Large-Scale Source Localization with a Wireless Sensor Network Application ⋆

James E. Weimer * **Bruno Sinopoli** ** **Bruce H. Krogh** **

* *Royal Institute of Technology (KTH), Stockholm, Sweden (e-mail: weimerj@kth.se)*
** *Carnegie Mellon University, Pittsburgh, PA USA (e-mail: {brunos | krogh}@ece.cmu.edu)*

Abstract: This paper concerns the problem of large-scale source localization arising when many potential sources must be classified as either active or inactive such that the probability of missing an active source is bounded. A new iterative heuristic called the Iterative Source Localization Procedure (ISLoP) is introduced that reduces the complexity of a source localization problem with J potential sources from 2^J to J per iteration, while also providing a local bounds on the maximum probability of a missed source. The ISLoP separates the source localization problem into a likelihood maximization problem followed by an active source localization problem. A diffusion example is used to demonstrate the performance of the ISLoP when compared to an estimation-based approach, where the heuristic is shown to have increasingly better performance as the bound on the maximum probability of a missed source is decreased. An experimental evaluation of the heuristic with respect to common wireless sensor networking errors is provided using a test bed implementation for a CO_2 sequestration site monitoring problem.

Keywords: source localization, detection, stochastic processes, wireless sensor networks

1. INTRODUCTION

In large-scale applications containing many spatially distributed sources, it is common to first detect whether any source exists before identifying the locations of the sources. The process of identifying the source locations is referred to as *source localization* and results in a multiple hypothesis testing problem consisting of only simple (non-composite) hypotheses. The problem of source localization has been addressed, in some form, in a variety of fields by many researchers, including in environmental monitoring by Bell (1962); Rossi et al. (2004); Saripalli et al. (2006), in communications by Ziskind and Wax (1988), in acoustic source localization by Sheng and Hu (2005); Zhang (2007), in fault detection by Lai (2000); Brumback and Srinath (1987), and in object tracking by Demetriou (2007); Liu et al. (2003).

With the recent advances in integrated circuit technology, miniaturized sensors with onboard wireless communication capabilities now exist, providing the means to perform monitoring in many situation for which monitoring was previously impractical due to hardware costs. Organized in networks, these wireless sensor-actuator devices provide unprecedented temporal and spatial sensing capabilities, which has allowed large-scale monitoring applications to flourish. When these monitoring applications concern the localization of many noisy sources, a large-scale source localization problem results.

The large-scale source localization problem is closely tied to the classical problem of signal detection, which is well studied and stems from the classical work of Neyman and Pearson (1933), Wald (1947), Cox and Anscombe (1952), and Armitage (1950). These seminal contributions have paved the way for many different approaches to source localization (e.g. see Willsky (1976) and Kailath et al. (1998)). The fundamental shortcoming of these approaches is that they are computationally infeasible when the number of sources is large (more than about 50 sources). In this paper, a new iterative heuristic is introduced that reduces the complexity of a source localization problem with J potential sources from 2^J to J per iteration, while providing a guaranteed bounds on the maximum probability of a missed source.

The following section introduces notation and formulates the source localization problem considered in this paper. Section 3 introduces the Iterative Source Localization Test (ISLoT) for large-scale source localization. Simulation results for large-scale source localization are provided in Section 4 for a diffusion example with a comparison to a feasible estimation-based test motivated by the work of Willsky and Jones (1974). Section 5 provides a robustness analysis of ISLoT with respect to common wireless sensor networking errors using a CO_2 sequestration site monitoring test bed. The concluding section summarizes the contributions of this paper and discusses directions for future work.

⋆ This work was performed in support of ongoing research in sensor systems and diagnostics at the National Energy Technology Laboratory under RDS contract DE-AC26-04NT41817, the Swedish Research Council, and the Knut and Alice Wallenberg Foundation.

