Metacognition for Radar Coexistence

Anthony F. Martone[1], Kelly D. Sherbondy[1], Jacob A. Kovarskiy[2], Benjamin H. Kirk[2], Charles E. Thornton[3], Jonathan W. Owen[4], Brandon Ravenscroft[4], Austin Egbert[5], Adam Goad[5], Angelique Dockendorf[5], R. Michael Buehrer[3], Ram M. Narayanan[2], Shannon D. Blunt[4], and Charles Baylis[5]

[1] U.S. Army Research Laboratory, Adelphi, MD
[2] Electrical Engineering Department, The Pennsylvania State University, University Park, PA
[3] Bradley Department of Electrical and Computer Engineering, The Virginia Polytechnic Institute and State University, Blacksburg, VA
[4] Radar Systems Lab, University of Kansas, Lawrence, KS
[5]Wireless and Microwave Circuits and Systems Program, Baylor University, Waco, TX 76798-7356

Abstract— **In this paper we investigate a metacognitive radar (MCR) model that comprehensively combines disparate cognitive radar (CR) strategies for advanced performance in congested electromagnetic environments (EME). This model changes CR strategies as the spectral environment and target evolve for efficient radar dynamic spectrum access (DSA). The model first implements spectrum sensing followed by spectrum classification to identify known EME scenarios. These spectral scenarios assess the congestion and complexity of time-frequency data collected by the passive sensing process. This evaluation prioritizes possible CR strategies that are effective for the given spectral conditions. The MCR model then evaluates different CR strategies via learning and selects the technique that provides the best radar performance.**

Keywords—cognitive radar, spectrum sharing, metacognition, spectrum sensing, waveform diversity, machine learning

I. INTRODUCTION

Since the seminal paper on "Cognitive Radar" by Haykin in 2006, there has been an explosion of associated research in the areas of adaptive signal processing and component technology [1,2]. From waveform design to machine learning, several cognitive radar (CR) strategies have been investigated and shown to greatly improve performance for challenging scenarios by introducing a feedback and control process, i.e. the perception action cycle (PAC), to adapt to the environment and change the behavior of the radar [3]. Although these CR strategies extend the state-of-the-art for radar, they are often narrow in scope for particular conditions and scenarios. Thus the efficacy of a single CR strategy can be limited for realistic environments that change dynamically. Furthermore, as software defined radio technologies evolve, it will soon be possible to integrate a wide variety of CR strategies into a single platform for multi-function capabilities. A more comprehensive feedback and control process is therefore required in order to utilize the appropriate CR strategy for changing scenarios, i.e. a best-of-breed approach that recognizes when to use certain CR strategies and when to change between them.

A prospective best-of-breed framework involves the bio-inspired concept of "metacognition," which in this context could monitor and self-regulate the operation of various CR strategies. Metacognition is a well-studied topic in the field of human learning [4] and has been applied to artificial intelligence models for a wide variety of topics [5]. Metacognition for radio and radar has been explored in a very limited quantity and is a relatively new topic for radar [6-8]. The metacognitive radio models use *Meta Knowledge*, *Meta Monitoring*, and *Meta Control* to select the appropriate cognitive radio engine, or agent, for changing scenarios. Meta Knowledge defines the performance and learning capabilities of each agent for particular scenarios. Meta Monitoring determines if/when the operating scenario changes. Meta Control selects different CE's based on radio performance.

In this paper we evolve the fundamental metacognitive radio model for radar by designing a model capable of selecting CR strategies in real-time using a software defined radar (SDRadar). The challenge is to develop a model that selects the most advantageous CR strategy based on the changing environment and real-time radar performance assessment. Here we address the application of metacognition for radar and communication systems spectrum sharing [9], a necessary consideration due to the ever-growing wireless communication industry and its need for more bandwidth. In particular, we focus on non-cooperative coexistence for radar to dynamically access the spectrum in real-time and develop spectral situational awareness about the EME. Solutions to this dynamic access challenge have been investigated using the SDRadar platform discussed in [10, 11]. This radar uses a Universal Software Radio Peripheral (USRP) and a host computer to sense, analyze, and decide an action within the coherent processing interval (CPI). The CR strategies considered in this paper have been previously investigated in [10-18] (further described in Section III) and referred to as the *DSA CR strategies*, which have advantages and disadvantages for different spectrum scenarios.

