HOW CAN A COGNITIVE RADAR MASK ITS COGNITION?

Kunal Pattanayak [*], *Vikram Krishnamurthy*[*] *and Christopher Berry*[†]

[*] Electrical and Computer Engineering, Cornell University, USA
[†] Lockheed Martin Advanced Technology Laboratories, USA.

ABSTRACT

We study how a cognitive radar can mask (hide) its cognitive ability from an adversarial jamming device. Specifically, if the radar optimally adapts its waveform based on adversarial target maneuvers (probes), how should the radar choose its waveform parameters (response) so that its utility function cannot be recovered by the adversary? This paper abstracts the radar's cognition masking problem in terms of the spectra (eigenvalues) of the state and observation noise covariance matrices, and embeds the algebraic Riccati equation into an economics-based utility maximization setup. Given an observed sequence of radar responses, the adversary tests for utility maximization behavior of the radar and estimates its utility function that rationalizes the radar's responses. In turn, the radar deliberately chooses sub-optimal responses so that its utility function almost fails the utility maximization test, and hence, its cognitive ability is masked from the adversary. We illustrate the performance of our cognition masking scheme via simple numerical examples. Our approach in this paper is based on revealed preference theory in microeconomics for identifying rationality.

Index Terms— Cognitive Radar, Revealed Preference, Adversarial Inverse Reinforcement Learning, Electronic Counter Countermeasures, Kalman Filter

1. INTRODUCTION

In abstract terms, a cognitive radar is a utility maximizer - it adapts its waveform, scheduling and beam by optimizing utility functions. Consider the scenario where an adversarial target probes a cognitive radar (to possibly degrade the radar's performance) and analyzes the radar's responses to estimate the radar's utility function. *How can the radar covertly mask its utility function by deliberately choosing responses that confuse the adversary?* In this paper, we propose a revealed preference-based approach to mask the radar's cognition with the working assumption that the cognitive radar satisfies economics-based rationality.

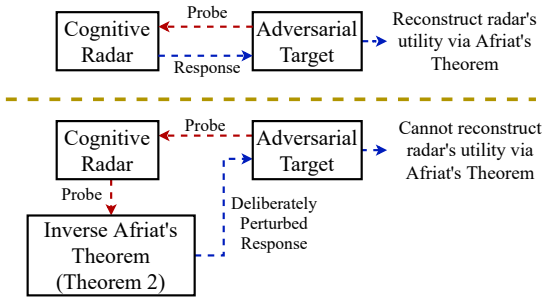

Fig. 1. Schematic of the Masking Cognition problem. The adversarial target sends a sequence of probe signals to the radar and records its responses. If the cognitive radar responds naively to the adversarial target's probes, its utility function can be recovered via Afriat's theorem (Top). If the radar deliberately perturbs its response using the *inverse* Afriat's theorem, the adversary fails to reconstruct its utility (Bottom).

Before going into the details, we emphasize that the problem formulation and algorithms developed here also apply to *adversarial inverse reinforcement learning*. In inverse reinforcement learning [1, 2, 3, 4], an inverse learner seeks to estimate the utility function of a decision maker by observing its decisions. A natural extension is: How can the decision maker hide its utility function by slightly perturbing the actions it takes in the presence of an adversary?

Related Work. This paper builds on our previous work [5] where the cognitive radar is not aware that it is being probed by the adversarial target. If the radar is aware of the adversary's motives, how to deliberately respond with sub-optimal responses to confuse the adversarial target (see Fig. 1)? In a companion paper [6] we formulate the electronic counter-countermeasure (ECCM) problem as a principal agent problem where the radar and adversary establish an information asymmetric contract. In comparison, the formulation in the current paper is adversarial where the radar seeks to confuse the adversary.

