This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

Tracking a Markov Target in a Discrete Environment with Multiple Sensors

Kevin Leahy and Mac Schwager

***Abstract*—In this work we consider using multiple noisy binary sensors to track a target that moves as a Markov Chain in a finite discrete environment, with symmetric probability of false alarm and missed detection. We study two policies. Firstly, we show that the greedy policy, whereby m sensors are placed at the m most-likely target locations, is one-step optimal in that it maximizes the expected maximum a posteriori (MAP) estimate. Secondly, we show that a policy in which the m sensors are placed in the second through $(m{+}1)^{st}$ most likely target locations achieves equal or slightly worse expected MAP performance, but leads to significantly decreased variance on the MAP estimate. The result is proven for $m = 1$, and Monte Carlo simulations give evidence for $m > 1$. Both policies are closed-loop, index-based active sensing strategies that are computationally trivial to implement. Our approach focuses on one-step optimality because of the apparent intractability of computing an optimal policy via dynamic programming in belief space. However, Monte Carlo simulations suggest that both policies perform well over arbitrary horizons.**

I. Introduction

This work focuses on finding a multi-agent search strategy that provides the best estimate of the location of a dynamic target that moves according to a Markov process in a discrete environment. Our work is inspired by [1], which proved that for a *single* sensor searching for a *stationary* target the optimal policy, in the sense of maximizing the expected maximum a posteriori (MAP) estimate, is to move to either of the first or second mostly likely target locations at each time step. The result is surprisingly simple, and holds over arbitrary time horizons. This motivates a natural next question: what is the optimal policy for *multiple* sensing agents tracking a *dynamic* target? We give two main results toward the conclusion of this question, under similar assumptions as [1]. Firstly, we prove that for a group of m sensing agents and a dynamic target, searching the m most likely locations at each time step is the optimal one-step policy with symmetric probability of false alarm and missed detection. Secondly, we show that searching the second through $(m+1)^{st}$ most likely locations at each step gives an equal or slightly worse expected MAP estimate, but also a significantly lower variance on the MAP estimate—that is, the second mostly likely locations provide a similar average result but with higher confidence. Specifically, for a single sensing agent and a dynamic target, we prove that searching the second most likely location at each step results in the *same* expected MAP estimate, while giving a *smaller* variance on the MAP estimate. For multiple sensing agents, we show through Monte Carlo simulations that the expected MAP estimate is slightly worse, but the variance is significantly lower. Finally, we also show through Monte Carlo simulations that these trends hold over arbitrary time horizons.

As a motivating example, consider the search for an endangered animal that has previously been tagged with a radio tracking device. The animal moves about its environment, and you know the areas it frequents (e.g. shady areas and waterholes), and how often it tends to move from area to area. You pilot a UAV, attempting to locate the animal via its tag, but given the animal's propensity for moving about and radio interference, it is difficult to be certain about where the animal is located. With all of this knowledge, where in the environment should you check to maximize your confidence about the animal's location?

Problems similar to the one we examine arise in a variety of domains, whenever sensing resources are limited and the motion of a phenomenon to be sensed is probabilistic. Such target tracking problems are of considerable interest to the robotics community, and have been well-studied recently, including gathering information about the environment, locating survivors at sea, general pursuit-evasion and search problems [2], object recognition and detection [3], and could be used in building an occupancy map of a dynamic environment. There is also a strong interest for the purpose of target localization with radar [4] or sensor network management [5], [6]. Even anomaly detection in cyberphysical systems has been approached as a problem of this form [7], [8]. For many of these applications, sensing resources are limited, and the computational complexity of optimal policies presents a significant barrier to practical implementation of such solutions. Our method is simple and efficient to implement, and provides a high degree of confidence in such scenarios.

We aim to track a target that is moving among a finite set of states as a Markov Chain (MC). At each time instant we have m sensors, each of which can look in a single state location for the target at each time step. The sensors return a binary result with false positives and false negatives. We suppose that the sensors do not have motion constraints, hence they can look at any location at any time step. As discussed previously, for a single sensor and a static target, it is known that searching either of the first or second most likely locations at each time step gives the maximum expected MAP estimate over an arbitrary time horizon. With a dynamic target and multiple sensors, the problem can be shown to be a high

K. Leahy is with the Department of Mechanical Engineering at Boston University, Boston MA 02215. Email: kjleahy@bu.edu

M. Schwager is with the Department of Aeronautics and Astronautics at Stanford University, Stanford CA 94305. Email: schwager@stanford.edu

This work was partially supported by NSF grants CNS-1330008 and IIS-1646921 and ONR grant N00014-12-1-1000.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

dimensional Partially Observable Markov Decision Process (POMDP), and hence a closed form solution is not available. Instead, we seek policies that optimize the expected MAP estimate over a single time step. We analyze the simple greedy strategy, in which the sensors look at the m most likely target locations. We prove this greedy policy is one-step optimal in the MAP sense. Further, we prove that for a single sensing agent and dynamic target, searching the second most likely location gives the same expected MAP estimate, but with lower variance. We conjecture that these results extend to the multi-agent case, as well. Monte Carlo simulations indicate that by *not* searching the most likely location, variance on the MAP estimate is significantly reduced, while the MAP estimate is slightly worse. This suggests a tradeoff between average estimation quality and estimation variability.

This problem falls generally within the class of problems known as *M*-ary Hypothesis Testing, in which a series of observations is used to decide which of M possible statistical situations give rise to those observations. These problems were first studied in the binary ($M = 2$) case in [9]. The presence of a moving target means our problem is a special case of *M*-ary Hypothesis Testing known as Sequential Hypothesis Testing. Early work for detecting moving targets in such sequential problems was performed in the Operations Research community, for both Markovian [10], [11] and non-Markovian targets [12], [13]. In these problems, over a finite horizon, the probability of detecting a target before the time horizon is maximized, using an exponential detection function.

Our work focuses instead on infinite horizon search problems, in which we attempt to maximize localization information about a target using binary detection. Many of the results from this area focus on an optimal strategy consisting of a finite sequence of search locations determined *a priori*, based on a prior distribution over possible target locations, such as [14] and [15]. In our work, however, we seek to evaluate the performance of a reactive strategy that can be used efficiently online: a simple sort operation is all that is required to determine which locations to search.

It is the closed-loop structure of the policy we find, as well as the generality of the problem formulation that make it most closely related to [1], although that work considers a stationary target rather than a moving target. Other work has considered index-based policies in this problem domain [16], including closed-loop index-based policies [17], but not for closed-loop policies or a moving target. Our reactive strategy provides a simple, computationally efficient feedback policy for searching for a target that is one-step optimal and performs well over an infinite horizon.

Some recent work has focused on problems similar to the one in this paper. Closely related is [18], which studies the case of a target moving on a one-dimensional graph, subject to time and energy constraints on the searcher. Our work studies the more general case of arbitrary discrete state spaces without constraints on the searcher, instead focusing on simple, efficient policies based on properties of the evolution of the belief about a target's location. Other work includes [19] and [20], both of which focus on probabilistic game theoretic approaches to pursuit evasion games with greedy policies. The policy in this paper is also greedy, and can be viewed as a pursuit evasion problem, but is not based on a game theoretic approach. Rather, we prove the one-step optimality of a greedy index-based approach, which is generally more computationally efficient than approaches rooted in game theory. A multi-agent search problem was considered in [21], but that work considered a stationary target and no false positive measurements. A preliminary version of this work, consisting of the special case for one sensor, was presented in [22].

The remainder of the paper is organized as follows. In Sec. II, we introduce the models used in this work, and formally state the problem. Next, we demonstrate the optimality of our index-based approach in Sec. III and we investigate policies with lower variance in Sec. IV. Simulation results are given in Sec V, and concluding remarks are in Sec. VI.

II. Problem Formulation

A. Preliminaries

In this section, we introduce notation and definitions that will be used throughout the text. For a set Σ, the cardinality and power set are written $|\Sigma|$ and 2^{Σ}, respectively.

Definition 1: *Partial Order* [23] For a set Σ, an *order*, sometimes called a *partial order*, on Σ is a binary relation $\leq$ on Σ such that $\forall x, y, z \in \Sigma$:

- $x \leq x$,
- $x \leq y$ and $y \leq x \implies x = y$, and
- $x \leq y$ and $y \leq z \implies x \leq z$.

Definition 2: *Total Order* [23] Let Σ be an ordered set. Then Σ is a totally ordered set if $\forall x, y \in \Sigma$, either $x \leq y$ or $y \leq x$.

Definition 3: *Join and Meet* [23] For a subset S of a partially ordered set $\langle \Sigma, \leq \rangle$, for any $x, y \in S$, z is the *meet* (or *greatest lower bound*) of x, y if:

- $z \leq x$ and $z \leq y$, and
- $\forall w$ such that $w \leq x$ and $w \leq y$, $w \leq z$.

The meet of x, y is written $x \wedge y$. Similarly, for any $x, y \in S$, z is the *join* (or *least upper bound*) of x, y if:

- $x \leq z$ and $y \leq z$, and
- $\forall w$ such that $x \leq w$ and $y \leq w$, $z \leq w$.

We denote the join of x, y as $x \vee y$.

Definition 4: *Lattice* [23] A non-empty ordered set Σ is a *lattice* if $x \wedge y$ and $x \vee y$ exist for all $x, y \in \Sigma$.

Definition 5: *Fence* [23] A *fence* is a partially ordered set $\Sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ such that $\sigma_1 > \sigma_2 < \sigma_3 > \ldots \sigma_n$ or $\sigma_1 < \sigma_2 > \sigma_3 < \ldots \sigma_n$ with no other comparabilites between elements in the set. Elements σ_1 and σ_n are called *endpoints* of the fence.

B. Target Model

For this problem, we consider a discrete environment consisting of a set of states S. There is a target moving in this environment whose location at time k is denoted as s^k, where $s^k \in S$. The target evolves as a discrete time MC whose transition probabilities are known a priori. We denote the probability of the target moving from a state s_1 to a state s_2 as $P(s_1, s_2)$.

C. Sensing Model

Our goal is to track the target as it moves about the environment. We do so by choosing a set of locations $q^k \subset S$ in which to take an observation y^k at each time step k. We let $m = |q^k| \leq |S|$ be the number of locations that may be searched at each time step. The number of locations does not change over time (i.e. m is a fixed value). We assume no restrictions on the motion of our sensors, although this is an interesting direction for future work. At each time instant q^k may be chosen to be any subset of S, regardless of the value of q^{k-1}. The measurement model is binary, with measurements $y^k \in Y$, where $Y = \{0,1\}^m$.

The set of locations being searched at each time are indexed $\{1, \ldots, i, \ldots, m\}$. The i^{th} location being searched at time k is denoted q_i^k, and its corresponding measurement is denoted y_i^k. The measurement y_i^k takes values in $\{0,1\}$. The measurement probabilities are:

$$P\left(y_i^k = 1 \mid q_i^k, s^k\right) = \begin{cases} 1-\beta & \text{if } q_i^k = s^k \\ \alpha & \text{otherwise} \end{cases}$$

$$P\left(y_i^k = 0 \mid q_i^k, s^k\right) = \begin{cases} \beta & \text{if } q_i^k = s^k \\ 1-\alpha & \text{otherwise} \end{cases},$$

where β is the probability of missed detection and α is the probability of false alarm. We assume $\alpha = \beta < 0.5$. Such an assumption is reasonable in many cases, such as obtaining a binary sensor from continuous measurements with additive Gaussian noise (or other symmetric noise) by using a threshold. We further assume the measurements of y_i^k are conditionally independent for all $i \in q^k$, i.e.,

$$\left(y_i^k \perp\!\!\!\perp y_j^k\right) | s^k , \tag{1}$$

for all $i, j \in q^k$.

