Strategy Concealment using Multi-Radar Data Fusion in Metacognitive Radars

Jyoti Bhatia*, Kunwar Pritiraj Rajput*, Bhavani Shankar M. R.*, Kumar Vijay Mishra[†], and Muralidhar Rangaswamy[‡]

* Interdisciplinary Center for Security, Reliability and Trust (SnT), University of Luxembourg
[†] United States DEVCOM Army Research Laboratory, Adelphi, Maryland, USA
[‡] United States Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio, USA

Abstract—This paper investigates a metacognitive radar scenario in which both radars and the adversary target possess cognitive capabilities, and target can infer radar's strategies defined by utility functions. In such environments, disruption of adversary cognition is achieved through smart interference design and purposeful slight variations in radar performance to hinder the target's ability to accurately infer radar operational strategies. In this work, we consider a multimetacognitive radar scenario in which multiple radars are trying to track an adversary target. The adversary has the capability to learn the individual radar's utility function using Afriat's theorem-based approach. After estimating the utility function of the radar, the adversary subsequently modify its probes to reduce the utility function of each radar. In response to this, the multiple radars collaborate through a fusion center (FC), which performs weighted utility maximization. The proposed collaborative utility maximization approach hides the individual radar strategies from the adversary, which is unaware of the FC. Simulations demonstrate that the collaborative strategy effectively masks the utility function, preventing the adversary target from accurately estimating it.

I. INTRODUCTION

Cognitive radars have seen significant advancements in recent years, driven by the integration of artificial intelligence, machine learning, and advanced signal processing techniques [1], [2]. The practical implementation aspects of cognitive radar techniques on software-defined radar platforms are thoroughly explored in [3]. An overview of current research on cognitive radars including architectures, mathematical tools and techniques for analysis, experimental setups, and avenues for research has been presented in [2].

On the other hand, the concept of metacognition, first introduced in [4], aims to further enhance the functionality of cognitive radars. This is undertaken by enabling a dynamic selection of the most suitable cognitive strategy in response to the changing environmental conditions and target actions. [5], [2]. While the cognitive strategies follow a framework of sense-learning adaptation, the metacognitive radars continuously adjust these strategies based on real-time feedback to optimize performance [6]. In a recent work [7], the authors investigated the sharing of spectrum for metacognitive radars,

highlighting the dynamic selection of cognitive strategies to optimize radar performance in congested spectral environments.

The cognitive strategy of radar can be seen as an utility optimizer, as it continuously optimizes its performance within predefined constraints [8], [9]. This perspective of seeing radar as a utility maximizer can be mapped to an economic utility maximization problem. It allows the researchers to use Afriat's theorem to radar scenarios. The authors in [8] have studied a scenario in which the cognitive radar does not know about the presence of an adversary. In [10], the author examined the radar and jammer as a principal agent problem, similar to those in economics, and proposed an approach for radar to confuse the adversary using an asymmetric information contract. The enemy target can be intelligent (e.g. an inverse reinforcement learner) and may attempt to understand the cognitive process. This estimation of cognitive abilities can be addressed within the economic theory framework using hypothesis testing based on Afriat's theorem. The enemy can explore radar cognitive strategies using inverse cognition approaches, as suggested in [8], [11]. The radar trying to mask their individual strategy has been suggested in the work using Afriat's method in [12]. This approach discussed in detail the utility maximization function in cognitive radars where Afriat's theorem can be used. They focused on masking individual radar strategy by compromising on utility maximization.