This paper can be viewed in the context of low-probability of intercept (LPI) radar as a countermeasure to electronic in-

[*]V. Krishnamurthy and K. Pattanayak are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 14853 USA. e-mail: vikramk@cornell.edu, kp487@cornell.edu. [†] C. Berry is with Lockheed Martin Advanced Technology Laboratories, Cherry Hill, NJ, 08002 USA. e-mail: christopher.m.berry@lmco.com. This research was supported in part by a research contract from Lockheed Martin and the Army Research Office grant W911NF-21-1-0093.

telligence (ELINT) gathering targets [7]. Masking cognition in the face of an adversarial target is closely related to the areas of ECCM and RF stealth [8] in electronic warfare. [9] provides a comprehensive list of ECCM techniques. [10, 11] propose waveform adaptation schemes to counter barrage jamming. [12] proposes time-frequency based ECCM solutions for deceptive jamming. [13, 14, 15] exploit frequency diversity for radio stealth in multi-target and moving target tracking. However, cognition masking strategies with minimal performance loss have not been explored previously.

Background. Revealed Preference and Afriat's Theorem

Our approach to masking cognition in radars is based on revealed preference in micro-economics. The area of revealed preference [16, 17, 18, 19] focuses on nonparametric detection of utility maximization behavior given a finite dataset of probe and response signals.

Definition 1. *A system is a utility maximizer if for a probe signal $\alpha \in \mathbb{R}^m_+$, the response signal $\beta \in \mathbb{R}^m_+$ satisfies:*

$$\beta = \text{argmax}_{\bar{\beta} \in \mathbb{R}^m_+} u(\bar{\beta}),\ \alpha' \bar{\beta} \leq 1, \tag{1}$$

where u is a monotone utility function.

In micro-economics, probe α is the vector of prices of a set of goods, and response β is the consumption vector. Hence, the constraint $\alpha' \bar{\beta} \leq 1$ is a budget constraint with total budget \$1. Indeed, this constraint can be replaced without loss of generality by $\alpha' \bar{\beta} \leq c$, where $c > 0$ is the actual budget. Given a finite time series of probes and responses from a system, how to test if the system is a utility maximizer (1)?

The key result in revealed preference is Afriat's theorem [17, 18]. A remarkable property of Afriat's theorem is that it gives testable conditions that are both necessary and sufficient conditions for a time series of probes and responses to be consistent with utility maximization behavior (1).

Theorem 1 (Afriat's Theorem [17])**.** *Given a sequence of probes and responses $\mathcal{D} = \{(\alpha_k, \beta_k), k \in \{1, 2, \ldots, K\}\}$, the following statements are equivalent:*

1. *There exists a monotone, continuous and concave utility function that satisfies* (1)*.*
2. Afriat's Test: *There exist reals $u_t, \lambda_t > 0,\ t = 1, 2, \ldots, K$, such that the following inequalities are feasible.*

$$u_s - u_t - \lambda_t \alpha_t' (\beta_s - \beta_t) \leq 0\ \forall t, s \in \{1, \ldots, K\}. \tag{2}$$

The monotone, concave utility function given by

$$u(\beta) = \min_{t \in \{1,2,\ldots,K\}} \{u_t + \lambda_t \alpha_t' (\beta - \beta_t)\} \tag{3}$$

constructed using u_t and λ_t (2) *rationalizes $\mathcal{D}$* (1)*.*

3. *The data set $\mathcal{D}$ satisfies the Generalized Axiom of Revealed Preference (GARP), namely for any $t \leq K$, $\alpha_t' \beta_t \geq \alpha_t' \beta_{t+1} \quad \forall t \leq k-1 \implies \alpha_k \beta_k \leq \alpha_k' \beta_1$.*

Afriat's theorem tests for economics-based rationality. In the radar context, the adversarial target uses Afriat's theorem to test for the radar's cognition. If Afriat's inequalities (2) have a feasible solution, then the adversary constructs a set of feasible utility functions (3) that rationalize the radar's responses. The estimated utility is set-valued since the reconstructed utility is ordinal - any positive monotone transformation of a feasible utility function rationalizes the radar's responses.

Outline: Sec. 2 below reconciles the abstract utility maximization setup of Definition 1 with the radar's cognitive behavior, specifically, waveform adaptation during target tracking. Sec. 3 proposes a cognition masking strategy for the radar when the radar knows an adversarial target is reconstructing its utility function. Finally, Sec. 4 illustrates the performance of the cognition masking scheme via two numerical examples.

