldots, \mathcal{Y}_n^c\}$ for $c \in \{0,1\}$ are i.i.d. samples which are drawn from a parametric distribution with probability density function (pdf) $p_\Theta(\boldsymbol{y})$ [33]-[37]. When there is a change in the statistical characteristics of the new batch of radar data, we assume that the pdf model of the family does not change, but the change is modeled as a transition from one parameter θ_0^c (the null hypothesis for class C_c, $\mathcal{H}_0^c$) to another θ_1^c (alternative hypothesis for class C_c, $\mathcal{H}_1^c$). Recall that for a specific class C_c, $c \in \{0,1\}$ represents the absence or presence of target in the area covered by the radar, and $\mathcal{H}_0^c$ and $\mathcal{H}_1^c$ are two hypotheses corresponding to 'no-change' and 'change' in the distribution of the radar data obtained for class C_c.

The optimum stopping time τ for the problem in (1) is found using CUmulative SUM (CUSUM) method which is calculated as follows [33]-[37]:

$$\tau = \inf\left\{ n \geq 1 : \left(g(\mathcal{Y}_0^c,\ \ldots,\ \mathcal{Y}_n^c) = R_n - \min_{1\leq k \leq n} R_k \right) \geq b \right\}, \tag{2}$$

where n is the batch index as defined before; b is a predefined threshold; and $R_k = \sum_{t=1}^{k} \ln \frac{p_{\theta_1^c}(\mathcal{Y}_t^c)}{p_{\theta_0^c}(\mathcal{Y}_t^c)}$ [31]-[33]. Before a change point, the accumulated log-likelihood sum, R_n, moves towards $-\infty$; whereas after the change point, if the change happened in the favor of θ_1^c, R_n starts to move towards ∞. Therefore, the condition $(R_n - \min_{1\leq k\leq n} R_k) \geq b$ optimally estimates if and when the change in the behavior of R_n occurs within $1 \leq k \leq n$.

Calculating τ from (2) requires the knowledge of the parameters θ_0^c, θ_1^c, and threshold value b. We use the calibration data $\mathcal{Y}_0^c$ for every class $c \in \{0, 1\}$, and the knowledge of the parametric distribution underlying the data to find an estimate of θ_0^c. We then compute the confidence intervals for the estimation of the parameters θ_0^c, and assign the confidence interval extremas to θ_1^c. Once we have the parameter sets θ_0^c and θ_1^c, we compute (2) using the calibration data and assign $b = \max_{1\leq t\leq T_0} g(\mathcal{Y}_0^c)$ recalling that $\mathcal{Y}_0^c = \boldsymbol{y}_t | r_t = c$ for $t = 1,\ \ldots,\ T_0$ [33]. Following the approach described above, when there is a change in the statistical characteristics of the environment from one known parametric distribution to another known parametric distribution, we illustrate the performance of CUSUM on change point detection and the effect of adapting the detection algorithm to the new environment model on a target detection problem in Section III.

Note that when there is no prior knowledge about the family of the distribution of the radar data in $\mathcal{Y}_0^c$ for $c \in \{0, 1\}$, we can employ an extended version of the CUSUM method [33]. In this approach, we divide the set $\mathcal{Y}_0^c = \{\boldsymbol{y}_t | r_t = c\}$ for $t = 1,\ \ldots,\ T_0$ into subsets with cardinality K and compute

$$\boldsymbol{z}_k = \frac{1}{K} \sum_{\nu=(k-1)K+1}^{kK} \boldsymbol{y}_\nu, \qquad k = 1,\ \ldots,\ T_0/K \tag{3}$$

to form $\mathcal{Z}_0^c = \{\boldsymbol{z}_k | r_k = c\}$ for $k = 1,\ \ldots,\ T_0/K$. For a sufficiently large K, $\boldsymbol{z}_k$ can be approximated as a Gaussian distributed random variable with an unknown parameter set θ_0^c, denoted as $p_{\theta_0^c}(\cdot)$. If there is a change in the statistical properties of the radar data, the change is from a Gaussian random variable with parameter set θ_0^c to another Gaussian random variable with parameter set θ_1^c, denoted as $p_{\theta_1^c}(\cdot)$. This type of change fits in the framework of CUSUM. Therefore, for extended CUSUM, we can utilize a method similar to (2) to estimate the change point