The aforementioned works are focused on single-radar versus adversary scenarios where cognition masking invariably leads to a significant decrease in performance. The literature is in debt regarding the use of multiple radars; this scenario overcomes the shortcomings of a single-radar setting. In a recent work [13], we studied a radar fusion scenario in which an intelligent target is trying to mitigate the performance of multiple radars. Additionally, existing works use the fusion of sensors for the detection and tracking of objects with radar or other sensors to obtain a better estimate of environment and improve the performance. In this work, we propose a fusion-based approach for metacognitive radars to mask strategies by combining radar response and probe data. This approach aims to conceal radar strategies from intelligent adversaries, such as inverse reinforcement learners, by maximizing common utility through collaborative efforts. These radars assess their vulnerability to adversarial inference attacks, dynamically se-

lect the fusion strategy to hide true utilities while maintaining operational effectiveness, and decide when and how to employ fusion-based masking approach on threat level assessments. This strategy enables radars to intelligently manage their own information footprint in adversarial settings, making them metacognitive. To our knowledge, the masking aspect of fusion in metacognitive radars has not yet been explored in the literature. For ease of understanding, we have considered a scenario of multiple radars and a single adversary in this case.

II. PROBLEM STATEMENT AND SYSTEM MODELING

This section states the problem considered and models the underlying system. The system involves N cognitive radars collaboratively tracking a single adversary, as shown in Fig.1. Each radar receives an input from the adversary and responds according to a utility-driven decision-making either individually (Fig.2) or in coordination through a Fusion Center (Fig.1). The fusion framework is detailed in Section II-C, while Section II-D extends the single-radar formulation in [12] to the multi-radar setting.
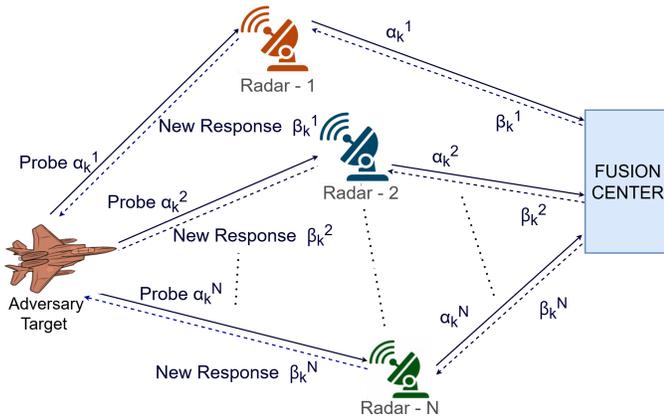

Fig. 1. Multiple radars working in fusion mode to mask their strategies.

A. Problem Statement

The proposed approach aims to protect N individual radar sensors from adversaries by modifying their responses in a network where multiple sensors share a common Fusion Center (FC). The key idea is to strategically mask each radar sensor's response when their data is being fused at the central processing unit. This ensures that the radar system's operations are less vulnerable to adversarial detection and interference during the fusion process.

To achieve this, radars adapt their responses based on a masking technique. This technique takes advantage of the insights gained from Afriat's Theorem, which has been explored to calculate utilities in economic contexts [12]. The individual radar send the response that is an optimized response for the fusion center. A simple fusion of radars does not ensure the masking of all radars. To protect each radar from adversarial target, the existing margin control-based masking approach is applied as a constraint. This ensures that each radar maintains

at least the same level of masking effectiveness as described in [12]. We get a new approach of masking by doing fusion with increasing mechanical design complexity. The metacognitive radars considered in this work conduct a continuous evaluation of their cognitive approach and modify their waveforms based on scenarios.This is a special case where the radars have to modify their waveforms and performance to mask their cognitive ability. The following are some operating conditions required for this work that can be extended later to a more general context.

B. Operating Conditions

1) In naive cognitive mode, the utility function of each of the N radars is monotonic and depends only on β^i (target probe to i^{th} radar), and the resource constraint is a function of both α^i and β^i (i^{th} radar response to the target), and in fusion the resources are shared and so the constraints are added.
2) It is assumed that the adversary is unaware of the existence of a fusion center. It can estimate the track of an individual sensor using IRL but is unaware of their collaborative decision-making capability.
3) In this work, it is assumed that multiple radar sensors are trying to track a single adversary target. It can be extended to multiple targets considering that cognitive radars are resource-constrained and need to allocate power and schedule their sensing actions dynamically in response to the environment [12], [14].