2. OPTIMAL WAVEFORM ADAPTATION AS UTILITY MAXIMIZATION

Waveform adaptation is a crucial functionality of a cognitive radar. In this section, we abstract optimal waveform adaptation of a cognitive radar using a Kalman filter for target tracking into the utility maximization setup of Definition 1. The key idea is to express the linear budget constraint of Definition 1 in terms of the eigenvalues (spectra) of the state and observation noise covariances of the radar's state space model.

Linear Gaussian dynamics for a target's kinematics [20] and linear Gaussian measurements at the radar are widely assumed as a useful approximation [21]. Accordingly, consider the following state space model for the radar:

$$\begin{aligned} x_{n+1} &= Ax_n + w_n(\alpha_k), \quad x_0 \sim \pi_0 \\ y_n &= Cx_n + v_n(\beta_k), \end{aligned} \tag{4}$$

where $x_n \in \mathcal{X} = \mathbb{R}^X$ is the target state with initial density $\pi_0 \sim \mathcal{N}(\hat{x}_0, \Sigma_0)$, $y_n \in \mathcal{Y} = \mathbb{R}^Y$ is the radar's observation, $w_n \sim \mathcal{N}(0, Q(\alpha_k))$ and $v_n \sim \mathcal{N}(0, R(\beta_k))$ are mutually independent, Gaussian noise processes.

The state noise covariance Q is parameterized by the adversarial target's probe α_k and the observation noise covariance R is parameterized by the radar's response β_k (see [5, Sec. III-B] for a detailed discussion on the relation between radar's waveform and observation noise covariance R). It is important to distinguish between the subscripts n, k in (4). The subscript n indicates system updates at the tracker level (faster timescale), and the subscript k indicates the epoch (slower timescale) for the probe and response. When state x_n represents the position and velocity in Euclidean space, A is a block diagonal constant velocity matrix [22]. The state noise

covariance $Q(\alpha)$ in (4) models acceleration maneuvers of the target parameterized by the probes α.

The radar estimates the target state $\hat{x}_n$ with covariance Σ_n from observations $y_{1:n}$. The posterior π_n is propagated recursively in time via the classical Kalman filter equations:

$$\Sigma_{n+1|n} = A\Sigma_n A' + Q(\alpha_k),\ K_{n+1} = C\Sigma_{n+1|n}C' + R(\beta_k)$$
$$\psi_{n+1} = \Sigma_{n+1|n}C'K_{n+1}^{-1},\ \hat{x}_{n+1} = A\hat{x}_n + \psi_{n+1}(y_{n+1} - CA\hat{x}_n)$$
$$\Sigma_{n+1} = (I - \psi_{n+1}C)\Sigma_{n+1|n}.$$

Assuming the model parameters (4) satisfy the conditions that $[A, C]$ is detectable and $[A, \sqrt{Q}]$ is stabilizable, the steady-state predicted covariance Σ_∞ is the unique positive semi-definite solution of the *algebraic Riccati equation* (ARE):

$$\mathcal{A}(\alpha_k, \beta_k, \Sigma) = -\Sigma + A(\Sigma - \Sigma C'[C\Sigma C' + R(\beta)]^{-1}C\Sigma)A' + Q(\alpha) = 0. \quad (5)$$

Denote $\Sigma^*(\alpha_k, \beta_k)$ as the solution of the ARE given probe α_k and response β_k at time k.

Our working assumption is that the radar maximizes a utility function u to choose its optimal waveform at the start of every epoch k. Hence, it only remains to justify the linear budget $\alpha_k'\beta_k \leq 1$ in Definition 1 to embed waveform optimization into the utility maximization setup. We suppose:

- the target probe α is the vector of eigenvalues of the positive definite matrix Q
- the radar response β is the vector of eigenvalues of the positive definite matrix R^{-1}.

The i^{th} component of β_k is the measurement precision (amount of energy) of the radar in the i^{th} mode. Similarly, the i^{th} component of α_k is the incentive for considering the i^{th} mode of the target. Put together, $\alpha_k'\beta_k$ measures the signal-to-noise ratio (SNR) of the radar. Thus, $\alpha_k'\beta_k \leq 1$ is effectively a bound on the radar's SNR. Hence, in the utility maximization context, the radar chooses the most precise observation noise covariance $R(\beta_k)$ such that its SNR lies below a particular threshold [1].

