Graphical Model-based Approaches to Target Tracking in Sensor Networks: An Overview of Some Recent Work and Challenges*

Murat Üney[†‡] and Müjdat Çetin[†]

{ muratuney, mcetin } @sabanciuniv.edu

[†]Sabancı University, Faculty of Engineering and Natural Sciences, İstanbul, Turkey
[‡]METU, Electrical and Electronics Eng. Dept., Ankara, Turkey

Abstract

Sensor Networks have provided a technology base for distributed target tracking applications among others. Conventional centralized approaches to the problem lack scalability in such a scenario where a large number of sensors provide measurements simultaneously under a possibly non-collaborating environment. Therefore research efforts have focused on scalable, robust, and distributed algorithms for the inference tasks related to target tracking, i.e. localization, data association, and track maintenance. Graphical models provide a rigorous tool for development of such algorithms by modeling the information structure of a given task and providing distributed solutions through message passing algorithms. However, the limited communication capabilities and energy resources of sensor networks pose the additional difficulty of considering the tradeoff between the communication cost and the accuracy of the result. Also the network structure and the information structure are different aspects of the problem and a mapping between the physical entities and the information structure is needed. In this paper we discuss available formalisms based on graphical models for target tracking in sensor networks with a focus on the aforementioned issues. We point out additional constraints that must be asserted in order to achieve further insight and more effective solutions.

1. Introduction

Sensor networks refer to a system comprised of networked sensor platforms with limited sensing, communication and computation capability acting together to extract information that describes some physical entity of interest. The processing requirements are underlined with high volumes of spatial observations in consecutive temporal steps. In cases where no infrastructure can be provided, the design of appropriate algorithms pose additional difficulties due to limited energy resources. These considerations motivate the processing of information on sensor networks in a collaborative and distributed manner.

Graphical Models can be used to represent the statistical structure of the system and measurement processes through the common language of graphs where each node represents a set of random variables and the edges between nodes represent the Markov properties of the random variables belonging to the nodes. Computationally efficient solutions are achieved by utilizing inference algorithms on a given graphical model representation [8]. This approach provides a rigorous basis for the treatment of data fusion problems characterized by multi-modal, spatially distributed observations. In the sensor network context, together with the class of message passing algorithms (MPAs) we acheive a means to handle the requirement of distributed computation under communication and energy constraints [2].

In a target tracking scenario, a number of moving objects are observed by sensors spread over a region. In order to estimate the positions and possibly some other attributes of the targets, inference tasks such as detection, source localization, data association, and estimation should be completed. Being hard problems even in their centralized versions, solutions suitable to sensor networks is possible through exploiting some prior information which describes the structure of the problem, i.e. the relationship among the variables, such as the knowledge of which targets are observed by which sensors. A graphical model representation of the so-called information structure yields scalable and distributed solutions utilizing MPAs.

The communication and energy bottleneck common to sensor networks require further modifications to be made on MPAs. Since most of the limited energy of sensor platforms are consumed by communication, energy/communication aware messaging schemes are of interest. Another aspect is the limited bandwidth neccessitating the use of finite representations for messages [2]. Also additional care should be taken when attacking an inference problem in sensor networks due to the fact that the information structure is not necessarily identical to the communication structure. Hence the messages of MPAs do not necessarily correspond to real communications.

In this paper, we overview the aforementioned concepts in the setting of target tracking in sensor networks. In sec-

tion 2 and 3 we present a brief introduction to graphical models together with implications of their use in sensor networks. Section 4 identifies the issues that arise in sensor network applications through a presentation of inference tasks in target tracking and graphical model solutions proposed in the literature. In section 5 we discuss the motivations provided by the communication constrained environment for extensions beyond existing graphical model based inference algorithms.