C. Fusion Model

The weighted sum method, a feature-in and feature-out data fusion type according to Dasarathy classification [15], is used in this work. The individual utilities are fused to obtain a common utility whose maximization at the fusion center drives the individual responses of the radars. Let the adversary probe during the kth epoch of the ith radar is denoted by α_k^i, while the corresponding radar response is represented by β_k^i. These responses are obtained by solving the convex optimization problem in (1) for the constraint given in (2), both given below:

$$U(\beta_k^i) = \sum_{i=1}^{N} \lambda^i \ u^i\left(\beta_k^i\right), \ k \in \{1, 2, \dots, K\} \qquad (1)$$

$$\text{Fusion constraint:} \sum_{i=1}^{N} \alpha_k^i \beta_k^i \leq N, \ k \in \{1, 2, \dots, K\} \quad (2)$$

The utility function of each radar $u^i(\beta_k^i)$ is assumed to be concave and monotonic. To analyze this set-up, we use Afriat's theorem, as outlined in Section II-D.

D. IRL based on Afriat's theorem for multiple radars

Afriat's theorem [16] provides conditions to reconstruct a utility function from the observed price and quantity data, ensuring the choices $\{\mathbf{p}_t, \mathbf{x}_t\}$, where $\mathbf{p}_t$ is the price vector and $\mathbf{x}_t$ is the consumption vector, aligned with the maximization of

utility. This holds if there exist utility levels $\{u_t\}$ and positive multipliers $\{\lambda_t\}$ satisfying the Afriat inequalities:

$$u_t - u_s \leq \lambda_s \mathbf{p}_s^T (\mathbf{x}_t - \mathbf{x}_s), \quad \forall s, t \tag{3}$$

where λ_s represents the marginal utility of income, and $\mathbf{p}_s^T(\mathbf{x}_t - \mathbf{x}_s)$ denotes expenditure differences. If satisfied, a nonsatiated, concave utility function can be constructed as:

$$u(x) = \min_t \left\{ u_t + \lambda_t \mathbf{p}_t^T (\mathbf{x} - \mathbf{x}_t) \right\}. \tag{4}$$

The Afriat theorem, central to the revealed preference theory, provides a non-parametric method to test utility maximization from observed choices. We apply the existing framework [12] to multiple cognitive radars, where N radar sensors track an adversary as shown in Fig. 2. Each radar optimizes a utility function $u^i(\alpha_k^i, \cdot)$ under constraints $c^i(\alpha_k^i, \cdot) \leq 0$, where the optimal radar response satisfies the following:

$$\beta_k^{*i} \in \arg\max u^i(\alpha_k^i, \beta^i) \quad \text{subject to} \quad c^i(\alpha_k^i, \beta^i) \leq 0. \tag{5}$$
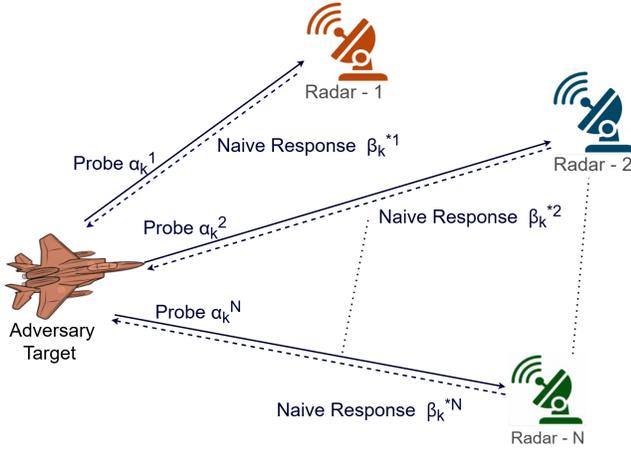

Fig. 2. Un-coordinated setting where different radars send their naïve responses to the adversary.

