Zero-Error Target Tracking with Limited Communication

Hua Li, Patricia R. Barbosa, *Student Member, IEEE,* Edwin K. P. Chong, *Fellow, IEEE,* Jan Hannig, and Sanjeev R. Kulkarni, *Fellow, IEEE*

***Abstract*—We study the problem of target tracking in a sensor network environment. In particular, we consider a target that moves according to a Markov chain, and a tracker that queries sets of sensors to obtain tracking information. We are interested in finding the minimum number of queries per time step such that a target is trackable under three different requirements. First we investigate the case where the tracker is required to know the exact location of the target at each time step. We then relax this requirement and explore the case where the tracker may lose track of the target at a given time step, but it is able to "catch-up" at a later time, regaining up-to-date information about the target's track. Finally, we consider the case where tracking information is only known after a delay of d time steps. We provide necessary and sufficient conditions on the number of queries per time step to track in the above three cases. These conditions are stated in terms of the entropy rate of the target's Markov chain.**

***Index Terms*—Target tracking, Markov chain, adaptive sensing, entropy rate, Huffman coding, causal source coding.**

I. Introduction

WE EXPLORE the similarities between the target tracking problem in a sensor network setting and Rényi-Ulam games [1], [2], [3], of which the game of "twenty questions" can be considered a subclass. Our goal is to find theoretical bounds on the number of queries per time step a tracker is required to ask a set of sensors to track a target. To the best of our knowledge, this work is the first to investigate the interplay between these two problems, providing the notable simplicity of our tracking scheme.

Perhaps the initial driving force in sensor networks research [4], the target tracking problem is revisited in this work. While much of recent research in sensor networks explores networking issues like energy consumption optimization [5], time synchronization [6], [7], sensor localization [8], [9], and routing [10], [11], additional communications problems such as data compression and message complexity have become increasingly important as the number of small and inexpensive networked sensing devices continues to grow. Moreover, the reliability and the capacity of the channel available for communication with the tracker lead to restrictions on the amount of data sent over such networks. As a consequence, the tracker needs to make judicious decisions when selecting sensors to send data, so that communication with the sensor network is kept to a minimum.

Related problems include those in the area of control under communication constraints. In particular, Tatikonda and Mitter [12], [13], [14] examined a control problem with a noiseless digital communication channel connecting a sensor to a controller, and provided upper and lower bounds on the channel rate required to achieve different control objectives, namely, asymptotic observability and asymptotic stabilizability [15]. Sahai and Mitter [16], [17], [18] investigated the problem of tracking and controlling unstable processes over noisy channels and demonstrated that Shannon's classical notion of capacity was insufficient to characterize noisy channels for this purpose. Furthermore, they identified a novel characterizing quantity called *anytime capacity* and showed that it is both necessary and sufficient for channels to have a certain amount of anytime capacity such that unstable processes can be tracked and stabilized.

We propose a sensor model in which sensors are capable of sending only one-bit messages to a tracker. These messages are used to gather tracking information about a moving target. In the literature, one-bit-message sensor networks are called *binary sensor networks* and have been previously considered for target tracking [19], [20], [21]. In [22], Evans et al. analyzed the problem of optimal sensor selection; their approach is to formulate the problem as a partially observed stochastic control problem, where sensors are not constrained to one-bit messages, and the tracker also controls the channel data rate so that mean squared errors are bounded.

The remainder of this paper is organized as follows. Section II formalizes the tracking problem under three different definitions. In Section III, the main theorems are stated; they are later proved in Section IV. Finally, Section V concludes this work.

II. Problem Formulation

Consider a target moving around in an area. Suppose that the area is partitioned into a number of non-overlapping regions, referred to as *locations*. The target motion can be described by a directed graph $G = (\mathcal{X}, E)$, where the set of nodes $\mathcal{X}$, with finite cardinality, represents target locations, and the set of edges E describes each neighborhood, that is, possible target motion. If there exists an edge $(i, j) \in E$,

Manuscript received May 15, 2007; revised November 11, 2007. The material in this paper was presented in part at the 44th Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, September 27–29, 2006.

H. Li, P. R. Barbosa, and E. K. P. Chong are with the Department of Electrical and Computer Engineering, Colorado State University, Ft Collins, CO 80523 USA (e-mail: {hua.li, patricia.barbosa, edwin.chong}@colostate.edu).

J. Hannig is with the Department of Statistics, Colorado State University, Ft Collins, CO 80523 USA (e-mail: jan.hannig@colostate.edu).

S. R. Kulkarni is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: kulkarni@princeton.edu).

Digital Object Identifier 10.1109/JSAC.2008.080510.

$i, j \in \mathcal{X}$, with associated transition probability $p_{ij} > 0$, connecting node $i \in \mathcal{X}$ to $j \in \mathcal{X}$, then given that the target is at node i, it moves to j with probability p_{ij}. Hence, we model the random motion of the target by an ergodic (time-homogenous) discrete Markov chain

$$\{X_t : t \in \mathbb{N}\}$$

with transition probabilities

$$p_{ij} = \Pr\{X_{t+1} = j | X_t = i\},\ i, j \in \mathcal{X}.$$

Associated with each target location is a sensor, which can sense if the target is at its location. Throughout this paper we use the term target location, sensor, and Markov chain state as synonyms.

Given the initial target location $X_1 = x_1 \in \mathcal{X}$, the history of locations visited by a target up to time step t is called a *track*.

Definition 1: The target *track* at time $t \geq 1$, denoted $X_{[1:t]}$, is defined as the following finite random sequence of states:

$$X_1\, X_2\, \ldots\, X_t.$$

The main goal of target tracking is to estimate tracks over time. We use the following notation for such estimates: at time step $\tau \geq t$, the *estimate* of $X_{[1:t]}$ is denoted by $\hat{X}^{\tau}_{[1:t]}$; likewise, at time step $\tau \geq t$, the *estimate* of state X_t is denoted by $\hat{X}^{\tau}_t$. In this work, we are interested in the problem of zero-error tracking.

A tracker estimates target tracks by querying subsets of sensors. At each time step t, the tracker may query the sensors a number of times, say K_t times. We denote the kth query at time t by $q_{t,k}$, $1 \leq k \leq K_t$. Furthermore, each query consists of a number of "questions", each of which addresses a particular sensor with a timestamp. Therefore, the query $q_{t,k}$ is characterized by a set of sensor-timestamp pairs:

$$q_{t,k} = \{(s^j_{t,k}, \tau^j_{t,k}) : 1 \leq j \leq J_{t,k}, s^j_{t,k} \in \mathcal{X}, \tau^j_{t,k} \leq t\},$$

where the sensor-timestamp pair $(s^j_{t,k}, \tau^j_{t,k})$ denotes the question "has the sensor $s^j_{t,k}$ detected the target at time $\tau^j_{t,k}$?", and $J_{t,k}$ denotes the number of "questions" in the query $q_{t,k}$.

In response, the queried set of sensors transmit a single bit to the tracker. Specifically, the response $r_{t,k}$ to query $q_{t,k}$ can be written as

$$r_{t,k} = \begin{cases} 1, & \text{if } X_{\tau^j_{t,k}} = s^j_{t,k} \text{ for some } j; \\ 0, & \text{otherwise,} \end{cases}$$

that is, $r_{t,k} = 1$ if and only if, for some $1 \leq j \leq J_{t,k}$, sensor $s^j_{t,k}$ has detected the target at time $\tau^j_{t,k}$.

