Abstract—Due to its compact and distributed nature, Loopy Belief Propagation has been proved to be theoretically appropriate as a basis to systematically deal with uncertainties with incomplete or corrupted observations in the broad applications of wireless sensor networks. However, the transmissions of massive belief messages of Loopy Belief Propagation could become a serious concern from the energy consumption perspective since sensor nodes are energy limited, and in many real-world wireless sensor network applications, the replacement of sensor node batteries is very difficult if not impossible. In this paper, we present a novel wavelet-based Loopy Belief Propagation, which can effectively reduce 50% of belief messages communicated among sensor nodes with only minimal degradation of the inference performance. Our preliminary results using real-world soil moisture data from an environmental monitoring application demonstrate the great promise of the proposed approach.

Keywords—wireless sensor networks; information inference; wavelet; loopy belief propagation

I. INTRODUCTION

With the advances of today’s technologies, it is possible to deploy wireless sensor networks (WSNs) consisting of thousands of tiny sensor nodes for various monitoring tasks. In such applications, sensor nodes need to process the information collected from each node jointly for information fusion, and for handling sensor/node failures resulting in contaminated or missing readings. Due to its compact representation, distributed propagation and robustness property, Loopy Belief Propagation (LBP) has been proved to be theoretically appropriate and naturally suitable for handling uncertainties on-line in WSNs through well organized belief message transmissions and simple local belief updates (e.g., [1], [2]). However, as individual sensor nodes only have limited battery power, the energy consumption resulting from frequent belief message exchanges in LBP tends to be a serious problem. In this paper, we present a wavelet-based LBP approach, referred to as W-LBP, to significantly reduce the transmissions of belief messages in the traditional LBP, making LBP-based in-network inference more energy efficient and thus more suitable for real-world WSN applications.

To illustrate our new approach, in this paper, we consider environmental monitoring wireless sensor networks where each sensor node has severe power limitation, and at the same time is not reliable, as environmental monitoring WSNs are usually deployed in harsh or even hostile environments. To address the two challenges of sensor/node failures and power limitation, we have proposed a unified approach of estimating missing observations based on LBP [3], where energy savings are achieved through putting sensor nodes into sleep mode periodically. In this work, we investigate the new W-LBP versus the traditional LBP for in-network estimation in the context of [3]. It is desirable for the W-LBP not only to achieve significant energy conservation but also to minimize any possible degradation of estimation performance at the same time.

A few studies to improve LBP’s energy efficiency in WSNs exist. One recent idea is to take advantage of multicasting in WSNs to multicast identical one-to-multi-target messages from a sensor node to its neighborhood, instead of distinct ones required in traditional LBP [4]. Another attempt is to schedule message passing based on whether an individual node has sufficient new information to warrant the transmission of a new message, to avoid the energy cost of transmission of un-informative messages [2]. Yet another promising approach is called Nonparametric Belief Propagation (NBP) (e.g., [5], [6], [7]), which reduces the communication volume of individual messages through sample-based approximation of each message. Unlike NBP targeted for non-Gaussian continuous random variables, the proposed W-LBP focuses on discrete random variable cases. While NBP and W-LBP are fundamentally different, both approaches share two basic ideas: (1) to approximate messages with less bits to conserve energy during message communication, and (2) to perform such approximation with only local information without causing extra communication. Our approach is applicable to various WSN applications for which LBP inference is suitable, including target tracking, hypothesis testing, self-calibration and network clustering.

The paper is organized as follows. Section II provides some brief background on Markov Random Field (MRF) for WSN modeling. Section III presents our W-LBP approach. Section VI describes our empirical study with real-world sensing data. Finally Section V gives conclusion and future work.