2. PROBLEM FORMULATION

2.1 Preliminaries

In the following we assume that there are J potentially active sources and N observations, $z \in R^N$. The active sources are indicated by the values of a binary vector $b \in \{0,1\}^J$, where a unit entry in the j^{th} component of b (i.e. $b_j = 1$) indicates that source j is active. We write $\mathbf{1}$ and $\mathbf{0}$ to denote the binary vectors of all ones and all zeros, respectively, and e_j to denote the elementary binary vector with only a single unit entry in the j^{th} component. The set of all possible source vectors is $\mathcal{B} = \{0,1\}^J$, and for two source vectors, $b, b' \in \mathcal{B}$, we write $b \leq b'$ to denote that the unit entries of b are a subset of the unit entries of b' (i.e. $b^T b' = b^T b$). Using the above notation we define the sets for $b \in \mathcal{B}$

$$\begin{aligned} \mathcal{B}_b &= \{b' | b' \in \mathcal{B} \wedge b \leq b'\} \setminus \{b\} \\ \mathcal{T}_b &= \{b' | b' \in \mathcal{B} \wedge ||b - b'|| = 1\} \\ \mathcal{T}_b^+ &= \{b' | b' \in \mathcal{B}_b \wedge b' \in \mathcal{T}_b\} \end{aligned} \quad (1)$$

to describe different subsets source vectors.

2.2 Binary Hypothesis Testing

This subsection provides a brief summary of the classical test for accepting the null hypothesis in a binary hypothesis testing problem developed by Wald (1947). A binary hypothesis testing problem between a null hypothesis, H_0, and an event hypothesis, H_1, is written as

$$H_0 : \tilde{z} : f_0(z) \;\; \text{vs.} \;\; H_1 : \tilde{z} : f_1(z), \quad (2)$$

where $f_0(z)$ and $f_1(z)$ are the distributions of the observation random variables, $\tilde{z}$, under the null and event hypotheses, respectively. Given an observation, z, a test, $\phi(z) \in \{H_0, H_1\}$, for deciding between the null and event hypotheses is required to satisfy the constraint

$$P\left[\phi(z) = H_0 | H_1\right] \leq \gamma. \quad (3)$$

A test for accepting the null hypothesis, such that the constraint is satisfied, results from a worst case analysis of Wald's approximation (as discussed in Wald (1947)) where

$$\frac{f_1(z)}{f_0(z)} \leq \gamma \longrightarrow \phi(z) = H_0. \quad (4)$$

The above test for accepting H_0 will be used later in this section to satisfy the source localization performance criteria defined later in this section.

2.3 M-ary Hypothesis Testing for Source Localization

This subsection reviews M-ary hypothesis testing (see e.g. Poor (1994); Trees (1968); Scharf (1991)) and defines useful terminology for discussing source localization problems. An M-ary hypothesis testing problem for source localization is formulated as

$$\begin{aligned} H_{\mathbf{0}} &: \; b = \mathbf{0} \\ &\;\; \vdots \qquad , \\ H_{\mathbf{1}} &: \; b = \mathbf{1} \end{aligned} \quad (5)$$

where each hypothesis assumes a unique combination of active and inactive sources. The distribution of the observations under hypothesis H_b is denoted as $f_b(z)$, and is gaussian such that

$$f_b(z) : N\left[\mu_b, \Sigma_b\right], \quad (6)$$

where, following the standard assumption that sources are additive,

$$\mu_b = m_0 + \sum_{j=1}^{J} b_j m_j \;\; \text{and} \;\; \Sigma_b = S_0 + \sum_{j=1}^{J} b_j S_j. \quad (7)$$

The mean and covariance of the observations when no sources are active is denoted by m_0 and $S_0 \succ 0$, respectively, while $m_j > 0$ and $S_j \succ 0$ represent the change in the mean and covariance, respectively, when source j is assumed active. For an M-ary hypothesis testing problem, the globally most-likely hypothesis and locally most-likely hypothesis, and the probability of a missed source are defined as follows.