At every time step t, a *querying scheme* is modeled by a finite binary *decision tree* T_t. Here we use the same notion of decision tree as used by Rivest et al. in [23]. Internal nodes in T_t corresponds to queries $q_{t,k}$, $1 \leq k \leq K_t$. The right and left children of each internal node represent queries following a "yes" (1) and "no" (0) response, respectively (unless they are leaves). Note that the root in T_t is considered an internal node. Finally, associated with each leaf in T_t is a track estimate up to time step t, i.e., $\hat{X}^t_{[1:t]}$.

To summarize, at each time t, sensors are queried by the tracker according to a querying scheme modeled by a binary decision tree T_t (following a path in the decision tree), resulting in a sequence of queries

$$(q_{t,1}, q_{t,2}, \cdots, q_{t,K_t})$$

and an associated sequence of responses

$$(r_{t,1}, r_{t,2}, \cdots, r_{t,K_t}),$$

and culminating with an estimate of the track up to time t, $\hat{X}^t_{[1:t]}$.

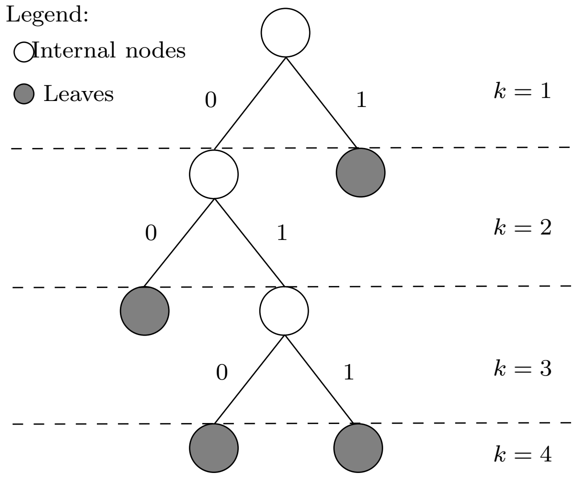


Fig. 1. Binary decision tree T_t with four levels

Let C denote the maximum number of queries allowed at each time step, called the *query quota*. Note that C is an integer. One could think of several scenarios where the number of queries at each time step is naturally limited in this way. First, the capacity of the communication channel between the tracker and the sensors is usually scarce. Also, given the limited processing capabilities of small and simple sensors, a constraint on how fast queries can be processed is expected. Moreover, if sensors are deployed on a hostile enemy environment, it is reasonable to limit the maximum number of responses sent by sensors at each time step to avoid being detected. Therefore, for every time step t, T_t has at most $m = C+1$ levels. Equivalently, we say that the height of tree T_t, denoted by $h(T_t)$, satisfies $h(T_t) \leq C+1$. Fig. 1 illustrates a binary decision tree with four levels.

We further make the following remarks. All sensors are fault-free, have memory, and are able to communicate without error or delay with the tracker, which knows the initial location of the target and its motion law. Sensor ranges do not overlap. Logarithms are taken to the base 2 and, by convention, $\log 0 = 0$ and $\log \frac{0}{0} = 0$.

In order to track the target, a tracker uses a *strategy* defined as follows:

Definition 2: A strategy $\mathcal{S}$ is a rule that, at each time step, maps the current querying scheme and its results to a next querying scheme and an estimate of the current target track. In other words, at each time step t, the tracker takes T_t and the sequence of responses at time t, and then, according to $\mathcal{S}$, generates T_{t+1} and outputs $\hat{X}^t_{[1:t]}$.

Fig. 2 depicts the mapping defined above.

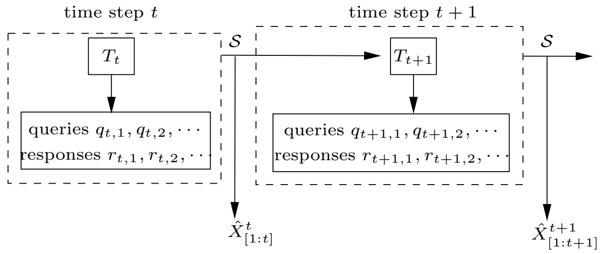


Fig. 2. Strategy mapping

We investigate three distinct "degrees" of target tracking, which we call *following*, *tracking*, and *d-tracking*, defined below. Whether or not these degrees hold depends on the target motion law, given by the tuple $(G, [p_{ij}], x_1)$, and the query quota C.

Definition 3: A target is *followable* if there exists a strategy such that, at each time step t, the target track estimate equals the true target track with probability one, that is,

$$\forall t \geq 1,\ \hat{X}^t_{[1:t]} = X_{[1:t]}\ a.s.$$

Definition 4: A target is *trackable* if there exists a strategy such that the target track estimate equals the true target track infinitely often with probability one, that is,

$$\forall t \geq 1,\ \exists \tau > t \text{ s.t. } \hat{X}^{\tau}_{[1:\tau]} = X_{[1:\tau]}\ a.s.$$

In other words, even if the tracker loses track of the target, if it is able to "catch-up" and regain current information about the target track at a later time step, tracking is still accomplished. On the other hand, if the target track is correctly estimated only after a fixed delay of d time steps, we define:

Definition 5: A target is *d-trackable* if there exists a strategy such that the target track estimate equals the true target track, after a delay of d time steps, infinitely often with probability one, that is,

$$\forall t \geq 1,\ \exists \tau > t \text{ s.t. } \hat{X}^{\tau+d}_{[1:\tau]} = X_{[1:\tau]}\ a.s.,$$

where $\hat{X}^{\tau+d}_{[1:\tau]}$ denotes the estimate of the target track $X_{[1:\tau]}$ at time $\tau + d$.

We are interested in answering the following questions:

1) What is the minimum query quota such that a target is followable?
2) What is the minimum query quota such that a target is trackable?
3) What is the minimum query quota such that a target is d-trackable?

III. Main Results

As we shall see in this section, there is an intimate connection between querying and coding. A binary sequential source code naturally arises from the use of binary decision trees to represent the querying scheme. Recently, Borkar et al. [24] revisited the problem of sequential source coding, formulated as a constrained optimization problem on a convex set of probability measures. Although we also impose a causality constraint on codewords, one key difference between our work and [24] is that we introduce a constraint *at each time step* (the query quota) on the number of bits of codewords.

We now state three theorems on the conditions under which a target is followable, trackable, and d-trackable. We defer the proofs of these theorems to Section IV.

A. Target Following

We use the notation E_i to denote the set of neighbors of state $i \in \mathcal{X}$, that is, $E_i = \{j \in \mathcal{X} : p_{ij} > 0\}$.

Theorem 1: A target is followable if and only if

$$C \geq \max_{i \in \mathcal{X}} \log |E_i|.$$

B. Target Tracking

As described in Section II, the target motion is modeled by an ergodic Markov chain with finite state space. Thus, the chain is positive recurrent [25], and its Shannon entropy rate can be calculated as [26], [27]:

$$H = -\sum_{i \in \mathcal{X}} \pi_i \sum_{j \in \mathcal{X}} p_{ij} \log p_{ij}, \tag{1}$$

where π_i is the stationary distribution of the Markov chain.