To that end, a metacognitive radar (MCR) model is developed in Section II that comprehensively combines the above disparate DSA CR strategies for advanced radar performance in congested spectral environments. This model changes CR strategies in real-time as the spectral environment and target evolve. The essential functions of this model utilize similar components of the metacognitive radio design. The model first implements a high-level case-based reasoning approach to identify known spectral scenarios of the EME. These spectral scenarios assess the congestion and complexity level of time-frequency data collected by passive sensing. This evaluation prioritizes possible CR strategies that are effective for the given spectral conditions. The MCR model then implements different CR strategies and selects the one that provides the best performance.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. Metacognitive Radar Model for DSA

The MCR model for DSA is shown in Fig. 1. The MCR Engine is formed using MCR Knowledge, MCR Monitoring, and MCR Control components in the same manner as [6,7]. The MCR Knowledge defines the learning rate and capabilities for each CR strategy for particular scenarios of interest, essentially highlighting *when* and *how* to implement the strategy for an observed scenario. The MCR Monitoring classifies scenarios according to CR strategies based on observations of the environment so that the appropriate CR strategy is implemented at the proper time. Finally, MCR Control defines the regulation of the learning process for implementing CR strategies. This task requires the metacognition to examine the effectiveness of a particular CR strategy during radar operation and determine when to switch to a new approach.

Radar DSA requires efficient implementation of the PAC with the overall goal of modifying its center frequency and bandwidth to operate within a sub-band β of the overall bandwidth B, where $\beta \le B$ so that mutual RFI is mitigated. This approach has been shown to jointly maximize the signal to interference plus noise ratio (SINR) and range resolution objective functions via multi-objective optimization [11]. Each CR strategy considers and implements a different approach to adjust the center frequency and bandwidth along with an associated waveform type and adaptation method. In this development, the linear frequency modulated (LFM) and spectrally-notched random FM waveforms are considered [14]. The adaptation methods are discussed in Section III.

As per the MCR block diagram in Fig. 1, passive spectrum sensing is first implemented to collect a sequence of I/Q time domain samples within the radar operating band B. Passive sensing requires that the radar remains inactive during the sensing period. A power spectrum discretized to size N, defined as $X = \{x_1, \ldots, x_N\}$, is then estimated using the Fourier transform for frequencies $F = \{f_1, \ldots, f_N\}$. This process is repeated M times on contiguous (non-overlapping) data to create a set of contiguous power spectra. The notation for X is expanded to $X_m = \{x_{m,1}, \ldots, x_{m,N}\}$ to represent the m^{th} power spectrum estimate and used to form the information matrix (i.e. the spectrogram) denoted as

$$\Theta = \begin{bmatrix} X_1 \\ . \\ X_M \end{bmatrix} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,N} \\ \vdots & \ddots & \vdots \\ x_{M,1} & \cdots & x_{M,N} \end{bmatrix}. \quad (1)$$

Note that Θ in (1) captures the time-frequency behavior of spectral activity, which is next classified into a category.

Define the set of C spectral categories as $\Omega = \{\omega_1, \ldots, \omega_C\}$, which are determined *a priori* to represent possible spectrum scenarios and activities. It is shown in Section IV that these categories are formulated based on the congestion and complexity of the RFI. The classification of Θ is defined as

$$\omega_c = \Omega(\Theta), \quad (2)$$

where $\Omega: \Theta \mapsto \omega_c$ is a generalized classification function that maps the spectral information in Θ to the category ω_c. The specific classification procedure examined here (Sect. III) is a case-based reasoning approach.

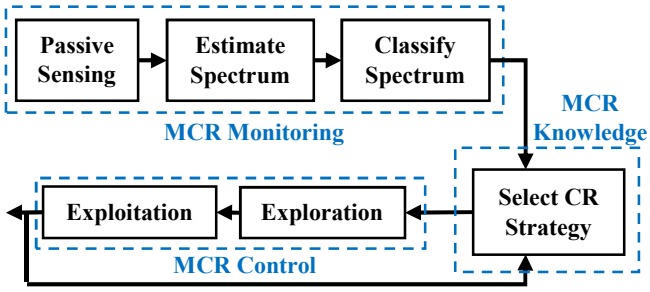

Fig. 1: Block diagram of MCR model for DSA

Once the spectral category is determined via (2), a subset of CR strategies are selected for radar implementation. The set of R CR strategies is defined as $\Psi = \{\Psi_1, \ldots, \Psi_R\}$. To achieve convergence to the "optimal" CR strategy, i.e., one that yields the greatest radar performance over time, we develop a formulation inspired by the multi-armed bandit problem [19]. This reinforcement learning problem balances the *exploration* and *exploitation* trade-off. Exploration samples the potential options to measure the likelihood of success, while exploitation uses the measured information to exploit the option that optimizes performance. This work uses the "explore-first" approach [19] where the MCR engine first sequentially samples CR strategies to evaluate performance before exploiting the "best" option.