D. Belief Model

The location of the target at time 0 is unknown, but we assume a prior belief at time 0, given as b^0. The prior belief for a specific state $s \in S$ is written $b^0(s)$. This belief evolves over time, and we denote the belief about state s at time k as $b^k(s)$. We update the belief after taking each measurement according to Bayes rule:

$$b^k(s) = \eta P\left(y^k \mid s, q^k\right) \sum_{s' \in S} P(s', s)\, b^{k-1}(s') , \tag{2}$$

where η is the appropriate normalization factor and $P(s', s)$ is the probability of the target transitioning from state s' to state s as describe in Sec. II-B.

Note that (2) can be written in matrix form as

$$b^k = \eta P\left(y^k = y | q^k, s\right) \Gamma b^{k-1}, \tag{3}$$

where Γ is a column stochastic matrix, whose i, j^{th} entry represents the probability of transitioning to state i from state j, b^k is a vector of length $|S|$ representing the belief from states 1 to $|S|$, and $P\left(y^k = y|q^k, s\right)$ is a matrix representing the measurement likelihood. Because of conditional independence, we may write

$$P\left(y^k = y|q^k, s\right) = \prod_{i=1}^{m} P\left(y_i^k = y|q_i^k, s\right) , \tag{4}$$

where $P\left(y_i^k = 1|q_i^k\right)$ and $P\left(y_i^k = 0|q_i^k\right)$ are diagonal matrices of the form

$$P\left(y_i^k = 1|q_i^k\right) = \begin{bmatrix} \beta & & & & & & \\ & \ddots & & & & \mathbf{0} & \\ & & \beta & & & & \\ & & & 1-\beta & & & \\ & & & & \beta & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & \beta \end{bmatrix} \tag{5}$$

$$P\left(y_i^k = 0|q_i^k\right) = \begin{bmatrix} 1-\beta & & & & & & \\ & \ddots & & & & \mathbf{0} & \\ & & 1-\beta & & & & \\ & & & \beta & & & \\ & & & & 1-\beta & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & 1-\beta \end{bmatrix}, \tag{6}$$

which we call *observation matrices*. These matrices, which we will henceforth identify with the shorthand P_{1i} and P_{0i}, sum to identity. Each matrix has a single entry that differs from the others, for the choice of q_i^k. Then the matrix representing $P\left(y^k = y|q^k, s\right)$ is simply the product of the matrices of $P\left(y_i^k = y_i|q_i^k, s\right)$.

It is worth noting that our problem can be seen as choosing the observation matrices at each time step that results in a long product of matrices with the initial belief of the form

$$\eta P\left(y^k|q^k\right) \Gamma P\left(y^{k-1}|q^{k-1}\right) \Gamma \ldots P\left(y^1|q^1\right) \Gamma b^0 . \tag{7}$$

In general, these matrices do not commute. If Γ is the identity matrix, the problem is reduced to the problem of locating a stationary target, which has been well studied in the literature [1], [24].

Notation: The search policy we propose in this work is index-based. This means we focus on the relative magnitude of the belief at each location at each time step. Such a focus allows us to avoid partitioning the belief space beyond a simple ordering, resulting in a highly computationally efficient policy. To formalize our argument, we will now introduce some notation about the sorted belief. Without loss of generality, we denote at any time the largest element of the belief as b_1^k. That is,

$$b_1^k = \max_{s \in S} b^k(s) .$$

Similarly, the next biggest element of the belief is denoted b_2^k, and so forth, with the smallest element being $b_{|S|}^k$. This notation allows us to think of the belief as though it were sorted each time it is updated, with b_i^k being the i^{th} biggest element of b^k at any given time step k, so that $b_1^k \geq b_2^k \geq \ldots \geq b_{|S|}^k$. We will call the location of b_i^k location i, so that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

$i = \left\{s|b^k(s) = b_i^k\right\}$. That way, we can conveniently refer to states in S according to the relative magnitude of their belief. For notational convenience, we will omit the time step k from the sorted belief except when necessary. That is, we will generally write b_i in place of b_i^k.

E. Problem Statement

Given the above target, sensor, and belief models, we can formally state our objective. We seek a policy for determining where to look at each time step as a function of the current belief to maximize the next step probability of finding the target.

Problem 1: Given a target modeled as a MC as described in Sec. II-B and a sensor with $\alpha = \beta$ modeled as described in Sec. II-C, find a policy $\mu : b \to S$ to maximize the function $J\left(b^k\right) = \max_{s \in S} b^k(s) = b_1^k$:

$$q^{k*} = \arg\max_{q^k} E_{y^k}[J(b^k)|b^{k-1}, q^k] \, . \tag{8}$$

To simplify our notation, we will henceforth omit the notation indicating that expectation of b^k and $J\left(b^k\right)$ is given b^{k-1} and q^k, but the reader is advised to keep in mind that all expectation is conditional on b^{k-1} and the choice of sensor location q^k. Further, we wish to note that the belief at time $k-1$ is equivalent to knowing $y^{1:k-1}$. In other words, b^{k-1} is a sufficient statistic for $y^{1:k-1}$.

Remark 1: We are restricting our focus to one-step optimal policies in this work, rather than using another approach such as dynamic programming solutions. Using such tools for a nonlinear objective function such as (8) is often intractable due to the nature of the evolution of belief. Our focus is on an efficient, tractable, one-step optimal approach.

Remark 2: The cost function we consider is the same cost function studied in [1] for a single static target and a single sensor. Other common choices in the related literature include entropy, mutual information, and time to detection. We focus on the probability of the target location because it permits an efficient solution.

III. Optimality of the Greedy Strategy

We wish to find a solution for choosing $q^k \subset S$ at each time to maximize (8). To accomplish this, we investigate the expectation of the belief b^k and the expectation of the MAP belief $\max b^k$, both with respect to y^k. We start with analyzing the expectation of b^k and its relationship to the choice of sensing locations.

A. Expectation of Belief

Starting from (3), the expectation of b^k is written

$$E_{y^k}\left[b^k\right] = \sum_{y \in Y^k} P\left(y^k = y|y^{1:k-1}\right) \eta P\left(y^k|q^k\right) \Gamma b^{k-1} \, . \tag{9}$$

We note that in this expression, the total probability of observing $y^k = y$, denoted by $P\left(y^k = y|y^{1:k-1}\right)$, is equal to $1/\eta$. This fact implies that each observation carries equal weight in expectation and lets us write (9) as

$$E_{y^k}\left[b^k\right] = \sum_{y \in Y^k} P\left(y^k|q^k\right) \Gamma b^{k-1} \, . \tag{10}$$

Henceforth, we will denote Γb^{k-1} as $\hat{b}^k$, the prior for b^k based on the transition probability for the target and the belief at time $k-1$. Then, our final expression for (9) becomes

$$E_{y^k}\left[b^k\right] = \sum_{y \in y^k} P\left(y^k|q^k\right) \hat{b}^k \, . \tag{11}$$

Because of conditional independence, we can write this as

$$E_{y^k}\left[b^k\right] = \left(\prod_{i=1}^{m} (P_{1i} + P_{0i})\right) \hat{b}^k \, . \tag{12}$$

Recalling from Sec. II-D that P_{1i} and P_{0i} sum to identity (see Sec. II-D), the expectation of b^k is simply $\hat{b}^k$. This rearrangement of (9) will be useful when we examine the expectation of the maximum next step probability of finding the target.

As mentioned in Sec. II-D, P_{1i} and P_{0i} are diagonal matrices with β and $1-\beta$ on the diagonal and zeros elsewhere. Therefore expanding the product in (12) means that each element in $\hat{b}^k$ is multiplied by each element in the set of outcomes

$$\binom{m}{n} \beta^{m-n} (1-\beta)^n \tag{13}$$

which is the binomial formula for $(\beta + (1-\beta))^m$. This result intuitively makes sense, given m conditionally independent sensors that each behave as a Bernoulli random variable with success probability $1-\beta$. We illustrate this property in Example 1 below.

Example 1: As an example, we consider $|S|= 3$, $m = 2$, and $q^k = \{1, 2\}$, with $q_1^k = 1$ and $q_2^k = 2$. Then

$$P\left(y_1^k = 1|q_1^k\right) = \begin{bmatrix} (1-\beta) & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \beta \end{bmatrix} \tag{14}$$

and

$$P\left(y_1^k = 0|q_1^k\right) = \begin{bmatrix} \beta & 0 & 0 \\ 0 & (1-\beta) & 0 \\ 0 & 0 & (1-\beta) \end{bmatrix} . \tag{15}$$

The corresponding matrices for q_2^k would be the same as those for q_1^k, except the first and second diagonal elements would

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

be switched. Equation (12) can then be written as

$$\begin{aligned} E_{y^k}\left[b^k\right] = &\begin{bmatrix} \beta\left(1-\beta\right) & 0 & 0 \\ 0 & \beta\left(1-\beta\right) & 0 \\ 0 & 0 & \beta^2 \end{bmatrix}\hat{b}^k \\ + &\begin{bmatrix} \left(1-\beta\right)^2 & 0 & 0 \\ 0 & \beta^2 & 0 \\ 0 & 0 & \beta\left(1-\beta\right) \end{bmatrix}\hat{b}^k \\ + &\begin{bmatrix} \beta^2 & 0 & 0 \\ 0 & \left(1-\beta\right)^2 & 0 \\ 0 & 0 & \beta\left(1-\beta\right) \end{bmatrix}\hat{b}^k \\ + &\begin{bmatrix} \beta\left(1-\beta\right) & 0 & 0 \\ 0 & \beta\left(1-\beta\right) & 0 \\ 0 & 0 & \left(1-\beta\right)^2 \end{bmatrix}\hat{b}^k \end{aligned} \quad (16)$$

Then, reading across a given row in the four matrices, it is clear that the four non-zero elements compose

$$\sum_{n=0}^{2}\binom{2}{n}\beta^{2-n}\left(1-\beta\right)^n\ , \quad (17)$$

but they are permuted differently depending on which locations are being searched. Therefore, although there are 2^m possible observations (because $y^k \in \{0,1\}^m$), there are only $m+1$ distinct values that might multiply each entry in $\hat{b}^k$. In this example, since $m=2$, there are 4 possible observations ($\{1,1\}$, $\{1,0\}$, $\{0,1\}$, and $\{0,0\}$), but $m+1=3$ distinct multipliers ($(1-\beta)^2$, $(1-\beta)\beta$, and β^2). Written in that order, these multipliers have multiplicity 1, 2, and 1, corresponding to $\binom{m}{0}$, $\binom{m}{1}$, and $\binom{m}{2}=\binom{m}{m}$.

The example above illustrates that each element in $\hat{b}^k$ gets multiplied by every term in the sum in (13), whose values range from a maximum of $(1-\beta)^m$ to a minimum of β^m. Let us denote, for each of the observations $y^k \in \{0,1\}^m$, the set of $m+1$ multipliers $\mathbf{C}=\{c_1,c_2,\ldots,c_i,\ldots,c_{m+1}\}$. Without loss of generality, we will assume these multipliers are ordered such that $c_1 \geq c_2 \geq \ldots \geq c_{m+1}$.

For m sensors, let n be the number of sensors measuring 1 at a given time step. Then we can say that there are $\binom{m}{n}$ ways of getting n measurements of 1. Likewise there $\binom{m}{m-n}$ different ways to receive $m-n$ measurements of 0. When computing the expectation, the belief at the n locations where sensors measure 1 are multiplied by

$$c_{y=1} = \left(1-\beta\right)^{m-n+1}\beta^{n-1}\ , \quad (18)$$

while at the $m-n$ locations where sensors measure 0 the belief is multiplied by

$$c_{y=0} = \left(1-\beta\right)^{m-n-1}\beta^{n+1}\ . \quad (19)$$

For locations where a sensor measures 1, the exponent $m-n+1$ corresponds to the $m-n$ sensors measuring 0, plus an additional $(1-\beta)$ term for the measurement at that location, while the exponent $n-1$ corresponds to the other $n-1$ locations that measure 1. The reverse logic holds for the exponents $m-n-1$ and $n+1$ for those locations where a 0 is measured. Finally, the $|S|-m$ locations where no measurement is made are each multiplied by

$$c_{other} = \left(1-\beta\right)^{m-n}\beta^{n}\ . \quad (20)$$

These three equations—(18), (19), and (20)—give analytical expressions for the multipliers for any given measurement outcome. Further, they show that for any number of sensors and any outcome, there are only 3 multipliers. This fact will be useful in computing the expectation of the maximum of our belief. We will refer to the generic multiplier c_i with the following form

$$c_i = \left(1-\beta\right)^{m+1-i}\beta^{i-1}\ , \quad (21)$$

which is valid for $i \in \{1,2,\ldots,m+1\}$. The relationship among these multipliers and the number and location of different measurements, is summarized in Table I, and a concrete example for the case of $m=3$ sensors is given in Table II.