We can then state our result on tracking as follows:

Theorem 2:

a) If a target is trackable, then $C \geq H$.
b) A target is trackable if $C \geq H + 1$.

C. Target d-tracking

Theorem 3:

a) If a target is d-trackable for some positive integer d, then $C \geq H$.
b) For any positive integer d, a target is d-trackable if $C \geq H + \frac{1}{d}$.

IV. Proofs

A. Proof of Theorem 1

We first prove necessity by contraposition. Assume $C < \max_{i \in \mathcal{X}} \log |E_i|$. Then, there must exist a state $i \in \mathcal{X}$ such that $C < \log |E_i|$. Since the Markov chain $\{X_t : t \geq 1\}$ is ergodic, thus irreducible, there exists a time step $t \geq 1$ such that the t-step transition probability from state x_1 to state i is strictly positive, i.e., $p^{(t)}_{x_1 i} > 0$ for some $t \geq 1$ (where $p^{(1)}_{x_1 i} = p_{x_1 i}$). From the definition of query quota, we know that at most C bits per time step can be transmitted to the tracker. Therefore, the number of choices for estimating X_{t+1} must be at most 2^C at each time step. But $2^C < |E_i|$, and hence no strategy is able to identify correctly all possible choices for X_{t+1}. Therefore,

$$\Pr\left\{\hat{X}^{t+1}_{[1:t+1]} \neq X_{[1:t+1]} \middle| X_t = i\right\} \geq \min_{j \in \mathcal{X}} p_{ij} > 0.$$

Hence, we have

$$\begin{aligned} &\Pr\left\{X^{t+1}_{[1:t+1]} \neq X_{[1:t+1]}\right\} \\ &\quad = \sum_{x_t \in \mathcal{X}} \Pr\left\{\hat{X}^{t+1}_{[1:t+1]} \neq X_{[1:t+1]} \middle| X_t = i\right\} p^t_{x_1\, i} > 0, \end{aligned}$$

contradicting the requirement for following strategies: $\Pr\left\{\hat{X}^t_{[1:t]} = X_{[1:t]}\right\} = 1$ for all $t \geq 1$.

To prove sufficiency, we show by induction that the simple and well-known binary search [28] yields a strategy $\mathcal{S}$ using which we can follow a target as long as $C \geq \max_{i\in\mathcal{X}} \log|E_i|$. Since the initial location of the target is known a priori, $\Pr\left\{\hat{X}^1_1 = X_1\right\} = 1$ trivially. For a fixed $t > 1$, assume $\Pr\left\{\hat{X}^t_{[1:t]} = X_{[1:t]}\right\} = 1$. It suffices to show that $\Pr\left\{\hat{X}^{t+1}_{t+1} = X_{t+1}\right\} = 1$. The binary decision tree T_{t+1} is generated as follows. An internal node at level k in T_{t+1} denotes a query of the form:

$$q_{t+1,k} = \left\{(s_1, t+1), (s_2, t+1), \ldots, (s_n, t+1)\right\},$$

where $|q_{t+1,1}| \leq \lceil |E_{x_t}|/2 \rceil$. The set of sensors to be queried is repeatedly reduced by about half until the target track is estimated with certainty. Hence, either

$$|q_{t+1,k}| = \left\lceil \frac{|q_{t+1,1}|}{2^{k-1}} \right\rceil \quad \text{or} \quad |q_{t+1,k}| = \left\lfloor \frac{|q_{t+1,k}|}{2^{k-1}} \right\rfloor.$$

Since we are interested in showing sufficiency, we consider the case where the largest number of queries is required, that is, the case where $|q_{t+1,k}| = \lceil |q_{t+1,k}|/2^{k-1} \rceil$. By definition, T_{t+1} has at most $C+1$ levels; thus, assuming the largest possible number of queries is required at time step $t+1$, we can write

$$|q_{t+1,C}| \geq \frac{|q_{t+1,1}|}{2^{C-1}}.$$

When using binary search, the last query asked is a singleton, therefore $|q_{t+1,C}| = 1$, and we have

$$1 \geq \frac{|E_{x_t}|/2}{2^{C-1}} \Rightarrow C \geq \log|E_{x_t}|.$$

Hence we are able to estimate correctly the state X_{t+1} among all possible choices, that is, $\Pr\left\{\hat{X}^{t+1}_{t+1} = X_{t+1}\right\} = 1$. Thus, it suffices to have $C \geq \max_{i\in\mathcal{X}} \log|E_i|$ to follow a target.

B. Proof of Theorem 2

Part (a) is proved by contraposition. The proof uses the idea of *strong typicality* and a result from large deviation theory.

First, we extend the concept of *strong typicality* [27, Ch.5] to finite state ergodic Markov chains $\{X_t : t \geq 1\}$. For every $t \geq 1$ and each state transition $(i,j) \in E$, we define the counting function N^t_{ij} on $x_{[1:t]}$ as

$$N^t_{ij}(x_{[1:t]} = \sum_{k=1}^{t} \mathbf{1}_i(x_k)\,\mathbf{1}_j(x_{k+1}),$$

where $\mathbf{1}_i(x_k)$ denotes the indicator function, that is,

$$\mathbf{1}_i(x_k) = \begin{cases} 1, & \text{if } x_k = i; \\ 0, & \text{otherwise.} \end{cases}$$

For a fixed $\delta > 0$ and every $t \geq 1$, the set

$$\Delta_{t,\delta} = \left\{ x_{[1:t]} : \left| \frac{N^t_{ij}}{t} - \pi_i p_{ij} \right| < \delta, \quad \forall (i,j) \in E \right\}$$

is called the *δ-strong typical set*, and sequences in this set are called the *δ-strong typical sequences*. Lemma 1 below states two important properties of strong typical sets and strong typical sequences.

Lemma 1: For a fixed $\delta > 0$, we have:

a) The probability of every δ-strong typical sequence $x_{[1:t]}$ satisfies

$$2^{-t(H+c_1\delta)} p_{x_1} < \Pr\left\{X_{[1:t]} = x_{[1:t]}\right\} < 2^{-t(H-c_1\delta)} p_{x_1},$$

where $c_1 = -\sum_{(i,j)\in E} \log p_{ij}$ and p_{x_1} is the initial distribution $\Pr\{X_1 = x_1\}$.

b) For sufficiently large t,

$$\Pr\left\{X_{[1:t]} \notin \Delta_{t,\delta}\right\} \leq c_2 2^{-c_3 t},$$

where c_2 and c_3 are positive constants.

Part (a) of the lemma is established by modifying the arguments for the case of i.i.d. sequences in [27, Ch.5]. Part (b) is proved using a result [29] from large deviation theory for finite Markov chains. To avoid interrupting the main idea, we defer the proof of Lemma 1 to the appendix.