Following [12], the radar constraint takes the following form, where it is monotone in β^i and linear in the adversary probe α_k^i:

$$c^i(\alpha_k^i, \beta^i) = \alpha_k^i \beta^i - 1, \quad u^i(\alpha_k^i, \beta^i) \equiv u^i(\beta^i). \tag{6}$$

Each radar maximizes its utility for accurate adversary tracking by cognitive waveform selection. The adversary is aware of this waveform and the constraint function $c^i(\alpha_k^i, \cdot)$ defined in (6) over K epochs. The data available with adversary to estimate the radars strategies, $\mathcal{D}_c^i$, using IRL takes the form:

$$\mathcal{D}_c^i = \{c^i(\alpha_k^i, \cdot), \beta_k^i\}_{k=1}^K \tag{7}$$

The adversary will estimate the feasibility region of the cognitive strategy of each radar using (8) defined by the nonnegative vector $\boldsymbol{\theta}$ for the set of inequalities [12]:

$$\mathcal{A}(\boldsymbol{\theta}, \mathcal{D}_c) \leq 0 \equiv \theta_s - \theta_k - \theta_{k+K}(c_k(\beta_s) - c_k(\beta_k)) \leq 0,$$
$$s, k \in \{1, 2, \ldots, K\}, s \neq k, \quad \boldsymbol{\theta} \in \mathbb{R}^{2K}. \tag{8}$$

Based on Theorem 1 in [12], the adversary estimates radar utilities from observations and selects α_k^i to degrade their performance. When radars operate independently, they return optimal yet naive responses, revealing their strategies to the adversary, as illustrated in Fig.2. The knowledge of the radar strategy to the adversary is measured with margin, a concept introduced in [12]. The margin μ^i is used to analyze the masking capability of radar sensors and can be calculated by (9) below:

$$\mu^i(\mathcal{D}_c^i) = \min_{\epsilon \geq 0} \epsilon \quad \text{s.t.} \quad \left(\mathcal{A}^i(u^i, \mathcal{D}_c^i) + \epsilon \mathbf{1}\right) \geq \mathbf{0}. \tag{9}$$

Here $\mathcal{A}^i$ represents the set of inequalities defined in (8) for the ith radar. The next section discusses the proposed algorithm in detail.

III. PROPOSED ALGORITHM

Previous research has primarily focused on concealing the strategies of individual radars from adversaries by estimating the adversary's understanding of the radar's tactics and trying to mitigate it by sending suboptimal responses by degrading their performance. [17]. In contrast, the proposed framework emphasizes optimizing a global utility function centered around an FC, rather than solely focusing on the utility functions of individual radars to mitigate the adversary's effect. This framework addresses a metacognitive scenario that involves two distinct cycles of cognition in radars.

Scenario 1: When radars are acting naive to the presence of the adversary as shown in Fig.2, and try to maximize their individual utility function [12].

Scenario 2: When an adversary attempts to degrade radar performance through perturbations, radars alter their responses to obscure their strategies. Instead of solely compromising their performance, the radars issue responses based on decisions made at the fusion center by optimizing the global utility function as shown in Fig.1.

Our primary focus is on the second scenario, as the first scenario has been extensively analyzed in previous works. The existing literature predominantly addresses individual optimization problems. In the second scenario, we assume that the adversary is aware of the resource constraints of each radar system [17]–[20]. Once radars recognize adversarial estimation attempts, they adopt FC-guided weighted utility maximization. The adversary, unaware of this fusion strategy, continues probing based on incorrect individual utility estimates. Although joint utility maximization may be suboptimal individually, it inherently masks radar strategies, and a constraint in fusion ensures masking for all sensors, enhancing robustness against inference attacks.

A notable aspect of this approach is that, when faced with a adversary trying to affect the radar performance, rather than compromising its performance for masking, the radar provides sub-optimal responses based on optimized decisions from the fusion center. This strategy allows individual radars to mask their strategies without sacrificing their operational effectiveness with some design modifications.

A. Fusion approach for masking utility of individual sensor

The fusion-based response will make the radar work in a new feasibility region obtained by (10):

$$\mathcal{A}(\widetilde{\boldsymbol{\theta}}, \mathcal{E}_c) \leq 0 \equiv \widetilde{\theta}_s - \widetilde{\theta}_k - \widetilde{\theta}_{k+K}(c_k(\widetilde{\beta}_s) - c_k(\widetilde{\beta}_k)) \leq 0,$$
$$s, k \in \{1, 2, \ldots, K\}, s \neq k, \quad \widetilde{\boldsymbol{\theta}} \in \mathbb{R}^{2K}, \quad (10)$$

where the new data set $\mathcal{E}_c$ for each sensor collected by adversary from its response is used to find these feasibility regions.

