Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Distributed Bayesian Hypothesis Testing in Sensor Networks

M. Alanyali S. Venkatesh O. Savas S. Aeron

Abstract— **We consider the scenario of N distributed noisy sensors observing a single event. The sensors are distributed and can only exchange messages through a network. The sensor network is modelled by means of a graph, which captures the connectivity of different sensor nodes in the network. The task is to arrive at a consensus about the event after exchanging such messages. The focus of this paper is twofold: a) characterize conditions for reaching a consensus; b) derive conditions for when the consensus converges to the centralized MAP estimate.**

The novelty of the paper lies in applying belief propagation as a message passing strategy to solve a distributed hypothesis testing problem for a pre-specified network connectivity. We show that the message evolution can be re-formulated as the evolution of a linear dynamical system, which is primarily characterized by network connectivity. This leads to a fundamental understanding of as to which network topologies naturally lend themselves to consensus building and conflict avoidance.

I. Introduction

Recent advances in sensor and computing technologies [11] provide impetus for deploying wireless sensor networks—a network of massively distributed tiny devices capable of sensing, processing and exchanging data over a wireless medium. Such networks are envisioned [11] to provide real-time information in such diverse applications as building safety, environmental control, power systems and manufacturing. For instance, as part of a building safety system an ad-hoc sensor network(SNET), may monitor in real time hot spots, smoke, biological and chemical contaminants, structural failures, interference sources to provide 3D building visualization to enable rapid evacuation and rescue of victims and personnel.

SNETs have received significant attention within the networking, signal processing and information-theory communities [2], [1]. The networking community has largely addressed the problem from the perspective of ad-hoc networks and generally ignored the distinction between data and information. This distinction is critical to efficient operation of a network as was pointed out earlier. The signal processing [5], [16], [1], [14], [13], [12] and network information theory [3], [15], [19], [8] communities have addressed the problem from the perspective of characterizing fundamental bounds for information quality in a distributed, bandwidth limited but yet highly coordinated and synchronized environments, i.e., as to how to transmit data from known sensors to known destinations.

This work was supported by ONR Young Investigator Award N00014-02-100362 and NSF CAREER Program under grant ANI-0238397.

Department of Electrical and Computer Engineering, and The Center for Information and Systems Engineering, Boston University, 8 Saint Mary's Street, Boston, MA 02215, USA. {alanyali, srv, savas, shuchin}@bu.edu

In this paper we deal with situations where such coordination and synchronization between different sensors is unavailable. We focus on the scenario of N distributed noisy sensors observing a single event. The sensors are distributed and can only exchange messages through a network. The sensor network is modelled by means of a graph, which captures the connectivity of different sensor nodes in the network. The task is to arrive at a consensus about the event after exchanging such messages. If the observations are centrally available there is a well-established solution methodology for such problems [20]. Fundamental problems arise when data is distributed and the centralized solutions are no longer feasible due to time and rate constraints(finite bit budget). This question of as to how to deal with distributed data from known sensors to known destinations has been an active topic of research within the control [23], [22], [4], [21], [9], [7] signal processing communities [6], [3], [6], [15], [19], [8], [18]. Unfortunately, despite best efforts a satisfactory theory for distributed problems is yet to emerge. The reasons can be broadly traced to conflicting requirements together with the generality of the setup. Information theoretic approaches can provide bounds in the limit of infinite delay and infinite number of sensors [3], [15]. However, real-time decisions and control cannot generally tolerate large delays leading to a fundamental gap between these different perspectives.

Motivated by these issues we focus on a scenario where N distributed sensors observe an event at a single location. The task is to classify this object among M different hypotheses. Unlike the traditional settings our attention is limited to: a) fixed network topology that provides a fixed routing mechanism between different sensor nodes; b) *informative* data as opposed to decisions are exchanged between different sensor nodes. A natural mechanism for exchanging informative data is provided by the "so called" belief propagation algorithm [17]. The main idea can be explained as follows: A sensor node j sending messages to node k summarizes information received from all the other nodes it is connected to and forwards this information to node k. The messages can be interpreted as a node's conditional marginal distribution. A node k upon reception of this message then updates its posterior probability, which is usually called as the belief. The process continues with node k updating its messages to be sent to its neighbors and so on. The organization of the paper is as follows. In Section II we provide a description of the problem setup. In Section III belief propagation algorithm is discussed and Section IV provides the main results.