Now, assume there exists a strategy $\mathcal{S}$ such that a target is trackable and $C < H$. Using Lemma 1, we get the following bound for the probability of the event $\left\{\hat{X}^t_{[1:t]} = X_{[1:t]}\right\}$:

$$\begin{aligned}
&\Pr\left\{\hat{X}^t_{[1:t]} = X_{[1:t]}\right\} \\
&= \sum_{x_{[1:t]}\in\mathcal{X}_{[1:t]}} q_{x_{[1:t]}} \Pr\left\{X_{[1:t]} = x_{[1:t]}\right\} \\
&= \sum_{x_{[1:t]}\in\Delta_{t,\delta}} q_{x_{[1:t]}} \Pr\left\{X_{[1:t]} = x_{[1:t]}\right\} \\
&\quad + \sum_{x_{[1:t]}\notin\Delta_{t,\delta}} q_{x_{[1:t]}} \Pr\left\{X_{[1:t]} = x_{[1:t]}\right\} \\
&\leq \sum_{x_{[1:t]}\in\Delta_{t,\delta}} q_{x_{[1:t]}} \Pr\left\{X_{[1:t]} = x_{[1:t]}\right\} + c_2 2^{-c_3 t} \\
&< \sum_{x_{[1:t]}\in\Delta_{t,\delta}} q_{x_{[1:t]}} 2^{-t(H-c_1\delta)} p_{x_1} + c_2 2^{-c_3 t},
\end{aligned}$$

where $q_{x_{[1:t]}} = \Pr\left\{\hat{X}^t_{[1:t]} = x_{[1:t]} | X_{[1:t]} = x_{[1:t]}\right\}$. Given the constraint on the height of the tree T_t for each time step $t \geq 1$, there are at most 2^{tC} choices for $\hat{X}^t_{[1:t]}$. Hence,

$$\sum_{x_{[1:t]}\in\Delta_{t,\delta}} \Pr\left\{\hat{X}^t_{[1:t]} = x_{[1:t]} | X_{[1:t]} = x_{[1:t]}\right\} \leq 2^{tC}.$$

Therefore,

$$\Pr\left\{\hat{X}^t_{[1:t]} = X_{[1:t]}\right\} < 2^{-t[H-C-c_1\delta]} p_{x_1} + c_2 2^{-c_3 t}.$$

Since we can always choose $\delta > 0$ such that $H - C - c_1\delta > 0$, and $c_3 > 0$, it is clear that

$$\sum_{t\geq 1} \Pr\left\{\hat{X}^t_{[1:t]} = X_{[1:t]}\right\} < \infty.$$

By the first Borel-Cantelli lemma [30], we have

$$\Pr\left\{\hat{X}^t_{[1:t]} = X_{[1:t]} \quad \text{i.o.}\right\} = 0,$$

thus contradicting the assumption that a tracking strategy $\mathcal{S}$ catches-up infinitely often with probability one. Hence, it is necessary that $C \geq H$ for a target to be trackable.

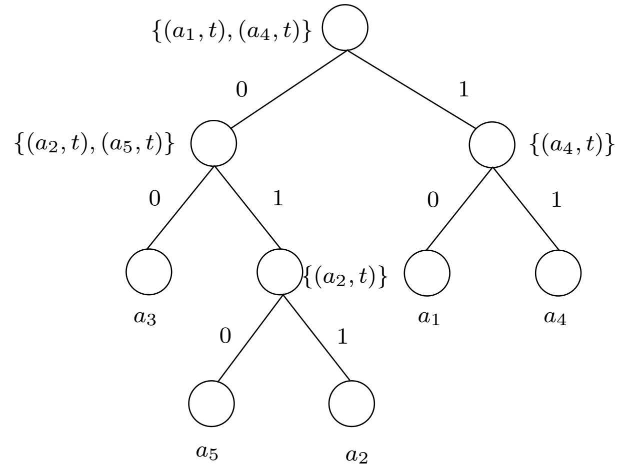


Fig. 3. Binary decision tree T_{x_1} derived from Huffman coding

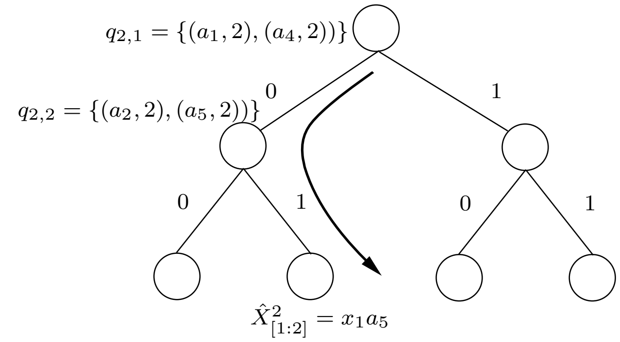


Fig. 4. Tree traversal on T_2

To prove part (b), assume $C \geq H + 1$. We describe an interactive tracking strategy which we refer to as the *catch-up* strategy, and we show (again, using induction) that a target is trackable when $C \geq H + 1$. Let $t = 1$. Given the initial target location $X_1 = x_1 \in \mathcal{X}$, we first show that there exists $\tau > 1$ such that $\hat{X}^{\tau}_{[1:\tau]} = X_{[1:\tau]}$ almost surely. In the catch-up strategy, each binary decision tree T_t is generated according to a scheme based on Huffman coding [26], where the tracker uses the Markov chain transition probabilities to get Huffman codewords. Traversing T_t (that is, querying) takes place as follows. The first nodes to be queried are those associated with codewords whose leftmost bit is 1 (equivalently, one could first query nodes whose leftmost bit is 0). If the response to this query is 1, the tracker would then query nodes whose corresponding codewords first two bits (from left to right) are 1, and so forth. On the other hand, if the response to the first query is 0, the following nodes queried would then be those associated with codewords whose first bit (from left to right) is 0 and second bit is 1.

From Huffman coding, we can directly derive a binary decision tree T_{x_1} from E_{x_1}, the set of neighbors of state x_1. Each internal node in T_{x_1} represents a possible query, and each leaf corresponds to a possible target location at $t = 2$. We maintain an auxiliary tree T^*, initialized with T_{x_1}. In order to generate T_2, we use the fact that at most C queries can be asked at any given time step, and prune T^* at level $k = \min(C + 1, h(T^*))$. The resultant tree is T_2, the binary decision tree for the querying scheme at time $t = 2$. Each leaf in T_2 is associated with an estimate $\hat{X}^2_{[1:2]}$. We now have two possibilities:

Case 1: The tracker reaches a leaf of T_{x_1} within C queries, the track $X_{[1:2]}$ is successfully determined, and the tracker has caught-up. Then, the procedure described above is repeated for $t \geq 3$, that is, we first derive tree T_{x_2} from E_{x_2} since location x_2 is successfully determined; we then set $T^* = T_{x_2}$, and generate T_3 with height $h(T_3) = \min(C + 1, h(T^*))$. At the next time step ($t = 3$), if the tracker reaches a leaf of T_3 within C queries, then it continues as in Case 1; otherwise, it proceeds as in Case 2.

Case 2: The tracker does not reach a leaf of T_{x_1} after C queries—it fails to "catch-up". In this case, the tracker takes a "best guess", for example, the maximum a posteriori (MAP) estimate of $X_{[1:2]}$ based on the query and response history. Then, at the next time step ($t = 3$), the tracker uses Huffman coding to derive trees T_i, where $i \in \mathcal{X}$ is associated with leaves of the subtree in T_{x_1}, whose root corresponds to the last node reached by the tracker in T_{x_1}. Each tree T_i is appended to T_{x_1}, replacing the leaf corresponding to location i in T_{x_1} by the root in T_i. All other nodes, except the last node reached and those descending from it, are then removed, thus yielding a "concatenated" tree $T_{x_1^+}$, "re-rooting" at the last node reached at the previous time step. We then update the auxiliary decision tree T^* with $T_{x_1^+}$ and prune T^* to get T_3 with its height $h(T_3) = \min(C + 1, h(T^*)))$. If, at the next step ($t = 3$), the tracker fails to "catch-up", then it derives T_4 in the same way as T_3. This process continues until the tracker "catches-up"—it reaches a leaf of $T_{x_{t-1}}$ at time step $t > 1$. Querying is then resumed as in Case 1.