We summarize the above with a proposition.

Proposition 1: For any measurement realization in $\{0,1\}^m$, there are three multipliers in $\mathbf{C}$ that multiply the entries in b, $c_{y=1}$ for entries with a measurement of 1, $c_{y=0}$ for entries with a measurement of 0, and c_{other} for the remaining entries. These coefficients have the relationship $c_{y=1} \geq c_{other} \geq c_{y=0}$. The exact values $c_{y=1}$, $c_{y=0}$, and c_{other} depend on the value of m and the number of sensors with a measurement of 1 in a given measurement realization.

B. Computing the Expected MAP Belief

To compute the expectation of the MAP belief $\max b^k$, we write a sum over all possible observations in $\{0,1\}^m$, written as

$$E_{y^k}\left[\max_{s\in S} b^k\left(s\right)\right] = \sum_{y^k\in\{0,1\}^m}\max_{s\in S} P\left(Y^k=y^k|q^k\right)\hat{b}^k\ , \quad (22)$$

where $P\left(Y^k=y^k|q^k\right)$ is the product of the m corresponding matrices P_{1i} and P_{0i}, for each observation and location in y^k and q^k. The matrix $P\left(Y^k=y^k|q^k\right)$ is a diagonal matrix, as are each of the P_{1i} and P_{0i} matrices. We continue Example 1 below to make computing (22) clearer.

Example 2: We consider the same setup as Example 1. Now, the expectation of the MAP belief as given by (22) is written as

$$\begin{aligned} E_{y^k}\left[\max_{s\in S} b^k\left(s\right)\right] = &\max_{s\in S}\left(\begin{bmatrix} \beta\left(1-\beta\right) & 0 & 0 \\ 0 & \beta\left(1-\beta\right) & 0 \\ 0 & 0 & \beta^2 \end{bmatrix}\hat{b}^k\right) \\ &+\max_{s\in S}\left(\begin{bmatrix} \left(1-\beta\right)^2 & 0 & 0 \\ 0 & \beta^2 & 0 \\ 0 & 0 & \beta\left(1-\beta\right) \end{bmatrix}\hat{b}^k\right) \\ &+\max_{s\in S}\left(\begin{bmatrix} \beta^2 & 0 & 0 \\ 0 & \left(1-\beta\right)^2 & 0 \\ 0 & 0 & \beta\left(1-\beta\right) \end{bmatrix}\hat{b}^k\right) \\ &+\max_{s\in S}\left(\begin{bmatrix} \beta\left(1-\beta\right) & 0 & 0 \\ 0 & \beta\left(1-\beta\right) & 0 \\ 0 & 0 & \left(1-\beta\right)^2 \end{bmatrix}\hat{b}^k\right). \end{aligned} \quad (23)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

# comb.	$c_{y=1}$	multiplicity	c_{other}	multiplicity	$c_{y=0}$	multiplicity
$\binom{m}{m}$	c_m	m	c_{m+1}	$\lvert S\rvert - m$	–	0
$\binom{m}{m-1}$	c_{m-1}	$m-1$	c_m	$\lvert S\rvert - m$	c_{m+1}	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\binom{m}{1}$	c_1	1	c_2	$\lvert S\rvert - m$	c_3	$m-1$
$\binom{m}{0}$	–	0	c_1	$\lvert S\rvert - m$	c_2	m

TABLE I: Analytical comparison of multipliers for m sensors. # comb. indicates how many different observations account for that arrangement of multiplier. $c_{y=1}$ is the multiplier for b_i getting a measurement of 1, and multiplicity is the number of such locations. The same applies for c_{other}—the unobserved locations—and $c_{y=0}$ for locations where 0 is measured.

# comb.	obs	$c_{y=1}$	multiplicity	c_{other}	multiplicity	$c_{y=0}$	multiplicity
1	$(1,1,1)$	$(1-\beta)\,\beta^2$	3	β^3	$\lvert S\rvert - m$	–	–
3	$(0,1,1);(1,0,1);(1,1,0)$	$(1-\beta)^2\,\beta$	2	$(1-\beta)\,\beta^2$	$\lvert S\rvert - m$	β^3	1
3	$(0,0,1);(0,1,0);(1,0,0)$	$(1-\beta)^3$	1	$(1-\beta)^2\,\beta$	$\lvert S\rvert - m$	$(1-\beta)\,\beta^2$	2
1	$(0,0,0)$	–	–	$(1-\beta)^3$	$\lvert S\rvert - m$	$(1-\beta)^2\,\beta$	3
$\binom{m}{n}$	$\{1\}^n\,\{0\}^{m-n}$	$(1-\beta)^{m-n+1}\,\beta^{n-1}$	n	$(1-\beta)^{m-n}\,\beta^n$	$\lvert S\rvert - m$	$(1-\beta)^{m-n-1}\,\beta^{n+1}$	$m-n$

TABLE II: Concrete example of multipliers for $m = 3$ sensors. # comb. indicates how many different observations account for that arrangement of multiplier, and *obs* is a list of those observations. $c_{y=1}$ is the multiplier for b_i getting a measurement of 1, and multiplicity is the number of such locations. The same applies for c_{other}—the unobserved locations—and $c_{y=0}$ for locations where 0 is measured.

This expression is simply (16) with the $\max$ operator in front of each term in the sum. Written with the P_{1i} and P_{0i} notation, this equation becomes

$$E_{y^k}\left[\max_{s\in S} b^k(s)\right] = \max_{s\in S}\left(P_{11}P_{12}\hat{b}^k\right) + \max_{s\in S}\left(P_{11}P_{02}\hat{b}^k\right) + \max_{s\in S}\left(P_{01}P_{12}\hat{b}^k\right) + \max_{s\in S}\left(P_{01}P_{02}\hat{b}^k\right) , \quad (24)$$

with each term in the sum corresponding to the measurement realizations $(1,1)$, $(1,0)$, $(0,1)$, and $(0,0)$, respectively. So it is straightforward to compute (22) for a given instantiation of this problem. Each sum over l_i accounts for $y_i^k = 0$ and $y_i^k = 1$, and the product is taken over the m agents, and therefore formula (22) accounts for all possible combinations of measurements.

It is worth noting that (8) is monotonic with respect to the number of sensors, regardless of their location. We formalize this notion with Theorem 1 below.

Theorem 1: For any q^k and $\{i|i \notin q^k\}$, (8) evaluated for $q^k \cup i$ is greater than or equal to (8) evaluated for q^k.

Proof: This theorem is best understood if we consider (8) in the form (22) for $m-1$ sensors

$$E_{y^k}\left[\max_{s\in S} b^k(s)\right] = \sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P\left(Y^k = y^k|q^k\right)\hat{b}^k . \quad (25)$$

Adding an m^{th} sensor changes the space of measurements from $\{0,1\}^{m-1}$ to $\{0,1\}^m$. The new measurements can be split into those in which the m^{th} sensor measures 1 and those in which it measures 0. Including this new sensor, we write (25) as

$$E_{y^k}\left[\max_{s\in S} b^k(s)\right] = \sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P_{1m}P\left(Y^k = y^k|q^k\right)\hat{b}^k + \sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P_{0m}P\left(Y^k = y^k|q^k\right)\hat{b}^k . \quad (26)$$

Recall that P_{1m} and P_{0m} sum to identity, so that any entry in the first term of (26) that is multiplied by β from P_{1m} is multiplied by $1-\beta$ by P_{0m} and vice-versa. Without loss of generality, assume that the original maximum from (25) was in the i^{th} entry—denoted $b(s_i)$. Then, depending on the location of the m^{th} sensor, if the i^{th} row of P_{1m} contains β we know that

$$\sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P_{1m}P\left(Y^k = y^k|q^k\right)\hat{b}^k \geq \beta b(s_i) \quad (27)$$

and

$$\sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P_{0m}P\left(Y^k = y^k|q^k\right)\hat{b}^k \geq (1-\beta)\, b(s_i) . \quad (28)$$

Otherwise, if the i^{th} row of P_{1m} contains $1-\beta$ then

$$\sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P_{1m}P\left(Y^k = y^k|q^k\right)\hat{b}^k \geq (1-\beta)\, b(s_i) \quad (29)$$

and

$$\sum_{y^k\in\{0,1\}^{m-1}} \max_{s\in S} P_{0m}P\left(Y^k = y^k|q^k\right)\hat{b}^k \geq \beta) b(s_i) . \quad (30)$$

Substituting either set of equations into (26) results in

$$E_{y^k}\left[\max_{s\in S} b^k(s)\right] \geq \beta(s_i) + (1-\beta)\, b(s_i) \quad (31)$$

$$\geq b(s_i) . \quad (32)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

■

Each $\max$ term in the sum in (22) is taking the maximum of a combination of $c_i b_j$ for $i \in \{1, \ldots, m+1\}$ and $j \in \{1, \ldots, |S|\}$. Then, when we consider what entry is the maximum, we can reason using only the indices i and j to draw some conclusions. Simply put, comparing $c_i b_j$ to $c_k b_l$, if $i \leq k$ and $j \leq l$, we know that $c_i b_j \geq c_k b_l$, since we assumed the indices on each c and b are ordered by their magnitude. But in the case where $i \geq k$ and $j \leq l$, or vice versa, our power to draw conclusions about the outcome is weakened, and we can only limit the set of possible maxima to some subset of entries in b^k. Despite the fact that both $\langle b, \geq \rangle$ and $\langle \mathbf{C}, \geq \rangle$ are *totally ordered sets* (see Def. 2), this ambiguity arises because of the nature of their product. The terms in (22) are from the product of b and $\mathbf{C}$, so we must determine the nature of to Cartesian product of these two sets, which form a *lattice* (Def. 4).

Before introducing and proving our main theorem (Theorem 2), we will now introduce two lemmas that aid in its proof. For these results we will use the following notation. Among those states $s \in S \setminus q^k$, i.e., locations without a sensor at time k, let o denote the location with the highest probability b_o.

Lemma 1: For any arrangement of m sensors, it is suboptimal to place any sensors beyond b_{m+1}.

Proof: From Prop. 1, each measurement outcome results in three distinct multipliers from $\mathbf{C}$ that multiply the entries in b: $c_{y=1}$ for those entries $s \in q^k$ with $y^k = 1$, $c_{y=0}$ for entries $s \in q^k$ with $y^k = 0$, and c_{other} for each of the entries where no measurement is taken. Our objective function (22) sums over each of the 2^m possible outcomes, taking the maximum of $\{c_{y=1} b_i | y_i^k = 1\}$, $\{c_{y=0} b_j | y_j^k = 0\}$, and $\{c_{other} b_l | l \notin q^k\}$ for each of those measurements.

Assume that two entries in b, i and j, with $i < j < m+1$, are not in the set q^k, such that

$$b_i > b_j \; . \tag{33}$$

Then

$$q^k = \{1, 2, \ldots, i-1, i+1, \ldots, j-1, j+1, \ldots, m+1, m+2\} \; , \tag{34}$$

the set $1, \ldots, m+2$, excluding i and j. Since there is no sensor at i or j, both b_i and b_j will be multiplied by c_{other} for all measurement realizations. Following from (33) we know that

$$c_{other} b_i > c_{other} b_j \; , \tag{35}$$

so the sensor at location j never contributes to the sum in (22).