2. Graphical Models

Graphical models consist of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ representing the set $X = \{x_i\}_{i=1:N}$ of random variables constituting the process to be modeled and Markov properties among them together wih functions defined on a subset of random variables, i.e. $\{\psi_j(X_j)\}$ where $X_j \subset X$. We focus on representations with undirected graphs, specifically Markov Random Fields. In this formalism, each node v of $\mathcal{G}$ is associated with a set of random variables $X_v \subset X$ and X is said to be Markov with respect to graph $\mathcal{G}$ if given disjoint sets $A, B, C \subset X$ satisfying the property that B separates A and C in $\mathcal{G}$, the random variables X_A and X_C are conditionally independent given X_B.

The Hammersley-Clifford Theorem guarantees that if X is Markov with respect to $\mathcal{G}$, then the joint probability distribution will be in the form of product of functions defined on cliques of the graph $\mathcal{G}$, i.e.

$$p(x) = \frac{1}{Z} \prod_{C \in \mathcal{C}} \psi_C(x_C) \qquad (1)$$

for strictly positive distributions, where Z is a normalization factor, referred to as the partition function, given by $Z = \sum_x \prod_{C \in \mathcal{C}} \psi_C(x_C)$ and $\psi_C(x_C)$ are called compatibility functions. A case of special consideration is when the factorization in Eq. (1) is composed of compatibility functions either on single nodes or edges, i.e.

$$p(x) = \frac{1}{Z} \prod_{v \in \mathcal{V}} \psi_v(x_v) \prod_{(s,t) \in \mathcal{E}} \psi_{s,t}(x_s, x_t) \qquad (2)$$

Through appropriate node aggregation, any graphical model can be put in this form [1]. Moreover, by aggregating the nodes constituting up a clique in Eq. (1), we construct the so-called junction tree representation. Tree structures are useful since there are efficient exact inference algorithms on trees. On a tree graph[1] $\mathcal{G}_\mathcal{T}$, the marginal probabilities at any node $v \in \mathcal{V}$ can be written in terms of the messages from neighboring nodes $ne(v)$ and the node compatibility function as

$$p_v(x_v) \propto \psi_v(x_v) \prod_{t \in ne(v)} m_{tv}(x_v) \qquad (3)$$

where the message from node t to v is in the form

$$m_{tv}(x_v) \propto \sum_{x_t} \psi_{v,t}(x_v, x_t) \psi_t(x_t) \prod_{u \in ne(t) \setminus v} m_{ut}(x_t) \qquad (4)$$

---

[1]For this case, the selection of the compatibility functions turn to $\psi_v(x_v) = p_v(x_v)$ (or $\psi_v(x_v) = p_v(x_v)p(z_v|x_v)$ when an observed variable z_v is associated with x_v) and $\psi_{s,t}(x_s, x_t) = p(x_s, x_t)/p(x_s)p(x_t)$.

and represent a set of fixed point equations due to the loop free structure of $\mathcal{G}_\mathcal{T}$ [18]. An exact solution to this set of equations defines a schedule for the messages as defined in Eq. (4), leading to a MPA. Similarly, the problem of finding the value of the variable associated with node v, i.e. x_v, at the maximum of the joint density is expressed by replacing the summations in Eq.s (3) and (4) by $\max$ operations and the solution has the same schedule as that in the previous problem.

Although Eq.s (3) and (4) describe the situation for a loop free graph $\mathcal{G}_\mathcal{T}$, they are also well defined for any MRF $\mathcal{G}$. Rather than performing inference through the junction tree representation, one may prefer to run (4) on $\mathcal{G}$ ignoring possible loops. Some reasons behind this approach are that aggregating the nodes to form a junction tree grows the domain set dimensions of the corresponding compatibility functions inhibiting distributedness and that construction of these functions is not always easy.

Loopy versions of the well-known sum-product algorithm have been successful in many applications including decoding of low-density parity check codes [12] which is characterized by graphs with long cycles. The error bounds and performance of loopy versions of inference algorithms have been studied further (see e.g. [20], [21]). Inference algorithms on loopy graphs together with assessment of their performance are of interest in the sensor network context where typically graphs with short cycles arise [2].