During the initial exploration phase each strategy in Ψ is sequentially selected for radar operation to measure performance. The radar uses each selected strategy for Q iterations, where each iteration implements a radar CPI. A function that estimates the *expected* radar performance is also initialized and adjusted for each implementation within the Q iterations. Define this heuristic function for $\Psi_r \in \Psi$ at time q as

$$\lambda(r, q) = \begin{cases} I_c(r) & q = 1 \\ \lambda(r, q-1) + \sigma_L[Z(q) - Z(q-1)] & q > 1, \end{cases} \quad (3)$$

where $0 \le \lambda(r, q) \le 1$ is the expected success, or reward, of the given CR strategy, $I_c(r)$ is the initial condition at $q = 1$ that is determined *a priori* based on MCR Knowledge, $0 \le Z(q) \le 1$ is some normalized measure of radar performance, and σ_L defines the learning rate.

Radar performance is measured by a set of objective functions. For example, in [11] performance was determined by jointly maximizing SINR and range resolution using multi-objective optimization. After (3) is determined for all the CR strategies in Ψ during the exploration phase, an exploitation phase is then implemented to select the best CR strategy Ψ_{r^*}, where

$$r^* = \arg\max_r \lambda(r, q). \quad (4)$$

The exploitation phase continuously monitors the spectrum for changing activity every CPI cycle. If the category selected by (2) changes, indicating different spectral conditions, then a new set of CR strategies are used and the exploration and exploitation phases are reiterated.

III. Overview of CR Strategies

There are several CR strategies that accomplish effective spectrum sharing. These techniques can be modeled as PACs with three steps: 1) sensing the RF spectrum, 2) processing the power spectrum, and 3) deciding upon and executing an appropriate action. Each of the CR strategies considered here operates on refined spectral information that is obtained by discretizing the power spectra into "sub-bands" using the fast spectrum sensing (FSS) technique outlined in [11]. This process occurs during MCR Monitoring, though it is used again for MCR Control during radar operation (discussed below).

Specifically, FSS identifies regions of low- and high-power RFI, where each region corresponds to a sub-band. The sub-bands constitute a new, refined power spectrum and are used to replace Θ as labeled input to each CR strategy. This process first calculates an average power spectrum from the M spectrograms in Θ as

$$\widehat{\Theta} = \{\hat{x}_1, \ldots, \hat{x}_N\} = (1/M) \sum_{m=1}^{M} X_m , \quad (5)$$

from which the available sub-bands are determined via

$$\Gamma = \square_{\phi}(\widehat{\Theta}). \quad (6)$$

Here $\square_{\phi}: \widehat{\Theta} \mapsto \Gamma$, where $\Gamma = \{\Gamma_1, \ldots, \Gamma_K\}$ is a set of $K \leq N$ sub-bands in the vector $\widehat{\Theta}$. The sub-band interference state $\Gamma_k \in \{0,1\}$ has an associated start frequency bin index $S_k \in \{1, \ldots, N\}$, end frequency bin index $E_k \in \{1, \ldots, N\}$, bandwidth β_k, number of frequency bins $l_k \in \{1, \ldots, N\}$, and RFI power

$$P_1(k) = \sum_{n=S_k}^{E_k} \hat{x}_n . \quad (7)$$

The sub-bands in (6) alternate between low power, $\Gamma_k = 0$ (RFI-free), and high power, $\Gamma_k = 1$ (RFI) sub-bands.

All CR strategy implementations occur during the exploration phase (Fig. 1). The first CR strategy is reacts to avoid RFI and uses the FSS algorithm directly during the radar operation to process sensed spectra [10,11]. This approach requires that the radar passively senses the EME between pulses within the CPI to produce a power spectrum, to which (6) is applied to determine the sub-bands. Then the instantaneous bandwidth and center frequency of the largest available sub-band is selected for radar operation.

Note that the passive sensing and FSS processing stage used for the radar operation (MCR Control) and spectrum classification (MCR Monitoring) are identical, but occur at different times. Using FSS during radar operation has the advantage of efficient processing and low latency, thereby facilitating rapid PAC implementation. This technique is not affected by the type of RFI signal modulation or the number of RFI sources, and is dependent on the rate at which the RFI spectrum changes.