$$\tau = \inf\left\{ n \geq 1 : \left(g(\mathcal{Z}_0^c,\ \ldots,\ \mathcal{Z}_n^c) = R_n - \min_{1\leq k \leq n} R_k \right) \geq b \right\} \tag{4}$$

where $R_k = \sum_{t=1}^{k} \ln \frac{p_{\theta_1^c}(\mathcal{Z}_t^c)}{p_{\theta_0^c}(\mathcal{Z}_t^c)}$; and $\mathcal{Z}_t^c$ is calculated from $\mathcal{Y}_t^c$ using the method in (3). Similar to the CUSUM method described earlier, to estimate the change point using (4), we estimate θ_0^c using the training data $\mathcal{Z}_0^c$ and estimate θ_1^c using the confidence interval extremas of $\hat{\theta}_0^c$ estimator. We also compute $b = \max_{1\leq k\leq T_0/K} g(\mathcal{Z}_0^c)$ recalling that $\mathcal{Z}_0^c = \boldsymbol{z}_k | r_k = c$ for $k = 1,\ \ldots,\ T_0/K$ [33].

III. Numerical Examples

In this section, we present simulation results of target detection in the presence of nonstationary interference. We show improved performance by using a detector that applies the CUSUM method to detect and estimate the change in the clutter distribution.

To setup the problem, we assume that the radar is collecting measurements from multiple range cells (indexed by j) over a sequence of coherent processing intervals (indexed by k). In each interval, it receives and processes N temporal samples. Without loss of generality, the target is assumed to be present in the $j = 1$ range cell, and it remains in that cell during the entire processing of $k = 1, 2, \ldots, K$ coherent intervals. Further, the target response is considered to be known and constant, which we denote as a. Then, the detection problem of the jth range cell at the k coherent interval can be expressed as

$$\begin{cases} \mathcal{H}_0: & \boldsymbol{y}_k^{(j)} = \boldsymbol{n}_k^{(j)} \\ \mathcal{H}_1: & \boldsymbol{y}_k^{(j)} = a\mathbf{1} + \boldsymbol{n}_k^{(j)} \end{cases}, \tag{5}$$

for $k = 1, 2, \ldots, K$, and $j = 0, 1, 2, \ldots$, where each vector is of dimension $N \times 1$. We model the nonstationarity in the clutter returns, $\boldsymbol{n}_k^{(j)}$, by representing a change of distribution from Gaussian to compound-Gaussian at the processing interval $k = k_0$. Therefore, for $k = 1, 2, \ldots, k_0$, we have $\boldsymbol{n}_k^{(j)} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$ with a known σ^2. However, for $k = k_0 + 1, \ldots, K$, the clutter distribution modifies to $\boldsymbol{n}_k^{(j)} \sim \mathcal{N}(\mathbf{0}, (1/\sqrt{v_k})\sigma^2 \boldsymbol{I})$, where $(1/v_k)$ follows a gamma distribution with unit mean and a known shape parameter $\nu > 0$.

To deal with such a nonstationary clutter characteristics, our proposed machine learning based radar first checks for any change in the clutter distribution by employing the CUSUM test, and then modifies the detection algorithm in accordance with the changed (if any) distribution.