3. Inference in Sensor Networks

The information structure of a sensing problem is determined by the relations within the variables constituting the problem where some of them correspond to state variables of concern and some of them correspond to observations. Graphical models provide a useful way of representation for the information structure revealing the interactions of locally collected measurements of sensor platforms within the overall picture. The introduction of observed variables in our problem setting by labeling them as z leads the form in Eq. (1) to turn to

$$p(x, z) = \frac{1}{Z} \prod_{C \in \mathcal{C}} \psi_C(x_C, z_C) \qquad (5)$$

The a-posteriori joint distribution of the variables constituting up x given all observations z is proportional to the right hand side of Eq. (5) and moreover, without loss of generality, equality is satisfied through appropriate scaling of the compatibility functions. Continuing with node aggregation we end up with

$$p(x|z) = \frac{1}{Z} \prod_{v \in \mathcal{V}} \psi_v(x_v, z_v) \prod_{(s,t) \in \mathcal{E}} \psi_{s,t}(x_s, x_t, z_{st}) \qquad (6)$$

where z_v and z_{st} are the set of noisy observations that only depend on the set of random variables x_v and $\{x_s, x_t\}$ respectively. When z satisfies the observation locality property, i.e. for each observation z_v, z_v is independent of $\{x_u\}_{u \in \mathcal{V}, u \neq v} \cup \{z_u\}_{u \in \mathcal{V}, u \neq v}$ given x_v, then the information structure given by Eq. (6) reduces to

$$p(x|z) = \frac{1}{Z} \prod_{v \in \mathcal{V}} \psi_v(x_v, z_v) \prod_{(s,t) \in \mathcal{E}} \psi_{s,t}(x_s, x_t) \qquad (7)$$

A collaborative and distributed processing in sensor networks is achieved through a mapping between real sensor platforms and partitions of the graph $\mathcal{G}$ which represents the information structure. In order to handle the general structure given in Eq.(5), we can start selecting a sensor platform for each factor $\psi_C(x_C, z_C)$ and construct a junction tree, such that every node of the tree is hosted by a physical sensor plaform which is responsible for the necessary computations involved with the corresponding compatibility function during steps of MPAs.

As aforementioned, one may prefer approximate inference and stay with the representation as given in Eq.(6), ignoring loops if any. It is clear that besides assigning factors on single nodes of $\mathcal{G}$, i.e. $\psi_v(x_v, z_v)$, to sensor platforms, two neighboring single node factor holders should be provided enough information to evaluate edge compatibility functions $\psi_{s,t}(x_s, x_t, z_{st})$. Note that, if the locality of observations is satisfied and if each factor $\psi_v(x_v, z_v)$ in Eq.(7) describes the contribution of a distinct sensor platform, an immediate mapping is implied.

After this mapping, messages of an inference algorithm between nodes of the information graph that are hosted by two distinct sensor platforms correspond to real communication over the network. In addition to the considerations on the effectiveness of selection of representation and inference algorithm, the implications of the communication constrained environment of sensor networks should be taken into account. We outline these issues in section 5.

4. Inference Tasks in Target Tracking

Target tracking deals with generating a list of tracks $\{\mathcal{T}_k\}_{k=1:M}$ from a list of observations $\{Z_i^t\}_{i=1:N}$ where each track identifies a target that has entered in the field of regard of the sensor network in the form of a list of sufficient statistics of the location. An example is as a list of location estimate and standard deviation tuples, i.e. $\mathcal{T}_k = \{(\hat{x}_k(\tau), \sigma(\tau))\}_{\tau=t_d:t_f}$ with t_d and t_f indicating time of detection and finalization time of track respectively such that $0 < t_d < t_f < t$. We may choose to embed any form of information useful for the identification of a target, such as signature, to the track in accordance with the capabilities of the sensor nodes.