II. GRAPHICAL MODEL

A. Modeling with MRF

Consider, for simplicity, the grid topology of WSN for environmental monitoring applications. MRF, an undirected Graphical Model, is an appropriate model on which
distributed inference in WSNs can be based. Graphical Model \( G(V, E) \) is a probabilistic network consisting of vertices \( V \), representing the set of random variables \( X = \{x_i\}_{i=1}^N \), and edges \( E \), each associated with a potential function, representing correlation relationships among connected subsets of \( X \). For directed Graphical Model, or Bayesian Network, edges \( E \) are encoded with causal relationships. For undirected Graphical Model, MRF, all \( X \) variables satisfy the Markov property: for any three disjoint sets \( A, B, C \subset V \), with respect to network \( G \), any path between \( A \) and \( B \) contains at least one node of \( C \), indicating \( X_A \) and \( X_B \) are conditionally independent given \( X_C \).

Given the Markov property with respect to graph \( G \), the Hammersley-Clifford theorem [8] says that the global posterior distribution of MRFs can be represented in a factorization format defined by cliques of \( G \):

\[
P(X | \theta) = \frac{1}{Z(\theta)} \prod_{c \in C} \psi_c(x_c | \theta_c), \tag{1}
\]

\[
Z(\theta) = \sum_{\psi_c} \prod_{c \in C} \psi_c(x_c | \theta_c), \tag{2}
\]

where \( C \) denotes cliques consisting of vertices in \( V \); \( \psi_c(x_c) \) denotes potential function associated with clique \( c \); and \( \theta_c \) denotes the current parameters. In this paper we consider pairwise MRF, which means each clique only includes two one-hop neighbors.

This factorized expression of global probabilistic distribution of MRF makes belief inference convenient. Furthermore, it is easy to see that learning the most possible parameters for this model is equal to getting the Maximum Likelihood (ML) of \( P(X | \theta) \) with a certain parameter set of \( \theta \).

### B. Loopy Belief Propagation

LBP is the loopy extension of Pearl’s Belief Propagation (BP) [9]. Although there is no mathematical proof for the convergence properties of LBP, excellent experimental results have been reported using LBP (e.g., [10]), especially in recent WSN research (e.g., [1], [4], [6], [11]). We now briefly describe how BP works in a pairwise MRF. Assume variables \( x_i \) and \( x_j \), with local effect \( \phi_i(x_i) \) and \( \phi_j(x_j) \), are connected through an edge potential (i.e., compatibility) function \( \psi(x_i, x_j) \). The joint probability of \( \{x_i\} \) is given by

\[
p_k(x_i) = \frac{1}{Z} \prod_{j \neq i} \psi(x_i, x_j) \prod_j \phi_j(x_j), \tag{3}
\]

and the messages and beliefs are then updated in the following way:

\[
m_{j,i}(x_i) = \sum_{x_j} \phi_j(x_j) \psi(x_i, x_j) \prod_{k \in N \setminus \{i,j\}} m_{k,j}(x_j), \tag{4}
\]

\[
b_{i,j}(x_i) = K \phi_i(x_i) \prod_{j \in N \setminus \{i,j\}} m_{j,i}(x_j), \tag{5}
\]

where \( b_{i,j}(x_i) \) is the belief (i.e., node potential) of site \( i \), \( K \) is the normalization constant, and \( m_{j,i} \) represents belief messages from \( j \) to \( i \). The general principle behind this message transmission scheme (4) is that any site \( i \) has to compute \( m_{j,i} \) based on the belief messages coming from its directly connected neighbors except for site \( j \), which it is going to send the updated messages to.

### III. Wavelet-based LBP

#### A. Wavelet Background

Wavelet theory provides a mathematical tool for hierarchically decomposing signals. Mathematically, the mother wavelet function \( \Psi \) satisfies

\[
\int \Psi(t) dt = 0. \tag{6}
\]

The wavelet basis functions which project the original signal to a wavelet coefficient domain are achieved by scale and shift operation on the mother wavelet function.

\[
\Psi(s, \tau) = \frac{1}{\sqrt{s}} \Psi(t - \tau) \tag{7}
\]

where \( \tau \) and \( s \) denote shift and scale factors, respectively. The one-dimensional (1-D) wavelet transformation \( W_j(s, \tau) \) is actually the inner production of signal \( f(t) \) and \( \Psi(s, \tau) \) as

\[
W_j(s, \tau) = \int f(t) \Psi(s, \tau)(t) dt = \langle f, \Psi(s, \tau) \rangle. \tag{8}
\]