The sensor at $m+2$ is multiplied by $c_{y=1}$ or $c_{y=0}$, depending on the measurement realization. Moving this sensor from location $m+2$ to location j can only improve the sum in (22), because $c_{y=1} b_j > c_{y=1} b_{m+2}$ and $c_{y=0} b_j > c_{y=0} b_{m+2}$. ■

Lemma 2: For a partially ordered set Σ forming a fence, the two maximum elements may be any two elements $\sigma_i \in \Sigma$ such that $\sigma_{i-1} < \sigma_i$ and $\sigma_i > \sigma_{i+1}$ or one such σ_i and its adjacent endpoint.

Proof: We first begin by ruling out any $\sigma_i \in \Sigma$ such that $\sigma_{i-1} > \sigma_i$ and $\sigma_i < \sigma_{i+1}$. Clearly, such an element is smaller than both σ_{i-1} and σ_{i+1} and so cannot be among the two maximum elements, as shown in Fig. 1a. This leaves the set σ_{i-k} and σ_{i+k}, where k is odd, as well as the endpoints as candidate maximum elements. If endpoint $\sigma_1 > \sigma_2$, then it belongs to the set σ_{i-k}, and is a candidate maximum element. Likewise, if $\sigma_{n-1} < \sigma_n$, then σ_n belongs to the set σ_{i+k}. In the event that either of these endpoints is not greater than its single neighbor, it is possible that either $\{\sigma_1, \sigma_2\}$ (Fig. 1b) or $\{\sigma_{n-1}, \sigma_n\}$ (Fig. 1c) are the two maximum elements. ■

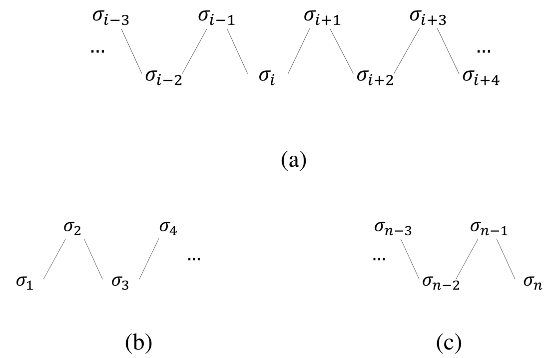


Fig. 1: Portions of fences. The middle portion of a fence is shown in Fig. 1a, demonstrating that σ_i cannot be among the two maximum elements. Endpoint σ_1 in Fig. 1b may be among the two maximum elements. Endpoint σ_n in Fig. 1c may also be among the two maximum elements.

We are now ready to state and prove our main theorem.

Theorem 2: For $m \geq 2$ identical, non-overlapping sensors with $\alpha = \beta$, the one-step optimal strategy with respect to (8) is to place sensors at $q^k = \{1, 2, \ldots, m\}$.

Proof: From Lemma 1, it is clear that the m sensors should be placed among the first $m+1$ entries in b. Viewed differently, there is a choice of which of the first $m+1$ locations should *not* be searched. We will demonstrate that $o = i+1$ results in higher expectation than $o = i$ for any choice of i.

For any measurement realization, $c_{y=1} \geq c_{other}$ from Prop. 1. If there is an index $j \in \{1, \ldots, i-1\}$ such that $y_j^k = 1$, it will contribute to the objective function whether $o = i$ or $o = i+1$, because

$$c_{y=1} b_j \geq c_{other} b_o . \tag{36}$$

But if all such $y_j^k = 0$, we must determine whether the expectation is higher for $o = i$ or $o = i+1$, since $c_{other} \geq c_{y=0}$ but $b_j \geq b_o$.

When considering switching o from $i+1$ to i, we must determine which outcomes are affected by this change. These measurement outcomes are those in which b_i potentially contributes to (22). We know that all sensors at indices prior to i must return a measurement of 0, but we must consider both possible outcomes for i: $y_i^k = 1$ and $y_i^k = 0$. For the other sensors $l \in \{i+2, \ldots, m\}$, let p denote the number such that $y_l^k = 1$. Then if $y_i^k = 1$ the contribution from location i is $c_{p+1} b_i$, using the notation from Table I for $p+1$ measurements of 1. The contribution from $o = i+1$ is $c_{p+2} b_{i+1}$, and finally the contribution from those sensors measuring 0 is $c_{p+3} b_1$,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

since b_1 is the biggest. We get one term in our objective function from these three items, namely

$$\max\{c_{p+1}b_i, c_{p+2}b_{i+1}, c_{p+3}b_1\} \ . \tag{37}$$

For $y_i^k = 0$, the analogous term is

$$\max\{c_p b_l, c_{p+1}b_{i+1}, c_{p+2}b_1\} \ , \tag{38}$$

where l is the index of some location measuring 1.

Changing o from $i+1$ to i changes these terms to

$$\max\{c_{p+1}b_{i+1}, c_{p+2}b_i, c_{p+3}b_1\} \tag{39}$$

and

$$\max\{c_p b_l, c_{p+1}b_i, c_{p+2}b_1\} \ , \tag{40}$$

respectively.

We note that there are four coefficients from $\mathbf{C}$ appearing in these $\max$ terms: $c_p, c_{p+1}, c_{p+2}, c_{p+3}$, which form a total order with $\geq$. Likewise, there are four elements of b in these $\max$ terms: b_1, b_i, b_{i+1}, b_l, also forming a total order with $\geq$. The ambiguity of the sensor outcomes is reflected in the fact that the Cartesian product of these elements forms a lattice, as shown in Fig. 2. Embedded in this lattice are seven points which are candidates for the terms that contribute to this sum, as displayed in Fig. 2 in bold. Since

$$c_{p+1}b_i \geq c_{p+2}b_{i+1} \tag{41}$$

in (37), we can ignore the point $c_{p+2}b_{i+1}$. The choice of $o = i+1$ vs. $o = i$ determines which subset of the remaining six points contributes to our sum. Figure 3 color codes how each choice of o might contribute, with Fig. 3a showing the results for $o = i+1$ and Fig. 3b showing the results for $o = i$. For each scenario, the largest red point and the largest green point are added to our sum.

Lemma 2 helps us see that of the five points making a fence, there are three possible sets that might be the two maximum points: $\{c_{p+3}b_1, c_{p+2}b_1\}$, $\{c_{p+1}b_i, c_{p+1}b_{i+1}\}$, and $\{c_{p+2}b_1, c_{p+1}b_i\}$. The first two results are an endpoint and one maximal element, while the third is the two maximal elements. Figure 3a, corresponding to $o = i+1$, shows that any of these three possibilities is realizable, since each involves one red and one green point. For $o = i$, however, the two maximal points are not realizable, meaning that if they were to have the highest sum, they could not both contribute to the objective function, as shown in Fig. 3b. Therefore in the event that these two points have the highest value, $o = i+1$ is a better choice than $o = i$. These results are summarized in the first three rows of Table. III.

Finally, we consider if $c_p b_l$, which is not part of the fence, is among the two highest valued points. In that case either $\{c_{p+2}b_1, c_p b_l\}$ or $\{c_{p+1}b_i, c_p b_l\}$ could be the two maximum points. In that case, Fig. 3b shows that the remaining choices (in red) to add to the objective function are each dominated by another unrealizable option (in green). For $o = i+1$, as shown in Fig. 3a, there is a red option that is not dominated by another point $c_{p+1}b_i$. The only other realizable result is $c_{p+3}b_1$, which is realizable for both $o = i+1$ and $o = i$. Therefore, again $o = i+1$ results in a higher valued objective

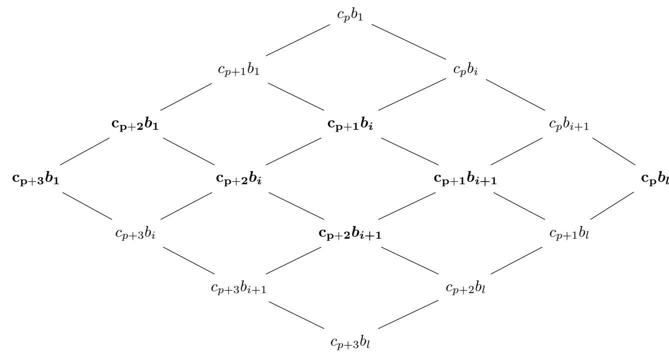


Fig. 2: Lattice of outcomes switching from $o = i+1$ to $o = i$. Outcomes appearing in (37)-(40) are indicated in bold. Lines between pairs of outcomes indicate the ordering relationship, with the upper member of each pair having a greater value. The maximum element in this lattice is $c_p b_1$ and the minium element is $c_{p+3}b_l$.

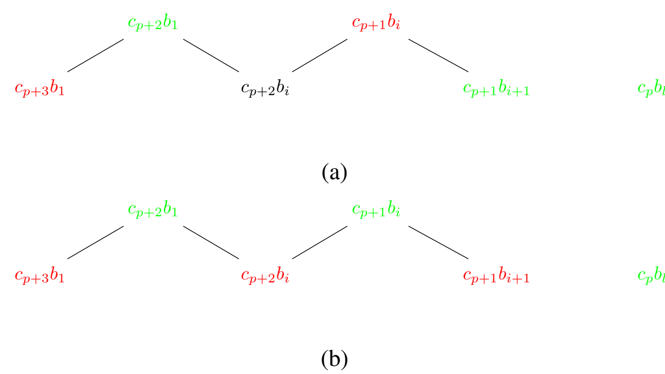


Fig. 3: Relevant outcomes for $b_o = i+1$ (a) with outcomes from (37) in red, while those from (38) are in green, and relevant outcomes for $b_o = i$ (b) with outcomes from (39) displayed in red, and those from (40) in green. The highest valued red and green point in each scenario contribute to the objective function.

function than $o = i$. These result are summarized in the last two rows of Table. III.

In either case (whether $c_p b_l$ is in the maximum outcome or not), Table III shows that $o = i+1$ permits all of the same outcomes to contribute to (8) as $o = i$, but not the other way around, implying that $o = i$ cannot be a better choice than $o = i+1$.

We have now shown that the optimal one-step strategy is to put m sensors in the first m locations. Further, we have shown that the result cannot improve, but can only get worse by changing the choice of locations without a sensor. ■

Theorem 2 shows the optimal strategy for any number of sensors in expectation. The expectation cannot increase with any other choice of sensor locations. We present, however, a special case for one sensor in which there is a second optimal choice.

Proposition 2: If $m = 1$, choosing either $q^k = 1$ or $q^k = 2$ maximizes (8).

Proof: This result follows directly from Proposition 8 in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

Max. points	Realizable $o = i+1$	Realizable $o = i$
$\{c_{p+2}b_1, c_{p+3}b_1\}$	Y	Y
$\{c_{p+1}b_i, c_{p+1}b_{i+1}\}$	Y	Y
$\{\boldsymbol{c_{p+2}b_1, c_{p+1}b_i}\}$	**Y**	**N**
$\{c_{p+2}b_l, c_p b_l\}$	N	N
$\{\boldsymbol{c_{p+1}b_i, c_p b_l}\}$	**Y**	**N**

TABLE III: Possible pairs of maximal points from (37)-(40) (See also Fig. 3). The left column lists the pair of maximum points, and the next two columns list whether those pairs are realizable with $o = i+1$ and $o = i$ respectively. Any outcome realizable at i is realizable at $i+1$, but not vice-versa. Outcomes that are different between the two choices are shown in bold.

[1]. Although that work does not consider a moving target, (8) is a one-step maximization. This means that once the belief has been updated to account for the target's motion, the target can be thought of as remaining stationary for that time step. If our horizon were longer than one time step, this proposition would not hold. ∎

IV. Policies with Lower Variance

Having two choices that lead to a one-step optimal result with respect to (8) leads to the natural question of whether there is a principled manner to distinguish between the two. This is resolved in Theorem 3 below.

Theorem 3: For $m = 1$, among policies that maximize $\mathbf{E}_{y^k}\left[\max_{s\in S} b^k(s)\right]$, the policy $q^k = 2$ minimizes the variance $var\left(\max_{s\in S} b^k(s)\right)$.