The list of observations Z_i^t indicates all observations of sensor platform i up to time t, i.e.

$$Z_i^t = \left\{ \{z_{ik}(\tau)\}_{k=1:M_i(\tau)} \right\}_{\tau=0:t} \qquad (8)$$

where $M_i(\tau)$ is the number of measurements taken at time τ. For a sensor platform bearing a single sensing element such as an acoustic amplitude sensor, $M_i(\tau) = 1$ for all τ whereas for multiple element cases such as Direction of Arrival and radar type sensor platforms $M_i(\tau)$ is related to the number of targets detected and false alarms.

In this section, we introduce tasks constituting target tracking in sensor networks such as source localization and data association. The graphical model representations that are exploited in the literature are presented along with the discussion of the nature of the problems.

4.1. Sensor Characterization

Although it is possible to use more general measurement models, additivity of noise is a suitable assumption for our discussion such that at a given time t, the measurement of a sensor platform is given by

$$z(t) = h_\theta(x(t)) + \omega(t)$$

where $\omega(t)$ is a Gaussian process with mean $\mu(t)$ and covariance matrix $\mathbf{K}(t)$, i.e. $\omega(t) \sim \mathcal{N}(\mu(t), \mathbf{K}(t))$, and h_θ is a memoryless and possibly a non-linear function on the state space $\mathcal{X}$, parameterized on $\theta \in \Theta$. Θ captures any parameters related to the sensor such as the modality, location and orientation.

A common ground for the characterization of sensors of different modalities can be provided through appropriate choice of observation likelihoods $p(z|x)$. It is reasonable to assume that the measurment noise processes of distinct sensors are independent. Hence given N observations $\{z_i(t)\}_{i=1:N}$, the joint likelihood takes the form of product of individual likelihoods, i.e.

$$p(\{z_i(t)\}_{i=1:N} | x(t)) = \prod_{i=1}^{N} p(z_i(t)|x(t)) \qquad (9)$$

making it possible to fuse information from distinct sensors of different modalities. An example for the computation of individual sensor likelihoods is given for an acoustic amplitude sensor located at x_θ as presented in [15]. The observation model is given by

$$z(t) = \frac{A}{\|x(t) - x_\theta\|} + \omega(t) \qquad (10)$$

and if the nominal amplitude A is known, the observation likelihood is given by

$$p(z|x; A) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2\sigma^2} \left(z(t) - A/r(t) \right)^2 \right\} \qquad (11)$$

where $r(t) = \|x(t) - x_\theta\|$. We have the option to choose a static model for A, i.e. we let A be a random variable with a known distribution and treat $p(z|x) = E_A\{p(z|x; A)\}$. Another approach is to aggregate any unknown random variable such as A to the parameter vector θ and characterize the sensors with the observation likelihood functions in the form $p(z|x, \theta)$ [7].

4.2. Source Localization

Consider a single target covered by some sensor platforms. We expect the sensor network to localize the source of the measurements even when the sensor platforms are not capable of reducing the uncertainty by themselves. In the above example, the individual likelihoods are functions of r yielding concentric equal-probability curves. Hence a

single sensor platform could reduce the set of possible locations of the target to a set of points on a circle. Also it is clear that this set is not compact. The main idea is to reduce the uncertainty on the target position through intersecting the location sets infered by different sensors. A natural way is to construct the posterior density function given the set of all observations $\{z_i(t)\}_{i=1:N}$ of the source. The posterior density $p(x(t)|\{z_i(t)\}_{i=1:N})$ is proportional to the joint likelihood and the prior, and the joint likelihood is in the form of product of individual likelihoods i.e.

$$p(x(t)|\{z_i(t)\}_{i=1:N}) \propto p(\{z_i(t)\}_{i=1:N}|x(t))p(x(t)) \quad (12)$$

$$\propto \prod_{i=1}^{N} p(z_i(t)|x(t))p(x(t))$$

The structure above implies a centralized procedure where any sensor platform provided with the knowledge of the prior and all individual likelihoods can perform the localization of the source with minimum possible uncertainty. The prior is inherited from the prediction step of a tracking algorithm run at the preceding time step. A natural graphical model representing the above factorization is a tree with the root node corresponding to x and hence the factor graph representation leads to the desired a posteriori distribution through the sum product algorithm [13].

