Manipulating the Posterior Support of a Discrete Bayesian Estimator Under Full Sensor Control

Daniel Selvaratnam * **Farhad Farokhi** ** **Iman Shames** ***
Henrik Sandberg *

* *Decision and Control Systems, School of Electrical Engineering & Computer Science, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: {selv,hsan}@kth.se).*
** *Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia (e-mail: farhad.farokhi@unimelb.edu.au)*
*** *College of Engineering & Computer Science, Australian National University, Canberra, ACT 2601, Australia (e-mail: iman.shames@anu.edu.au)*

Abstract: The asymptotic implausibility problem is introduced from the perspective of an adversary that seeks to drive the belief of a recursive Bayesian estimator away from a particular set of parameter values. It is assumed that the adversary controls all sensors informing the estimator, and can transmit false measurements stochastically according to a fixed distribution of its choice. First, we outline a method for verifying whether a given distribution solves the problem. We then consider the class of spoofing attacks, and show that the asymptotic implausibility problem has a solution if and only if it can be solved by a spoofing attack. Attention is restricted to finite parameter and observation spaces.

Keywords: Network Security, Bayesian Estimation, Cyber-Physical Systems, Network Systems, Resilient Control Systems, Signal Processing

1. INTRODUCTION

The posterior of a Bayesian estimator evolves in discrete-time, driven by the measurements that it receives. Its steady-state value is of particular interest, because it captures the final belief of the estimator. Here, we adopt the perspective of an adversary that transmits false measurements to drive the belief of the estimator away from a given subset of the parameter space. Mathematically, its goal is to ensure certain parameter values are excluded from the asymptotic support of the posterior, which renders them increasingly implausible to the estimator with time. We call this the *asymptotic implausibility* problem. The adversary is assumed to control all the sensors informing the estimator, and can thus choose the distribution of measurements arbitrarily. A stochastic policy is desirable for avoiding detection because randomness makes the measurement sequence seem less contrived. For example, the measurement sequence may be required to pass a randomness test for it to be trusted by the estimator. Moreover, if the sensors are being manipulated by the adversary at the physical layer, their outputs may be inherently stochastic. We therefore restrict the adversary to only transmit measurements stochastically according to a static distribution. Attention is also restricted to finite parameter and observation spaces for simplicity. Future work will address more complex scenarios involving the presence of truthful sensors, time-varying measurement distributions, and continuous parameter and observation spaces. But in this work, the adversary must design a single finite-dimensional measurement distribution to achieve its objective.

The goal of asymptotic implausibility is related to the *Byzantine generals problem*, in which a group of agents (the generals) must reach a consensus based on messages received from each general, while a subset of them work actively to disrupt this process. The original formulation by Lamport et al. [1982] considers distributed decision-making, with each general sending messages deterministically to all others. The paper develops algorithms enabling the loyal generals to robustly reach consensus, and quantifies the minimum proportion of them required to do so. The signal processing literature considers a variant of this problem involving stochastic messages and centralised decision-making by a fusion centre (see the survey paper [Vempaty et al., 2013] and the references therein). The goal of the fusion-centre is either estimation or detection. For the former, the parameter belong to $\mathbb{R}^n$, and the adversary attempts to induce a large estimation error. For the latter, the parameter belongs to a binary set: the hypothesis is either true or false. Our work differs from both of these by considering finite-cardinality parameter spaces. Also, rather than attempting to design robust estimation/detection strategies, we assume a recursive Bayesian estimator, and treat the problem from the adversary's perspective. To the best of the authors' knowledge, the asymptotic implausibility problem has not previously appeared in the literature. However, for the special case of binary parameter spaces, the setup reduces to that of the aforementioned detection problem, because there are only two possible parameter values.