0-7803-8335-4/04/$17.00 ©2004 AACC

II. SETUP

Consider a Bayesian hypothesis testing problem in which $\mathcal{H} = \{H_1, H_2, \cdots, H_M\}$ denotes the set of hypotheses and π_o is the prior probability distribution on $\mathcal{H}$. We are interested in the MAP estimate of the true hypothesis based on a collection $(X_v : v \in V)$ of observations that are indexed by a set V of sensors. For each $m = 1, 2, \cdots, M$, let $f_m : \mathbf{R}^V \mapsto \mathbf{R}_+$ be the conditional probability density function of $(X_v : v \in V)$ given that H_m is the true hypothesis. We shall assume that observations are conditionally independent given the true hypothesis. That is,

$$f_m(x) = \prod_{v \in V} f_m^v(x_v), \qquad x = (x_v : v \in V) \in \mathbf{R}^V$$

for marginal densities $f_m^v : \mathbf{R} \mapsto \mathbf{R}_+,\ v \in V$. Given $X_v = x_v$ for $v \in V$, the posterior distribution π of the true hypothesis is identified uniquely by the relation

$$\pi(H_m) \propto \pi_o(H_m) \prod_{v \in V} f_m^v(x_v), \quad m = 1, 2, \cdots, M. \quad (1)$$

In particular hypothesis H_{m^*} is a MAP estimate if

$$\pi_o(H_{m^*}) \prod_{v \in V} f_{m^*}^v(x_v) = \max_m \left\{ \pi_o(H_m) \prod_{v \in V} f_m^v(x_v) \right\}.$$

We concentrate on distributed applications in which a single decision maker that has access to all observations $(X_v : v \in V)$ is not available. Instead, it is assumed that each sensor can communicate with a certain subset of other sensors, and thereby forms an estimate of the posterior distribution π based on both its own observation and its prior correspondence with its neighbors. The objective of the paper is to identify communication schemes which guarantee that each sensor eventually identifies a MAP estimate. Furthermore applications of interest concern vast numbers of sensors; in turn non-scalable schemes such as simple flooding of observations are excluded from the present discussion. Specifically, we examine the performance of Pearl's belief propagation algorithm [17], which is subject to recent interest in similar statistical inference problems.

The communication structure among the sensors is represented via a directed graph $G = (V, E)$. The vertices V of this graph correspond to sensors, and an ordered pair (v', v) of vertices belongs to the edge set E if there exits a unidirectional communication link from sensor v' to sensor v. We will identify each edge $e \in E$ with its source vertex $s(e)$ and its destination vertex $d(e)$ so that $e = (s(e), d(e))$. Sensor v' is referred to as a *neighbor* of sensor v if there is a link from vertex v' to vertex v in G, so that sensor v' can send a message to sensor v. Let $N(v)$ denote the set of neighbors of sensor v so that

$$N(v) = \{v' \in V : (v', v) \in E\}, \qquad v \in V.$$

The communication graph G is not required to bear any relationship to the underlying statistical model.

III. BELIEF PROPAGATION

We start with a brief digression to statistical inferencing via belief propagation [17] in order to motivate the distributed message passing algorithm adopted here. Let $(Y_v : v \in V)$ be a random vector with values in $\mathcal{Y}^V$, and for certain mappings $\phi_v : \mathcal{Y} \mapsto \mathbf{R}_+,\ v \in V$, and $\psi_e : \mathcal{Y}^2 \mapsto \mathbf{R}_+,\ e \in E$, let the distribution of $(Y_v : v \in V)$ satisfy

$$P(Y_v = y_v : v \in V) \propto \prod_{v \in V} \phi_v(y_v) \prod_{e \in E} \psi_e(y_{s(e)}, y_{d(e)}), \quad (2)$$