Proof: Choosing $q^k = 2$ has the same expectation as $q^k = 1$ as shown in Corollary 2, and both have higher expectation than any other choice for q^k. Further, the variance of the estimate of the target location for $q^k = 2$ is always less than or equal to the variance when $q^k = 1$. To prove this, we write the variance

$$var\left(\max_{s\in S} b^k(s)\right) = \mathbb{E}\left[\left(\max_{s\in S} b^k(s)\right)^2\right] - \left(\mathbb{E}\left[\max_{s\in S} b^k(s)\right]\right)^2 . \tag{42}$$

We have already computed $\mathbb{E}\left[\max b^k(s)\right]$, and hence we know the second term in the variance equation, which is the same for both strategies. What remains is to compute the first term,

$$\begin{aligned}\mathbb{E}&\left[\left(\max b^k(s)\right)^2\right]\\ &= P\left(Y^k = 1\right)\max_{s\in S}\left(\eta P\left(Y^k = 1|q^k, s\right)\hat{b}^k\right)^2 \\ &\quad + P\left(Y^k = 0\right)\max_{s\in S}\left(\eta P\left(Y^k = 0|q^k, s\right)\hat{b}^k\right)^2 .\end{aligned} \tag{43}$$

By noticing that $\frac{1}{\eta}$ is $P\left(Y^k = y\right)$, we can write

$$\begin{aligned}\mathbb{E}&\left[\left(\max b^k(s)\right)^2\right]\\ &= \frac{1}{P\left(Y^k = 1\right)}\max_{s\in S}\left(P\left(Y^k = 1|q^k, s\right)\hat{b}^k\right)^2 \\ &\quad + \frac{1}{P\left(Y^k = 0\right)}\max_{s\in S}\left(P\left(Y^k = 0|q^k, s\right)\hat{b}^k\right)^2 .\end{aligned} \tag{44}$$

Then, since $P\left(Y^k = 1\right) = (1-\beta)\hat{b}^k\left(q^k\right) + \beta\left(1 - \hat{b}^k\left(q^k\right)\right)$ and $P\left(Y^k = 0\right) = \beta\hat{b}^k\left(q^k\right) + (1-\beta)\left(1 - \hat{b}^k\left(q^k\right)\right)$ we can evaluate the variance for a given choice of q^k.

There are two cases for which we need to prove the theorem: $(1-\beta)b_2 > \beta b_1$ and $\beta b_1 \geq (1-\beta)b_2$.

Case 1: $(1-\beta)b_2 > \beta b_1$

For strategy $q^k = 1$, (44) evaluates to

$$\begin{aligned}\mathbb{E}&\left[\left(\max b^k(s)\right)^2 |q^k = 1\right] = \\ &\frac{(1-\beta)^2 b_1^2}{(1-\beta)b_1 + \beta(1-b_1)} + \frac{(1-\beta)^2 b_2^2}{\beta b_1 + (1-\beta)(1-b_1)} ,\end{aligned} \tag{45}$$

while for strategy $q^k = 2$, it evaluates to

$$\begin{aligned}\mathbb{E}&\left[\left(\max b^k(s)\right)^2 |q^k = 2\right] = \\ &\frac{(1-\beta)^2 b_2^2}{(1-\beta)b_2 + \beta(1-b_2)} + \frac{(1-\beta)^2 b_1^2}{\beta b_2 + (1-\beta)(1-b_2)} .\end{aligned} \tag{46}$$

Case 2: $\beta b_1 \geq (1-\beta)b_2$

For strategy $q^k = 1$, (44) evaluates to

$$\begin{aligned}\mathbb{E}&\left[\left(\max b^k(s)\right)^2 |q^k = 1\right] = \\ &\frac{(1-\beta)^2 b_1^2}{(1-\beta)b_1 + \beta(1-b_1)} + \frac{\beta^2 b_1^2}{\beta b_1 + (1-\beta)(1-b_1)} ,\end{aligned} \tag{47}$$

while for strategy $q^k = 2$, it evaluates to

$$\begin{aligned}\mathbb{E}&\left[\left(\max b^k(s)\right)^2 |q^k = 2\right] = \\ &\frac{\beta^2 b_1^2}{(1-\beta)b_2 + \beta(1-b_2)} + \frac{(1-\beta)^2 b_1^2}{\beta b_2 + (1-\beta)(1-b_2)} .\end{aligned} \tag{48}$$

The strategy that minimizes variance can be found by determining which is smaller (45) or (46) for Case 1, and (47) or (48) for Case 2. For simplicity of notation, we will replace $(1-\beta)b_1 + \beta(1-b_1)$ with P_{b_1} and $(1-\beta)b_2 + \beta(1-b_2)$ with P_{b_2}.

For Case 1, we will demonstrate that choosing $q^k = 2$ is the optimal choice by showing that

$$\frac{b_1^2}{P_{b_1}} + \frac{b_2^2}{1-P_{b_1}} \geq \frac{b_2^2}{P_{b_2}} + \frac{b_1^2}{1-P_{b_2}} . \tag{49}$$

Denote $\frac{b_2}{b_1}$ with γ, so $0 \leq \gamma \leq 1$ covers all valid b_1, b_2 pairs. The domain of valid b_1, γ pairs is shown in Fig. 4. We will begin by demonstrating that

$$P_{b_2}\left(1 - P_{b_1}\right) - \gamma^2 P_{b_1}\left(1 - P_{b_2}\right) \geq 0 , \tag{50}$$

which is a simple rearrangement of the inequality in (49). Expanding this inequality and replacing b_2 with γb_1 yields

$$\begin{aligned}&b_1\gamma(1-2\beta) + \beta - \left(b_1(1-2\beta) + \beta\right)\left(b_1\gamma(1-2\beta) + \beta\right)\\ &- \gamma^2 b_1(1-2\beta) + \beta - \left(b_1(1-2\beta) + \beta\right)\left(b_1\gamma(1-2\beta) + \beta\right)\\ &\geq 0 .\end{aligned} \tag{51}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

We will call the left-hand expression of (51) $f(b_1, \gamma)$. The domain of this function is displayed in Fig. 4. Then, the inequality in (51) will be demonstrated by showing that $f(b_1, \gamma)$ is non-negative along the boundaries of its domain and concave in b_1 thus proving its overall non-negativity, which proves the inequality in (49).

Starting with the boundary for $b_1 = b_2$, (i.e. $\gamma = 1$), clearly, $f(b_1, \gamma)$ evaluates to 0. This further implies that the variance for choosing b_1 and b_2 is the same when $b_1 = b_2$.

We also wish to check the behavior at the other boundaries. Next, we examine the boundary $b_1 + b_2 = 1$ for $0.5 \leq b_1 \leq 1$. Then, $b_2 = 1 - b_1$, which means $0 \leq b_2 \leq 0.5$. We note that for $b_1 \geq 0.5$, $b_1 \geq P_{b_1}$. Thus we may write

$$b_1^2 \left(1 - 2P_{b_1}\right) \geq P_{b_1}^2 \left(1 - 2b_1\right) \ , \tag{52}$$

which we rearrange as follows:

$$b_1^2 \left(1 - 2P_{b_1}\right) + b_1^2 P_{b_1}^2 \geq P_{b_1}^2 \left(1 - 2b_1\right) + b_1^2 P_{b_1}^2 \tag{53}$$

$$b_1^2 \left(1 - P_{b_1}\right)^2 \geq P_{b_1}^2 \left(1 - b_1\right)^2 \tag{54}$$

$$\frac{\left(1 - P_{b_1}\right)^2}{P_{b_1}^2} \geq \frac{\left(1 - b_1\right)^2}{b_1^2} \ . \tag{55}$$

By noting that $1 - b_2 = b_1$ implies $P_{b_2} = 1 - P_{b_1}$, we may then further rearrange the expression to

$$\frac{P_{b_2}\left(1 - P_{b_1}\right)}{P_{b_1}\left(1 - P_{b_2}\right)} - \frac{b_2^2}{b_1^2} \geq 0 \, , \tag{56}$$

and finally

$$P_{b_2}\left(1 - P_{b_1}\right) - \frac{b_2^2}{b_1^2} P_{b_1}\left(1 - P_{b_2}\right) \geq 0 \, , \tag{57}$$

which is (50). Hence we have demonstrated the non-negativity of $f(b_1, \gamma)$ along the boundary $b_1 + b_2 = 1$.

Finally, we must check the behavior along the boundary where $b_2 = \frac{1-b_1}{|S|-1}$. We observe the following relationships:

$$\left(b_1^2 - b_2^2\right) \geq 0 \tag{58}$$

$$P_{b_2}\left(1 - P_{b_2}\right) \geq 0 \, , \tag{59}$$

which therefore imply

$$\left(b_1^2 - b_2^2\right) P_{b_2}\left(1 - P_{b_2}\right) \geq 0 \, . \tag{60}$$

We also note that since $\beta \leq 0.5$, then $1 - 2\beta \geq 0$. Then we may also write that

$$b_1^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \geq 0 \tag{61}$$

and

$$b_2^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \geq 0 \, , \tag{62}$$

since each of the terms in these two expressions is non-negative. The fact that $1 - |S| b_2$ is non-negative follows from the fact that this boundary represents a regime in which b_2 is no bigger than $\frac{1}{|S|}$. The sum is of the previous in equalities is similarly non-negative:

$$b_1^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) + b_2^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \geq 0 \, , \tag{63}$$

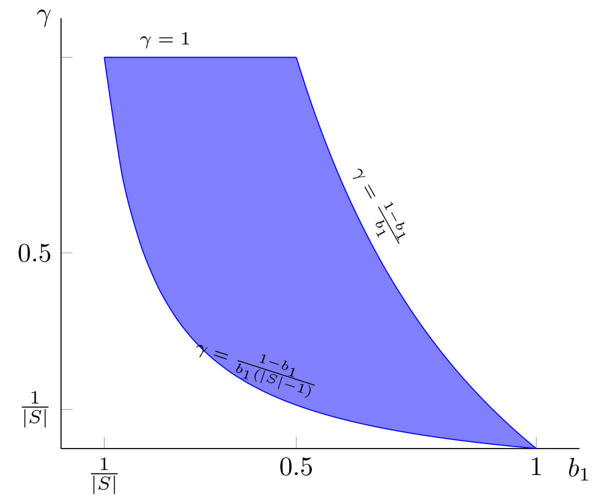


Fig. 4: Colored region marks the b_1, γ pairs constituting the domain of $f(b_1, \gamma)$. Corresponding constraints on b_1 and γ are labeled for the borders of the region.

and negating that expression flips the inequality to

$$-b_1^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) - b_2^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \leq 0 \, . \tag{64}$$

Combining (60) and (64), we find

$$\begin{aligned}\left(b_1^2 - b_2^2\right) P_{b_2}\left(1 - P_{b_2}\right) \\ \geq -b_1^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \\ - b_2^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \ .\end{aligned} \tag{65}$$

Now we rearrange this expression as follows:

$$\begin{aligned} & b_1^2 P_{b_2}\left(1 - P_{b_2}\right) + b_1^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \\ & \geq b_2^2 P_{b_2}\left(1 - P_{b_2}\right) - b_2^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \end{aligned} \tag{66}$$

$$\begin{aligned} & b_1^2 P_{b_2}\left(1 - P_{b_2} + \left(1 - |S| b_2\right)\left(1 - 2\beta\right)\right) \\ & \geq b_2^2\left(P_{b_2} - \left(1 - |S| b_2\right)\left(1 - 2\beta\right)\right)\left(1 - P_{b_2}\right) \end{aligned} \tag{67}$$

$$\frac{P_{b_2}\left(1 - P_{b_2} + \left(1 - |S| b_2\right)\left(1 - 2\beta\right)\right)}{\left(P_{b_2} - \left(1 - |S| b_2\right)\left(1 - 2\beta\right)\right)\left(1 - P_{b_2}\right)} - \frac{b_2^2}{b_1^2} \geq 0 \, . \tag{68}$$