We detect the change point of the clutter distribution by testing for the confidence intervals for the sample mean and sample variance. Specifically, we assume the availability of a batch of N_t training measurements $\{\mathcal{Y}_0^c,\ \ldots,\ \mathcal{Y}_{N_\mathrm{t}}^c\}$, where $c \in \{0, 1\}$ denote the two classes that represent the absence and presence of the target, respectively; and $\mathcal{Y}_n^c = \{\boldsymbol{y}_n^c, r_n^c\}$ with $r_n^0 = 0$ when $\boldsymbol{y}_n^0$ belongs to the target-absent class $\mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$, and $r_n^1 = 1$ when $\boldsymbol{y}_n^1$ belongs to the target-present class $\mathcal{N}(a\mathbf{1}, \sigma^2 \boldsymbol{I})$. Then, for each class $c \in \{0, 1\}$, we compute the sample mean and sample variances as

$$\begin{aligned} \hat{\mu}^0 &= \frac{1}{N_\mathrm{t} N} \sum_{n=1}^{N_\mathrm{t}} \mathbf{1}^T \boldsymbol{y}_n^0 \\ \hat{\mu}^1 &= \frac{1}{N_\mathrm{t} N} \sum_{n=1}^{N_\mathrm{t}} \mathbf{1}^T \boldsymbol{y}_n^1 \\ {s^2}^0 &= \frac{1}{N_\mathrm{t} N - 1} \sum_{n=1}^{N_\mathrm{t}} (\boldsymbol{y}_n^0 - \hat{\mu}^0 \mathbf{1})^T (\boldsymbol{y}_n^0 - \hat{\mu}^0 \mathbf{1}) \\ {s^2}^1 &= \frac{1}{N_\mathrm{t} N - 1} \sum_{n=1}^{N_\mathrm{t}} (\boldsymbol{y}_n^1 - \hat{\mu}^1 \mathbf{1})^T (\boldsymbol{y}_n^1 - \hat{\mu}^1 \mathbf{1})\ . \end{aligned} \tag{6}$$

Subsequently, noting that the sample mean and variance follow Gaussian distribution and chi-square distribution with N_tN-1 degrees of freedom, respectively, the lower and upper limits on the 95% confidence intervals of mean and variance under $\mathcal{H}_1$ are calculated as

$$a_{\rm ll} = \hat{\mu}^1 - 1.96\frac{\sigma}{\sqrt{N_{\rm t}N}}, \qquad a_{\rm ul} = \hat{\mu}^1 + 1.96\frac{\sigma}{\sqrt{N_{\rm t}N}}\ , \quad (7)$$

$$\sigma^2_{\rm ll} = \frac{(N_{\rm t}N-1){s^2}^1}{\chi^2_{N_{\rm t}N}(\alpha/2)}, \qquad \sigma^2_{\rm ul} = \frac{(N_{\rm t}N-1){s^2}^1}{\chi^2_{N_{\rm t}N}(1-\alpha/2)}\ , \quad (8)$$

where $\alpha = 0.05$. Next, we evaluate $R_k = \sum_{n=1}^{k} \ln \frac{p_{\theta_1^1}(\boldsymbol{y}_n^1)}{p_{\theta_0^1}(\boldsymbol{y}_n^1)}$ and $g_n = R_n - \min_{1\leq k\leq n} R_k$, and use the CUSUM method of (2) to determine the change in the distribution with parameters from $\theta_0^1 = [a, \sigma^2]$ to either $\theta_1^1 = [a_{\rm ll}, \sigma^2_{\rm ll}]$, or $\theta_1^1 = [a_{\rm ll}, \sigma^2_{\rm ul}]$, or $\theta_1^1 = [a_{\rm ul}, \sigma^2_{\rm ll}]$, or $\theta_1^1 = [a_{\rm ul}, \sigma^2_{\rm ul}]$. The change in the distribution is declared when $g_n \geq b$, where b is chosen from the training data as $b = \max_{1\leq n\leq N_{\rm t}} g_n$. Similar training is also done with the data obtained for $\mathcal{H}_0$. A change is declared in the statistical conditions of environment when change is detected in the distributions under $\mathcal{H}_0$ and/or $\mathcal{H}_1$.

The performance characteristic of the change point detection and estimation is shown in Fig. 1 in terms of the cumulative distribution function of the estimation delay. It clearly demonstrates that for more than 96.5% of the time the change in the clutter distribution is detected by the proposed radar within just one processing interval. In other words, this implies that our method can detect and estimate a change in the clutter characteristics almost as soon as it happens.