for $y_v \in \mathcal{Y},\ v \in V$. Such graphical models arise in a variety of contexts where efficient computation of marginal distributions $P(Y_v = y_v),\ v \in V$, is of interest. Let the undirected graph $\tilde{G} = (V, \tilde{E})$ be defined so that the unordered pair $[v, v'] \in \tilde{E}$ if and only if $(v, v') \in E$ or $(v', v) \in E$. It is well-known that if $\tilde{G}$ is a tree, then local message passing via Pearl's sum-product algorithm [17] results in distributed, local computation of the marginal distributions. Namely, let the kth message sent from sensor $v' \in N(v)$ to sensor v be the vector $m_k^{(v',v)} = (m_k^{(v',v)}(y) : y \in \mathcal{Y})$ defined by

$$\begin{aligned} m_0^{(v',v)}(y) &= 1 \\ m_k^{(v',v)}(y) &= \sum_{y' \in \mathcal{Y}} \psi_{(v',v)}(y', y)\phi_{v'}(y') \prod_{\hat{v} \in N(v') - \{v\}} m_{k-1}^{(\hat{v},v')}(y'), \end{aligned}$$

$k \geq 1$, and suppose that upon receiving the kth messages from all of its neighbors, each sensor $v \in V$ constructs an estimate $\hat{\pi}_k^v$ of the local marginal distribution by setting

$$\hat{\pi}_k^v(y) \propto \phi_v(y) \prod_{v' \in N(v)} m_k^{(v',v)}(y), \qquad y \in \mathcal{Y}. \quad (3)$$

Then $\hat{\pi}_k^v$ converges to the correct marginal distribution $(P(Y_v = y) : y \in \mathcal{Y})$ within a finite number of steps, provided that $\tilde{G}$ is a tree.

Consider now the special case

$$\begin{aligned} \mathcal{Y} &= \mathcal{H} \\ \phi_v(H_m) &= f_m^v(x_v) \sqrt[|V|]{\pi_o(H_m)}, \quad (4) \\ \psi_e(H_j, H_m) &= \mathbf{1}\{j = m\}, \qquad j, m = 1, 2, \cdots, M, \end{aligned}$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function whose value is 1 if its argument is correct and is 0 otherwise. It is straightforward to verify that equality (2) reduces to

$$P(Y_v = y_v : v \in V) \propto \mathbf{1}\{y_v = y_{v'} \text{ for all } v, v' \in V\}\pi(y_{v^*}),$$

where π is given by relation (1) and $v^* \in V$ is arbitrary. In particular $P(Y_v = Y_{v'} \text{ for all } v, v' \in V) = 1$ and each marginal Y_v has distribution π. The sum-product algorithm now prescribes the message composition

$$\begin{aligned} m_0^{(v',v)}(h) &= 1 \quad (5) \\ m_k^{(v',v)}(h) &= \phi_{v'}(h) \prod_{\hat{v} \in N(v') - \{v\}} m_{k-1}^{(\hat{v},v')}(h), \quad (6) \end{aligned}$$

for each $h \in \{H_1, H_2, \cdots, H_m\}$, $k \geq 1$. From the prior discussion it is clear that if $\tilde{G}$ is a tree, then the algorithm assures that $\hat{\pi}_k^v = \pi$ for large enough k, hence each sensor can identify a MAP estimate based on the global observation set. The message passing algorithm (5)–(6) has an evident practical appeal: Each message is determined locally by the observation at the sensor and the prior messages received from neighboring sensors. Furthermore, the algorithm entails a relaxed synchronization among sensors, as it can be implemented by programming each sensor to send out initial messages immediately and to send out its kth messages only after receiving $(k-1)$th messages from all of its neighbors. However, asymptotic features of the sum-product in general topologies of G are not well-understood. In fact, there is ample evidence that the algorithm may, in general, fail to converge, or may converge to an inaccurate estimate of the marginal distributions. We next address these issues in the particular instantiation that pertains to the present hypothesis testing problem, and give an account of the asymptotic behavior of the estimates $\hat{\pi}_k^v : v \in V$ for general graphs.