Our earlier work [Farokhi et al., 2018] and [Selvaratnam, 2019, Chapter 3] consider attacks on a single-step discrete Bayesian estimator with finite parameter and observation spaces. There, the adversary seeks to make the posterior equal to a desired distribution by sending a single stochastic measurement. Since the value of the posterior at any point depends on that measurement, the goal is to choose a measurement distribution that achieves equality in expectation. The obvious limitation is that equality need not hold for a particular measurement realisation, which is what the estimator actually experiences. Neither does it hold as the number of measurements approaches infinity. The asymptotic implausibility problem, by contrast, explicitly aims to manipulate the steady-state posterior support, and its intended effect is achieved almost surely.

Our contributions here are as follows. First, the asymptotic implausibility problem is formulated mathematically in Section 3, and a procedure for testing candidate solutions is given. Then, a particular class of solutions based on *spoofing* is proposed, for which the adversary mimics a false parameter value by sending measurements according to the corresponding likelihood function. This strategy is straightforward to implement, and hard to detect, because the distribution of measurements is consistent with what the estimator expects according to its own model. Since the posterior then concentrates about the false value, it necessarily vanishes at all other values, including those the adversary seeks to render implausible. However, the false value cannot be chosen arbitrarily. In Section 4, we characterise the parameter values that can be mimicked to solve the asymptotic implausibility problem. Conversely, it is also shown that if a spoofing solution does not exist, then no solution exists at all. Finally, a numerical example is given in Section 5, and concluding remarks are made in Section 6.

2. PRELIMINARIES

2.1 Set theory

Let $\mathbb{R}$ denote the real numbers, and $\mathbb{N} := \{0, 1, 2, ...\}$ the natural numbers. The cardinality of a set A is denoted by $|A|$. The (non-strict) subset relation is denoted by $\subset$, and the strict relation by $\subsetneq$. Given $i, j \in \mathbb{N}$, let $[i : j] := \{k \in \mathbb{N} \mid i \leq k \leq j\}$, and $x_{i:j} := (x_i, x_{i+1}, \ldots, x_{j-1}, x_j)$.

Lemma 1. If $A \subset B$, then $A = B \iff |A| = |B|$.

Proof. Let $A \subset B$. If $A = B$, then clearly $|A| = |B|$. Conversely, if $A \neq B$, then $A \subsetneq B$, by which $|A| < |B|$.

2.2 Kullback-Leibler Divergence

Given a finite set $\mathcal{D}$, the Kullback-Leibler (KL) Divergence of probability distribution $g : \mathcal{D} \to [0, 1]$ from probability distribution $f : \mathcal{D} \to [0, 1]$ is

$$\boldsymbol{\mathcal{K}}(f \| g) = \sum_{d \in \mathcal{D}} f(d) \ln \frac{f(d)}{g(d)} \in [0, \infty], \tag{1}$$

where by convention [Cover and Thomas, 2005, Chapter 2.3],

$$0 \ln 0 := 0, \quad 0 \ln \frac{0}{0} := 0, \quad \forall p > 0,\ p \ln \frac{p}{0} := \infty. \tag{2}$$

Remark 1. The divergence $\boldsymbol{\mathcal{K}}(f \| g) < \infty$ if and only if

$$g(d) = 0 \implies f(d) = 0.$$

3. PROBLEM FORMULATION

Let $\mathbb{X}$ and $\mathbb{O}$ be finite sets, the *parameter space* and *observation space*, respectively. A discrete Bayesian estimator maintains a posterior distribution $p_k : \mathbb{X} \times \mathbb{O}^k \to [0, 1]$, which describes the estimators belief about the value of an unknown parameter $X \in \mathbb{X}$ based on the $k \in \mathbb{N}$ measurements it has received so far. According to the estimator, $p_k(x \mid z_{1:k})$ is the probability that $X = x$, given the sequence of measurements $z_1, ..., z_k \in \mathbb{O}$. Its measurement model consists of a family of *likelihood functions* $\{\ell_x : \mathbb{O} \to [0, 1] \mid x \in \mathbb{X}\}$, where $\ell_x(z)$ is the probability (assumed by the estimator) of receiving the measurement $z \in \mathbb{O}$ given that $X = x$. The estimator models X as a random variable drawn from a *prior* distribution $p_0 : \mathbb{X} \to (0, 1]$, and updates its posterior in time according to Bayes' rule:

$$p_{k+1}(x \mid z_{1:k+1}) = \frac{\ell_x(z_{k+1}) p_k(x \mid z_{1:k})}{\displaystyle\sum_{x' \in \mathbb{X}} \ell_{x'}(z_{k+1}) p_k(x' \mid z_{1:k})}, \tag{3}$$

for all $k \in \mathbb{N}$.

