A Message Passing Approach To Multiagent Gaussian Inference for Dynamic Processes

(Extended Abstract)

Stefano Ermon
Department of Computer Science
Cornell University
Ithaca, New York
ermonste@cs.cornell.edu

Carla Gomes
Department of Computer Science
Cornell University
Ithaca, New York
gomes@cs.cornell.edu

Bart Selman
Department of Computer Science
Cornell University
Ithaca, New York
selman@cs.cornell.edu

ABSTRACT

In [1], we introduced a novel distributed inference algorithm for the multiagent Gaussian inference problem, based on the framework of graphical models and message passing algorithms. We compare it to current state of the art techniques and we demonstrate that it is the most efficient one in terms of communication resources used. Moreover, we show experimentally that it outperforms the other methods in terms of estimation error on a general class of problems, even in presence of data loss.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems

General Terms

Algorithms

Keywords

Distributed problem solving, Reasoning

1. INTRODUCTION

Distributed inference tasks are becoming more and more important as myriads of tiny inexpensive sensing devices are being deployed. In many such problems, a network $G = (V, E)$ of sensing devices that are capable of local communication is used to collect information about the state of the world, that is then used as evidence to solve inference problems according to a known global probabilistic model.

In many real world problems, it is fundamental for such probabilistic models to capture the spatio-temporal dynamics of the system, for example in the case of tracking a moving target or monitoring the temperature of an environment over time. In this work we consider the case of linear Markovian dynamics, where the global state $x \in \mathbb{R}^n$ changes over time according to the following difference equation:

$$x_{k+1} = A_k x_k + w_k \ , \tag{1}$$

Cite as: A Message Passing Approach To Multiagent Gaussian Inference for Dynamic Processes (Extended Abstract), S. Ermon, C. Gomes and B. Selman, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. XXX-XXX.

where w_k is a white Gaussian noise. This type of model is often used as a first order approximation (by linearization) of more general nonlinear dynamics. We also assume that each sensing agent $i \in V$ obtains at each time step k an observation $y_k(i)$, that is a linear combination of the state variables x_k corrupted by additive Gaussian noise.

The inference problem we consider is that of computing at each node $i \in V$ and for each time step k the minimum mean square error estimate of the global state x_k given all the evidence up to time k available at node i, assuming latencies in the communication links. In our jointly Gaussian setting, it corresponds to a complete characterization of the posterior probability distribution of the state given the evidence.

Given the severe communication and energy restrictions of many real world networks, centralized solutions where a single node receives and elaborates all the information are not sufficiently scalable so that there is a need for distributed solutions. In [1], we introduced a novel distributed inference algorithm (`BP-approx`) based on the framework of graphical models and message passing algorithms, where inference is performed locally at each node on the basis of information that is retrieved both locally and by communication with neighboring nodes. By using Belief Propagation (BP) inspired updates, nodes locally elaborate and fuse the information they receive before transmitting it again, thus reducing the total number of messages needed and distributing the computational burden over the network.

In `BP-approx`, each message represents a Gaussian probability distribution, that can be completely described using a mean and covariance pair. The size of each exchanged message is therefore proportional to $n^2 + n$, where n is the dimensionality of the hidden state space. A key feature of the protocol is that to enforce an ordered flow of information it imposes an hierarchy among the nodes by using a *spanning tree* of the network. As a consequence of the hierarchical structure, only $2(N - 1)$ messages are exchanged every time step in a network with $|V| = N$ nodes.

In contrast, a standard centralized Kalman filter (`CKF`) requires to exchange messages of size $n^2 + n$ from every node in the network to every other node, so that a total of N^2 messages are exchanged every time step. The most popular alternative distributed approach (`DKF`) introduced in [2]) is based on *consensus filters* and exchanges messages of size $n + n^2 + n$. In that case, every node sends a message to each of its neighbors, so that $2|E|$ (that is $O(N^2)$) messages are exchanged every time step.

As we can see from our analysis, thanks to the spanning tree infrastructure used `BP-approx` is the most efficient solution in terms of communication resources used.

In our simulation experiments, we compare the performance of these algorithms in terms of the average empirical variance of the estimation error, defined as $||\widehat{x}(k) - x(k)||_2$, where $\widehat{x}(k)$ and $x(k)$ are respectively the estimated and true state of the world at time k. A network is generated by randomly scattering 50 sensing devices in a target area and assuming that they can communicate if their distance is smaller than a threshold r. Moreover, we assume that there is fixed constant probability of loosing a data packet over each communication link, independently of the distance r. We also assume that each communication link has a latency of 1 time step associated with it. To fix a baseline in our experiments, we assume that `CKF` is not affected by any data loss and does not experience any communication latency. We also introduce a baseline for the latency constrained case called `KF-delayed`, a version of `CKF` that is affected by latencies but not by data-losses.