The second CR strategy predicts to avoid RFI via a stochastic approach. This approach involves a model that is trained relatively quickly and assumes that the RFI statistics over a short interval are representative of a longer interval. Specifically, this approach employs an alternating renewal process (ARP) model to measure the likelihood of available frequency sub-bands [13,20]. The ARP model is trained using the information in Θ to develop statistics on the sub-bands from (6) based on either parametric or non-parametric variants [13]. Then, given the current spectral environment at time v, the likelihood of available frequency sub-bands at a later time Δ_v is estimated. This likelihood function estimates the probability of unused frequency sub-bands at $(v + \Delta_v)$ based on the detected RFI states from the training information. A threshold is next used to determine the edges of the unused frequency sub-bands. Contiguous sub-bands above the threshold are subsequently concatenated to determine the widest available bandwidth (and associated center frequency).

Another CR strategy uses a Reinforcement Learning (RL) approach that treats the CR's environment as a Markov Decision Process (MDP). An MDP is a mathematical model for reinforcement learning (RL); a field that uses human-defined rewards to guide a machine's learning process [21]. A particular MDP is characterized by the set of states and actions the radar observes, the state transitions, and the reward function that maps pairs of states and actions to numeric rewards. The goal of RL is to find a policy π, which is a rule that the radar utilizes to select actions given observed states, and can be found directly with a dynamic programming approach or indirectly via function approximation with deep neural networks [21]. In [17], the former was demonstrated for CR simulations, although the MDP dimensionality was high due to the large state space. Conversely, in [12] the latter approach was shown to reduce the dimensionality of the simulated learning problem, but still required a significant training period. However, after this training period has ended, learned behavior can guide CR waveform selection with minimal complexity.

The fourth technique spectrally notches the transmit waveform rather than adapting the center frequency and bandwidth. This technique relies on spectral shaping optimization to generate instantiations of random FM waveforms that are constant modulus in the time domain and possess an enforced spectral mask containing notches based on the FSS response [14], where notch depth depends on the particular shaping optimization and corresponding number of iterations. The main advantage of this approach is that the RF spectrum can be utilized at the full bandwidth, without notched regions. It has also very recently been shown that SRN can likewise be combined with prediction [18] to facilitate yet another CR strategy.

The above CR strategies modify the radar waveform from CPI to CPI and possibly from pulse to pulse. The proposed approach of Fig. 1 considers a high-power tunable matching network [22] to maximize the Power Added Efficiency (PAE) or output power of the transmitter front-end power amplifier. The goal of the technique is to retune a matching network to maintain the best amplifier efficiency and spectral containment of the transmitted waveform as the transmit parameters are varied by the above DSA implementations. It has been shown that the output power of the radar transmission is significantly improved with this procedure [16]. The disadvantage that this technique may have is that the digital DSA approaches described above can adapt much faster than the physical high-power tuner, which could cause sub-optimal matching. The present high-power tuner technology relies on mechanical tuning techniques, so it requires at least tens of milliseconds to perform a tuning operation. Tuning within the metacognitive process thus focuses on maximizing the overall performance, rather than instantaneous performance [23].

Given this complex space of CR strategies to accomplish DSA, there is a need to deconstruct these approaches into simpler categories for easy integration. We also adopt a more structured nomenclature to explain how these techniques are used in conjunction with one another. The various steps in the methods described above can be allocated into the following generic categories: Front-End, Adaptation Method, Waveform Synthesis, and Radar Processing. Consequently, Fig. 2 shows a merged system concept in which each aspect of these CR strategies is extracted, categorized, and combined for use on the SDRadar.

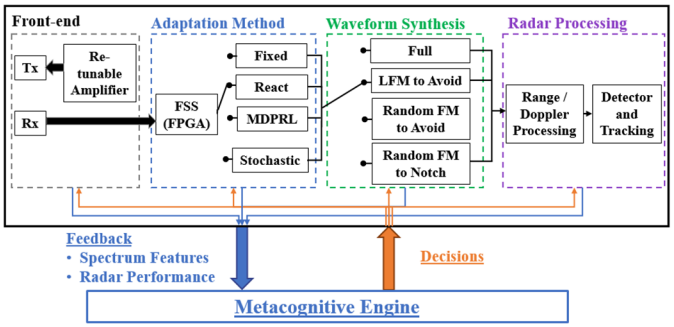

Fig. 2: Efficient implementation of CR strategies on the SDRadar.

The Front-End contains the tunable RF matching network to maximize the output power of the transmit waveform. It is possible to complement this matching network with additional tuning components, e.g. an adaptable filter bank. The Adaptation Method aspect in Fig. 2 determines the method for adaptation, which consists of Fixed, React, MDP with RL (MDPRL), and Stochastic. Note that these Adaptation Method elements require spectrum refinement using the FSS method of (6).