Remark 2. It can be assumed, without loss of generality, that $p_0(x) > 0$ for all $x \in \mathbb{X}$, because any parameter values with zero prior can simply be excluded from consideration by removing them from $\mathbb{X}$.

In adopting the recursion (3), the estimator makes the standard assumptions

$$X \sim p_0, \tag{4}$$

$$z_1, z_2, ... \sim \ell_X \text{ i.i.d..} \tag{5}$$

These need not be true. Suppose that, in fact,

$$z_1, z_2, ... \sim q \text{ i.i.d.,} \tag{6}$$

where the *observed distribution* $q : \mathbb{O} \to [0, 1]$ is the distribution that the received measurements actually follow. In this case, the posterior p_k decays to zero outside the set

$$\mathcal{P}_\infty := \arg\min_{x \in \mathbb{X}} \boldsymbol{\mathcal{K}}(q \| \ell_x),$$

regardless of the estimators choice of prior [Walker, 2013]. Specifically, if $\mathcal{P}_\infty \subset \mathcal{P} \subset \mathbb{X}$, then

$$\lim_{k \to \infty} \sum_{x \in \mathcal{P}} p_k(x \mid z_{1:k}) = 1 \text{ a.s.} \tag{7}$$

under certain regularity assumptions [Berk, 1966, Section 4]. This implies that

$$\forall x \in \mathbb{X} \setminus \mathcal{P}_\infty,\ \lim_{k \to \infty} p_k(x \mid z_{1:k}) = 0 \text{ a.s..} \tag{8}$$

This well-known result can be summarised as follows.

Remark 3. The only parameter values that remain plausible in the long run, are those that minimise the Kullback-Leibler divergence from the observed distribution to distributions in the measurement model.

Suppose the observed distribution q can be chosen by the adversary. In Problem 1, below, it seeks to make the posterior decay on the set $\mathcal{E} \subsetneq \mathbb{X}$, thereby driving the belief of the estimator away from $\mathcal{E}$ as $k \to \infty$. For example, one might expect the adversary to include $X \in \mathcal{E}$ if it knows the true parameter value X. Moreover, the adversary could choose $\mathcal{E} = \mathbb{X} \setminus \{x^\star\}$ to convince the estimator that $X = x^\star \in \mathbb{X}$.

Problem 1. (Asymptotic Implausibility). Given some non-empty $\mathcal{E} \subsetneq \mathbb{X}$, find a probability distribution $q : \mathbb{O} \to [0,1]$ such that

$$\mathcal{E} \cap \arg\min_{x\in\mathbb{X}} \mathcal{K}(q\|\ell_x) = \emptyset. \tag{9}$$

Remark 4. In general, there is a distinction between the distribution of measurements received by the estimator (the observed distribution) and the measurement distribution chosen by the adversary (the *adversarial distribution*), because there may be multiple sources of information. Here, however, the two are equal, because the adversary has full control over the sensors informing the estimator.

The objective (9) can be written in an alternative form.

Proposition 1. The probability distribution $q : \mathbb{O} \to [0,1]$ solves Problem 1 if and only if

$$\min\{\mathcal{K}(q\|\ell_x) \mid x \in \mathbb{X}\} < \min\{\mathcal{K}(q\|\ell_e) \mid e \in \mathcal{E}\}. \tag{10}$$

Proof. The distribution q satisfies (9) if and only if $e \notin \arg\min_{x\in\mathbb{X}} \mathcal{K}(q\|\ell_x)$ for every $e \in \mathcal{E}$. That is,

$$\forall e \in \mathcal{E},\ \exists x \in X,\ \mathcal{K}(q\|\ell_x) < \mathcal{K}(q\|\ell_e.) \tag{11}$$