Definition 1. A hypothesis, $H_{b'}$, is the *globally most-likely* (GML) hypothesis for source localization if $\forall b \in \mathcal{B}$,

$$f_{b'}(z) \geq f_b(z) \quad (8)$$

Definition 2. A hypothesis, $H_{b'}$, is a *locally most-likely* (LML) hypothesis for source localization if $\forall b \in \mathcal{T}_{b'}$,

$$f_{b'}(z) \geq f_b(z) \quad (9)$$

Definition 3. The *probability of a missed source* (P_{MS}) of an M-ary hypothesis test, $\phi(z)$, for source localization is defined as

$$P_{MS}(b', \hat{b}) = \max_{b \in \mathcal{B}_{b'}} P\left[\phi(z) = H_{\hat{b}} | H_b\right]. \quad (10)$$

In words, the probability of a missed source is the maximum, over the hypotheses assuming additional active sources when compared to b', of the integral over the observation space of the corresponding distribution where $H_{\hat{b}}$ is accepted.

The classical approach to solving an M-ary hypothesis testing problem is to select the GML hypothesis (e.g. see Poor (1994)). To identify the GML hypothesis in (5) requires, in general, calculating the *log-likelihood* for each hypothesis H_b, $l_b(z)$, which can be written assuming gaussian distributions as

$$l_b(z) = -\frac{1}{2}(z - \mu_b)^T \Sigma_b^{-1} (z - \mu_b)^T + \ln \det \Sigma_b. \quad (11)$$

For large J, calculating the 2^J different likelihoods is impractical; moreover, simply selecting the GML hypothesis does not ensure the probability of a missed source is bounded.

2.4 Problem Statement

Given that $\gamma \in [0, 1]$ denotes the maximum probability of a missed source,[1] we consider the problem of identifying a *prominent source vector*, $b^P \in \mathcal{B}$, and an *active source vector*, $b^A \in \mathcal{B}$, such that

[1] The error occurring when too many sources are decided to be active will be evaluated as a performance measure for different source localization strategies in Section 4.

1. H_{b^P} is GML ; and
2. $b^A = \arg\min_{b \in \mathcal{A}_{b^P}} ||b||,$

$$\tag{12}$$

where

$$\mathcal{A}_{b^P} = \{b | b, b^P \in \mathcal{B} \wedge b \leq b^P \wedge P_{MS}(b, b^P) \leq \gamma\}. \tag{13}$$

The set $\mathcal{A}_{b'}$ denotes active source vectors which ensure that both the probability of a missed source is bounded and that the active sources in the prominent source vector are active. While part one in (12) seeks to find the sources that best describe the observation, part two tries to select the minimum number of active sources such that the sources that best describe the observations are active and the probability of a missed source is bounded. The following section, introduces a new source localization procedure that is feasible for large J and bounds the probability of a missed source.

3. ITERATIVE SOURCE LOCALIZATION PROCEDURE

This section introduces the Iterative Source Localization Procedure (ISLoP) as a two-part heuristic for identifying active sources consisting of *iterative likelihood maximization* (ILM) followed by *active source localization* (ASL), where ILM is concerned with identifying the prominent source vector, while ASL identifies the active source vector. The following subsections describe ILM and ASL, respectively.

3.1 Iterative Likelihood Maximization (ILM)

ILM is concerned with identifying the prominent source vector corresponding to the GML hypothesis. As discussed in Section 2, finding the GML hypothesis requires calculating the likelihood of each hypothesis. For a source localization problem containing 100 sources, calculating the GML requires $2^{100} \sim 1.26 \times 10^{30}$ different likelihood ratio calculations corresponding to each hypothesis, which is infeasible.[2] In this subsection, an iterative heuristic requiring J likelihood calculations at each iteration is introduced to identify a LML hypothesis as an approximation for the GML hypothesis to be used as the prominent source vector, b^P.