The discrete wavelet transformation (DWT) was developed to apply the wavelet transform to digital signals. Mallat introduced a tree algorithm for computing DWT by using filter banks [12], [13], in which any original digital signal is decomposed into approximated signal and the corresponding detail signal through low-pass (h) and high-pass filters (g), related as quadrature mirror filters. Since each filter halves the frequencies of the signal, the filter outputs are subsampled by 2. For one level decomposition, the transform coefficients, \( a_k \) and \( d_k \), have the following expression:

\[
a_k^{l+1} = \sum_n h_{n-2k} a_k^l, \tag{9}
\]

\[
d_k^{l+1} = \sum_n g_{n-2k} a_k^l,
\]

where \( j \) denotes the resolution and \( k \) is the index for the samples. For a signal of length \( N \) where \( N \) is a power of 2, we have \( j = \log_2 N \). For single level signal decomposing and reconstructing, it can be illustrated in Figure 1.

In Figure 1, high-pass and low-pass analysis filters, indicated with dotted squares, are denoted as \( H \) and \( L \) respectively, whereas the corresponding synthesis filters in

618
the reconstruction process are denoted as \( H^* \) and \( L^* \) (the transposed matrix of \( H \) and \( L \), respectively). In fact, this decomposition process can be applied recursively to the approximate coefficients \( A \) until the desired result is reached. The finest resolution level \( A' \) is the original signal \([14]\).

**B. W-LBP**

In the case of discrete random variables, the belief message is a vector of numbers. Our idea is to adopt wavelet methodology to compress the belief message by dropping its details at the sender site, and thus only to transmit the approximation of the original belief message to its one-hop neighbor. On the receiver side, the details of the local belief are then used to reconstruct and estimate the original belief message, before further operation with potential function. In W-LBP, the “message” transmitted from site \( i \) to \( j \) is

\[
o_{ij} \leftarrow \sum_{n} \phi_n(x_i) \prod_{j \in N(i)} \psi_{ij}(x_i, x_j) \omega_{ij}(x_j).
\]

(10)

Accordingly, the expression for local belief will be

\[
b_j(x_i) = K \phi_n(x_i) \prod_{j \in N(i)} \psi_{ij}(x_i, x_j) \omega_{ij}(x_j).
\]

(11)

To decompose \( \omega_y \) at site \( i \), as shown in Figure 1, we have

\[
A_j = L \omega_{ij} \quad ;
\]

(12)

\[
D_j = H \omega_{ij} \quad .
\]

(13)

As original \( D_j \) is not available at site \( j \), to reconstruct and estimate \( \omega_{ij} \) at site \( j \), we have

\[
\hat{D}_j = D_j \quad ;
\]

(14)

\[
\hat{\omega}_{ij} = L^* A_{ij} + H^* \hat{D}_j.
\]

(15)

A block diagram of W-LBP with one level wavelet decomposition is illustrated in Figure 2.

Although one level wavelet decomposition is depicted in Figure 2, multilevel wavelet decomposition may be preferred in some situations due to severe energy limitation, as the size of an approximated message with DWT will be reduced to \( 2^{-m} \) of the original message size, where \( m \) indicates the level of wavelet decomposition. To reduce wavelet operations at sensor nodes, Haar wavelet is adopted in this paper, which is defined as step functions:

\[
S_j: \text{ site } i
\]

DW: wavelet decomposition
RW: wavelet reconstruction

Adjustment: adjusting operations on \( \hat{D}_{ij} \) if any element of \( \hat{\omega}_{ij} \) is

\[
\Psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \end{cases}
\]

(16)