Once such a change in the clutter distribution is detected at $k = k_0$, the proposed radar accordingly modifies the log-likelihood ratio computation from $\ln[p(\boldsymbol{y}_k^{(j)}|\mu, \sigma^2; \mathcal{H}_1)/p(\boldsymbol{y}_k^{(j)}|\sigma^2; \mathcal{H}_0)]$, for $k = 1, 2, \ldots, k_0$, to $\ln[p(\boldsymbol{y}_k^{(j)}|\mu, \nu, \sigma^2; \mathcal{H}_1)/p(\boldsymbol{y}_k^{(j)}|\nu, \sigma^2; \mathcal{H}_0)]$ for $k = k_0 + 1, \ldots, K$. This corresponds to the modification of the test statistic from $\{\boldsymbol{1}^T \boldsymbol{y}_k^{(j)}, k = 1, 2, \ldots, k_0\}$ to $\{\ln[{\boldsymbol{y}_k^{(j)}}^T \boldsymbol{y}_k^{(j)} + \sigma^2\nu] - \ln[(\boldsymbol{y}_k^{(j)} - a\boldsymbol{1})^T(\boldsymbol{y}_k^{(j)} - a\boldsymbol{1}) + \sigma^2\nu], k = k_0+1, \ldots, K\}$ for every range cell j.

The detection performance of the proposed algorithm is shown in Fig. 2 in terms of the receiver operating characteristics (ROCs) at two different SNR values. This plot additionally includes the ROCs of a conventional radar detector that does not understand the change in the clutter distribution and applies the standard log-likelihood ratio $\ln[p(\boldsymbol{y}_k^{(j)}|\mu, \sigma^2; \mathcal{H}_1)/p(\boldsymbol{y}_k^{(j)}|\sigma^2; \mathcal{H}_0)]$ for all the processing intervals $k = 1, 2, \ldots, K$. Comparing the blue with red curves, particularly in lower probability of false alarm regions where a radar system typically operates, a substantial improvement in the detection performance of the proposed radar due to the incorporation of the modified clutter characteristics is clearly evident from these ROC plots.

IV. Conclusions

In this paper we developed a machine learning based detection algorithm to detect a target in the presence of nonstationary environment (clutter). It is not possible to devise a supervised learning algorithm to model infinitely many representations of the nonstationary characteristics. Therefore, we employed the incremental learning and drift detection algorithms for building a detection algorithm that incrementally learns the environment and updates the system parameters on the fly. In addition to incremental learning, we used an active drift learning technique to detect and estimate any change-point (if present) in the environment distribution. Our numerical examples showed that the proposed method is able to quickly detect a change in the underlying clutter distribution, and as a consequence produced a substantially improved detection performance compared to a conventional algorithm that was not aware of any environmental change. In our future work, we will extend our model to incorporate a Bayesian formulation of the change point parameter. We will also explore the active drift learning under noisy labels and passive drift leaning methodologies. Additionally, we will validate the performance of our proposed technique with real data.

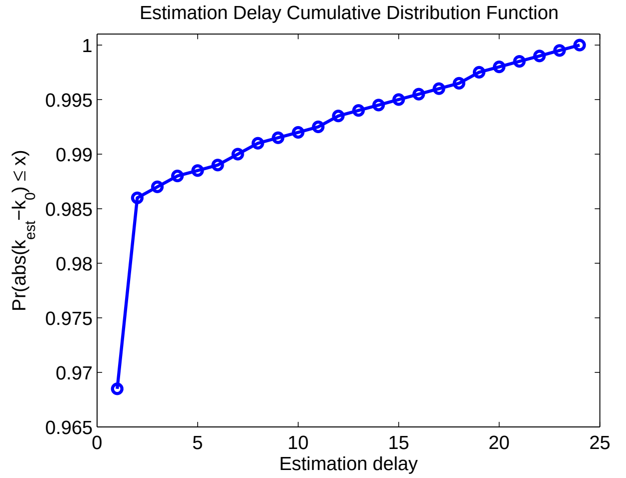

Fig. 1. Performance of the change point detection test in terms of the cumulative distribution function of the estimation delay.