To summarize, we have justified how the cognitive radar's waveform adaptation can be cast as the constrained utility maximization problem of Definition 1.Hence, the adversarial target can now use Afriat's Theorem 1 to reconstruct the radar's utility. How should the radar react so that its utility function is not be recovered accurately? The rest of the paper focuses on a cognition masking strategy for the radar. The key idea is for the radar to deliberately choose sub-optimal waveforms so that the radar's utility u satisfies the Afriat's inequalities (2) by a small margin, thus confusing the adversarial target at the cost of performance degradation.

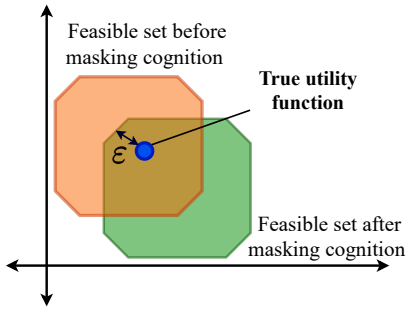

Fig. 2. Masking Cognition by Performance Degradation. If the radar responds naively to the adversary target's probes, its utility passes the utility maximization test by a large margin and hence, is close to the center of the feasible set (orange region) computed by the adversary. By deliberately perturbing its response and degrading its performance, the radar shifts the feasible set (green region) so that the true utility is within ϵ distance from the edge of the set.

3. INVERSE REVEALED PREFERENCE FOR MASKING UTILITY FUNCTION

We now present the main result of this paper, namely, inverse Afriat's Theorem. If the adversary uses Afriat's theorem to reconstruct the radar's utility function, the radar uses the inverse Afriat Theorem below to deliberately perturb its responses and mask its utility function. Put differently, the radar deliberately compromises on its performance to prevent the reconstruction of its utility function.

In terms of the radar's choice of waveform parameters, the radar chooses sub-optimal sensing modes (observation noise covariance) given the adversarial target's maneuvers (state noise covariance) so that the radar's utility function is masked from the adversary.

Theorem 2 (Inverse Afriat's Theorem to Mask Cognition)**.** *Suppose the radar optimizes a monotone, continuous utility function u, and the adversary uses Theorem 1 to estimate the radar's utility function. Given the adversary's probe sequence $\{\alpha_k\}_{k=1}^K$, the radar's response sequence $\{\beta_k\}_{k=1}^K$ that masks its utility function is given by:*

$$\beta_{1:K} = \text{argmin}_{\hat{\beta}_{1:K}} \sum_{k=1}^{K} \|\hat{\beta}_k - \beta_k^*\|_2^2, \quad (6)$$

$$u(\hat{\beta}_s) \geq u(\hat{\beta}_t) - \nabla_\beta u(\hat{\beta}_t^*)'(\hat{\beta}_s - \hat{\beta}_t) + \epsilon,\ \forall\, s, t, \quad (7)$$

$$\hat{\beta}_t \geq 0,\ \alpha_k'(\beta_k^* - \hat{\beta}_k) = 0,\ \forall\, t. \quad (8)$$

In (6), $\beta_k^* = \text{argmax}_{\{\beta \in \mathbb{R}_+^m : \alpha_k'\beta \leq 1\}} u(\beta)$ is the naive response that maximizes utility u given probe signal α_k. The variable ϵ parametrizes the margin with which the radar's response passes Afriat's test for utility maximization.

[1] see [5] for a more detailed discussion on the linear budget in terms of the solution to the ARE (5).