It is important to note that, as a natural property of belief distribution, there should not appear any negative element at any moment during belief inference. This property can be exploited to improve our estimation on the dropped details \( D_j \) at site \( i \) during the reconstruction process on receiver site \( j \). Whenever any negative element occurs in the estimated \( \hat{\omega}_{ij} \) by (15), due to the inappropriate \( \hat{D}_{ij} \), some adjusting operations are conducted to correct \( \hat{D}_{ij} \) and get a better estimation of \( \hat{\omega}_{ij} \), so that the previous negative elements in \( \hat{\omega}_{ij} \) become zeros after adjustment. The adjusting operations depend on the type of wavelet employed in W-LBP. Let \( \hat{\omega}_{ij} = \{ s_k \}_{k=0}^{N} \), \( \hat{D}_{ij} = \{ d_k \}_{k=0}^{N/2} \), and \( \hat{\omega}_{ij} = \{ \hat{d}_k \}_{k=0}^{N/2} \). For Haar, we have

\[
d_k = \begin{cases} s_k \frac{s_i}{\sqrt{2}} & s_i < 0 \quad \text{and} \quad i = 2k - 1 \\ 0 & \text{otherwise} \end{cases}
\]

(17)

where \( s_i \) is the level of wavelet decomposition.

\[
\hat{\omega}_{ij} = L^* A_{ij} + H^* \hat{D}_{ij}.
\]

(18)

During wavelet transform of a signal, the approximation part always conveys the most important information and thus is kept untouched during the process of estimation. On the other hand, due to the missing of the original details, it can be expected that the accuracy of the estimation using W-LBP would not be as good as that using traditional LBP.

The motivation for such estimation is that LBP requires massive data communication which is very energy consuming in a wireless network. Studies have indicated that about 3000 instructions could be executed for the same energy cost as sending a bit for 100 meters by radio \([17]\). For one level Haar transformation and corresponding estimation,
the computation is very simple, which only need a few additions and multiplications as shown in (17), (18). The energy consumption for such low workload on the processor is negligible compared with the energy savings of the reduced communications.

IV. EMPIRICAL STUDY

C. Experiment Setup

We studied the application of estimating missing observations in environmental monitoring using 32×32 lattice real-world soil moisture sensing data from the Southern Great Plains Hydrology experiment of 1997 (SGP97) in Oklahoma, in which volumetric soil moisture content was recorded for the top 5 cm of soil. We obtained data of 15 days without rain fall. The values of soil moisture ranged from 0 to 30, where a small number indicates high moisture. This dataset was divided into two subsets: data of July 14 and July 16 used as the testing data and the rest of the data used for parameter learning of the MRF model. The measurement range of soil moisture was discretized to 14 levels to form the local vector of 14 discrete states. We counted the frequency of each state’s appearance and normalized it to a probability distribution at each site.

To simulate the missing readings, we randomly designated a certain percentage of broken sites, increasing from 5% to 50% out of a total of 1024 sites. Such partial observations for the MRF model come from two test days, July 14, the dry day and July 16, the wet day. For each site with missing readings, an “agent” among its one-hop neighbors was selected to perform distributed fusion for the missing site. Agent selection protocol was out of the scope of this paper and will be discussed in another paper. In the LBP/W-LBP process, site readings were used as the initial priors for those working sites, while for each missing site, the average value over all local beliefs from its direct neighbors (one-hop neighbors) was obtained as its initial prior.

The purpose of distribution fusion based on LBP/W-LBP on MRF is to handle the uncertainty problem through fusing the information from partial observation and the correlation information encoded in each potential associated with edges of an MRF model. The statistical relationships (i.e., spatial correlations) are embedded in historical readings and thus can be obtained through a learning process. We iteratively applied a mathematical scaling procedure, Iterative Proportional Fitting (IPF) [15], [16], to extract edge potentials from training data, as follows:

\[ \psi^{t+1}_c(x_c) = \psi^t_c \frac{\hat{P}(x_c)}{P'(x_c)}, \]  

where the \( t \) superscript denotes the time of iterations, \( P'(x_c) \) the estimated marginal, \( \hat{P}(x_c) \) the empirical marginal and \( \psi^t_c(x_c) \) the estimated potential at iteration \( t \).