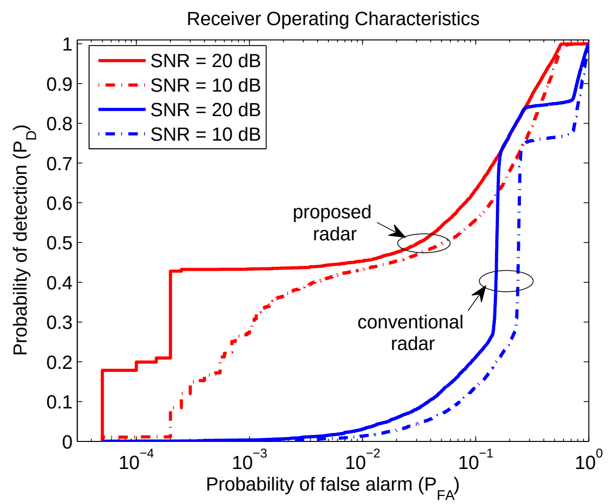

Fig. 2. Improved detection performance using the proposed radar.

References

[1] S. Haykin, "Cognitive radar: A way of the future," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 30–40, Jan. 2006.

[2] M. Wicks, "Cognitive radar: A way forward," in *2011 IEEE Radar Conference*, May 2011, pp. 012–017.

[3] J. Guerci, R. Guerci, M. Rangaswamy, J. Bergin, and M. Wicks, "Cofar: Cognitive fully adaptive radar," in *2014 IEEE Radar Conference*, May 2014, pp. 0984–0989.

[4] A. Ozturk, "An application of a distribution identification algorithm to signal detection problems," in *The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, Nov 1993, pp. 248–252 vol.1.

[5] J. Metcalf, S. Blunt, and B. Himed, "A machine learning approach to distribution identification in non-gaussian clutter," in *2014 IEEE Radar Conference*, May 2014, pp. 0739–0744.

[6] P. Chen, W. L. Melvin, and M. Wicks, "Screening among multivariate normal data," *Journal of Multivariate Analysis*, vol. 69, no. 1, pp. 10 – 29, 1999.

[7] P. Wang, H. Li, and B. Himed, "Knowledge-aided parametric tests for multichannel adaptive signal detection," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5970–5982, Dec 2011.

[8] ——, "Bayesian parametric GLRT for knowledge-aided space-time adaptive processing," in *2011 IEEE Radar Conference*, May 2011, pp. 329–332.

[9] P. Wang, Z. Sahinoglu, M.-O. Pun, H. Li, and B. Himed, "Knowledge-aided adaptive coherence estimator in stochastic partially homogeneous environments," *IEEE Signal Processing Letters*, vol. 18, no. 3, pp. 193–196, March 2011.

[10] P. Wang, H. Li, Z. Wang, and B. Himed, "Knowledge-aided parametric adaptive matched filter with automatic combining for covariance estimation," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6067–6071.

[11] J. R. Guerci and E. J. Baranoski, "Knowledge-aided adaptive radar at DARPA: An overview," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 41 – 50, Jan. 2006.

[12] W. Melvin and J. Guerci, "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 3, pp. 983–996, July 2006.

[13] P. Setlur and M. Rangaswamy, "Signal dependent clutter waveform design for radar STAP," in *2014 IEEE Radar Conference*, May 2014, pp. 1311–1316.

[14] M. Rangaswamy, J. Michels, and B. Himed, "Statistical analysis of the non-homogeneity detector for STAP applications," *Digital Signal Processing*, vol. 14, no. 3, pp. 253 – 267, 2004.

[15] K. Bell, C. Baker, G. Smith, J. Johnson, and M. Rangaswamy, "Fully adaptive radar for target tracking part I: Single target tracking," in *2014 IEEE Radar Conference*, May 2014, pp. 0303–0308.

[16] ——, "Fully adaptive radar for target tracking part II: Target detection and track initiation," in *2014 IEEE Radar Conference*, May 2014, pp. 0309–0314.