By noting that $b_2 = \frac{1-b_1}{|S|-1}$ implies $P_{b_1} = P_{b_2} - \left(1 - |S| b_2\right)\left(1 - 2\beta\right)$ and $b_1 = 1 - \left(|S| - 1\right) b_2$, we finally write (68) as

$$\frac{P_{b_2}\left(1 - P_{b_1}\right)}{P_{b_1}\left(1 - P_{b_2}\right)} - \frac{b_2^2}{b_1^2} \geq 0 \tag{69}$$

and then

$$P_{b_2}\left(1 - P_{b_1}\right) - \frac{b_2^2}{b_1^2} P_{b_1}\left(1 - P_{b_2}\right) \geq 0 \tag{70}$$

which is again (50). We note that the term $P_{b_1}\left(1 - P_{b_2}\right)$ is always positive, hence the direction of the inequality in (68) is always correct. This demonstrates the non-negativity of $f(b_1, \gamma)$ along the boundary $b_2 = \frac{1-b_1}{|S|-1}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

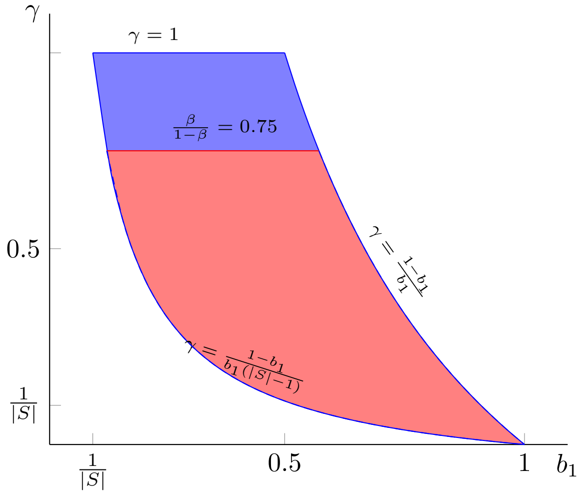


Fig. 5: Red region marks the b_1, γ pairs constituting the domain of $g\left(b_1, \gamma, \beta\right)$ for a given value of β. We include the boundary for $\beta = 0.75$ as an example. Corresponding constraints on b_1 and γ are labeled for the borders of the region.

To complete our proof, we show the concavitiy of $f\left(b_1, \gamma\right)$ with respect to b_1. Taking the second derivative of $f\left(b_1, \gamma\right)$ yields

$$\frac{\partial^2 f\left(b_1, \gamma\right)}{\partial b_1^2} = 2\gamma\left(\gamma^2 - 1\right)\left(1 - 2\beta\right)^2 . \tag{71}$$

It is clear that the term $\left(1 - 2\beta\right)^2$ and 2γ are non-negative. Since $0 \leq \gamma \leq 1$, then $\left(\gamma^2 - 1\right) \leq 0$. Hence, the second derivative of our function is non-positive and therefore the function is concave with respect to b_1.

Now, we have demonstrated the function is equal to zero for $b_1 = b_2$, and non-negative along the rest of the boundaries. Further, by demonstrating that the function is concave in b_1, we know the values of the function between these boundaries must also be non-negative. Hence, the function is greater than or equal to zero for the entire feasible set of values of b_1, b_2, and β.

For Case 2, to show that the optimal choice of q^k is 2, we must show that

$$\frac{\left(1-\beta\right)^2 b_1^2}{P_{b_1}} + \frac{\beta^2 b_1^2}{1 - P_{b_1}} \geq \frac{\beta^2 b_1^2}{P_{b_2}} + \frac{\left(1-\beta\right)^2 b_1^2}{1 - P_{b_2}} . \tag{72}$$

We will follow the same process that was followed in Case 1 to prove this inequality. We begin by rearranging the inequality to

$$\frac{\left(1-\beta\right)^2}{P_{b_1}} + \frac{\beta^2}{1 - P_{b_1}} \geq \frac{\beta^2}{P_{b_2}} + \frac{\left(1-\beta\right)^2}{1 - P_{b_2}} . \tag{73}$$

Now we make common denominators:

$$\begin{aligned} &\frac{\left(1 - P_{b_2}\right)\left(1-\beta\right)^2}{P_{b_1}\left(1 - P_{b_2}\right)} + \frac{P_{b_2}\beta^2}{P_{b_2}\left(1 - P_{b_1}\right)} \\ &\qquad \geq \frac{\left(1 - P_{b_1}\right)\beta^2}{P_{b_2}\left(1 - P_{b_1}\right)} + \frac{P_{b_1}\left(1-\beta\right)^2}{P_{b_1}\left(1 - P_{b_2}\right)} . \end{aligned} \tag{74}$$

We can group terms to get

$$\frac{\left(1 - P_{b_1} - P_{b_2}\right)\left(1-\beta\right)^2}{P_{b_1}\left(1 - P_{b_2}\right)} \geq \frac{\left(1 - P_{b_1} - P_{b_2}\right)\beta^2}{P_{b_2}\left(1 - P_{b_1}\right)} . \tag{75}$$

For the boundary where $1 = b_1 + b_2$, we note that

$$1 = P_{b_1} + P_{b_2} , \tag{76}$$

which implies that both sides of (75) evaluate to 0. To evaluate along the other two boundaries, we note that elsewhere, $1 - P_{b_1} - P_{b_2} > 0$ and we continue simplifying to get

$$\frac{P_{b_2}\left(1 - P_{b_1}\right)}{P_{b_1}\left(1 - P_{b_2}\right)} \geq \frac{\beta^2}{\left(1-\beta\right)^2} , \tag{77}$$

and finally rearranging to get

$$P_{b_2}\left(1 - P_{b_1}\right) - \frac{\beta^2}{\left(1-\beta\right)^2} P_{b_1}\left(1 - P_{b_2}\right) \geq 0 . \tag{78}$$

We will call the left-hand side of (78) $g\left(b_1, \gamma, \beta\right)$. We will demonstrate that $g\left(b_1, \gamma, \beta\right)$ is non-negative for the remaining two boundaries, thus proving the inequality in (72).

We note that we have shown the non-negativity of $f\left(b_1, \gamma\right)$ over its entire domain in Case 1. Then the boundary of $g\left(b_1, \gamma, \beta\right)$ where $\frac{\beta}{1-\beta} = \gamma$ is $f\left(b_1, \gamma\right)$ with $\beta = \frac{\gamma}{1+\gamma}$. This boundary has therefore already been shown to be non-negative, since it is contained in Case 1.

Therefore, we must only check the behavior along the boundary where $b_2 = \frac{1-b_1}{|S|-1}$. As in Case 1, we observe:

$$\left(\left(1-\beta\right)^2 - \beta^2\right) \geq 0 \tag{79}$$

$$P_{b_2}\left(1 - P_{b_2}\right) \geq 0 , \tag{80}$$

which therefore imply

$$\left(\left(1-\beta\right)^2 - \beta^2\right) P_{b_2}\left(1 - P_{b_2}\right) \geq 0 . \tag{81}$$

We also note that since $\beta \leq 0.5$, then $1 - 2\beta \geq 0$. Then we may also write that

$$\left(1-\beta\right)^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \geq 0 \tag{82}$$

and

$$\beta^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \geq 0 , \tag{83}$$

since each of the terms in these two expressions is non-negative. As before, the fact that $1 - |S| b_2$ is non-negative follows from the fact that this boundary represents a regime in which b_2 is no bigger than $\frac{1}{|S|}$. The sum is of the previous in equalities is similarly non-negative:

$$\left(1-\beta\right)^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) + \beta^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \geq 0 , \tag{84}$$

and negating that expression flips the inequality to

$$-\left(1-\beta\right)^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) - \beta^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \leq 0 . \tag{85}$$

Combining (81) and (85), we find

$$\begin{aligned} &\left(\left(1-\beta\right)^2 - \beta^2\right) P_{b_2}\left(1 - P_{b_2}\right) \\ &\qquad \geq -\left(1-\beta\right)^2 P_{b_2}\left(1 - |S| b_2\right)\left(1 - 2\beta\right) \\ &\qquad\quad - \beta^2\left(1 - P_{b_2}\right)\left(1 - |S| b_2\right)\left(1 - 2\beta\right) . \end{aligned} \tag{86}$$

Now we rearrange this expression as follows:

$$(1-\beta)^2 P_{b_2}(1-P_{b_2}) + (1-\beta)^2 P_{b_2}(1-|S|b_2)(1-2\beta) \geq \beta^2 P_{b_2}(1-P_{b_2}) - \beta^2(1-P_{b_2})(1-|S|b_2)(1-2\beta) \quad (87)$$

$$(1-\beta)^2 P_{b_2}(1-P_{b_2}+(1-|S|b_2)(1-2\beta)) \geq \beta^2 (P_{b_2}-(1-|S|b_2)(1-2\beta))(1-P_{b_2}) \quad (88)$$

$$\frac{P_{b_2}(1-P_{b_2}+(1-|S|b_2)(1-2\beta))}{(P_{b_2}-(1-|S|b_2)(1-2\beta))(1-P_{b_2})} - \frac{\beta^2}{(1-\beta)^2} \geq 0\,. \quad (89)$$

By noting that $b_2 = \frac{1-b_1}{|S|-1}$ implies $P_{b_1} = P_{b_2} - (1-|S|b_2)(1-2\beta)$ and $b_1 = 1-(|S|-1)\,b_2$, we finally write (89) as

$$\frac{P_{b_2}(1-P_{b_1})}{P_{b_1}(1-P_{b_2})} - \frac{\beta^2}{(1-\beta)^2} \geq 0 \quad (90)$$

and then

$$P_{b_2}(1-P_{b_1}) - \frac{\beta^2}{(1-\beta)^2} P_{b_1}(1-P_{b_2}) \geq 0 \quad (91)$$

which is again (78). We note that the term $P_{b_1}(1-P_{b_2})$ is always positive, hence the direction of the inequality in (89) is always correct. This demonstrates the non-negativity of $g(b_1,\gamma,\beta)$ along the boundary $b_2 = \frac{1-b_1}{|S|-1}$.

Finally, we compute the second derivative of $g(b_1,\gamma,\beta)$

$$\frac{\partial^2 g(b_1,\gamma,\beta)}{\partial b_1^2} = \frac{2\gamma(2\beta-1)^3}{(1-\beta)^2}\,. \quad (92)$$

We know that $\gamma \geq 0$ and $(1-\beta)^2 > 0$. Further, $(2\beta-1) < 0$, since $\beta < 0.5$, hence (92) is negative, and $g(b_1,\gamma,\beta)$ is concave in b_1. We have then demonstrated that $g(b_1,\gamma,\beta)$ is non-negative along its boundaries and concave in b_1, and is therefore non-negative over its entire domain.

The proofs for Case 1 and Case 2 together show that choosing $q^k = 2$ always minimizes $var\left(\max_{s\in S} b^k(s)\right)$ for a one-step optimization using a single sensor.

■

Intuitively, Theorem 3 holds because of the likelihood of *not* detecting the target. Each missed detection at a location $i > 1$ reinforces the belief that location 1 is the true location of the target. On the other hand, a detection (true or false) at location $i > 1$ results in a new higher belief at location i. Either way, the belief maintains a "peak" at location 1 or location i. If location 1 is searched instead, a missed detection results in the belief for location 1 being redistributed among the other locations, rather than having the bulk of its mass move to location i. Thus the strategy of search location 1 can result in a transition from a high-confidence state to a nearly uniform confidence state with one missed detection—resulting in higher variance.

The result of Theorem 3 leads us to make Conjecture 1 below.

Conjecture 1: For $m \geq 2$ identical, non-overlapping sensors with $\alpha = \beta$, placing sensors at $q^k = \{2, 3\ldots, m+1\}$ results in lower variance $var\left(\max_{s\in S} b^k(s)\right)$ than placing sensors at $q^k = \{1, 2, \ldots, m\}$.