If (11) holds, then choosing $e^\star \in \arg\min_{e\in\mathcal{E}} \mathcal{K}(q,\ell_e)$,

$$\exists x \in X,\ \mathcal{K}(q\|\ell_x) < \mathcal{K}(q,\ell_{e^\star}) = \min\{\mathcal{K}(q,\ell_e) \mid e \in \mathcal{E}\},$$

which in turn implies (10). Conversely, if (10) holds, then

$$\forall e \in \mathcal{E},\ \min\{\mathcal{K}(q\|\ell_x) \mid x \in \mathbb{X}\} < \mathcal{K}(q,\ell_e),$$

which in turn implies (11).

3.1 Testing candidate solutions

By definition, q is a solution to Problem 1 if and only if none of the minimisers of $\mathcal{K}(q\|\ell_x)$ with respect to x lie in $\mathcal{E}$. Thus, in order to test whether a given q is a solution, the adversary must construct and sort the set $\{\mathcal{K}(q\|\ell_x) \mid x \in \mathbb{X}\}$. The construction requires $|\mathbb{X}|$ KL divergence evaluations, each of which is $O(|\mathbb{O}|)$, and the sorting is $O(|\mathbb{X}|\log|\mathbb{X}|)$ [Cormen et al., 2001]. The overall time-complexity of verifying a solution is therefore $O\left(|\mathbb{X}||\mathbb{O}| + |\mathbb{X}|\log|\mathbb{X}|\right)$.

The size of the solution set for Problem 1 depends on the sizes of $\mathcal{E}$ and $\arg\min_{x\in\mathbb{X}} \mathcal{K}(q\|\ell_x)$. Typically $\mathcal{K}(q\|\ell_x)$ has a unique minimiser x for any q, which makes $\arg\min_{x\in\mathbb{X}} \mathcal{K}(q\|\ell_x)$ a singleton. If $\mathcal{E}$ is also small, the two sets are unlikely to intersect for random choices of q. In such cases, a conceivable strategy is for the adversary to sample q randomly from the probability simplex and accept the first candidate that is a solution. But this is a poor strategy when $\mathcal{E}$ is large and $\mathcal{K}(q\|\ell_x)$ has multiple minimisers. Moreover, some instances of Problem 1 do not have solutions. Section 4 provides a test for whether solutions exist, and a systematic strategy to find them when they do.

4. SOLUTIONS VIA SPOOFING

The existence of solutions to Problem 1 depends on a property called *distinguishability*. If $\ell_x = \ell_y$ for parameter values $x, y \in \mathbb{X}$, then the estimator cannot distinguish between the possibilities $X = x$ and $X = y$, because they give rise to identical measurement distributions. Accordingly, we define the distinguishability relation between two points in the parameter space as follows.

Definition 1. (Point-point distinguishability). Points $x \in \mathbb{X}$ and $y \in \mathbb{X}$ are *distinguishable* if $\ell_x \neq \ell_y$. Otherwise, they are *indistinguishable*.

Next, we extend this definition by declaring a point in the parameter space to be distinguishable from a subset of the space if it is distinguishable from every element of that subset.

Definition 2. (Point-set distinguishablity). Point $y \in \mathbb{X}$ is distinguishable from set $\mathcal{E} \subsetneq \mathbb{X}$ if

$$\ell_y \notin \{\ell_e \mid e \in \mathcal{E}\}. \tag{12}$$

Otherwise, y and $\mathcal{E}$ are indistinguishable.

This leads to a direct solution to Problem 1.