ILM identifies a LML hypothesis H_{b^P} by iteratively maximizing the likelihood over an evolving subset of the possible hypotheses according to

$$\begin{aligned}
& b^P := \mathbf{0} \\
& \text{while } (1) \\
& \quad b' := \arg\max_{b \in \mathcal{T}_{b^P}} l_b(z) \\
& \quad \text{if: } \quad l_{b^P}(z) \geq l_{b'}(z) \\
& \quad \text{then: } \text{return} \\
& \quad \text{else: } b^P := b' \\
& \text{end}
\end{aligned} \quad . \tag{14}$$

In words, ILM is initialized by assuming no active sources in the prominent source vector, $b^P = \mathbf{0}$. After initialization, ILM identifies the most likely source vector differing from the guessed prominent source vector by exactly one active source (either one more or one less active source), denoted by b' in (14). The resulting likelihood, $l_{b'}(z)$, is compared to the likelihood of the guessed prominent source vector, $l_{b^P}(z)$. If the prominent source vector is more likely, $l_{b^P}(z) \geq l_{b'}(z)$, ILM terminates and H_{b^P} is LML by Definition 2. Otherwise, ILM updates the guess of the prominent source vector with the maximum likelihood source vector, $b^P := b'$, and the process continues. Since ILM only calculates J likelihoods at each iteration, it is feasible for source localization problems with many sources. Once H_{b^P} is identified as LML, the ISLoP proceeds to active source localization.

3.2 Active Source Localization (ASL)

While ILM identifies the prominent source vector that best describes the observations, it does not identify the active source vector, b^A. Before introducing our approach for ASL, we observe that by applying the results for accepting the null hypothesis in a binary hypothesis testing problem from Section 2, which is based on the constraint

$$P_{MS}(b, b^P) \leq \gamma, \tag{15}$$

then

$$\max_{b \in \mathcal{B}_{b^A}} l_b(z) \leq \eta \longrightarrow \phi(x) = b^P \tag{16}$$

where $\eta = \ln \gamma + l_{b^P}(z)$. Since the active source vector norm is minimized in (12), ASL iteratively searches the source vectors in order from those with the smallest norms to largest according to

$$\begin{aligned}
& b^A := b^P \\
& \text{while } (1) \\
& \quad b' := \arg\max_{b \in \mathcal{T}^+_{b^A}} l_b(z) \\
& \quad \text{if: } \eta \geq l_{b'}(z) \\
& \quad \text{then: } \text{return} \\
& \quad \text{else: } b^A := b' \\
& \text{end}
\end{aligned} \tag{17}$$

In words, initially ASL assumes the active source vector, b^A, to be exactly the prominent source vector, b^P. For each hypothesis assuming an additional active source, the log-likelihood is calculated and compared to the guessed active source vector. For each hypothesis assuming an additional active source that prevents the decision to accept H_{b^P} from satisfying the constraint, the corresponding source is added to the ASL guess as an active source. This process continues until no additional source hypothesis (when compared to the current ASL assumption) violates the constraint on the probability of missed source. Once the ASL terminates, the unit entries of the active source vector denote which sources are assumed active by the ISLoP. The ISLoP developed in this section is evaluated through simulation and experimentation in the follow two sections.

4. SIMULATION RESULTS

The MSL strategy introduced in this work was simulated using a diffusion example motivated by a large area CO_2 sequestration site monitoring Weimer et al. (2010). In this

[2] Assuming $N = 10$, a 3.6 GHz machine with 2.9 GB of RAM requires $\sim 4 \times 10^{18}$ years to calculate the 2^{100} likelihoods.

example, 100 sensors and 100 sources are co-located in a grid formation. We assume a linear model of the form

$$\begin{bmatrix} x_{k+1} \\ z_{k+1} \end{bmatrix} = \begin{bmatrix} A & B\Gamma \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ z_k \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} w_k \\ h_k \end{bmatrix}$$
$$y_k = [\, I \;\; 0 \,] \begin{bmatrix} x_k \\ z_k \end{bmatrix} + v_k. \tag{18}$$

where

$$A = e^{300 A_c}$$
$$B = 10^{-6} \int_0^{300} e^{A_c \tau} \partial\tau \tag{19}$$