The number of iterations depends on the size of MRF and prior inputs.

D. Result Analysis

As shown in Section III, the communication load of W-LBP is reduced to \( 2^{-m} \) of original LBP, where \( m \) is the level of signal decomposition. If only one level of wavelet decomposition is employed in the W-LBP process, such as in this empirical study, then 50% energy conservation compared to the original LBP can be achieved. However, such energy savings come at the expense of some minor degradation of estimation performance. To understand the tradeoff between the energy conservation and estimation performance offered by W-LBP, we conducted comparison experiments between traditional LBP and our proposed W-LBP to infer missing observations given identical partial observations on the same MRF constructed via IPF. For both test cases, 20 trails of distributed inference on each different percentage of randomly assigned missing sites were performed, and the average performance on inference accuracy is reported in Figures 3 and 4 for dry day test data and wet day test data, respectively. Two observations can be obtained: 1) In general, there is only a slight degradation on the estimation performance of W-LBP compared to traditional LBP, which is under 5% for the dry day and 9% for the wet day; 2) W-LBP inherits the robustness property from LBP: the accuracy rate decreases gradually as the percentage of missing reading sites increases, and there is no sudden performance drop even when half of the monitoring sites are missing (either broken or in sleep). Furthermore, in addition to the comparison of accuracy rates, we also analyzed the inference errors in our experiments when accurate estimation was not achieved with LBP or W-LBP. As listed in TABLEs I and II, estimation error severity has been classified into three levels: level one indicating the inference error bounded by \( \pm 1 \) discrete state, level two indicating the inference error bounded by \( \pm 2 \) discrete states, and level three indicating the inference error bounded by \( \pm 3 \) discrete states, respectively. As we can see, more than 96% and 93% of estimation errors with W-LBP fall into level one even when the missing readings exceed 50% for dry day test cases and wet day test cases, respectively. Overall, the error severity distribution achieved with W-LBP is comparable with that achieved with traditional LBP.
Note that the small fluctuation in the error distribution with the percentage of missing sites is due to the different geographical distributions of missing sites randomly selected in the grid in our experimental trials. To illustrate, Figure 5 shows the spatial distribution of estimation errors by W-LBP and LBP, respectively, on wet day test cases with the same random distribution of an initial 50% missing readings. As we can see, W-LBP actually produces an estimation error pattern very close to that which results from traditional LBP except for several more level 2 and level 3 errors. These results indicate that, even with substantial missing observations, W-LBP inference still inherits robustness property from LBP regarding the accuracy rate and error bounds, but with dramatic energy savings.

V. CONCLUSION AND FUTURE WORK

In this paper, we present a new W-LBP, a modified version of LBP based on wavelet methodology, to significantly reduce the communication volume during distributed belief inference, with very minimal degradation of estimation performance. The proposed W-LBP thus could become a better and more realistic communication basis to support distributed inference in WSNs for various applications where energy saving is crucial due to sensor nodes’ severe energy limitation. We demonstrate our approach through in-network estimation application using real-world sensing data. Haar wavelet was chosen due to its simplicity to implement in sensor node. Although only one level of wavelet decomposition is illustrated in our empirical study, multilevel decomposition can be adopted to achieve more substantial energy conservation. Therefore, the proposed W-LBP provides full flexibility to tradeoff inference performance with energy efficiency and opens up a new design and operational space to optimally match the specific objectives of WSNs under resource constraints. Also, due to the nature of localized communications of distributed belief inference in WSNs, the W-LBP inherits the scalability of the original LBP. We note that W-LBP is aimed for distributed belief inference at transport layer, and thus does not address MAC layer issues such as idle listening. Our future work includes systematically investigating W-LBP with multilevel decomposition and studying the relationship between the decomposition levels and their performance characteristics. We also plan to investigate W-LBP with other wavelet bases.

ACKNOWLEDGMENT

This work is supported in part by National Science Foundation under grant CNS-0758372.

REFERENCES