Lemma 2. If $y \in \mathbb{X}$ is distinguishable from $\mathcal{E} \subsetneq \mathbb{X}$, then

$$\mathcal{E} \cap \arg\min_{x\in\mathbb{X}} \mathcal{K}(\ell_y\|\ell_x) = \emptyset.$$

Proof. The KL divergence between two discrete distributions is zero if and only if they are equal [Cover and Thomas, 2005, Theorem 2.6.3]. Thus,

$$\min\{\mathcal{K}(\ell_y\|\ell_x) \mid x \in \mathbb{X}\} = \mathcal{K}(\ell_y\|\ell_y) = 0.$$

Choose any $e \in \mathcal{E} \subset \mathbb{X}$. Equation (12) implies $\ell_y \neq \ell_e$, and therefore $\mathcal{K}(\ell_y\|\ell_e) > 0$, by which $e \notin \arg\min_{x\in\mathbb{X}} \mathcal{K}(\ell_y\|\ell_x)$.

Corollary 1. If $y \in \mathbb{X}$ is distinguishable from $\mathcal{E} \subsetneq \mathbb{X}$, then $q = \ell_y$ is a solution to Problem 1.

The strategy proposed in Corollary 1 is a spoofing attack: it requires the adversary to mimic a point y that is distinguishable from $\mathcal{E}$. If no such point exists, we declare $\mathcal{E}$ to be *simply indistinguishable* (without reference to any other point), and obtain a converse result.

Definition 3. (Simply indistinguishable set). A set $\mathcal{E} \subsetneq \mathbb{X}$ is simply indistinguishable if $\{\ell_e \mid e \in \mathcal{E}\} = \{\ell_x \mid x \in \mathbb{X}\}$. Otherwise, it is simply distinguishable.

Corollary 2. The subset $\mathcal{E} \subsetneq \mathbb{X}$ is simply indistinguishable if and only if $|\{\ell_e \mid e \in \mathcal{E}\}| = |\{\ell_x \mid x \in \mathbb{X}\}|$.

Proof. Since $\mathcal{E} \subset \mathbb{X}$, this follows from Lemma 1.

Lemma 3. If $\mathcal{E} \subsetneq \mathbb{X}$ is simply indistinguishable, then no solution to Problem 1 exists.

Proof. If $\mathcal{E}$ is simply indistinguishable, then

$$\{\mathcal{K}(q\|\ell_x) \mid x \in \mathbb{X}\} = \{\mathcal{K}(q\|\ell_e) \mid e \in \mathcal{E}\}$$

for every probability distribution q, implying that (10) cannot hold. The result follows from Proposition 1.

We have shown that, if a solution to Problem 1 exists, then it can be solved via a spoofing attack, and it suffices for the adversary to mimic a parameter value that is distinguishable from those it seeks to make the estimator disbelieve.

Remark 5. The results of this section are parallel to those of Kosut et al. [2011] for estimation over a finite-dimensional (uncountable) parameter space. The measurement model there is linear, and the adversary seeks to induce a large estimation error, as opposed to asymptotic implausibility. The attack strategy is deterministic: inject false measurements that lie in the column space of the measurement matrix. This is a type of spoofing, because the received measurements correspond to a false parameter value. Although the definitions differ mathematically, the notion of *observability* in Kosut et al. [2011] plays a similar role to that of *distinguishability* here.

4.1 Time-complexity of finding spoofing solutions

Lemma 3 allows the adversary to reduce its search space from the $|\mathbb{O}|$-dimensional probability simplex to the finite set of spoofing solutions. If a spoofing solution cannot be found, then the lemma guarantees no other solutions exist. The search for a spoofing solution requires a comparison between the elements of $\{\ell_e \mid e \in \mathcal{E}\}$ and $\{\ell_x \mid x \in \mathbb{X} \setminus \mathcal{E}\}$. If an element ℓ_y of the latter is not present in the former, then the corresponding $y \in \mathbb{X} \setminus \mathcal{E}$ is distinguishable from $\mathcal{E}$, and $q = \ell_y$ solves Problem 1. If no such element is found, then $\mathcal{E}$ is simply indistinguishable, and no solution exists. Since each likelihood function is a $|\mathbb{O}|$-dimensional vector, the worst-case time-complexity of this procedure is $O(|\mathcal{E}|(|\mathbb{X}| - |\mathcal{E}|)|\mathbb{O}|) = O(|\mathcal{E}||\mathbb{X}||\mathbb{O}|)$.