We further illustrate the catch-up strategy with the following instructive example.

Example 1: Let $X_1 = x_1 \in \mathcal{X}$, $C = 2$, and assume $X_2 = a_2 \in \mathcal{X}$. Also, let $E_{a_0} = \{a_1, a_2, a_3, a_4, a_5\}$ with transition probabilities $\{0.2, 0.1, 0.3, 0.2, 0.2\}$. Thus, one possible set of Huffman codewords is $\{10, 011, 00, 11, 010\}$, and the binary decision tree T_{x_1} is illustrated in Fig. 3.

Since $t = 2$, T^* is the same as T_{x_1}. Moreover, since $C = 2$, T_2 can have at most three levels, as shown in Fig. 4, where the tracker traverses two levels (corresponding to the quota constraint of two queries allowed at each time step), fails to "catch-up", and takes $x_1 a_5$ as the "best guess" for the track $X_{[1:2]}$.

In this case, T_{a_2} and T_{a_5} are generated using the set of neighbors E_{a_2} and E_{a_5}, respectively. Let $E_{a_2} = \{a_1, a_3, a_4, a_5\}$ with transition probabilities $\{0.4, 0.1, 0.1, 0.4\}$, and $E_{a_5} = \{a_4, a_3\}$ with transition probabilities $\{0.5, 0.5\}$. T_{a_2} and T_{a_5} are directly derived from Huffman coding and illustrated in Fig. 5. Trees T_{a_2} and T_{a_5} are appended to T^* (T_{x_1}) as follows: the root in T_{a_2} takes place of the leaf corresponding to location a_2 in T^*; likewise, the root in T_{a_5} takes place of the leaf corresponding

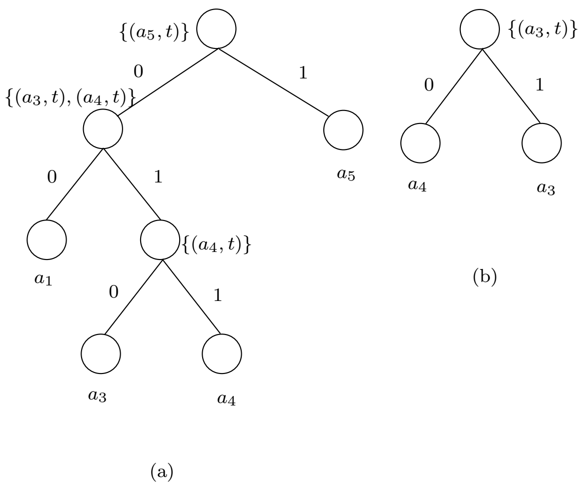


Fig. 5. Binary decision tree T_{a_2} (a) and T_{a_5} (b)

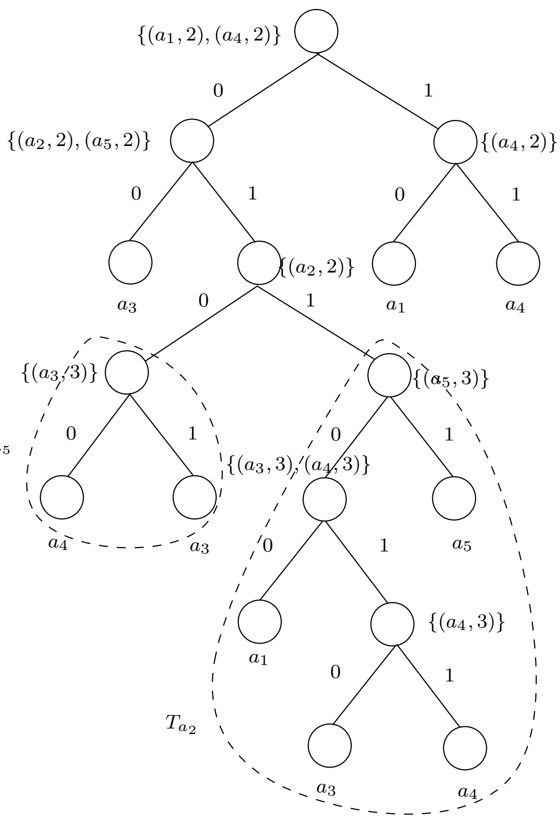


Fig. 6. Concantenating T_{a_5} and T_{a_2}

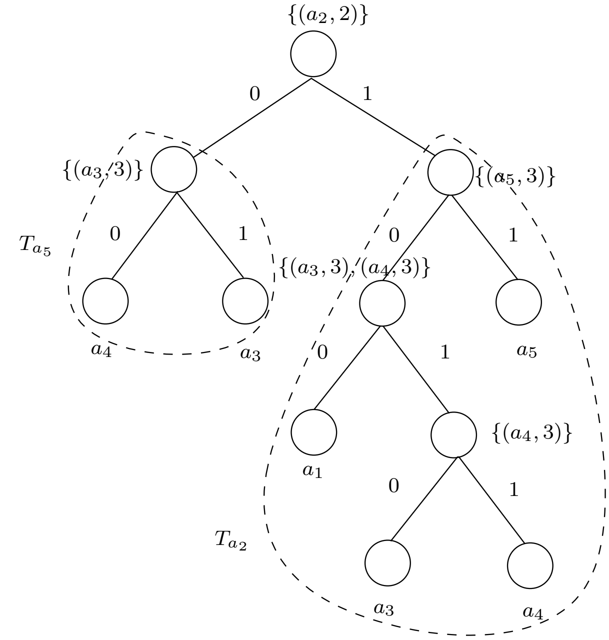


Fig. 7. Updating T^*

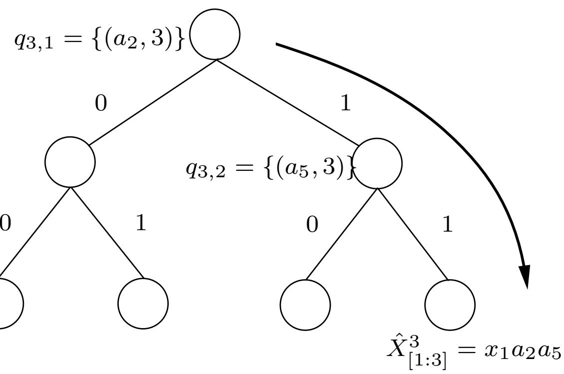


Fig. 8. Tree traversal on T_3

to location a_5 in T^*, resulting in the concatenated tree shown in Fig. 6. Removing all nodes except the last node reached, which is associated with the query $\{(a_2, 2)\}$, and the nodes descending from it, we obtain the updated auxiliary tree T^* ($T_{x_1^+}$), shown in Fig. 7. The decision tree T_3 is obtained by pruning the auxiliary tree T^* at level $C + 1 = 3$, shown in Fig. 8. Assuming $X_3 = a_5$, the tracker is able to reach a leaf of T_{x_2} at time 3, i.e., $\hat{X}^3_{[1:3]} = X_{[1:3]}$. In this case, although the tracker lost track at time 2, it catches-up at time 3.