As a benchmark application, we consider a second-order ODE system of the form $\ddot{x} = n$ where velocity $\dot{x}$ is modeled as a Brownian motion. The system is discretized with time step ϵ to

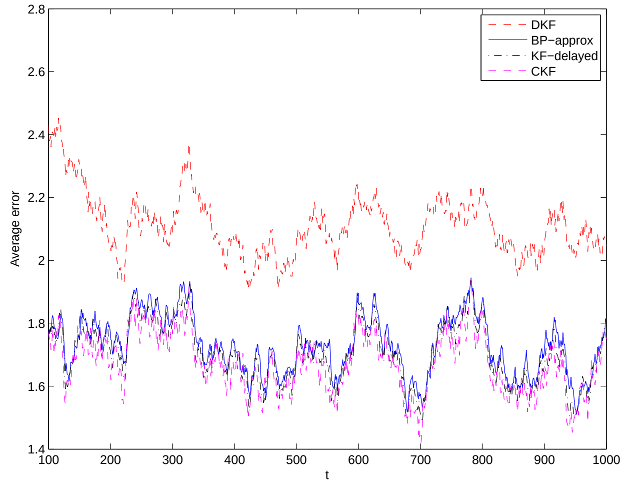
$$\begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} 1 & \epsilon \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w$$

where w is white Gaussian noise. As shown in [3], this model can be used for monitoring the temperature at several locations in en environment using a network of sensors. However, the same equations can be used to model the dynamics of a moving object and electrical networks.

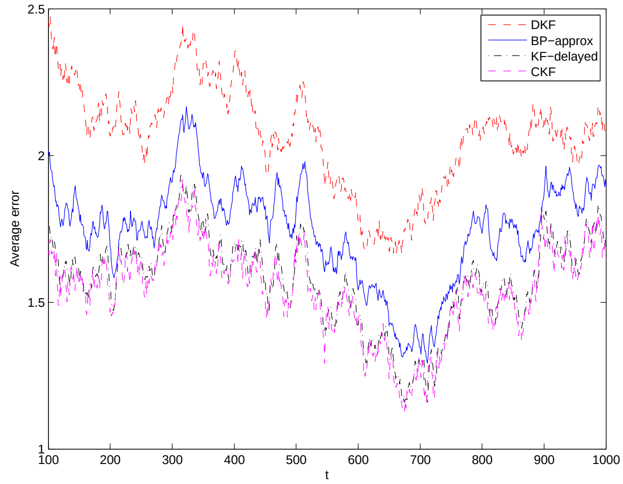
The first experiment shown in Figure 1(a) is performed without any data-loss. The improvement of `BP-approx` over `DKF` on the average error is of about 19%. Empirically we have also seen that the performance gap tends to increase with higher noise levels, measured by a larger variance of w. Moreover we can see that the approximation given by `BP-approx` is almost as good as the theoretical optimum in presence of latencies given by `KF-delayed`. An intuitive explanation of the performance gap is that `DKF` uses a "loopy" inference method and therefore it might overcount information significantly, despite its attempt to reduce the effect of these errors using a consensus or a high-pass filter.

We study the effect of a 5% data-loss in the communication packages in Figure 1(b). While the performance of both methods decreases, the improvement of `BP-approx` over `DKF` is still over 15%. In practice, the gap would be even larger because `BP-approx` is allowed to organize the nodes of the network into a spanning tree using the best quality communication links. With the `BP-approx` method, information about the past history of the process is always maintained locally by the nodes but never exchanged using the messages. This fact ensures a high-level of tolerance against communication losses. Moreover it greatly reduces the risk of double counting information when nodes drop out and then join the network again, a common scenario in wireless sensor networks caused by frequent temporary communication failures.

In conclusion, the hierarchy among the nodes imposed by the spanning tree plays a key role in this approach, both because it enforces an ordered flow of information and because it greatly reduces the communication requirements.



(a) Performance comparison without data loss.



(b) Performance comparison with 5% data loss.

Figure 1: Simulative comparison between the algorithms.

2. ACKNOWLEDGMENTS

This research is funded by NSF grant 0832782.

3. REFERENCES

[1] S. Ermon, C. Gomes, and B. Selman. Collaborative multiagent Gaussian inference in a dynamic environment using belief propagation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 1419–1420. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[2] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *Proc. of the 46th IEEE Conference on Decision and Control*, 2007.

[3] N. Vaswani. Particle filtering for large-dimensional state spaces with multimodal observation likelihoods. *Signal Processing, IEEE Transactions on*, 56(10):4583–4597, 2008.