The situation is complicated when there are more than one target covered by some sensors. The acoustic amplitude measurement example renders the difficulty of resolving multiple targets since $M_i(\tau) = 1$ in (8) regardless of the number of targets in the coverage of the individual platform. The factorized form of the joint likelihood is in accordance with sensor-target coverage relations implying a graph. This representation helps to construct collaborative algorithms for localization in a message passing fashion. For a demonstrating example, assume that two targets A and B with corresponding parameter sets θ_A and θ_B are moving in the coverage of the sensor network and also that sensor measurements on platforms 1 and 3 are affected only by target A whereas 2 and 4 by both targets. The corresponding MRF and the factor graph representation are given in Figure 1(a), (b). Note that due to the common targets covered by more than one sensor, the resulting graphs contain loops. For this case, a loopy version of the sum-product algorithm can be used leading to approximate but successful results [16].

Another issue is that the higher order conditionals (e.g. $p_{z_2|x_A,\theta_A,x_B,\theta_B}(.)$) appear in accordance with the number of targets covered by an individual sensor platform. Hence the knowledge of these functions for any possible number of targets covered is required which may not feasible for densely located targets. Therefore it is preferable that this requirement be removed. This is possible through hypothesizing intermediate level random variables z_i^js indicating the contribution of target j to sensor i [5]. This contribution is characterized by nothing but the observation likelihood derived from the individual observation model, i.e. $p(z_i^j|x_j) = p_{z|x}(z_i^j|x_j)$ and all contributions are coupled to the sensor measurement z_i through a transition density
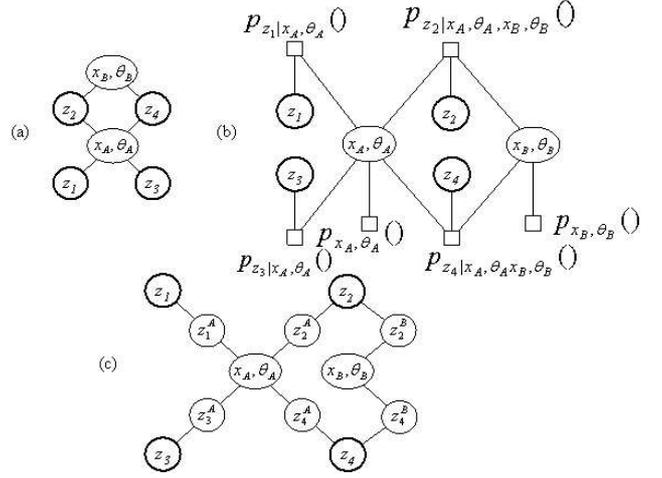
Figure 1. Graph representations of the example; (a) MRF, (b) factor graph, (c) extended MRF.

$p(z_i|\{z_i^j\}_{j=1:K_i})$ where K_i is the number of targets covered by sensor platform i. The resulting MRF representation for the above case is presented in Figure 1(c). Rather than using an inference algorithm on loopy graphs, an approximate finite step solution is proposed in [7] leading to a distributed algorithm of message passing structure where each step implies assignment of a sub-task to a sensor platform.