We now show that the catch-up strategy can indeed be used to track a target. In other words, we show that the tracker regains current information about the target location infinitely often with probability one when the catch-up strategy is applied. Let $l_{[1:\tau]}$ be the average number of queries asked up to time τ, that is,

$$l_{[1:\tau]} = \frac{\sum_{k=1}^{\tau} l_k}{\tau}, \quad \tau > 1,$$

where l_k is the number of queries asked to estimate X_k with certainty. Clearly, l_k is a bounded function of the Markov chain $\{X_t : t \geq 1\}$, for every $k \geq 1$. By the Generalized Convergence Theorem for bounded functions of discrete-time and ergodic Markov chains with finite state space [31], we have $l_{[1:\tau]} \xrightarrow{a.s.} L$, where the limit L (according to the Source Coding Theorem [26], [27]) satisfies $H \leq L < H + 1$, and H is given by (1). It suffices to show that

$$\Pr\{l_{[1:\tau]} < C, \quad \text{for some } \tau > 1\} = 1,$$

and we show it by contradiction. Assume

$$\Pr\{l_{[1:\tau]} \geq C, \quad \forall \tau > 1\} = p > 0.$$

Then, since $C \geq H+1$ and $H \leq L < H+1$ hold, we have

$$\Pr\{l_{[1:\tau]} \geq L,\ \forall \tau > 1\} \geq \Pr\{l_{[1:\tau]} \geq C,\ \forall \tau > 1\} = p > 0.$$

Therefore, we get

$$\Pr\left\{\lim_{\tau\to\infty} l_{[1:\tau]} = L\right\} \leq 1-p,$$

a contradiction. Hence, there exists $\tau > 1$ such that $\hat{X}^{\tau}_{[1:\tau]} = X_{[1:\tau]}$ holds almost surely. By induction, now assume the above is true for $t = t'$, that is, there exists $\tau' > t'$ such that $\hat{X}^{\tau'}_{[0:\tau']} = X_{[0:\tau']}$ holds almost surely. Using the same reasoning as above (for the case $t = 1$), we can show that there exists $\tau'' > \tau'$ such that $\hat{X}^{\tau''}_{[1:\tau'']} = X_{[1:\tau'']}$ holds almost surely. Hence, it suffices to have $C > H+1$ to track a target.

C. Proof of Theorem 3

Part (a) is proven once again using contradiction and strong typicality. Similarly to the tracking case, we assume that $C < H$ and that there exists a strategy $\mathcal{S}$ such that a target is d-trackable. The track estimate $\hat{X}^{t+d}_{[1:t]}$ has at most $2^{(t+d)C}$ choices, since $\hat{X}^{t+d}_{[1:t]}$ can be true at most $2^{(t+d)C}$ times. Hence, the probability of the event $\left\{\hat{X}^{t+d}_{[1:t]} = X_{[1:t]}\right\}$ is bounded by

$$\Pr\left\{\hat{X}^{t+d}_{[1:t]} = X_{[1:t]}\right\} < 2^{-(t+d)[(H-C)-c_1\delta]} + c_2 2^{-c_3(t+d)}.$$

Again, by the first Borel-Cantelli lemma, we have

$$\Pr\left\{\hat{X}^{t+d}_{[1:t]} = X_{[1:t]} \quad \text{i.o.}\right\} = 0,$$

a contradiction. Thus, $C \geq H$.

We show part (b) using a block version of the catch-up strategy described in the proof of Theorem 2. Consider the sequence of random variables $\{W_n : n > 0\}$, where $W_n = (X_{d(n-1)+1}, \ldots, X_{dn})$, $d > 0$, that is, each random variable W_n is a segment of length d of the sequence $\{X_t : t \geq 1\}$. We call the sequence $\{W_n : n > 0\}$ a *block Markov chain* taking values in the state space $\mathcal{X}^d$. Assuming $C \geq H + 1/d$, and given the initial target location x_0, we skip querying during the first d time steps. For each time step t from $t = d+1$ to $t = 2d$, we apply the catch-up strategy to get $\hat{X}^{2d}_{[1:d]}$. This is done using the transition probabilities of the Markov chain $\{W_n : n > 0\}$ to generate Huffman codewords. Thus, at $t = 2d$, we have the estimate $\hat{W}_1 = \left(\hat{X}^{2d}_1, \hat{X}^{2d}_2, \ldots, \hat{X}^{2d}_d\right)$. This procedure is repeated for every "block" of d time steps, hence if C_W is the maximum number of queries allowed during each block, $C_W = dC$. Moreover, the entropy H_W of the Markov chain $\{W_n : n > 0\}$ can be calculated in terms of the entropy rate H of the original Markov chain as

$$\begin{aligned} H_W &= \lim_{n\to\infty} \frac{-\log\Pr\left\{X_{[1:nd]} = x_{[1:nd]}\right\}}{n} \\ &= \lim_{n\to\infty} d\left\{\frac{-\log\Pr\left\{X_{[1:nd]} = x_{[1:nd]}\right\}}{nd}\right\} \\ &= -d\sum_{i,j\in\mathcal{X}} \pi_i p_{ij}\log p_{ij}, \end{aligned}$$

that is, $H_W = dH$.

By Theorem 2, if $C_W \geq H_W + 1$, that is, if $C \geq H + \frac{1}{d}$ and $d > 0$, a target is d-trackable.

V. Concluding Remarks

In this paper, we characterized the number of queries required to follow, track, and d-track a target that moves according to a Markov chain. Necessary and sufficient conditions have been presented for all cases, as well as corresponding following, tracking, and d-tracking strategies. Throughout this paper we have associated the state of the Markov chain with the target *location*. This association was merely for ease of presentation. Indeed, in addition to location, the state of the Markov chain could include information about the velocity of the target, as well as other dynamic quantities needed to be tracked (so long as their evolution can be represented as a Markov chain).

It is of interest to extend our results to the multi-target scenario, as well as to consider the case where sensors are faulty (i.e., their query responses may be wrong), and where noise is present in the communication between sensors and tracker. In this direction, it would be natural to introduce the notion of distance between sensors (states) and analyze tracking performance under criteria such as the mean squared error. We conjecture that results related to rate-distortion theory are possible. Another interesting variation is to take sensor responses to be the number of sensors that reply "yes" to a query. Future work also includes investigating the mean number of time steps (in terms of number of queries) involved in the *catch-up* strategy before the target track can be successfully determined, i.e., the mean *lag* time. Although the simplicity of the catch-up strategy is particularly attractive, it is of interest to find the strategy that incurs the minimum lag time.

Appendix
Proof of Lemma 1

Proof: To show part (a), consider $-\log p(x_{[1:t]})$. It can be bounded from above as follows:

$$\begin{aligned} -\log p(x_{[1:t]}) &= \sum_{k=1}^{t} -\log p(x_{k+1}|x_k) - \log p_{x_1} \\ &= \sum_{i,j} N_{ij}(x_{[1:t]})(-\log p_{ij}) - \log p_{x_1} \\ &< t\sum_{i,j}(\pi_i p_{ij} + \delta)(-\log p_{ij}) - \log p_{x_1} \\ &= t\Big[\sum_{i,j}(\pi_i p_{ij})(-\log p_{ij}) - \log p_{x_1} \\ &\quad + \sum_{i,j}\delta(-\log p_{ij})\Big] - \log p_{x_1} \\ &= t(H + c_1\delta) - \log p_{x_1} \end{aligned}$$

where $c_1 = \sum_{x,y} -\log p_{ij}$.