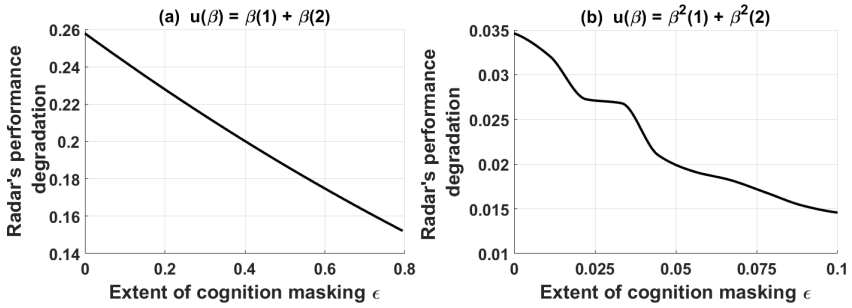

Fig. 3. Deliberate performance loss (vertical-axis) of the cognitive radar (6) as a function of ϵ (horizontal-axis) which measures the extent of cognition masking on the adversary's side (6). (i) $\epsilon = 0$ corresponds to maximum cognition masking and hence results in maximum performance loss. (ii) Due to larger local variation, for a fixed value of ϵ, the quadratic utility (sub-figure (b)) requires smaller perturbation (≈ 10 times) from the optimal response compared to the linear utility (sub-figure (a)).

Theorem 2 masks the cognitive radar's utility function by deliberately perturbing its responses so that the responses almost fail the adversary's test for utility maximization (Theorem 1). If the radar naively responds to the adversary's probes ($\eta_k^* = 0,\ \forall k$), the radar's utility function passes the Afriat's test (2) by a large margin and is thus a high-confidence utility estimate for the adversary[2]. Since the radar's utility passes the Afriat's test by a small margin due to the masking scheme of Theorem 2 (7), it now lies close to the edge of the feasible set of utilities[3], and is no more a high-confidence utility estimate for the adversary. The constraint $\alpha_k' \eta_k = 0$ ensures the radar does not violate its resource constraint $\alpha_k' \beta_k \leq 1$.

Extent of cognition masking ϵ. A smaller value of ϵ implies better cognition masking and higher performance degradation of the radar. Setting ϵ to 0 in (7) completely masks the radar's utility function (u lies on the edge of the feasible set), but requires maximum degradation of radar performance (large perturbation (6) from the optimal response β_k^*). On the other extreme, a large value of ϵ results in zero performance loss of the radar, but exposes the radar's utility function to the adversary since it lies very close to the center of the feasible set. Finally, Theorem 2 provides a scheme for the radar to mask its utility function. Masking cognition is simply a special case of Theorem 2 where λ is set to 0.

4. NUMERICAL EXAMPLES

Theorem 2 specified the procedure for a cognitive radar to effectively mask its cognition from an adversarial target. Below, we illustrate via simple numerical examples the masking performance of Theorem 2 for two different utility functions.

We chose $K = 50$ and $m = 2$, the dimension of adversarial target's probe and radar's response. The elements of the adversarial target's probe signals are generated randomly and independently over time as $\alpha_k(i) \sim \text{Unif}(0.2, 2.5)$ for all $i = 1, 2$ and time $k = 1, 2, \ldots, K$,where Unif(a, b) denotes uniform pdf with support (a, b). Recall that the probe signal α_k is the diagonal of the state noise covariance matrix: $Q_k = \text{diag}[\alpha_k(1), \alpha_k(2)]$.

Given the probe sequence $\{\alpha_k, k = 1, 2, \ldots, K\}$, the cognitive radar chooses its response sequence $\{\beta_k, k = 1, 2, \ldots, K\}$ via (6) in Theorem 2. Recall from Sec. 2 that response β_k is the diagonal of the inverse of radar's observation noise covariance matrix: $R_k^{-1} = \text{diag}[\beta_k(1), \beta_k(2)]$. We generate two separate sequences of responses for the same probe sequence, but for two different utility functions:

$$(a)\ u(\beta) = \beta(1) + \beta(2), \qquad (b)\ u(\beta) = \beta^2(1) + \beta^2(2)$$

Figure 3 shows the loss in performance (minimum perturbation from optimal response (6)) of the cognitive radar as a function of ϵ (extent of cognition masking), for both choices of utility functions. From Fig. 3, we see that for both utility functions, the radar's performance decreases with increasing ϵ (larger extent of utility masking). This is expected since a larger ϵ implies a larger shift of the feasible set of utilities constructed by the adversarial target.

5. CONCLUSION AND EXTENSIONS

This paper focuses on masking a radar's cognition when probed by an adversarial target. Our main result is Theorem 2 that describes the radar's cognition masking strategy. The radar deliberately chooses sub-optimal responses at the cost of its performance, but prevents its utility function from being recovered by the adversary.