IV. Main Result

For each pair of edges $e, e' \in E$ let

$$a_{e,e'} = \mathbf{1}\{d(e') = s(e),\ s(e') \neq d(e)\}.$$

Note that $a_{e,e'} = 1$ if and only if edge e' leads to the origin of edge e but the ordered pair (e', e) is not a directed cycle. For each hypothesis $h \in \{H_1, H_2, \cdots, H_m\}$ let

$$\begin{aligned} u^h(v) &= \log(\phi_v(h)), \qquad v \in V \\ x_k^h(e) &= \log(m_k^e(h)), \qquad e \in E. \end{aligned}$$

Taking the logarithm of both sides in equalities (5)–(6) leads to the linear system

$$x_k^h(e) = u^h(s(e)) + \sum_{e' \in E} a_{e,e'} x_{k-1}^h(e'), \qquad x_0^h(e) = 0. \quad (7)$$

Define the vector $u^h = (u^h(s(e)) : e \in E)$ and define the binary matrix $A = [a_{e,e'}]_{E \times E}$, so that equality (7) takes the vector form

$$x_k^h = u^h + A x_{k-1}^h, \qquad x_0^h = 0.$$

Note in particular that

$$x_k^h = \sum_{j=0}^{k-1} A^j u^h, \qquad k \geq 1 \quad (8)$$

and $A^j = [a_{e,e'}^j]_{E \times E}$ where $a_{e,e'}^j$ is the number of directed paths of length j that start with edge e', end with edge e, and that do not have any 2-hop cycles.

Suppose that A is primitive. For each $e, e' \in E$ there exists a directed path that starts with edge e', ends with edge e and that does not have any 2-hop cycles. The spectral radius of A, denoted here by $\rho(A)$, is then strictly larger than 1. Let $(r_e : e \in E)$ and $(l_e : e \in E)$ be respectively a right and a left eigenvector of A corresponding to the eigenvalue $\rho(A)$, suitably normalized so that $r_e > 0$, $l_e > 0$ for $e \in E$ and $\sum_{e \in E} r_e l_e = 1$. Define the *weighted in-degree* $i(v)$ and the *weighted out-degree* $o(v)$ of each sensor $v \in V$ as

$$\begin{aligned} i(v) &= \sum_{v' \in N(v)} r_{(v',v)} \\ o(v) &= \sum_{v' : v \in N(v')} l_{(v,v')}. \end{aligned}$$

Theorem 4.1: If A is primitive then for $v \in V$

$$\lim_{k \to \infty} \hat{\pi}_k^v(H_m) = 0, \quad (9)$$

for each hypothesis $m \in \{1, 2, \cdots, M\}$ such that

$$\prod_{v \in V} \phi_v(H_m)^{o(v)} < \max_{m'} \left\{ \prod_{v \in V} \phi_v(H_{m'})^{o(v)} \right\}.$$

Proof. Define the matrix $W = [w_{e,e'}]_{E \times E}$ by setting $w_{e,e'} = r_e l_{e'}$ for $e, e' \in E$. Let

$$\alpha(k) = \sum_{j=0}^{k-1} \rho(A)^j.$$

By equality (8)

$$\lim_{k \to \infty} \frac{x_k^h}{\alpha(k)} = \lim_{k \to \infty} \sum_{j=0}^{k-1} \left(\frac{A^j}{\rho(A)^j} \right) \frac{\rho(A)^j}{\alpha(k)} u^h = W u^h,$$

where the last equality follows since $\rho(A) > 1$ and

$$\lim_{k \to \infty} \left\| \rho(A)^{-k} A^k - W \right\|_\infty = 0$$

due to [10, Theorem 8.5.1]. The estimate $\hat{\pi}_k^v(h)$ of sensor v at step k therefore satisfies

$$\begin{aligned} \hat{\pi}_k^v(h) &\propto \phi_v(h) \exp\left(\sum_{v' \in N(v)} x_k^h(v', v) \right) \\ &= \phi_v(h) \exp\left(\alpha(k) \left(i(v) \sum_{v' \in V} u^h(v') o(v') + \varepsilon(k) \right) \right), \end{aligned}$$

where $\varepsilon(k) \to 0$ as $k \to \infty$. The conclusion of the theorem now follows since

$$\phi_v(h) \exp\left(\alpha(k) i(v) \sum_{v' \in V} u^h(v') o(v') \right) = \phi_v(h) \left(\prod_{v' \in V} \phi_{v'}(h)^{o(v')} \right)^{\alpha(k) i(v)}$$

and $\lim_{k \to \infty} \alpha(k) = \infty$ owing to $\rho(A) > 1$.