Similarly, we can bound it from below:

$$t(H - c_1\delta) - \log p_{x_1} < -\log p(x_{[1:t]}).$$

Therefore, we have

$$2^{-t(H+c_1\delta)}p_{x_1} < \Pr\{X_{1:t]} = x_{[1:t]}\} < 2^{-t(H-c_1\delta)}p_{x_1}.$$

This concludes the proof of part (a).

To prove part (b), we use a result from large deviation theory as stated in the following proposition [29].

Proposition 1: Suppose that $\{X_t\}$ is an ergodic finite state Markov chain with state space $\mathcal{X}$, and let b_t denote its L^1 convergence parameter:

$$b_t = \sup_i \sup_j |p_{ij}^{(t)} - \pi_j|.$$

Then, the series $b = \sum_{t\geq 0} b_t$ converges, and for any bounded function $F : \mathcal{X} \to \mathbb{R}$ and any $\delta > 0$ we have,

$$\log \Pr\Big\{\frac{\sum_{k=1}^t F(X_k)}{t} - \pi(F) \geq \delta\Big\} \leq -\frac{t-1}{2}\Big(\frac{\delta}{b\bar{F}} - \frac{3}{t-1}\Big)^2,$$

as long as $t \geq 1 + 3b\bar{F}/\delta$, where $\bar{F} = \max_x |F(x)|$ and $\pi(F)$ is the mean of the function F with respect to the stationary distribution π of the Markov chain, that is, $\pi(F) = \sum_{i\in\mathcal{X}} F(i)\pi_i$.

To apply Proposition 1 and prove part (b), we first construct an ergodic finite state Markov chain $\{Y_t\}$ from the original Markov chain $\{X_t\}$ by taking $Y_t = (X_{t-1}, X_t)$. It is well known that $\{Y_t\}$ is again an ergodic finite state Markov chain with state space $\mathcal{Y} = \mathcal{X}^2$. It is easy to verify that the stationary distribution λ_{ij} of $\{Y_t\}$ equals to $\pi_i p_{ij}$, $(i,j) \in \mathcal{X}^2$. Let b_t be the L^1 convergence parameter sequence of $\{Y_t\}$ and $b = \sum_{t\geq 0} b_t$. Let function $F(Y_t)$ be the indication function $\mathbf{1}_{ij}(Y_t)$. Note that F is bounded and $\bar{F} = \sup_j |F(j)| \leq 1$. Applying Proposition 1 to function F, we have,

$$\begin{aligned}\log \Pr\Big\{\frac{\sum_{k=1}^t \mathbf{1}_{ij}(Y_t)}{t} - \lambda_{ij} \geq \delta\Big\} &\leq -\frac{t-1}{2}\Big(\frac{\delta}{b} - \frac{3}{t-1}\Big)^2 \\ &\leq -\frac{t-1}{2}\Big(\frac{\delta}{b} - 3\Big)^2\end{aligned}$$

holding for all $(i,j) \in \mathcal{X}^2$ for t large ($\geq 1+3b/\delta$). Writing the above inequality in terms of the counting function $N_{ij}(X_{[1:t]})$ and substituting λ_{ij} by $\pi_i p_{ij}$, we get

$$\log \Pr\Big\{\frac{N_{ij}(X_{[1:t]})}{t} - \pi_i p_{ij} \geq \delta\Big\} \leq -\frac{t-1}{2}\Big(\frac{\delta}{b} - 3\Big)^2, \quad (2)$$

holding for t large.

Similarly, applying Lemma 1 to the function $F' = 1 - F$, we get the following bound

$$\log \Pr\Big\{\frac{N_{ij}(X_{[1:t]})}{t} - \pi_i p_{ij} \leq -\delta\Big\} \leq -\frac{t-1}{2}\Big(\frac{\delta}{b} - 3\Big)^2, \quad (3)$$

holding for t large.

Combining (2) and (3), we have

$$\log \Pr\Big\{\Big|\frac{N_{ij}(X_{[1:t]})}{t} - \pi_i p_{ij}\Big| \geq \delta\Big\} \leq -(t-1)\Big(\frac{\delta}{b} - 3\Big)^2, \quad (4)$$

holding for t large. Using (4), we can now bound the probability of the complement of the strong typical set $\Delta_{t,\delta}$ as follows:

$$\begin{aligned}&\Pr\Big\{X^t \notin \Delta_{t,\delta}\Big\} \\ &= \Pr\Big\{\bigcup_{(i,j)\in\mathcal{X}^2}\Big\{\Big|\frac{N_{ij}(X_{[1:t]})}{t} - \pi_i p_{ij}\Big| \geq \delta\Big\}\Big\} \\ &\leq \sum_{(i,j)\in\mathcal{X}^2} \Pr\Big\{\Big|\frac{N_{ij}(X_{[1:t]}}{t} - \pi_i p_{ij}\Big| \geq \delta\Big\} \\ &\leq |\mathcal{X}|^2 2^{-(t-1)(\frac{\delta}{b}-3)^2} \\ &\leq c_2 2^{-c_3 t}\end{aligned}$$

for t large, where the constants $c_2 = |\mathcal{X}|^2 2^{(\frac{\delta}{b}-3)^2}$ and $c_3 = (\frac{\delta}{b} - 3)^2$. This concludes the proof of the lemma. ■

REFERENCES

[1] S. M. Ulam, *Adventures of a Mathematician.* New York: Scribner, 1976.

[2] A. Rényi, "On a problem of information theory," *MTA Mat. Kut. Int. Kozl.*, vol. 6B, pp. 505–516, 1961.

[3] P. Erdös and A. Rényi, "On two problems of information theory," *MTA Mat. Kut. Int. Kozl.*, vol. 8A, pp. 229–243, 1963.

[4] M. Hewish, "Reforming fighter tactics," *Jane's International Defense Review*, June 2001.

[5] J. Carle and D. Simplot-Ryl, "Energy-efficient area monitoring for sensor networks," *IEEE Computer*, vol. 37, no. 2, pp. 40–46, 2004.

[6] S. Ganeriwal, R. Kumar, and M. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. ACM Conference on Embedded Networked Sensor Systems (SENSYS)*, Los Angeles, California, November 2003, pp. 138–149.

[7] J. Elson and D. Estrin, "Time synchronization for wireless sensor networks," in *Proceedings of the* 15^{th} *International Parallel and Distributed Processing Symposium*, San Francisco, California, April 2001, p. 186.

[8] R. Iyengar and B. Sikdar, "Scalable and distributed GPS free positioning for sensor networks," in *Proc. IEEE International Conference on Communications(ICC)*, Anchorage, Alaska, May 2003, pp. 338–342.

[9] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low-cost outdoor localization for very small devices," *IEEE Pers. Commun.*, vol. 7, no. 5, pp. 28–34, 2000.

[10] D. Tian and N. Georganas, "Energy efficient routing with guaranteed delivery in wireless sensor networks," in *Proc. IEEE Wireless Communications and Networking Conference(WCNC)*, New Orleans, Louisiana, March 2003, pp. 1923–1929.

[11] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, "Highly-resilient, energy-efficient multipath routing in wireless sensor networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 4, pp. 11–25, 2001.