Hence in general, the source localization task produces a list of target position estimates $\{\hat{x}_j(t)\}_{j=1:M}$ given the set of sensor measurements $\{z_i(t)\}_{i=1:N}$. Although the sensor platforms are not capable of resolving multiple targets individually, it is possible to localize multiple targets in the coverage of the network by integrating the prior information from target tracks, with the observations from sensor platforms. The next step to maintain tracks is to assign each computed location to a track in the list. Without any other discriminating information, this task is carried out by assessing the consistency between the predicted locations calculated using tracks and estimated locations. Extending the state space of targets to include a parameter set θ, which carries some discrimating information, also helps this procedure [19].

A similar problem arises when sensor platforms are capable of resolving multiple targets in the coverage region individually, such as Direction of Arrival sensors and radar type sensors. At each step, the ith platform receives a list of measurements $\{z_{ik}(t)\}_{k=1}^{M_i(t)}$. Ideally each $z_{ik}(t)$ belongs to a target in the coverage, however naturally there are chances of missed detections and false alarms. Therefore before taking any further step, we should determine for each sensor's measurement list, which measurements belong to which target, which measurements are false alarms, and which targets in the coverage it has failed to detect. This problem is referred to as data association and for the multiple-target multiple-sensor case, it is a multi-dimensional assignment problem in general.

4.3. Data Association

The Data Association problem in multi-sensor multi-target tracking is as follows: We have N sensors and for a given time instant, i^{th} sensor has made M_i measurements, i.e. there are N lists of size M_i in the form

$$\left\{ \{z_{ik}\}_{k=1:M_i} \right\}_{i=1:N} = \qquad (13)$$
$$\left\{ \{z_{1k}\}_{k=1:M_1}, \{z_{2k}\}_{k=1:M_2}, ..., \{z_{Nk}\}_{k=1:M_N} \right\}$$

Sensor measurements must be assigned to targets taking into account possible false alarms and missed detections as well.

The possible choice for the semantics of assignment variables vary. From a target centric view, target $j \in 1, ..., M$ is associated with an element from the Cartesian product of the measurement set $\{z_{ik}\}_{k=1:M_i}$. Hence the result would be an association hypothesis in the form of a set of N-tuples, i.e. $\{(j_1, j_2, ..., j_N)\}_{j=1:M}$ for each target, indicating that the measurement z_{ij_i} of the ith sensor, is most likely originated from target j. The constraint is that no two targets can be assigned to the same measurement, i.e. $j_i \neq j'_i$ for $j \neq j'$ and $i = 1, ..., N$. Similarly, a sensor platform oriented choice of the assignment variables leads to a set of M_i-tuples from the permutation set of $(1, ..., M_i)$, i.e. $\{(i_1, i_2, ..., i_M)\}_{i=1:N}$ indicating that the measurement z_{ii_k} of the ith sensor, is most likely originated from target k.

The assesment of a given association hypothesis is similarly through a probability measure given the set of observations. Since all hypotheses are equally likely, the MAP estimation over the set of valid hypotheses is equivalent to maximizing the joint likelihood (of hypotheses). Next we observe that the knowledge of which targets are covered by which sensor platforms reduce the dimension of association variables. Moreover this also implies a certain factorization of the joint likelihood where in the case of sensor centric hypothesis construction, the association variables corresponding to sensors measuring the same target(s) (and similarly in a target oriented formulation targets measured by the same sensor) appear in the same factor. The graphical model approach exploits these facts by selecting the vertex set $\mathcal{V}$ as the association hypothesis set and the edge set $\mathcal{E}$ with respect to the appropriate binary relation of "measure-same-target" or "measured-by-same-sensor" with clique potentials ψ_C accordingly [4].