Algorithm 1 Ensure Masking of Each Radar using a Constrained Fusion

---

1: **Compute Naive Response Sequence**: Calculate the radar's naive response sequence $\beta_{1:K}^{*i}$ by solving the convex optimization problem:

$$\beta_k^{*i} = \arg\max u^i(\beta^i), \quad c^i(\alpha_k^i, \beta^i) \leq 0, \quad \beta^i \geq 0,$$
$$k \in \{1, 2, \ldots, K\}, \quad i \in \{1, 2, \ldots, N\} \quad (11)$$

where u^i are concave monotone functions in β^i and $c^i(\alpha_k, \beta^i)$ are convex monotone in β^i.

2: **Weights Allocation in the Fusion**: Find the weights by weighted sum of utilities for naive responses of N sensors:

$$U(\beta_k^{*i}) = \sum_{i=1}^{N} \lambda^i \, u^{*i}\left(\beta_k^{*i}\right), \forall k \in \{1, 2, \ldots, K\} \quad (12)$$

$$\text{where} \sum_{i=1}^{N} \lambda^i = 1 \quad (13)$$

3: **Generate the New Responses from the Fusion Center**: Determine the fused responses for each sensor by solving a convex optimization problem. The constrained margin is used to ensure the masking of each radar by controlling the margin in a new region formed by the adversary target for the updated feasibility region.

$$U(\widetilde{\beta}_k^i) = \sum_{i=1}^{N} \lambda^i \, u^i\left(\widetilde{\beta}_k^i\right), k \in \{1, 2, \ldots, K\} \quad (14)$$

$$\text{Fusion Constraint} = \sum_{i=1}^{N} \alpha_k^i \cdot \widetilde{\beta}_k^i \leq N, k \in \{1, 2, \ldots, K\} \quad (15)$$

$$\text{Margin Constraints} = \mathcal{A}^i(\widetilde{u}^i, \mathcal{D}_c^i) + m^i \geq 0, m^i \leq M_{\text{thresh}}^i \quad (16)$$

Here, $\widetilde{u}^i$ represents the feasibility set for new constraints. The M_{thresh}^i is obtained using Eq. 18.

4: **Compare the New Margins with the Naive Margins**: To compare the new margins, use the new feasibility region calculated using Eq. (10) and the naive margins using Eq.(9).

To conceal the utility functions of the radars from an adversary without the latter's knowledge, it is essential that the margin of each radar is constrained to be less than or

Parameters	Numerical Values
N	3,5
α_k^1	Unif(0.2, 2.5)
α_k^2	Unif(0.3, 2.6)
α_k^3	Unif(0.4, 2.7)
α_k^4	Unif(0.5, 2.8)
α_k^5	Unif(0.6, 3.0)
K	50
$u^i(\beta_k^i)$	$\sqrt{\beta_k^i}$
η value	unif(0, 0.1...0.7)

equal to the threshold margin (maximum) determined using Afriat's theorem, i.e., the following constraints are satisfied:

$$\mathcal{A}^i(\widetilde{u}^i, \mathcal{E}_c^i) + \mathbf{m}^i \geq 0, \quad \mathbf{m}^i \leq M_{\text{thresh}}^i \mathbf{1}. \quad (17)$$

$$M_{\text{thresh}}^i = (1 - \eta) M_u^i\left(\{\alpha_k, \beta_k^*\}_{k=1}^K\right);$$
$$i \in \{1, 2, \ldots, N\}, \quad \eta \in [0, 1]. \quad (18)$$

Here, M_u^i defines the naive margins calculated using (9) when a radar responds naively to an adversary's probe. M_{thresh}^i represents the threshold margins calculated from the naive margins by a masking control parameter η. As the value of η increases from 0 to 1 the threshold margin value decreases. Masking is carried out by keeping the operating point closer to the boundary of the feasible region, preventing the exposure of true radar strategies. When the margin is smaller than M_{thresh}^i, the masking capability of the radar is higher. Margin conditions in (18) ensure that the radars will work closer to the boundary of the new feasibility region, and this margin of fused sensors is controlled through η. Here, η is the same for each sensor and sets the threshold margin for each sensor as shown in (18).