We now turn to distributed hypothesis testing and interpret Theorem 4.1 in terms of the posterior distribution π. Consider first a symmetric structure for the graph G so that $o(v) = o(v') = \mu > 0$ for all $v, v' \in V$. It can be verified, for example, that the torus of Figure 1 provides one such

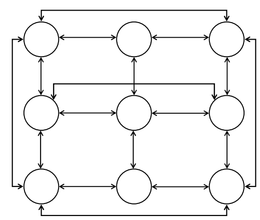


Fig. 1. The 3×3 torus in which $o(v) = o(v')$ for all nodes v, v'.

structure. The definition (4) of node potentials $\phi_v : v \in V$ then leads to

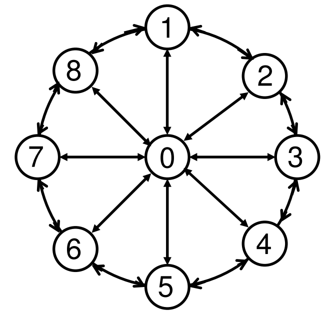
$$\prod_{v \in V} \phi_v(H_m)^{o(v)} = \left(\pi_o(H_m) \prod_{v \in V} f_m^v(x_v) \right)^{\mu},$$

for $m = 1, 2, \cdots, M$, and equality (9) indicates that for each sensor v, $\lim_{k \to \infty} \hat{\pi}_k^v(H_m) = 0$ for all m such that H_m is not a MAP estimate with respect to π. In particular if the MAP estimate is unique, then the sensors unanimously identify it. This conclusion does not hold for general graphs whose vertices may have different weighted out-degrees. In fact it is easy to see that observations made at sensors with larger weighted out-degrees have more influence on the collective opinion. In the general case the final consensus reflects the right MAP estimate for certain values of observations $(x_v : v \in V)$, but identifies wrong choices for others.

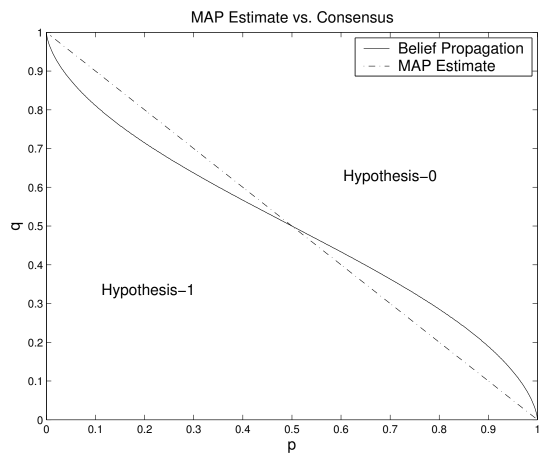
Example 4.1: The example illustrates a 9-node asymmetric communication graph as depicted in Figure 2(a). Here node 0 communicates with all other nodes, whereas every other node communicates with three nodes. The Figure 2(b) illustrates how a single highly connected node in an asymmetric graph can bias the consensus decision away from the optimal.

Each edge in the graph represents two directed edges in opposite directions. The weighted out-degrees of the sensors with respect to this graph satisfy $o(1) = o(2) = \cdots = o(8)$ and $o(0)/o(1) = 1.5091$. Suppose that the observations $(x_v : v \in V)$ translate to node potentials $\phi_0 = [q, 1-q]$, $\phi_1 = [p, 1-p]$ and $\phi_v = [0.5, 0.5]$ for $v = 2, 3, \cdots, 8$, where $p, q \in [0, 1]$. Figure 2(b) illustrates the true MAP estimate and the final consensus due to belief propagation for different values of p and q. Note that the consensus is determined to a larger extent by the value of q rather than the value of p. Note also that the consensus reflects a flawed estimate if (p, q) lies in the area between the solid and dashed lines. However, it is also surprising that the bias is small even with a high degree of asymmetry. □

We next illustrate the asymptotic behavior of the estimates $\hat{\pi}_k^v : v \in V$ in two topologies for which A is reducible. In the scope of the following three examples, it is understood that for vertices $v, v' \in V$, $(v, v') \in E$ if and only if $(v', v) \in E$. Note that the undirected graph $\tilde{G} = (V, \tilde{E})$ is such that the unordered pair $[v, v'] \in \tilde{E}$ if and only if $(v, v') \in E$ and $(v', v) \in E$. The first example concerns the case when $\tilde{G}$ is a tree, and it is well-known that in this case belief propagation leads to the true posterior distributions for general Markov fields [17].