$A_c \in R^{100\times100}$ represents the continuous time dynamics of $\dot{x}_t = A_c x_t$ where the dynamics of all the interior locations evolve according to

$$\begin{aligned} \dot{x}_t(i,j) =& 0.1x_t(i-1,j) + 0.1x_t(i,j-1) - 0.4x(i,j) \\ &+ 0.1x(i+1,j) + 0.1x(i,j+1) \end{aligned} \tag{20}$$

where $x_t(i,j)$ denotes the value at position i, j. The exterior dynamics evolve assuming the far-field condition that $x_t(i,j) = 0$ for any i and j outside the field. The noise is distributed as

$$\begin{bmatrix} \tilde{w}_k \\ \tilde{v}_k \\ \tilde{h}_k \end{bmatrix} : N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10^{-4}\times I & 0 & 0 \\ 0 & 2(10^{-3})\times I & 0 \\ 0 & 0 & 10^2 \times I \end{bmatrix} \right). \tag{21}$$

The initial distribution on the process state and the source state is written as

$$\begin{aligned} \tilde{x}_0 &: N\left(\mathbf{0}, 10^{-3}\times I\right) \\ \tilde{z}_0 &: N\left(100\times\mathbf{1}, 2500\times I\right) \end{aligned}. \tag{22}$$

Using the system described above, the observation set for source localization is defined as

$$z = \begin{bmatrix} y_0 \\ \vdots \\ y_K \end{bmatrix}, \tag{23}$$

where the value of K is determined using aggregate source detection according to Weimer (2010). Aggregate source detection is a test, $\phi'(z) \in \{H_0, \neg H_0, H_{-1}\}$, which sequentially decides whether no sources are active ($\phi'(z) = H_0$), *some* source is active ($\phi'(z) = \neg H_0$), or more observations are needed to make a decision ($\phi'(z) = H_{-1}$). Aggregate source detection uses the SPRT introduced by Wald (1947) to implement a threshold test according to

$$\phi'(z) = \begin{cases} H_0 & \text{if } \max\limits_{b\in\mathcal{T}_\mathbf{0}} l_b(z) - l_\mathbf{0}(z) \le \eta_\beta \\ \neg H_0 & \text{if } \max\limits_{b\in\mathcal{T}_\mathbf{0}} l_b(z) - l_\mathbf{0}(z) \ge \eta_\alpha \\ H_{-1} & \text{otherwise} \end{cases}, \tag{24}$$

where η_β and η_α are functions of thresholds the maximum probability of *false alarm*, α, and maximum probability of *missed alarm*, β, as

$$\eta_\beta = \frac{1-\beta}{\alpha} \quad \text{and} \quad \eta_\alpha = \frac{\beta}{1-\alpha}. \tag{25}$$

In the following simulations, aggregate source detection is performed until it is decided that some source is active ($\phi'(z) = \neg H_0$), then ISLoP is performed using the same observations used to decide some source is active.

The ISLoP introduced in the previous section is compared to an estimation-based test motivated by the work of Willsky and Jones (1974) where by observing that inactive sources are equivalently active sources of zero magnitude, the source magnitudes are estimated assuming all the sources are active. Based on the source magnitude estimates, a threshold test is applied such that a source is considered active if the corresponding estimate exceeds the threshold and inactive if it does not. For comparison, we evaluate the performance of both the ISLoP and estimation-based tests in terms of the probability of a missed source and the percentage of correctly localized sources, where the percentage of correctly localized sources is the number of active sources correctly localized divided by the total number of sources decided to be active. In the following discussion, the strategy that correctly localizes a larger percentage of active sources for the same probability of missed source is considered to have better performance.

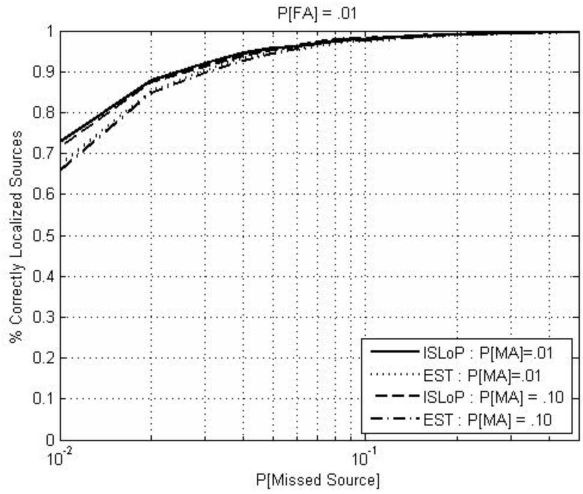

Fig. 1. Percentage of identified sources that are active vs. probability of type III error vs. probability of miss.