Finally, a useful extension of this paper would be to study more general game-theoretic settings where even the adversary knows the radar is trying to mask its cognition. *How to detect play from the Nash equilibrium of a game between the radar and adversary?*

[2] In inverse RL learning [1, 23], a popular choice of a point utility estimate is the point that satisfies optimality conditions with the largest margin.

[3] It follows from simple observation that utilities that pass Afriat's test with zero margin form the edge of the set of utilities for which Afriat's inequalities are feasible. Hence, the margin by which a utility function passes Afriat's test is proportional to its distance from the edge of the feasible set.

6. REFERENCES

[1] A. Ng and S. Russell. Algorithms for inverse reinforcement learning. In *Proc. 17th International Conf. Machine Learning*, pages 663–670, 2000.

[2] P. Abbeel and A. Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1, 2004.

[3] B. D. Ziebart, A. L. Maas, J. A. Bagnell, A. K. Dey, et al. Maximum entropy inverse reinforcement learning. In *Aaai*, volume 8, pages 1433–1438. Chicago, IL, USA, 2008.

[4] V. Krishnamurthy and G. Yin. Langevin dynamics for inverse reinforcement learning of stochastic gradient algorithms. *arXiv preprint arXiv:2006.11674*, 2020.

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

[6] A. Gupta and V. Krishnamurthy. Principal agent problem as a principled approach to electronic counter-countermeasures in radar. *arXiv preprint arXiv:2109.03546*, 2021.

[7] R. G. Wiley. *ELINT: The interception and analysis of radar signals*. Artech House, 2006.

[8] D. Lynch. Introduction to rf stealth. *Raleigh, NC: Scitech Publishing Inc, 2004. 573*, 2004.

[9] L. Neng-Jing and Z. Yi-Ting. A survey of radar ecm and eccm. *IEEE Transactions on Aerospace and Electronic Systems*, 31(3):1110–1120, 1995.

[10] C. Shi, F. Wang, M. Sellathurai, and J. Zhou. Low probability of intercept-based distributed mimo radar waveform design against barrage jamming in signal-dependent clutter and coloured noise. *IET Signal Processing*, 13(4):415–423, 2019.

[11] F. A. Butt, I. H. Naqvi, and U. Riaz. Hybrid phased-mimo radar: A novel approach with optimal performance under electronic countermeasures. *IEEE Communications Letters*, 22(6):1184–1187, 2018.

[12] S. Gong, X. Wei, and X. Li. Eccm scheme against interrupted sampling repeater jammer based on time-frequency analysis. *Journal of Systems Engineering and Electronics*, 25(6):996–1003, 2014.

[13] W.-Q. Wang. Moving-target tracking by cognitive rf stealth radar using frequency diverse array antenna. *IEEE Transactions on Geoscience and Remote Sensing*, 54(7):3764–3773, 2016.

[14] W.-Q. Wang. Adaptive rf stealth beamforming for frequency diverse array radar. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, pages 1158–1161. IEEE, 2015.

[15] Z. Zhang, S. Salous, H. Li, and Y. Tian. Optimal coordination method of opportunistic array radars for multi-target-tracking-based radio frequency stealth in clutter. *Radio Science*, 50(11):1187–1196, 2015.

[16] P. Samuelson. A note on the pure theory of consumer's behaviour. *Economica*, pages 61–71, 1938.

[17] S. Afriat. The construction of utility functions from expenditure data. *International economic review*, 8(1):67–77, 1967.

[18] W. Diewert. Afriat and revealed preference theory. *The Review of Economic Studies*, pages 419–425, 1973.

[19] H. Varian. The nonparametric approach to demand analysis. *Econometrica*, 50(1):945–973, 1982.

[20] X. R. Li and V. P. Jilkov. Survey of maneuvering target tracking. part i. dynamic models. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1333–1364, 2003.

[21] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with applications to tracking and navigation*. John Wiley, New York, 2008.

[22] S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[23] N. D. Ratliff, J. A. Bagnell, and M. A. Zinkevich. Maximum margin planning. In *Proceedings of the 23rd international conference on Machine learning*, pages 729–736, 2006.