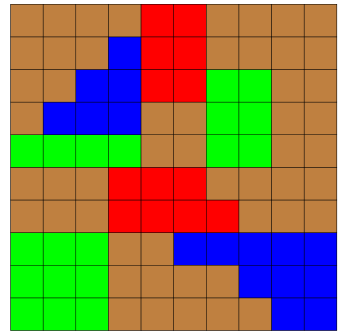


Fig. 6: Environment used in the case study. Green indicates grassland, blue indicates water sources, red indicates predator habitat, and brown indicates plains.

Although we do not prove it, Conjecture 1 suggests that there is a trade-off when choosing a surveillance strategy. The greedy strategy ($q^k = \{1, 2, \ldots, m\}$) maximizes the MAP belief, but results in an estimate with higher variance. This result is supported in Sec. V below.

V. Simulation Results

To demonstrate our results from Secs. III and IV, we simulate the wildlife tracking problem introduced in Sec. I. The environment we consider is a discrete 10×10 grid, consisting of four types of cells: grassland, water sources, predator habitat, and plains. The probability of the animal staying in each of these types of cells is 0.9, 0.75, 0.1 and 0.4, respectively. The total probability of moving to a neighboring cell is therefore 0.1, 0.25, 0.9, and 0.6, respectively. Thus, the animal prefers to hide in grassland and spend time near water, while spending less time in open areas and predator territory. If the animal does not stay in its current cell, it is equally likely to move to any of its neighboring cells. The value of α and β was set to 0.05 for each simulation.

A. *One Step Results*

To confirm our theoretical results, we first present results for one time step. We randomly sampled the belief space and true target location, and ran each strategy repeatedly to compute the mean and variance of the belief at the next step. Figure 7 presents multi-agent results for $m = 3$ agents for six-randomly-sampled belief states, with $10,000$ trials for each strategy at each belief state comparing $q^k = \{1, 2, 3\}$ and $q^k = \{2, 3, 4\}$. From the figure it is clear that the mean performance is higher for the greedy strategy, but the variance also tends to be higher for the greedy strategy. This figure, along with the results in Sec. V-B, confirms the results of Theorem 2 and supports Conjecture 1. Figure 8 presents results for the single agent case for the strategies $q^k = 1$ and $q^k = 2$. From the figure, it is clear that the means are the same, but the variance for $q^k = 2$ is always smaller than or equal to the variance for $q^k = 1$, confirming Theorem 3.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

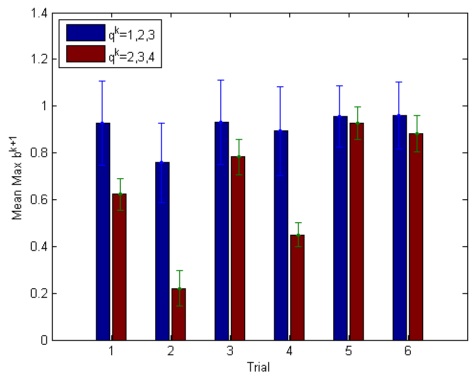


Fig. 7: Comparison of multi-agent strategies $q^k = \{1,2,3\}$ and $q^k = \{2,3,4\}$ for randomly sampled regions of the belief space. For each of the six trials, the strategies were repeated on the same randomly-sampled belief and true target location $10,000$ times. The strategy $q^k = \{1,2,3\}$ performs better on average but the variance for $q^k = \{2,3,4\}$ never exceeds that of $q^k = \{1,2,3\}$.

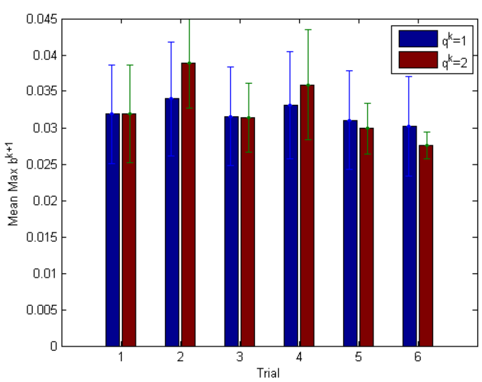


Fig. 8: Comparison of single agent strategies $q^k = 1$ and $q^k = 2$ for randomly sampled regions of the belief space. For each of the six trials, the strategies were repeated on the same randomly-sampled belief and true target location $10,000$ times. The strategies perform about the same on average but the variance for $q^k = 2$ never exceeds that of $q^k = 1$.

B. Results Over Arbitrary Horizons

To examine the performance of our search strategies over longer horizons, we simulated a search problem over a horizon of 100 time steps. To validate Theorem 2 from Sec. III-B, we present simulation results for $m = 3$ agents. Two strategies were compared: always choose $q^k = \{1,2,3\}$ and always choose $q^k = \{2,3,4\}$. For each strategy, 1000 simulations were performed. The target location was initialized randomly based on the steady state distribution for the target MC for each simulation, and the initial belief was set to the steady state distribution.

Figure 9 shows the results comparing the two strategies, including the mean and standard error for $\max_{s\in S} b^k(s)$. The mean for $q^k = \{1,2,3\}$ was 0.67 compared to a mean of 0.52 for $q^k = \{2,3,4\}$, meaning the lower variance strategy achieved 78% of the performance of the $q^k = \{1,2,3\}$ strategy. Although we have not proved it, moving a sensor from location 1 to location 4 appears to reduce the variance in the multi-agent case, supporting Conjecture 1. But unlike the single agent case, the mean performance is worse when there is no sensor at location 1. Nonetheless, these results suggest that for the multi-agent setting—as in the single agent results in Theorem 3—omitting 1 from the set of sensor locations results in an estimate with lower variance, which supports Conjecture 1. Additionally, these results lend further validation to Theorem 2.

To validate our single agent results, simulations were performed to compare three strategies: always choose $q^k = 1$, and always choose $q^k = 2$, and minimize the expected one-step entropy. We include the expected one-step entropy minimization to demonstrate the efficacy of both strategies with respect to a different choice of simple one-step strategy. Entropy is a common choice of objective function in search problems, but is harder to compute than the objective function we consider in this work.

Figure 10 shows the mean and standard error for $max_{s\in S} b^k(s)$ for each of the three strategies. In the mean, the performance of all three scenarios is comparable. The $q^k = 1$ strategy had a mean performance of 0.36 compared to a mean of 0.31 for the $q^k = 2$ strategy. Thus the $q^k = 2$ strategy achieved 84% of the performance of the $q^k = 1$ strategy. However $q^k = 2$ performs much better than $q^k = 1$ in terms of variance, as is expected. These results suggest a much more confident estimate of target location for strategies that result in minimum variance. Figure 10 suggests that although our policy is based on a one-step optimization, it performs well over arbitrary horizons. The mean for all three strategies was lower than the mean for any of the multi-agent simulations. This result follows from the monotonicity of adding additional sensors (see Thm. 1).

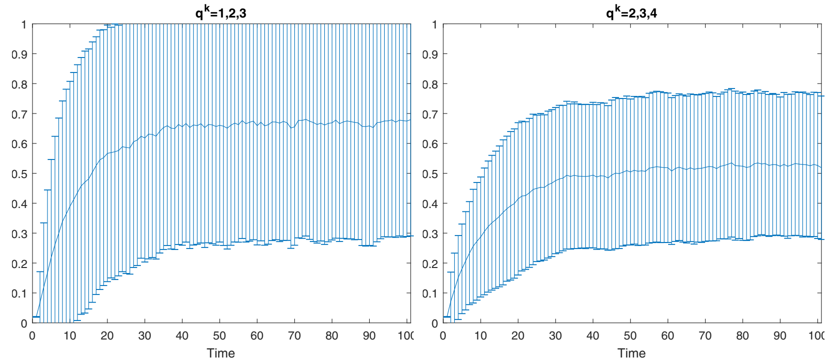
There are small differences in the mean performance of the $q^k = 1$ and $q^k = 2$ strategies visible in Fig. 10. These differences arise because our policy is designed to be one-step optimal, but our simulation is run for many steps. The two strategies end up with slightly different performance because they enter different regions of the belief space over the horizon of the simulation.

C. Sensitivity of Results

In this section, we examine the sensitivity of our results to the assumptions of our simulation. Namely, we investigate the effects of changing the size of the search area, the number of agents, and the values of α and β. For each of the sensitivity results, the probability of staying in each cell or moving in any of the four compass directions was 0.2. The choice of probabilities was made to remove any effect of a non-uniform stationary Markov distribution. For each analysis, 1000 simulations were performed over a horizon of 100 time steps.

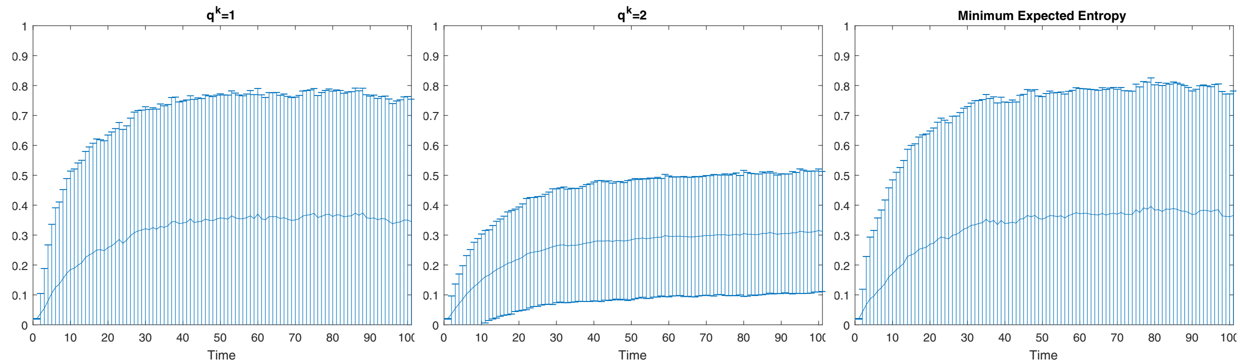
For the first analysis, we varied the number of agents from 2 to 10, and the grid size from 5×5 to 20×20. The results are presented in Fig. 11. The results show that as the environment increases, the performance of both strategies declines

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control



(a) (b)

Fig. 9: Multi-agent simulation results showing mean and standard error results for maximum belief from 1000 simulations over 100 time steps. Strategy $q^k = \{1, 2, 3\}$ (9a) has a higher value and higher variance than strategy $q^k = \{2, 3, 4\}$ (9b), consistent with Theorem 2, and suggesting the variance result from Theorem 3 extends to the multiple agent case.



(a) (b) (c)

Fig. 10: Single agent simulation results showing mean and standard error results for maximum belief from 1000 simulations over 100 time steps. Strategies $q^k = 1$ and $q^k = 2$ perform comparably in expectation (Figs. 10a&10b), with the latter demonstrating lower variance, consistent with Theorem 3. Both are more computationally simple than minimizing entropy (Fig. 10c).

(11a, 11b). Likewise as the number of agents increases, the perfomance of both strategies improves (11c, 11d). Figures 11e and 11f show that in either case, $q^k = \{1, 2, 3\}$ tends to outperform the other strategy, resulting in a higher result for 11 out of 16 data points.

For the second analysis, the values of α and β were both varied from 0.01 to 0.2 (12). Increasing values of α and β result in worse performance from both strategies, with a slightly stronger effect from α. The ratio of the two strategies (12e, 12f) does not conclusively show one strategy outperforming the other.

VI. Conclusion and Future Work

The results of this work demonstrate that we have found an effective, computationally efficient method for tracking a moving target with high confidence. There is great opportunity for future work in this area. For example, considering scenarios with multiple targets as well as motion constraints on the searcher are natural extensions of this problem. It would be interesting to extend this work to an examination of non-binary sensors. Further, looking at variations in false alarm and missed detection rates may yield different policies. Likewise, restrictions on the types of target motion and environment topology may allow even simpler policies to be found, as in the special case for searching on a grid. The great variety of potential applications for such work suggests that pursuing these research avenues will prove worthwhile.