In Fig 1, the percentage of correctly localized sources is plotted against the probability of a missed source for tests assuming constant probability of false alarm, $\alpha = 0.01$, and varying probability of missed alarm, β, from 0.01 to 0.10 in aggregate source detection. These results suggest that there is not a significant difference between the percentage of correctly localized sources for the ISLoP and estimation-based approaches, with the ISLoP performing only marginally better. These results also indicate that the performance of both strategies is not significantly affected by changes in the probability of a missed alarm. This is a direct result of the threshold used for rejecting the hypothesis, η_α, in (25), not being significantly affected by small changes in the probability of a missed alarm.

In Fig. 2 the percentage of correctly localized sources is plotted against the probability of a missed source for tests assuming varying probability of false alarm from 0.01 to 0.10 and constant probability of a missed alarm, $\beta = 0.01$, in aggregate source detection. The results in Fig. 2 suggest that the percentage of correctly localized sources can be increased by reducing the probability of a missed alarm, where for larger probabilities of missed alarm, the ISLoP strategy increasingly outperforms the estimation-based strategy. Comparing the results in Figs. 1 and 2, we observe that a change in the probability of a

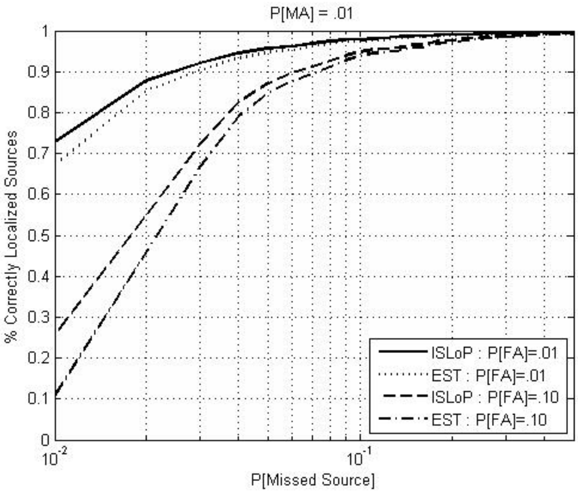

Fig. 2. Percentage of identified sources that are active vs. probability of type III error vs. probability of false alarm.

false alarm has a more significant effect on the percentage of correctly localized sources as compared to the same change in the probability of a missed alarm. This is due to the fact that when the probability of a false alarm is high, the number of observations required to reject the null hypothesis decreases (in general), and results in less information available to perform source localization.

Using a 3.6 GHz machine with 2.9 GB of RAM, the ISLoP was completed in 50 seconds, where approximately equal time (about 25 seconds) is consumed performing both ILM and ASL. These results illustrate that the ISLoP is a suitable test for source localization problems containing a large number of sources.

5. EXPERIMENTAL RESULTS

Aggregate source detection and the ISLoP have been applied to a CO_2 sequestration monitoring problem described in Weimer et al. (2010). In this problem, observations are gathered sequentially and sources can become active at any time (not just initially). To address these *emergent* sources, detection and localization is performed by iteratively applying the aggregate source detection and the ISLoP described in the previous section as illustrated in Fig. 3. In Fig. 3, the x-axis denotes the actual time and