[17] J. Schlimmer and R. Granger, "Incremental learning from noisy data," *Machine Learning*, vol. 1, no. 3, pp. 317–354, 1986.

[18] G. Widmer and M. Kubat, "Learning in the presence of concept drift and hidden contexts," *Machine Learning*, vol. 23, no. 1, pp. 69–101, 1996.

[19] L. Kuncheva, "Classifier ensembles for changing environments," in *Multiple Classifier Systems*, ser. Lecture Notes in Computer Science, F. Roli, J. Kittler, and T. Windeatt, Eds. Springer Berlin Heidelberg, 2004, vol. 3077, pp. 1–15.

[20] P. Vorburger and A. Bernstein, "Entropy-based concept shift detection," in *Proc. Sixth International Conference on Data Mining ICDM*, Dec 2006, pp. 1113–1118.

[21] J. Z. Kolter and M. A. Maloof, "Dynamic weighted majority: an ensemble method for drifting concepts," *The Journal of Machine Learning Research*, vol. 8, pp. 2755 – 2790, Dec. 2007.

[22] R. Elwell and R. Polikar, "Incremental learning of concept drift in nonstationary environments," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1517–1531, Oct. 2011.

[23] R. Polikar, "Ensemble learning," in *Ensemble Machine Learning: Methods and Applications*, C. Zhang and Y. Ma, Eds. Springer, 2012.

[24] A. Tsymbal, "The problem of concept drift: definitions and related work," Trinity College, Dublin, Ireland, Tech. Rep. TCD-CS-2004-15, 2004.

[25] E. A. de Oliveira, "The rosenblatt bayesian algorithm learning in a nonstationary environment," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 584 – 588, Mar. 2007.

[26] S. Hoeglinger and R. Pears, "Use of hoeffding trees in concept based data stream mining," in *Third International Conference on Information and Automation for Sustainability (ICIAFS)*, Dec 2007, pp. 57–62.

[27] A. Tsymbal, M. Pechenizkiy, P. Cunnigham, and S. Puuronen, "Dynamic integration of classifiers for handling concept drift," *Information Fusion*, vol. 9, no. 1, pp. 56 – 68, Jan. 2008.

[28] M. M. Masud, J. Gao, L. Khan, J. Han, and B. Thuraisingham, "Integrating novel class detection with classification for concept-drifting data streams," in *Machine Learning and Knowledge Discovery in Databases*, ser. Lecture Notes in Computer Science (LNCS), vol. 5782, 2009, pp. 79–94.

[29] D. R. Lowne, S. J. Roberts, and R. Garnett, "Sequential non-stationary dynamic classification with sparse feedback," *Pattern Recognition*, vol. 43, pp. 897 – 905, 2010.

[30] M. M. Masud, J. Gao, L. Khan, J. Han, and B. Thuraisingham, "Classification and novel class detection in concept-drifting data streams under time constraints," *IEEE Trans. Knowledge and Data Engineering*, vol. 23, no. 6, pp. 859 – 874, Jun. 2011.

[31] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge University Press, 2008.

[32] V. Veeravalli and T. Banarjee, "Quickest change detection," Elsevier: E-reference Signal Processing, 2013, available at [arXiv:1210.5552].

[33] C. Alippi and M. Roveri, "Just-in-time adaptive classifiers—Part I: Detecting nonstationary changes," *IEEE Transactions on Neural Networks*, vol. 19, no. 7, pp. 1145–1153, July 2008.

[34] G. Lorden, "Procedures for reacting to a change in distribution," *Ann. Math. Statist.*, vol. 42, no. 6, pp. 1897–1908, Dec. 1971.

[35] M. Pollak, "Optimal detection of a change in distribution," *Ann. Statist.*, vol. 13, pp. 206–227, Mar. 1985.

[36] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Statist.*, vol. 14, pp. 1379–1387, Dec. 1986.

[37] Y. Ritov, "Decision theoretic optimality of the CUSUM procedure," *Ann. Statist.*, vol. 18, pp. 1464–1469, Nov. 1990.