Remark 6. For fixed $|\mathbb{O}|$ and $|\mathcal{E}|$, searching for a spoofing solution has lower asymptotic complexity in $|\mathbb{X}|$ than verifying an arbitrary candidate via the method of Section 3.1.

Remark 7. If the aim is only to test for the existence of solutions to Problem 1, and not to find a solution, $O(1)$ set equality tests have been developed [Yellin, 1992]. Such may be of use to the adversary when designing its set $\mathcal{E}$, to avoid making its own asymptotic implausibility problem infeasible.

5. NUMERICAL EXAMPLE

5.1 Problem setup

Consider a source localisation problem over the occupancy-grid map in Figure 1, taken from MIT's Radish robotics research dataset [Howard, 2010]. It is a 573 × 645 occupancy grid, so the parameter space cardinality is 369585. A Bayesian estimator receives binary measurements from two sensors, each reporting either 'hit' or 'miss' at every time-step. The observation space is therefore

$$\mathbb{O} := \{\text{hit}, \text{miss}\}^2,$$

which has cardinality 4. The probability of detection (i.e., of measuring 'hit') for a given sensor is a decreasing function of its distance from the source. Information from the map is incorporated by setting the prior to zero at all occupied locations. Once the estimator has sufficiently narrowed down the possible source locations, a search team is to be deployed to physically locate and extract the source. An adversary knows the source location and seeks to capture the source for itself. To do so, it must enter the building at the entry point, capture the source, and return to exit the building, without encountering the search team. Its goal is therefore to direct the belief of the estimator away, not just from the source, but from its entire planned entry/exit route. It constructs the exclusion zone in Figure 1 accordingly. The adversary controls both sensors, and its task is to devise a measurement distribution that renders the exclusion zone asymptotically implausible.

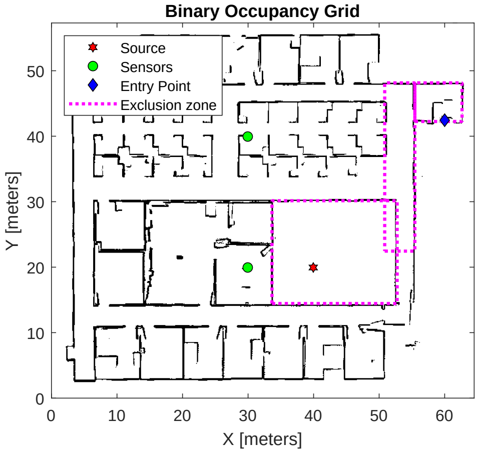


Fig. 1. Source localisation layout

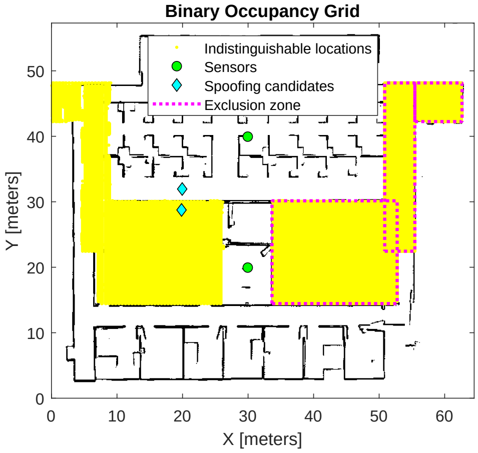


Fig. 2. Points that are indistinguishable from the exclusion zone are plotted in yellow.

5.2 Spoofing solutions

Since the prior vanishes at the occupied locations, the adversary must restrict itself to spoofing only unoccupied locations, as per Remark 2. The set of locations indistinguishable from the exclusion zone are plotted in Figure 2. In this problem instance, the likelihood functions depend only on distance from source, and so the set of indistinguishable locations possesses symmetry. In order to induce asymptotic implausibility, the adversary may therefore spoof any unoccupied location not plotted in yellow.