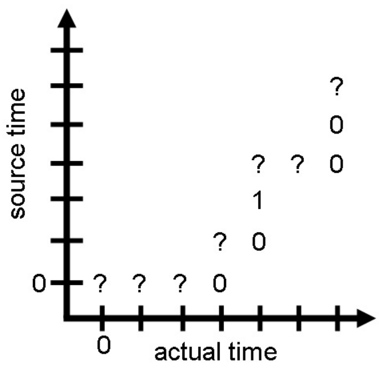

Fig. 3. Sequential aggregate source detection

the y-axis represents the source time, which is the time when sources are assumed to potentially become active. In the figure, a "?" means that more observations are needed to make a decision, a "0" equates to deciding no sources are active, and "1" denotes the decision that some sources are active. In the xcase illustrated in Fig. 3, it is assumed that sources can only become active at time zero (source time equals zero) and it was decided that more observations were needed. This continues until time step three when it is decided that no sources became active at time zero. Once a decision is made about whether sources became active at time zero, aggregate source detection immediately begins testing whether sources became active at time one. This process continues until a decision is made at time four that some sources became active at time two. After deciding some sources are active, the ISLoP is performed to localize the active sources. After localizing the active sources, aggregate source detection continues, but assumes that the localized sources are active. This process continues indefinitely.

To evaluate the performance of ISLoP strategy for the CO_2 sequestration monitoring application when a wireless sensor network (WSN) is used to gather the observations, a test bed consisting of 22 sensor nodes was developed. In the test bed setup, overhead light intensities are projected onto the test bed, where *dark* and *light* corresponds to areas of high and low CO_2 concentrations, respectively. The sensors measure the light intensity and transmit their measurements over the wireless network. Aggregate source detection and ISLoP use the observations to determine which (if any) sources are active. Figure 4 illustrates the how the CO_2 concentration evolves over time for 3 emergent sources. Experiments were performed using the

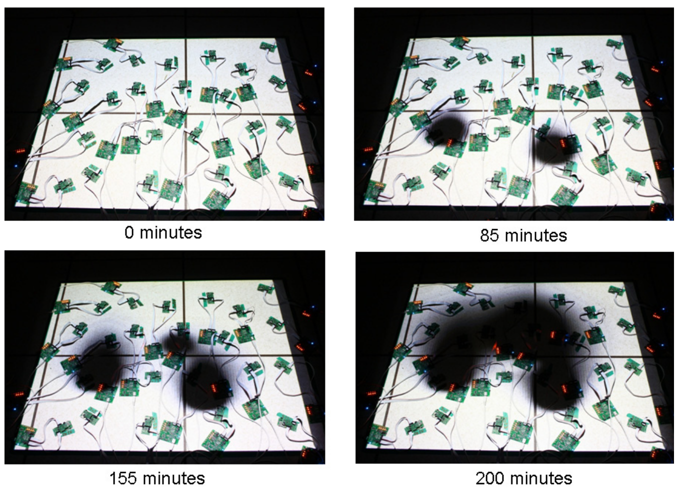

Fig. 4. Test bed experiment for 3 emergent sources at times 0, 85, 155, and 200 minutes

test bed to evaluate the ISLoP strategy in the presence of different errors common to WSNs: sensor localization error (sensor in the wrong position), model parameter error, and sensor death (sensor falls off the network). One thousand $(1,000)$ simulations were performed for different combinations of active source consisting of: synchronous distributed sources (S-D), synchronous clustered sources (S-C), asynchronous distributed sources (A-D), and asynchronous clustered sources (A-C). Sources are synchronous/asynchronous if they become active at the

Table 1. Percentage of correctly localized sources

	None	Sensor Location	Model Parameter	Sensor Death
S-D	89	80	71	69
S-C	93	85	78	75
A-D	88	80	72	69
A-C	75	71	58	49

same/different time(s). Sources are clustered/distributed if they are spatially near/far from one another. Table 1 illustrates the percentage of correctly localized sources for different source configurations and WSN error combinations assuming the maximum probability of a missed source is 0.10.

The results indicate that for all source scenarios, sensor failures and model parameter errors have the largest effects on the localization accuracy. In general asynchronous clustered sources result in the worst performance. This is a result of there not being enough information to distinguish which source, located near an already active source, has become active. The effects of the newly activated source are dampened by the already active source.