It is interesting to notice that the target centric view has been favored in the centralized multiple-sensor multiple-target tracking research. On the other hand, the sparse nature of the sensor networks due to the limited coverage region of individual sensors implies that the search space can be reduced through sensor centric hypothesis construction. This formulation utilizes target-sensor coverage information and in the graph representation, each vertex is associated with the association hypothesis of a sensor platform. However in the case of targets sensed by more than 2 sensors, potentials for cliques on more than 3 nodes appear in the factorization of the joint likelihood. In [4], such cliques are replaced with target centric association variables

in order to keep the pairwise factorization yielding a hybrid modeling approach. Also the target-sensor coverage information used in data association is extracted through a graphical model where each node represents a partition of the field of regard determined by sensor coverage regions with random variables this time in order to associate targets with these sections. Proposing an N-scan data association scheme based on the above approach is possible in order to achieve robustness in ambiguous cases often caused by crossing targets. Interested reader is referred to [3] for details.

Exact inference over $\mathcal{G}$ is costly due to the inherent loopy nature. Therefore inference on loopy graphical models, specifically MAP estimation procedures, are of interest as well as communication-sensitive message passing.

4.4. Track Maintenance

At the begining of this section, we have discussed the form of the target track information. Maintenance of tracks is comprised of initiating a new track upon detection of a new target, update existing tracks using output from the lower level processes including the localization process, and finalize an existing track inferred to be corresponding to a target not located in the field of regard of the network anymore. Also the lower tasks exploit the information embedded in the tracks yielding the information structure.

In order to keep the collaborative nature of the network, global track maintenance tasks can be distributed to individual sensor platforms. The sparse nature of problem is exploited in order to minimize resource usage by leader based schemes where each target is associated with a track holder which is reponsible for the completion of the localization task and hand-off the track information to the new leader through a selection scheme [14]. The new leader can be decided based on expected information utility measures and costs such as described in [6] and [22].

When the sensor platforms are capable of resolving multiple targets, localization is performed through distributed estimation following data association. In [3], this is carried out in the framework of graphical model based data association by using a multiple hypothesis tracking-like approach.

The important point is to provide some mechanisms for resolution of possible inconsistencies. For example duplicate performance of track initiation and in general maintenance tasks should be detected and merged. Also closely moving targets may lead to ambiguous track segments which could be joined after some discriminating information is provided [7].

5. Discussion

We have mentioned that sensor networks should work in a collaborative and distributed manner. We also outlined why graphical models are convenient tools to achieve such solutions for inference problems in sensor networks, together with applications on target tracking. When inference in sensor networks is considered, due to the inherent limitations in communication bandwidth and energy resources,

additional aspects should be considered from a graphical model perspective.

First of all, limited communication bandwidth implies finite representations of messages. Also as discussed in section 2 sensor network applications often lead to loopy graph representations and loopy versions of MPAs perform in an iterative manner and converge asymptotically if they do. This nature together with message errors yield approximate results. Second, being the most energy consuming action, communication should be performed in an efficient manner. During iterations of a loopy MPA, in order to reduce or at least upper bound the number of messages for a given task, censoring of messages which has not been changed more than an amount with respect to a measure is utilized [3]. Decision of sending or censoring a message is given individually by sensor platforms regardless of any other information such as states of the neighboring platforms or penalty for communication, resulting a myopic behavior.

The effects of the above issues on the results provide an analysis framework for communication constrained inference such as presented in [10]. As we point out in section 3, a step in the design perspective is involved with mapping the information structure to the network structure, preferably in consideration of the cost of communications. As an example, in the case of a junction tree representation and a communication network which provides links between any two sensor platforms with appropriately assigned costs, it is possible to apply the communication constrained inference paradigm through handing-off variables associated with nodes of the tree while keeping the information structure valid and reducing the communication cost [17].

The statement of a general design problem in a communication constrained environment is involved with defining costs for communication and estimation errors and then search individual actions for platforms which satisfy optimality in some sense subject to the communication constraints. Hence we could achieve an overall behavior which can adapt computations local to sensor platforms as well as the structure of messages between physical nodes to the underlying communication structure. In [11], this approach is utilized for a self-organizing decentralized detection network. However it is not straightforward to apply the results on an estimation problem due to lack of a finite parameterization of individual actions in the estimation problem setting. Therefore we are still in need of methods which can explicitly handle the trade-off between communication cost and estimation accuracy or decrease the communication cost in some sense in a less myopic manner.