[12] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1196–1201, July 2004.

[13] S. Tatikonda, A. Sahai, and S. Mitter, "Control of LQG systems under communication constraints," in *Proc. 1999 American Control Conference*, vol. 4, San Diego, CA, June 1999, pp. 2778–2782.

[14] ——, "Control of LQG systems under communication constraints," in *Proc.* 37^{th} *IEEE Conference on Decision and Control*, vol. 1, Tampa, FL, December 1998, pp. 1165–1170.

[15] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1056–1068, July 2004.

[16] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link—Part I: Scalar systems," *IEEE Trans. Inform. theory*, vol. 52, pp. 3369–3395, 2006.

[17] ——, "Source coding and channel requirements for unstable processes," submitted. [Online]. Available: http://arxiv.org/abs/cs.IT/0610143

[18] ——, "The necessity and sufficient of anytime capacity for stabilization of a linear system over a noisy communication link: Part II: vector systems," Submitted. [Online]. Available: http://arxiv.org/abs/cs.IT/0610146

[19] J. Aslam, Z. Butler, F. Constantin, V. Crespi, G. Cybenko, and D. Rus, "Tracking a moving object with a binary sensor network," in *Proc. ACM Conference on Embedded Networked Sensor Systems (SENSYS)*, Los Angeles, California, November 2003, pp. 150–161.

[20] J. Liu, P. Cheung, L. Guibas, and F. Zhao, "A dual-space approach to tracking and sensor management in wireless sensor networks," in *Proc. 1^{st} ACM International Workshop on Wireless Sensor Networks and Applications*, Atlanta, Georgia, April 2002, pp. 131–139.

[21] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for target tracking," *EURASIP JASP: Special Issues on Sensor Networks*, vol. 2003, no. 4, pp. 378–391, 2003.

[22] R. Evans, V. Krishnamurthy, G. Nair, and L. Sciacca, "Networked sensor management and data rate control for tracking maneuvering targets," *IEEE Trans. Signal Processing*, vol. 53, no. 6, pp. 1979–1991, 2005.

[23] R. L. Rivest, A. R. Meyer, D. J. Kleitman, K. Winkimann, and J. Spencer, "Coping with errors in binary search procedures," *J. Computer and System Sciences*, vol. 20, no. 3, pp. 396–404, 1980.

[24] V. S. Borkar, S. K. Mitter, A. Sahai, and S. Tatikonda, "Sequential source coding: an optimization viewpoint," in *Proc. 44^{th} IEEE Conference on Decision and Control, and the European Control Conference*, Seville, Spain, December 2005, pp. 1035–1042.

[25] E. Çinlar, *Introduction to Stochastic Processes*. Prentice-Hall, 1975.

[26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

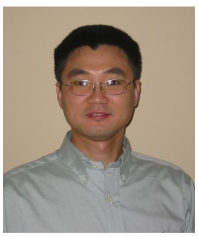
[27] R. W. Yeung, *A First Course in Information Theory*. Kluwer Academic, 2002.

[28] D. Knuth, *The Art of Computer Programming, Vol. 3: Sorting and Searching*. Addison-Wesley, 1997.

[29] L. L. M. I. Kontoyiannis and S. Meyn, "Relative entropy and exponential deviation bounds for general markov chains," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Adelaide, Australia, December 2005, pp. 1563–1567.

[30] P. Billingsley, *Probability and Measure*. Wiley-Interscience, 1995.

[31] S. Meyn and R. Tweedie, *Markov Chain and Stochastic Stability*, ser. Communications and Control Engineering Series. Springer-Verlag, 1993.



Hua Li received the B.A. from Civil Aviation University of China in 1996, the M.S. from Beijing University of Posts and Te lecommunications in 1999, and the Ph.D. degree from Colorado State University in 2007, all in electrical engineering. From 2007, he joined Fair Isaac Corporation as a scientist. His research interests include information theory, Markov deci sion process, partially observable Markov decision process, modeling, and machine learning algorithms.

Patricia Barbosa received the B.Sc. and M.Sc. degrees from the State University of Campinas, Brazil in 1999 and 2000, respectively, and the M.S. degree from the University of Southern California in 2002. She is currently a Ph.D. candidate at the Department of Electrical and Computer Engineering, Colorado State University. She is also with Intermap Technologies, Denver, Colorado, working on remote sensing.

Edwin K. P. Chong received the B.E.(Hons.) degree with First Class Honors from the University of Adelaide, South Australia, in 1987; and the M.A. and Ph.D. degrees in 1989 and 1991, respectively, both from Princeton University, where he held an IBM Fellowship. He joined the School of Electrical and Computer Engineering at Purdue University in 1991, where he was named a University Faculty Scholar in 1999, and was promoted to Professor in 2001. Since August 2001, he has been a Professor of Electrical and Computer Engineering and a Professor of Mathematics at Colorado State University. His current interests are in communication networks and optimization methods. He coauthored the recent best-selling book, An Introduction to Optimization, 2nd Edition, Wiley-Interscience, 2001. He was on the editorial board of the IEEE Transactions on Automatic Control, and is currently an editor for Computer Networks and the Journal of Control Science and Engineering. He is a Fellow of the IEEE, and served as an IEEE Control Systems Society Distinguished Lecturer. He received the NSF CAREER Award in 1995 and the ASEE Frederick Emmons Terman Award in 1998. He was a co-recipient of the 2004 Best Paper Award for a paper in the journal Computer Networks. He has served as Principal Investigator for numerous funded projects from DARPA and other defense funding agencies.

Jan Hannig received the Mgr. (M.S. equivalent) degree in mathematics Cumma Sum Laude from Charles University, Prague, Czech Republic in 1996; and the Ph.D. degree in statistics from Michigan State University in 2000. He joined the Department of Statistics at Colorado State University in 2000, and was promoted to Associate Professor in 2006. He will be joining the Department of Statistics and Operations Research at the University of North Carolina at Chapel Hill as an Associate Professor in the Summer of 2008. He is an elected member of the International Statistical Institute. He has served as Principal Investigator for several funded projects from National Science Foundation.

Sanjeev Kulkarni received the B.S. in Mathematics, B.S. in Electrical Engineering, M.S. in Mathematics from Clarkson University in 1983, 1984, and 1985, respectively, the M.S. degree in Electrical Engineering from Stanford University in 1985, and the Ph.D. in Electrical Engineering from M.I.T. in 1991. From 1985 to 1991 he was a Member of the Technical Staff at M.I.T. Lincoln Laboratory. Since 1991, he has been with Princeton University where he is currently Professor of Electrical Engineering and Master of Butler College. Prof. Kulkarni received an ARO Young Investigator Award in 1992, and an NSF Young Investigator Award in 1994. He has also received several teaching awards at Princeton University, including four awards from the Undergraduate Engineering Council for a Machine Vision course taught in Fall 1991, an Image Processing course taught in Spring 1992, and an Introduction to Electrical Signals and Systems course taught in Fall 1999 and Fall 2000. Prof. Kulkarni is a Fellow of the IEEE, and he has served as an Associate Editor for the IEEE Transactions on Information Theory. Prof. Kulkarni's research interests include statistical pattern recognition, nonparametric estimation, learning and adaptive systems, information theory, wireless networks, signal/image/video processing, and econometrics and finance.