For purposes of comparison, Figure 3 plots the temporal evolution of the integral of the posterior over the exclusion zone, for the two candidate spoof locations depicted in Figure 2. Although both locations are physically close to each other, one is distinguishable from the exclusion zone, and the other is not. Figure 3 confirms that the first induces asymptotic implausibility, while the other fails to.

Finally, Figure 4 plots various snapshots of the posterior over time, to illustrate the changing belief of the estimator.

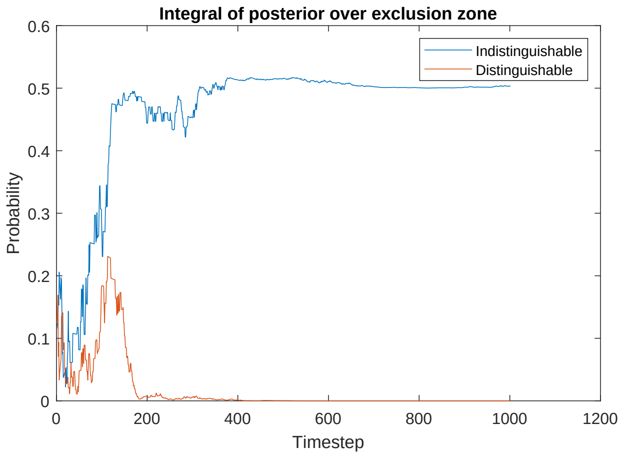


Fig. 3. Probability that source lies in exclusion zone according to the estimator, for two different spoof locations.

6. CONCLUSION

This paper poses the asymptotic implausibility problem, wherein an adversary seeks to drive the belief of a Bayesian estimator away from a given region of its parameter space. The adversary controls all the sensors informing the estimator, but must transmit measurements according to a fixed probability distribution, which it must design. We first demonstrate how to verify a candidate solution. We then consider the class of spoofing solutions, which require the adversary to mimic a false parameter value. Not just any false parameter value will do. To solve the asymptotic implausibility problem, the adversary can only mimic parameter values that are distinguishable from those it seeks to make the estimator disbelieve. It is enough for the adversary to restrict attention to spoofing attacks, because if no spoofing solution exists, then no solution exists at all. In some cases, it is more efficient to search for a spoofing solution than to verify an arbitrary solution candidate.

The assumptions herein endow the adversary with great power. In addition to controlling all the sensors, the adversary knows the estimator's measurement model and parameter space. These results are therefore indicative of the worst damage that an adversary could do. Future work will relax these assumptions by allowing for additional honest sources of information, and for a mismatch between the estimator's measurement model and the adversary's. Extensions to continuous parameter and observation spaces, and to time-varying adversarial distributions are also of interest. A special case of the latter is when the adversary is permitted to send a deterministic time-varying measurement sequence.

REFERENCES

Berk, R.H. (1966). Limiting Behavior of Posterior Distributions when the Model is Incorrect. *The Annals of Mathematical Statistics*, 37(1), 51–58.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. (2001). *Introduction To Algorithms*. MIT Press.

Cover, T.M. and Thomas, J.A. (2005). *Elements of Information Theory*. Hoboken, N.J. : J. Wiley, 2005.

Farokhi, F., Selvaratnam, D.D., and Shames, I. (2018). Security Analysis of Quantized Bayesian Estimators.

Howard, A. (2010). Sdr_site_b. http://hdl.handle.net/1721.1/62245.

Kosut, O., Jia, L., Thomas, R.J., and Tong, L. (2011). Malicious Data Attacks on the Smart Grid. *IEEE Transactions on Smart Grid*, 2(4), 645–658. doi:10.1109/TSG.2011.2163807.