(a)



(b)

Fig. 2. The figure illustrates how a single highly connected node in an asymmetric graph can bias the consensus decision away from the optimal. (a) The 9-node asymmetric communication graph where node 0 communicates with all other nodes, whereas every other node communicates with three nodes, (b) Decision regions for the MAP estimate and the final consensus of belief propagation, delineated respectively by the dashed line and the solid line. Observations $(x_v : v \in V)$ correspond to node potentials $\phi_0 = [q, 1-q]$, $\phi_1 = [p, 1-p]$ and $\phi_v = [0.5, 0.5]$ for $v = 2, 3, \cdots, 8$.

Example 4.2: (Trees) Suppose that $\tilde{G}$ is a tree, so that A is nilpotent since $A^j = 0$ for all integers j larger than the diameter of $\tilde{G}$. Equality (8) therefore indicates that the messages are guaranteed to converge within a number of steps no larger than the diameter of $\tilde{G}$. Note that for $e, e' \in E$

$$\sum_{j=0}^{\infty} a_{e,e'}^j = \begin{cases} 1 & \text{if there exists a simple directed path} \\ & \text{in } G \text{ with first edge } e' \text{ and last edge } e \\ 0 & \text{else,} \end{cases}$$

hence equality (8) leads to

$$\lim_{k\to\infty} x_k^h(e) = \sum_{v\in V} \mathbf{1}\{\text{dist}(v, s(e)) < \text{ dist}(v, d(e))\} u^h(v),$$

for $e \in E$, where $\text{dist}(v, v')$ represents the length of the unique path between vertices $v, v' \in V$. It now follows by equality (3) that the limit of the estimate $\hat{\pi}_k^v(h)$ at each sensor $v \in V$ is equal to the posterior distribution (1). □

Example 4.3: (Rings) Suppose that $\tilde{G}$ is a simple cycle, so that for $e, e' \in E$ the sequence $(a_{e,e'}^j : j = 0, 1, 2\cdots)$ has period $|V|$. In particular $A^j = A^{j+|V|}$ and thus A is idempotent. Equality (8) then leads to

$$\lim_{k\to\infty} \frac{x_k^h}{k} = \frac{1}{|V|} \sum_{j=0}^{|V|-1} A^j u^h.$$

It is not difficult to see that $\sum_{j=0}^{|V|-1} a_{e,e'}^j = 1$ for all edges $e, e' \in E$ that have a common orientation (that is, clockwise or counter-clockwise) and that $\sum_{j=0}^{|V|-1} a_{e,e'}^j = 0$ otherwise; in turn

$$\lim_{k\to\infty} \frac{x_k^h(e)}{k} = \frac{1}{|V|} \sum_{v\in V} u^h(v), \qquad e \in E.$$

Therefore for large k the estimate $\hat{\pi}_k^v(h)$ of each sensor $v \in V$ at step k satisfies

$$\begin{aligned}\hat{\pi}_k^v(h) &\propto \phi_v(h) \exp\left(\sum_{v'\in N(v)} x_k^h(v', v)\right) \\ &\approx \phi_v(h) \exp\left(\frac{2k}{|V|} \sum_{v'\in V} u^h(v')\right) \\ &= \phi_v(h) \left(\prod_{v'\in V} \phi_{v'}(h)\right)^{\frac{2k}{|V|}}.\end{aligned}$$

Since

$$\prod_{v\in V} \phi_v(H_m) = \pi_o(H_m) \prod_{v\in V} f_m^v(x_v), \qquad m = 1, 2, \cdots, M,$$

it follows that if H_m is not a MAP estimate with respect to π, then $\hat{\pi}_k^v(H_m) \to 0$ as $k \to \infty$. Note that if π leads to a unique MAP estimate m^* then the estimate distribution $\hat{\pi}_k^v$ of each sensor $v \in V$ converges so that

$$\lim_{k\to\infty} \hat{\pi}_k^v(H_m) = \mathbf{1}\{m = m^*\}, \qquad m = 1, 2, \cdots, M.$$

In other words each sensor identifies the MAP estimate, although the limit of $\hat{\pi}_k^v$ is not necessarily the correct posterior distribution. If the MAP estimate is not unique, then convergence of $\hat{\pi}_k^v$ may not hold, as illustrated by Example 4.4. □

Example 4.4: (4-ring) Consider a binary hypothesis testing problem involving 4 sensors arranged on a ring as shown in Figure 3.