The Waveform Synthesis component of Fig. 2 determines the maneuverability of the waveform, the elements of which include full-band transmission, partial band transmission with LFM to avoid, random FM to avoid, and random FM to notch. The maneuverability for avoidance and notching is a form of frequency hopping dependent on the RFI. Finally, the Radar Processing is used to estimate range-Doppler information and to detect targets. Each of these four categories switch between their constituent elements to provide ample combinations so that the appropriate processing path for a given RFI scenario can be chosen to maximize performance.

IV. Case Based Reasoning for Spectrum Classification

The classification mapping function of (2) is defined using the tiered knowledge-based classifier described in Fig. 3. Classification and CR strategy selection are based on two features estimated by the spectrum data in Θ for each sub-band determined by (6). These features are designed to provide a "generalized classification" to down-select the CR strategies in Ψ. The first feature measures the *average congestion level* defined as

$$\overline{C}_{\mathrm{G}} = \frac{1}{B}\sum_{k=1}^{K} \Gamma_k \beta_k \,. \tag{8}$$

Note that $\overline{C}_{\mathrm{G}}$ is the percent bandwidth of occupied sub-bands as indicated by Γ_k. The congestion level provides knowledge of how to synthesize the waveform. For example, a very low $\overline{C}_{\mathrm{G}}$ indicates that a Full-Bandwidth transmission is possible, while a very high $\overline{C}_{\mathrm{G}}$ indicates that narrowband transmissions are necessary. Furthermore, moderately low $\overline{C}_{\mathrm{G}}$ implies low bandwidth interference, which suggests a notching scenario.

The second feature measures the *average complexity* of the RFI for the occupied sub-bands indicated by Γ:

$$C_{\mathrm{x}} = \frac{1}{N_\eta}\sum_{k=1}^{K} \Gamma_k \left[\frac{\sigma_{\mathrm{on}}(k)}{\mu_{\mathrm{on}}(k)} + \frac{\sigma_{\mathrm{off}}(k)}{\mu_{\mathrm{off}}(k)}\right] / 2, \tag{9}$$

where $N_\eta \le K$ is the number of sub-bands with RFI, $\mu_{\mathrm{on}}(k)$ is the average on-time of the RFI in sub-band Γ_k for $\sigma_{\mathrm{on}}(k)$ is its standard deviation, and $\mu_{\mathrm{off}}(k)$ is the average off-time of the RFI in Γ_k for $\sigma_{\mathrm{off}}(k)$ its standard deviation. The development of these statistics require modeling the RFI as a binary, alternating renewal process and determination of how long the RFI is "on" and how long it stays "off" over the observation of the M collected spectra [13]. It is further required that (9) is normalized using

$$\overline{C}_{\mathrm{X}} = C_{\mathrm{x}} / C_{\mathrm{x,MAX}} \,, \tag{10}$$

where $C_{\mathrm{x,MAX}}$ is the maximum obtainable C_{x} value from (9) and is determined *a priori* based on the RFI type. Here complexity is defined in terms of the randomness of the occupied sub-bands and is used to identify the appropriate Adaptation Method. For example, if several sub-bands are occupied with highly random RFI then the MDPRL and Stochastic methods are less effective than the React approach.

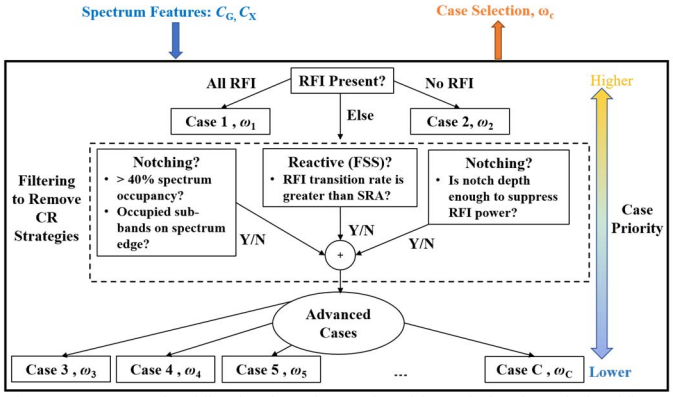

Fig. 3: Spectrum classification based on a tiered knowledge-based classifier to identify the categories in Ω.

The classifier in Fig. 3 first examines the higher priority cases, i.e. the simple cases, by using $\overline{C}_{\mathrm{G}}$ to identify either a fully occupied spectrum, denoted as Case 1 (category ω_1 from (2)), or an unoccupied spectrum, denoted as Case 2 (category ω_2). Case 1 requires no action by the radar because it simply waits until an operation opportunities exist. In contrast, Case 2 allows the radar to transmit a Full-band waveform. These high priority cases require a minimal amount of exploration of CR strategies since the solutions are straightforward.