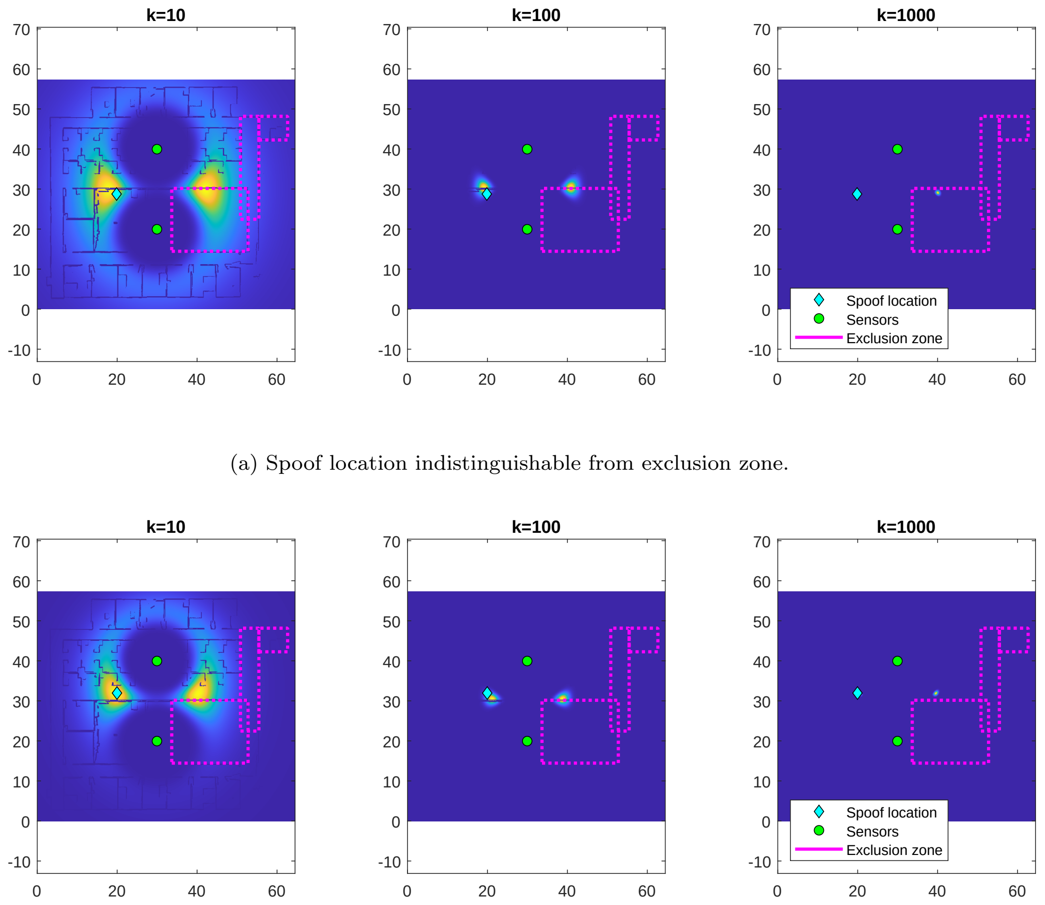
Lamport, L., Shostak, R., and Pease, M. (1982). The Byzantine Generals Problem. *ACM Transactions on Programming Languages and Systems*, 4(3), 382–401. doi:10.1145/357172.357176.

Selvaratnam, D.D. (2019). *Efficient Algorithms for Autonomous Agents Facing Uncertainty*. Ph.D. thesis, University of Melbourne.

Vempaty, A., Tong, L., and Varshney, P.K. (2013). Distributed Inference with Byzantine Data: State-of-the-Art Review on Data Falsification Attacks. *IEEE Signal Processing Magazine*, 30(5), 65–75. doi:10.1109/MSP.2013.2262116.

Walker, S.G. (2013). Bayesian inference with misspecified models. *Journal of Statistical Planning and Inference*, 143(10), 1621–1633. doi:10.1016/j.jspi.2013.05.013.

Yellin, D.M. (1992). Representing sets with constant time equality testing. *Journal of Algorithms*, 13(3), 353–373. doi:10.1016/0196-6774(92)90044-D.



(a) Spoof location indistinguishable from exclusion zone.

(b) Spoof location distinguishable from exclusion zone.

Fig. 4. Posterior at time-step k for different spoof locations. Dark blue regions have low probability, and yellow regions high probability. The prior is assigned zero probability at occupied cells.