References

[1] D. Castanon. Optimal search strategies in dynamic hypothesis testing. *Systems, Man and Cybernetics, IEEE Transactions on*, 25(7):1130–1138, 1995.

[2] T. H. Chung, G. A. Hollinger, and V. Isler. Search and pursuit-evasion in mobile robotics. *Autonomous robots*, 31(4):299–316, 2011.

[3] N. Atanasov, B. Sankaran, J. Le Ny, G. J. Pappas, and K. Daniilidis. Nonmyopic view planning for active object classification and pose estimation. *Robotics, IEEE Transactions on*, 30(5):1078–1090, 2014.

[4] A. Abdel-Samad, A. H. Tewfik, et al. Search strategies for radar target localization. In *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, volume 3, pages 862–866. IEEE, 1999.

[5] S. Aeron, V. Saligrama, and D. Castañón. Energy efficient policies for distributed target tracking in multihop sensor networks. In *Decision and Control, 2006 45th IEEE Conference on*, pages 380–385. IEEE, 2006.

[6] S. Aeron, V. Saligrama, and D. Castanon. Efficient sensor management policies for distributed target tracking in multihop sensor networks. *IEEE Transactions on Signal Processing*, 56(6):2562–2574, 2008.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2018.2857760, IEEE Transactions on Automatic Control

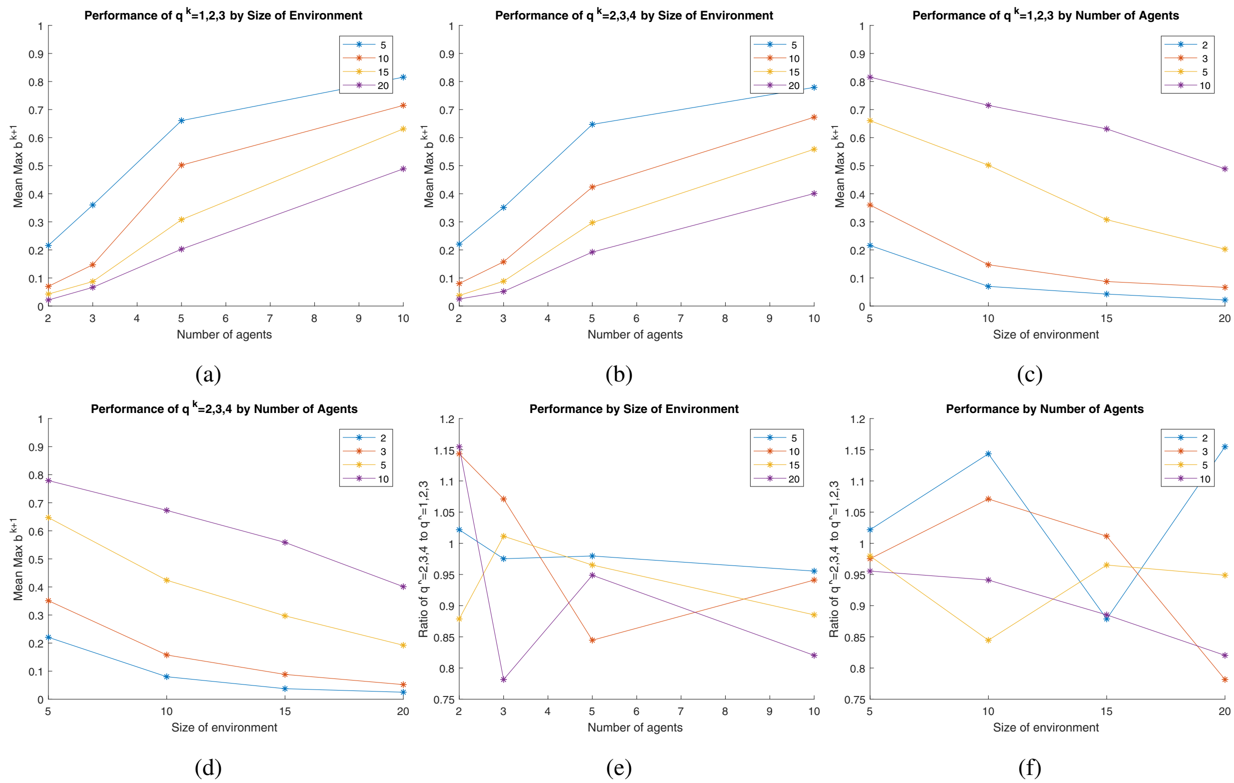


Fig. 11: Analysis of the effect of the size of the environment compared to the number of agents. Each data point represents 1000 simulations over 100 time steps. Strategies are grouped by the size of the environment (11a, 11b), by the number of agents (11c, 11d), and the ratio of $q^k = \{2,3,4\}$ to $q^k = \{1,2,3\}$ is presented by size of environment (11e) and by number of agents (11f). Fewer agents and a larger environment decrease both strategies, but $q^k = \{1,2,3\}$ generally performs better than $q^k = \{2,3,4\}$.

[7] K. Cohen and Q. Zhao. Quickest anomaly detection: a case of active hypothesis testing. In *Information Theory and Applications Workshop (ITA), 2014*, pages 1–5. IEEE, 2014.

[8] K. Cohen and Q. Zhao. Active hypothesis testing for anomaly detection. *IEEE Transactions on Information Theory*, 61(3):1432–1450, 2015.

[9] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, pages 326–339, 1948.

[10] S. S. Brown. Optimal search for a moving target in discrete time and space. *Operations Research*, 28(6):1275–1289, 1980.

[11] A. R. Washburn. Search for a moving target: The fab algorithm. *Operations research*, 31(4):739–751, 1983.

[12] L. D. Stone. Necessary and sufficient conditions for optimal search plans for moving targets. *Mathematics of Operations Research*, 4(4):431–440, 1979.

[13] W. R. Stromquist and L. D. Stone. Constrained optimization of functionals with search theory applications. *Mathematics of Operations Research*, 6(4):518–529, 1981.

[14] J. B. Kadane. Optimal whereabouts search. *Operations Research*, 19(4):894–904, 1971.

[15] D. Assaf and S. Zamir. Optimal sequential search: a bayesian approach. *The Annals of Statistics*, pages 1213–1221, 1985.

[16] K. Cohen, Q. Zhao, and A. Swami. Optimal index policies for anomaly localization in resource-constrained cyber systems. *Signal Processing, IEEE Transactions on*, 62(16):4224–4236, 2014.

[17] K. Cohen and Q. Zhao. Asymptotically optimal anomaly detection via sequential testing. *IEEE Trans. Signal Processing*, 63(11):2929–2941, 2015.

[18] N. Noori, A. Renzaglia, J. Vander Hook, and V. Isler. Constrained probabilistic search for a one-dimensional random walker. *IEEE Transactions on Robotics*, 32(2):261–274, 2016.

[19] J. P. Hespanha, H. J. Kim, and S. Sastry. Multiple-agent probabilistic pursuit-evasion games. In *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, volume 3, pages 2432–2437. IEEE, 1999.

[20] J. P. Hespanha, M. Prandini, and S. Sastry. Probabilistic pursuit-evasion games: A one-step nash approach. In *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, volume 3, pages 2272–2277. IEEE, 2000.

[21] N. Song and D. Teneketzis. Discrete search with multiple sensors. *Mathematical Methods of Operations Research*, 60(1):1–13, 2004.

[22] K. Leahy and M. Schwager. Always choose second best: Tracking a moving target on a graph with a noisy binary sensor. In *Proc. of the European Control Conference (ECC 16)*, June 2016.

[23] B. A. Davey and H. A. Priestley. *Introduction to lattices and order*. Cambridge university press, 2002.

[24] L. D. Stone. *Theory of optimal search*, volume 118. Elsevier, 1976.

Kevin Leahy obtained his PhD degree from the Department of Mechanical Engineering at Boston University, where he was a member of the Hybrid and Networked Systems (HyNeSs) and Multi-Robot Systems Laboratories. He received his BA degree in Economics from Boston University in 2009. His research interests include formal methods, persistent surveillance, and information-based planning for teams of robots.

Mac Schwager is an assistant professor in the Department of Aeronautics and Astronautics at Stanford University. He obtained his BS degree in 2000 from Stanford University, his MS degree from MIT in 2005, and his PhD degree from MIT in 2009. He was a postdoctoral researcher working jointly in the GRASP lab at the University of Pennsylvania and CSAIL at MIT from 2010 to 2012. His research interests are in distributed algorithms for control, perception, and learning in groups of robots and animals. He received the NSF CAREER

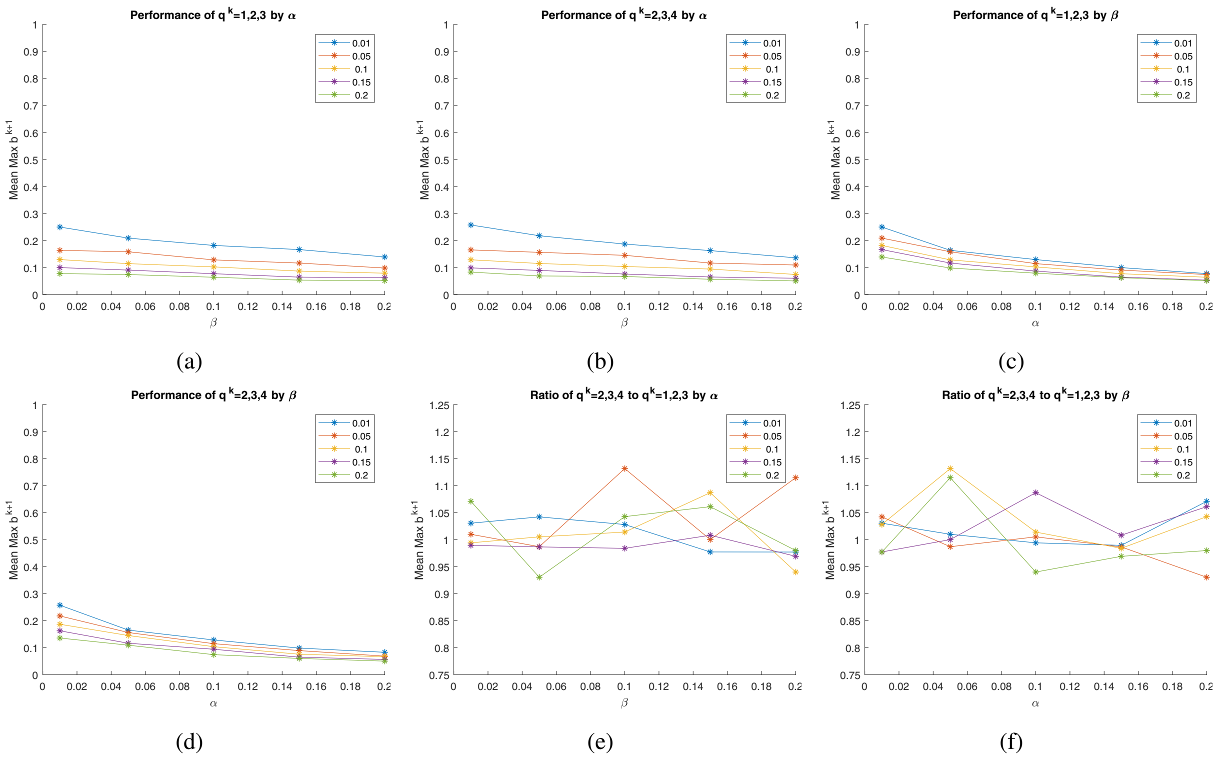


Fig. 12: Analysis of the effect of the value of α and β. Each data point represents 1000 simulations over 100 time steps. Strategies are grouped by α (12a, 12b), by β (12c, 12d), and the ratio of $q^k = \{2,3,4\}$ to $q^k = \{1,2,3\}$ is presented by α (12e) and by β (12f). Higher values of α and β decrease the performance of both strategies, with a slightly stronger effect from α.

award in 2014.