The controlled fusion results in a controlled masking of all sensors. One can ensure the masking of each sensor with certain compromise in the Fusion utility. However, despite loss in the Fusion utility, there is a gain with respect to performance compared to individual utilities. Algorithm 1 describes the details for performing this fusion.

IV. NUMERICAL EXAMPLE

The experimental values for this application are given in Table I. The fusion of multiple radars leads to a gain in overall utility when compared to individual utility. The gain in utility achieved with fusion will reduce slightly as the margin value is restricted by constraining the operating point closer to the feasibility region boundary, and so with masking (η) the gain in utility will reduce. The parameter values are shown in table I. The results are shown in Fig. 3 and Fig. 4 for N = 3 and 5, respectively.

V. CONCLUSION

This work demonstrates that multiple cognitive radars can effectively mask their individual strategies from an adversary
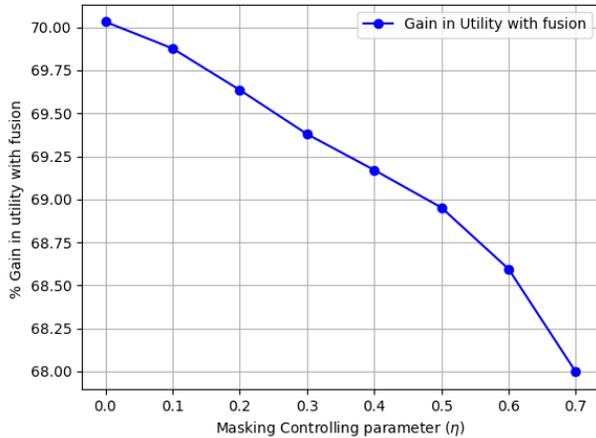
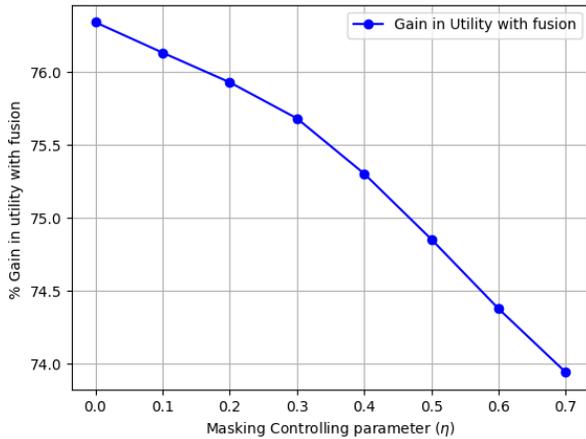Fig. 3. For 3 sensors gain in utility and its variation with masking parameter η.

Fig. 4. For 5 sensors gain in utility and its variation with masking parameter η.

by fusing their responses through a coordinated decision-making framework. While simple fusion offers limited protection, incorporating a margin constraint ensures utility masking across all sensors without significant degradation in overall fusion performance. The proposed weighted-sum fusion approach shows substantial gains in both utility and adversarial resilience. Future work will explore more sophisticated fusion strategies, including Dempster-Shafer theory and entropy-based methods, to further enhance masking and adaptability.

REFERENCES

[1] S. Z. Gurbuz, H. D. Griffiths, A. Charlish, M. Rangaswamy, M. S. Greco, and K. Bell, "An overview of cognitive radar: Past, present, and future," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 12, pp. 6–18, Dec. 2019.

[2] K. V. Mishra, B. Shankar M.R., and M. Rangaswamy, Eds., *Next-Generation Cognitive Radar Systems*, SciTech Publishing, 2024.