The classifier next attempts to eliminate the obvious notching and avoidance cases. As shown in Fig. 3, a poor candidate for notching arises when the spectral occupancy of RFI is greater than 40% and near a band edge. In addition, if the RFI power is greater than the notch depth, a mutual RFI residue remains. If either of these cases are true, then waveform notching is eliminated as a candidate waveform. Finally, if the

transition rate of the RFI is greater than the adaptation rate of React, then the React approach is eliminated as a possible Adaptation Method.

The advanced cases require further examination of C_G and C_X as described in Fig. 4. The congestion measure of Fig. 4a determines the Waveform Synthesis category and is uniformly partitioned into 5 levels of congestion, where each level corresponds to possible synthesis techniques. The complexity measure of Fig. 4b is likewise uniformly partitioned into 5 levels to identify the Adaptation Method. Note that overlap exists between levels, which indicates that multiple CR strategies are used in Ψ. For example, the advanced case $\overline{C}_G$ = 0.5 and $\overline{C}_X$ = 0.5 results in six CR strategies within Ψ: React-LFM-Avoid, React-RandomFM-Avoid, MDPRL-LFM-Avoid, MDPRL-RandomFM-Avoid, Stochastic-LFM-Avoid, Stochastic-RandomFM-Avoid.

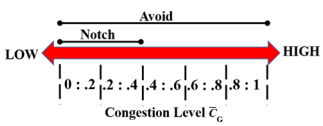

Fig. 4a: Congestion level for Waveform Synthesis.

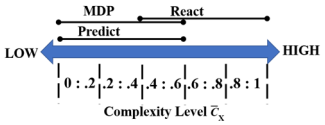

Fig. 4b: Complexity level for Adaptation Method.

Fig. 4: Congestion and complexity measures to determine the CR strategies for advanced cases of the spectrum classifier.

V. Simulations and Results

Consider a proof-of-concept simulation to illustrate the effectiveness of the proposed MCR model for DSA. The simulation compares the Stochastic-LFM-Avoid and React-LFM-Avoid CR strategies (uses only the LFM waveform) in the presence of narrowband, periodic RFI (without the notch, RandomFM, or MDPRL options). The simulation involves a time sequence of events to form a vignette of changing RFI conditions for a stationary radar that implements range-Doppler processing to detect a moving point target. The vignette considers 35 events that occur on a CPI-to-CPI timescale (one event per CPI). The radar CPI is 81.92 ms long with a PRI of 409.6 μs. The target is located at a range of R_T = 1.5 km and is traveling away from the radar with a velocity of V_R = –54 (m/s), which corresponds to a Doppler shift of f_D = –1250 Hz. The SINR is selected as the radar performance metric:

$$Z = P_r \tau \beta / I_N, \qquad (11)$$

where P_r is the radar receive power, τ is the pulse width, and I_N is the interference plus noise power. Peak P_r and I_N are estimated from the range-Doppler information per event. The interference is modeled as a sinusoidal signal that occupies a single frequency bin (bin 25) at 25 MHz within a B = 100 MHz bandwidth. The target detection threshold requires that $Z \geq 0$ dB for the constant false alarm rate (CFAR) detection criteria. Analysis of this bandwidth using (6) produces K = 3 sub-bands corresponding to $\eta_1 = \eta_3 = 0$ (low power sub-band, i.e. the noise floor) and η_2 = 1 (high power sub-band, i.e. interference). The interference is modeled as an alternating renewal process with $\mu_{on}(2) = \mu_{off}(2)$ = 4.1 ms, $\sigma_{off}(k)$ = 0, and $\sigma_{on}(2)$ is varied throughout the experiment. The congestion level is negligible in this analysis due to the narrowband signal.

The vignette in Fig. 5 illustrates the RFI events and the impact on the complexity metric in (10). The scenario starts with the no RFI for 10 CPI cycles and the MCR model identifies the ω_2 category (no RFI), thereby selecting the fixed Adaptation Method for the full band LFM waveform, i.e. $\Psi = \Psi_1$ = Fixed-LFM-Full. Fig. 6 shows that the radar achieves a high SINR (20 dB) during this time using this method.

A periodic sinusoid RFI with $\sigma_{on}(2)$ = 0 is then introduced at CPI 11 and remains active for 10 cycles. The complexity measure remains 0 during this time period, though SINR is significantly decreased since the radar continues to operate in the full band mode. The next evaluation of the RFI scenario at CPI 11 causes the classifier to select the Ψ_1 = Stochastic-LFM-Avoid CR strategy, which greatly improves the SINR.