Suppose that the observations $(x_v : v \in V)$ translate to node potentials $\phi_0 = [0.65, 0.35]$, $\phi_1 = [0.3, 0.7]$, $\phi_2 = [0.5, 0.5]$, $\phi_3 = [0.55, 0.45]$, so that H_1 is the unique MAP estimate in the centralized solution. The decentralized solution identifies the same MAP estimate since the estimate of the posterior distribution at each sensor converges to $(0, 1)$ as illustrated in Figure 4(a). If the node potentials are $\phi_0 = [0.7, 0.3]$, $\phi_1 = [0.4, 0.6]$, $\phi_2 = [0.6, 0.4]$, $\phi_3 = [0.3, 0.7]$, then the posterior distribution π assigns equal probabilities to both hypotheses. In this case the decentralized beliefs $\hat{\pi}_k^v$ display oscillations around the correct probabilities as shown in Figure 4(b). □

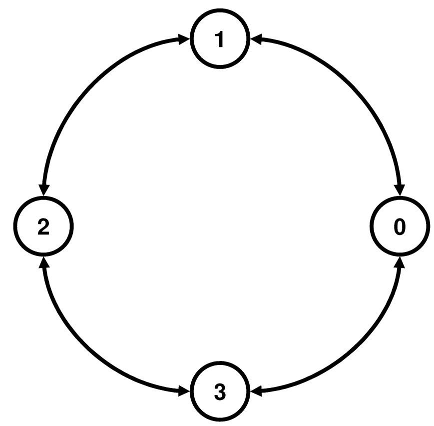


Fig. 3. A ring topology of a sensor network

V. Conclusion

We have considered the scenario of N distributed noisy sensors observing a single event. The sensors are distributed and can only exchange messages through a network. The sensor network is modelled by means of a graph, which captures the connectivity of different sensor nodes in the network. The task is to arrive at a consensus about the event after exchanging such messages. The paper focuses on characterizing the fundamental conditions required to reach a consensus.

The novelty of the paper lies in applying belief propagation as a message passing strategy to solve a distributed hypothesis testing problem for a pre-specified network connectivity. We show that the message evolution can be reformulated as the evolution of a linear dynamical system, which is primarily characterized by network connectivity. This leads to a fundamental understanding of as to which network topologies naturally lend themselves to consensus building and conflict avoidance.

References

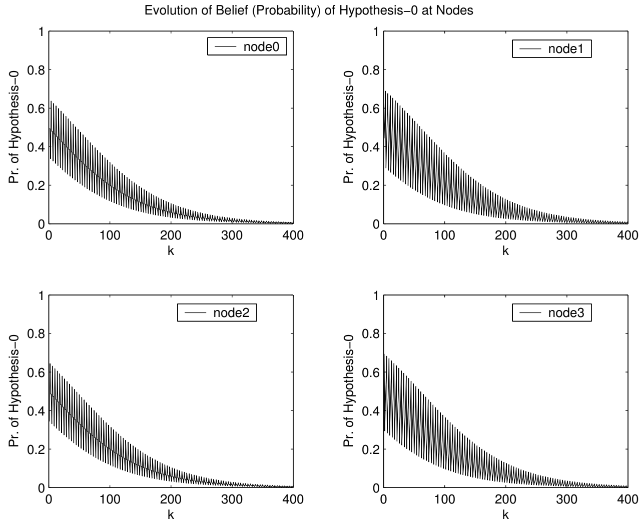
[1] Collaborative information processing. *IEEE Signal Processing Magazine*, 2002.

[2] I.F. Akyildiz, W. Su, and Y. Sankarasubramaniam. A survey on sensor networks. *IEEE Communications Magazine*, August 2002.