[3] A. F. Martone, K. D. Sherbondy, J. A. Kovarskiy, B. H. Kirk, J. W. Owen, B. Ravenscroft, A. Egbert, A. Goad, A. Dockendorf, C. E. Thornton, R. M. Buehrer, R. M. Narayanan, S. Blunt, and C. Baylis, "Practical aspects of cognitive radar," in *2020 IEEE Radar Conference (RadarConf)*, Florence, Italy, 2020, pp. 1–6, IEEE.

[4] Gerard T. Capraro and Michael C. Wicks, "Knowledge Based Systems and Metacognition in Radar," in *Knowledge Science, Engineering and Management*, vol. 6291, pp. 573–578. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, Series Title: Lecture Notes in Computer Science.

[5] Anthony F. Martone, Kelly D. Sherbondy, Jacob A. Kovarskiy, Benjamin H. Kirk, Charles E. Thornton, Jonathan W. Owen, Brandon Ravenscroft, Austin Egbert, Adam Goad, Angelique Dockendorf, R. Michael Buehrer, Ram M. Narayanan, Shannon D. Blunt, and Charles Baylis, "Metacognition for radar coexistence," in *2020 IEEE International Radar Conference (RADAR)*, 2020, pp. 55–60.

[6] K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "Toward metacognitive radars: Concept and applications," in *2020 IEEE International Radar Conference (RADAR)*, Washington, DC, USA, 2020, pp. 77–82, IEEE.

[7] Anthony F. Martone, R. Michael Buehrer, and David M. McNamara, "Emerging trends in radar: Metacognitive radar networks for the next generation of intelligent sensing," *IEEE Aerospace and Electronic Systems Magazine*, pp. 1–5, 2025.

[8] V. Krishnamurthy, D. Angley, R. Evans, and B. Moran, "Identifying cognitive radars—inverse reinforcement learning using revealed preferences," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4529–4542, 2020.

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

[10] Anurag Gupta and Vikram Krishnamurthy, "Principal-agent problem as a principled approach to electronic counter-countermeasures in radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 4, pp. 3223–3235, 2022.

[11] Vikram Krishnamurthy, Kunal Pattanayak, Sandeep Gogineni, Bosung Kang, and Muralidhar Rangaswamy, "Adversarial radar inference: Inverse tracking, identifying cognition, and designing smart interference," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2067–2081, 2021.

[12] K. Pattanayak, V. Krishnamurthy, and C. M. Berry, "Metacognitive radar: Masking cognition from an inverse reinforcement learner," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 6, pp. 8826–8844, 2023.

[13] Bhavani Shankar Mysore Rama Rao, Jyoti Bhatia, Kunwar Pritiraj Rajput, and Bjorn Ottersten, "Intelligent target maneuverability in presence of tracking with multiple radars," in *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025, pp. 1–5.

[14] Phani Chavali and Arye Nehorai, "Scheduling and power allocation in a cognitive radar network for multiple-target tracking," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5667–5679, 2013.

[15] Federico Castanedo, "A review of data fusion techniques," *The Scientific World Journal*, vol. 2013, pp. Article ID 704504, 19 pages, 2013.

[16] Sydney N. Afriat, "The construction of utility functions from expenditure data," *International Economic Review*, vol. 8, no. 1, pp. 67–77, 1967.

[17] Kunal Pattanayak, Vikram Krishnamurthy, and Christopher M. Berry, "Metacognitive radar: Masking cognition from an inverse reinforcement learner," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 6, pp. 8826–8844, 2023.

[18] Kunal Pattanayak, Vikram Krishnamurthy, and Christopher Berry, "Meta-cognition. an inverse-inverse reinforcement learning approach for cognitive radars," in *2022 25th International Conference on Information Fusion (FUSION)*, 2022, pp. 01–08.

[19] Kunal Pattanayak, Vikram Krishnamurthy, and Christopher Berry, "How can a cognitive radar mask its cognition?," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5897–5901.

[20] Kunal Pattanayak, Vikram Krishnamurthy, and Christopher Berry, "Inverse-inverse reinforcement learning. how to hide strategy from an adversarial inverse reinforcement learner," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 3631–3636.