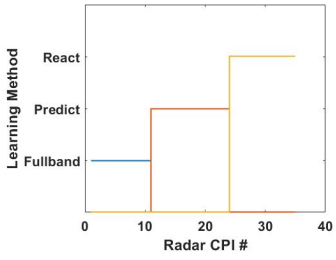

Fig. 5a: RFI Events

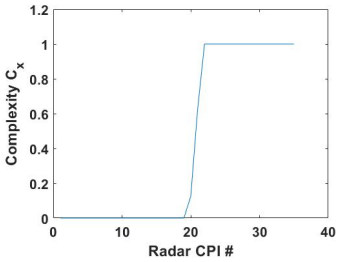

Fig. 5b: Complexity measure

Fig. 5: Illustrations of RFI occurrence and the resulting complexity. The MCR model is updated every CPI cycle and adapts to the varying RFI conditions. The RFI is a narrowband sinusoid that exhibits periodic and random behaviors.

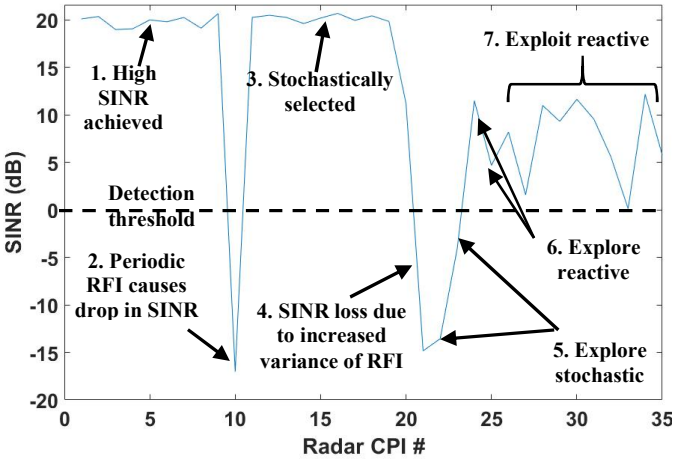

Fig. 6: Radar performance measure for various changes of RFI throughout 35 CPI cycles.

The next event occurs between CPIs 20 to 22 when the variance $\sigma_{on}(2)$ of the RFI increases, thus increasing complexity as $\overline{C}_X$ = 0.125 for CPI 20, $\overline{C}_X$ = 0.625 for CPI 21, and $\overline{C}_X$ = 1 for CPI 22. Further SINR loss is observed during this transition. Once $\overline{C}_X > 0.4$ (Fig. 4b) at CPI 21, the classifier selects a new category that evaluates two CR strategies: Stochastic-LFM-Avoid (Ψ_1) and React-LFM-Avoid (Ψ_2). The exploration approach is then implemented for Q = 2 iterations. Approach Ψ_1 is first explored for CPIs 22 and 23, followed by an exploration of Ψ_2 during CPIs 24 and 25. As observed in Fig.

6, the SINR is improved using Ψ_2. Subsequently, exploitation occurs at CPI 26 and selects Ψ_2 = React-LFM- Avoid using (4).

VI. Conclusions

A metacognitive radar (MCR) model was introduced for DSA in congested environments. This model has the advantage of leveraging multiple CR strategies using MCR Knowledge, MCR Monitoring, and MCR Control. A proof-of-concept simulation was implemented to evaluate the effectiveness of this proposed model, where it was shown that the proposed technique demonstrates the capability to accurately identify and adapt to varying RFI scenarios to improve performance. Future work will focus on evaluation of the proposed model in real-time using the SDRadar of [10,15].

References

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

[2] M.S. Greco, F. Gini, P. Stinco K. Bell, "Cognitive radars: on the road to reality: progress thus far and possibilities for the future," *IEEE Signal Processing Mag.*, vol. 35, no. 4, pp. 112-125, July 2018.

[3] A.F. Martone, "Cognitive radar demystified," *URSI Bulletin,* no. 350, pp. 10-22, Sept. 2014.

[4] C. Tas, E.C. Brown, A. Esen-Danaci, P.H. Lysaker, M. Brüne, "Intrinsic motivation and metacognition as predictors of learning potential in patients with remitted schizophrenia," *J. Psychiatric Research*, vol. 46, no. 8, pp. 1086-1092, May 2012.

[5] Z. Yang, K. Merrick, L. Jin, H.A. Abbass, "Hierarchical deep reinforcement learning for continuous action control," *IEEE Trans. Neural Networks & Learning Systems*, vol. 29, no. 11, pp. 5174-5184, Nov. 2018.

[6] M. Gadhiok, A. Amanna, M.J. Price, J.H. Reed, "Metacognition: enhancing the performance of a cognitive radio," *IEEE Intl. Conf. Cognitive Methods in Situation Awareness & Decision Support*, Miami Beach, FL, Feb. 2011, pp. 198-203.