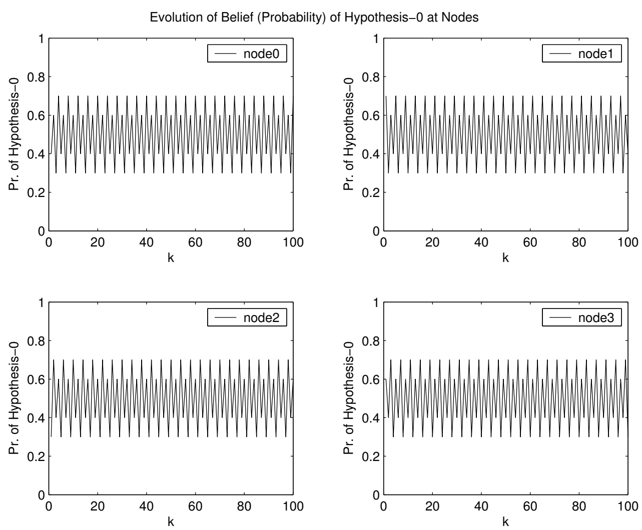
[3] T. Berger, Z. Zhang, and H. Vishwanathan. The ceo problem. *IEEE Transactions on Information Theory*, May 1996.

[4] V. Borkar and P. Varaiya. Asymptotic agreement in distributed estimation. *IEEE Transactions on Automatic Control*, 27, 1982.

[5] J. F. Chamberland and V. V. Veeravalli. Decentralized detection in sensor networks. *IEEE Transactions on Signal Processing*, 2003.



(a)



(b)

Fig. 4. Evolution of beliefs in a ring of 4 sensors. (a) Depicts scenario where there is a unique MAP (centralized) estimate; (b) Depicts scenario when multiple MAP estimates exist.

[6] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley Series in Communications, 1991.

[7] Editor F. A. Sadjadi. *Selected Papers on Sensor and Data Fusion*. SPIE Milestone Series, 1996.

[8] Michael Gastpar, Bixio Rimoldi, and Martin Vetterli. To code, or not to code: Lossy source channel communication revisited. *IEEE Transactions on Information Theory*, 49(5):1147–1158, May 2003.

[9] Y. C. Ho. Team decision theory and information structures. *Proc. IEEE*, 68, 1980.

[10] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[11] G. T. Huang. Casting the wireless sensor net. *MIT Technology Review*, July-August 2003.

[12] P. Ishwar, R. Puri, S. S. Pradhan, and K. Ramchandran. On rate-constrained estimation in unreliable sensor networks. In *IPSN 2003*, Information Processing in Sensor Networks, pages 178–192, Palo Alto, CA, USA, April 2003.

[13] R. Nowak. Distributed em algorithms for density estimation and clustering in sensor networks. IEEE Transactions on Signal Processing, Special Issue on signal Processing in Networking, 2003.

[14] R. Nowak and U. Mitra. Boundary estimation in sensor networks. In *2nd International Workshop on Information Processing in Sensor Networks*, Palo Alto, CA, April 2003.

[15] Yasutada Oohama. The rate distortion function for the quadratic gaussian ceo problem. *IEEE Transactions on Information Theory*, 44(3):1057–1070, May 1998.

[16] N. Patwari and A. Hero. Reducing transmissions from wireless sensors in distributed detection networks using hierarchical censoring. In *ICASSP*. IEEE, 2003.

[17] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Interference*. Morgan-Kauffman, 1988.

[18] Sandeep S. Pradhan, Jim Chou, and Kannan Ramachandran. Duality between source coding and channel coding and its extension to the side information case. *IEEE Transactions on Information Theory*, 49(5):1181–1203, May 2003.

[19] Sandeep S. Pradhan and Kannan Ramchandran. Distributed source coding using syndromes (discus): Design and construction. *IEEE Transactions on Information Theory*, 49(3):626–643, March 2003.

[20] H. L. Van Trees. *Detection Estimation and Modulation Theory*. John Wiley and Sons, 1968.

[21] J. N. Tsitsiklis. Decentralized detection. *in Advances in Statistical Signal Processing, H. V. Poor and J. B. Thomas Eds*, 2.

[22] J. N. Tsitsiklis and M. Athans. Convergence and asymtotic agreement in distributed decision problems. *IEEE Trans. Automatic Control*, 1984.

[23] H. Witsenhausen. A counterexample appearing in stochastic optimal control. *SIAM J. Contr*, 1968.