References

[1] P. Brèmaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulations and Queues*, Springer-Verlag, 1991.

[2] M. Cetin, L. Chen, J. W. Fisher III, A. T. Ihler, R. L. Moses, M. J. Wainwright, and A. S. Willsky, "Distributed fusion in sensor networks: A graphical models perspective", *IEEE Signal Processing Magazine*, 23(4), July 2006, pp. 42–55.

[3] L. Chen, M. Cetin, and A. S. Willsky, "Distributed data association for multi-target tacking in sensor networks", Int. Conf. on Information Fusion, July 2005.

[4] L. Chen, M. Wainwright, M. Cetin, and A. S. Willsky, "Data association based on optimization in graphical models with application to sensor networks", *Mathematical and Computer Modelling, Special Issue on Opt. and Control for Military Applications*, 43(9-10), May 2006, pp. 1114–1135.

[5] M. Chu, *A Hierarchical Framework for Constructing Computationally Efficient Algorithms for Distributed Inference Problems*, PhD thesis, EECS Dept., Massachusetts Institute of Technology, 2003.

[6] M. Chu, H. Haussecker, and F. Zhao, "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks", Technical Report P2001-10113, Xerox Palo Alto Research Center, May 2001.

[7] M. Chu, S. Mitter, and F. Zhao, "Distributed multiple target tracking and data association in ad hoc sensor networks", Int. Conf. on Information Fusion, July 2003.

[8] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter, *Probabilistic Networks and Expert Systems*, Springer, 1999.

[9] C. Crick and A. Pfeffer, "Loopy belief propagation as a basis for communication in sensor networks", UAI'03, Aug. 2003.

[10] A. Ihler, J. W. Fisher III, and A. S. Willsky, "Loopy belief propagation: Convergence and effects of message errors", *J. of Machine Learning Research*, 6, May 2006, pp. 905–936.

[11] P. Kreidl and A. S. Willsky, "An efficient message-passing algorithm for optimizing decentralized detection networks", Technical Report LIDS-P-2726, MIT LIDS, 2007.

[12] F. R. Kschischang, "Codes defined on graphs", *IEEE Signal Proc. Mag.*, 41(4), Aug. 2003, pp. 118–125.

[13] F. R. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm", *IEEE Trans. on Info. Theory*, 47(2), February 2001, pp. 498–519.

[14] J. Liu, J. Liu, and P. Cheung, "Distributed group management in sensor networks: Algorithms and applications to localization and tracking", *Telecom. Systems*, 26(2-4), June 2004, pp. 235–251.

[15] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for taget tracking", *EURASIP Journal on Applied Signal Processing*, 23(4), March 2003, pp. 378–391.

[16] J. M. F. Moura, J. Lu, and M. Kleiner, "Intelligent sensor fuison: A graphical model approach". ICASSP03, April 2003.

[17] M. Paskin, C. Guestrin, and J. McFadden, "A robust architecture for distributed inference in sensor networks", IPSN'05, April 2005.

[18] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufman, 1988.

[19] J. Shin and Z. F. Guibas, Leonidas J., "A distributed algorithm for managing multi-target identities in wireless ad-hoc sensor networks", IPSN'03, 2003.

[20] M. J. Wainwright, T. Jaakkola, and A. Willsky, "Tree consistency and bounds on the performance of the max-product algorithm and its generalizations", *Statist. Computing*, 14, April 2004, pp. 143–166.

[21] J. Yedidia, W. Freeman, and Y. Weiss, "Constructing free energy approximations and generalized belief propagation algorithms", *IEEE Trans. on Info. Theory*, 51(7), July 2005, pp. 2282–2312.

[22] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration", *IEEE Signal Proc. Mag.*, 19(2), March 2002, pp. 61–72.