[7] H. Asadi, A.H. Volos, M.M. Marefat, T. Bose, "Metacognition and the next generation of cognitive radio engines," *IEEE Communications Mag.*, vol. 54, no. 1, pp. 76-82, Jan. 2016.

[8] G.T. Capraro, M.C. Wicks, "Metacognition for waveform diverse radar," *Intl. Waveform Diversity & Design Conf.*, Kauai, HI, Jan. 2012.

[9] S.D. Blunt, E.S. Perrins, *Radar & Communication Spectrum Sharing*, SciTech Publishing, 2018.

[10] B.H. Kirk, R.M. Narayanan, K.A. Gallagher, A.F. Martone, K.D. Sherbondy, "Avoidance of time-varying radio frequency interference with software-defined cognitive radar," *IEEE Trans. Aerospace & Electronic Systems*, vol. 55, no. 3, pp. 1090-1107, June 2019.

[11] A.F. Martone, K.I. Ranney, K. Sherbondy, K.A. Gallagher, S.D. Blunt, "Spectrum allocation for noncooperative radar coexistence," *IEEE Trans. Aerospace & Electronic Systems*, vol. 54, no. 1, pp. 90-105, Feb. 2018.

[12] M. Kozy, J. Yu, R. M. Buehrer, A. Martone, K. Sherbondy, "Applying deep-Q networks to target tracking to improve cognitive radar," *IEEE Radar Conf.*, Boston, MA, Apr. 2019.

[13] J.A. Kovarskiy, R.M. Narayanan, A.F. Martone, K.D. Sherbondy, "A stochastic model for prediction and avoidance of RF interference to cognitive radars," *IEEE Radar Conf.*, Boston, MA, Apr. 2019.

[14] B. Ravenscroft, J.W. Owen, J. Jakabosky, S.D. Blunt, A.F. Martone, K.D. Sherbondy, "Experimental demonstration and analysis of cognitive spectrum sensing and notching for radar," *IET Radar, Sonar & Navigation*, vol. 12, no. 12, pp. 1466-1475, Dec. 2018.

[15] J.W. Owen, et al., "Demonstration of real-time cognitive radar using spectrally-notched random FM waveforms," *IEEE Intl. Radar Conf.*, Washington, DC, Apr. 2020.

[16] Alcala-Medel, A. Egbert, C. Calabrese, A. Dockendorf, C. Baylis, G. Shaffer, A. Semnani, D. Peroulis, A. Martone, E. Viveiros, and R.J. Marks II, "Fast frequency-agile real-time optimization of high-power tuning network for cognitive radar applications," IEEE MTT-S International Microwave Symposium, Boston, MA, June 2019.

[17] E. Selvi, R.M. Buehrer, A.F. Martone, K.D. Sherbondy, "On the use of Markov decision processes in cognitive radar: an application to target tracking," *IEEE Radar Conf.*, Oklahoma City, OK, Apr. 2018.

[18] J.A. Kovarskiy, J.W. Owen, R.M. Narayanan, S.D. Blunt, A.F. Martone, K.D. Sherbondy, "Spectral prediction and notching of RF emitters for cognitive radar coexistence," *IEEE Intl. Radar Conf.*, Washington, DC, Apr., 2020.

[19] A. Slivkins, "Introduction to multi-armed bandits," *Foundations and Trends in Machine Learning*, vol. 12 no. 1-2, Nov. 2019.

[20] M. Sharma, A. Sahoo, "Stochastic model based opportunistic channel access in dynamic spectrum access networks," *IEEE Trans. on Mobile Computing*, vol. 13, no. 7, pp. 1625-1639, July 2014.

[21] K. Arulkumaran, M.P. Deisenroth, M. Brundage A.A. Bharath, "Deep reinforcement learning: a brief survey," *IEEE Signal Processing Mag.*, vol. 34, no. 6, pp. 26-38, Nov. 2017.

[22] A. Semnani, G.S. Shaffer, M.D. Sinanis, and D. Peroulis, "High-Power Impedance Tuner Utilising Substrate-Integrated Evanescent-Mode Cavity Technology and External Linear Actuators," *IET Microwaves, Antennas & Propagation*, Vol. 13, No. 12, 2019, pp. 2067-2072.

[23] A. Goad, A. Egbert, A. Dockendorf, C. Baylis, and A. Martone, "Optimizing Transmitter Amplifier Load Impedance for Tuning Performance in a Metacognition-Guided, Spectrum Sharing Radar," *IEEE Intl. Radar Conf.*, Washington, DC, Apr., 2020.