6. DISCUSSION AND FUTURE WORK

In this paper, the problem of large-scale source localization constrained by a maximum probability of a missed source is addressed. Using a new iterative heuristic, the source localization problem with J potential sources is reduced from 2^J to J per iteration while maintaining the desired bounds on the probability of a missed source. It is shown through simulation results that the iterative heuristic performs increasingly better than a feasible estimation-based approach as the probability of a missed source is decreased. An experimental evaluation of the heuristic with respect to common wireless sensor networking errors is provided using a test bed implementation for a CO_2 sequestration site monitoring problem illustrating that modeling errors and sensor failures significantly affect source localization performance.

Future work on this problem includes an evaluation of the performance on the localization heuristic when incorporated with a fault detection strategy. Based on the findings herein, further investigation is warranted as to the effect of the fault detector's performance on source localization and vice versa when observations are gathered sequentially.

REFERENCES

Armitage, P. (1950). Sequential analysis with more than two alternative hypotheses and its relation to discriminant function analysis. *Journal of the Royal Statistical Society*, Ser. B(12), 137–144.

Bell, T.G. (1962). Submarine and sonar detection. Technical Report 545, U.S. Navy Underwater Sound Lab.

Brumback, B. and Srinath, M. (1987). A chi-suare test for fault-detection in kalman filters. *IEEE Transactions on Automatic Control*.

Cox, D.R. and Anscombe, F.J. (1952). Sequential tests for composite hypotheses. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48, 290–+. doi: 10.1017/S030500410002764X.

Demetriou, M.A. (2007). Process estimation and moving source detection in 2-d diffusion processes by scheduling of sensor networks. *American Control Conference, 2007. ACC '07*, 3432–3437. doi:10.1109/ACC.2007.4282247.

Kailath, T., Fellow, L., and Poor, H.V. (1998). Detection of stochastic processes. *IEEE Trans. Inform. Theory*, 44, 2230–2259.

Lai, T.L. (2000). Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems. *Information Theory, IEEE Transactions on*, 46(2), 595–608. doi:10.1109/18.825826.

Liu, J., Reich, J., and Zhao, F. (2003). Collaborative in-network processing for target tracking. *Journal on Applied Signal Processing*, 4, 378–391.

Neyman, J. and Pearson, E.S. (1933). On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Royal Society of London Philosophical Transactions Series A*, 231, 289–337.

Poor, H.V. (1994). *An introduction to signal detection and estimation (2nd ed.)*. Springer-Verlag New York, Inc., New York, NY, USA.

Rossi, L.A., Krishnamachari, B., and Kuo, C.C.J. (2004). Monitoring of diffusion processes with pde models in wireless sensor networks. In *Defense and Security Symposium 2004*.

Saripalli, P., Amonette, J., Rutz, F., and Gupta, N. (2006). Design of sensor networks for long term monitoring of geological sequestration. *Energy Conversion and Management*, 47(13-14), 1968 – 1974.

Scharf, L.L. (1991). *Statistical Signal Processing, Detection, Estimation, and Time Series Analysis*. Addison-Welsley Publishing Company Inc., Reading, Massachusetts.

Sheng, X. and Hu, Y.H. (2005). Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *Signal Processing, IEEE Transactions on*, 53(1), 44–53. doi: 10.1109/TSP.2004.838930.

Trees, H.L.V. (1968). *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, Inc., New York.

Wald, A. (1947). *Sequential Analysis*. John Wiley & Sons, Inc., New York.

Weimer, J. (2010). Large-scale multiple source detection using wireless sensor networks. diss.

Weimer, J., Sinopoli, B., and Krogh, B. (2010). Wireless leak detection for carbon sequestration site surface monitoring. *submitted to International Journal of Greenhouse Gas Control*.

Willsky, A. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12, 601–611.

Willsky, A.S. and Jones, H.L. (1974). A generalized likelihood ratio approach to state estimation in linear systems subjects to abrupt changes. volume 13, 846–853. doi:10.1109/CDC.1974.270554.

Zhang, C. Zhang, Z.F.D. (2007). Maximum likelihood sound source localization for multiple directional microphones. *In ICASSP*.

Ziskind, I. and Wax, M. (1988). Maximum likelihood localization of multiple sources by alternating projection. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(10), 1553–1560. doi:10.1109/29.7543.