DETECTION AND LOCALIZATION OF MATERIAL RELEASES WITH SPARSE SENSOR CONFIGURATIONS

Emily B. Fox *Jason L. Williams* *John W. Fisher III* *Alan S. Willsky*

Massachusetts Institute of Technology Cambridge, MA 02139

1. ABSTRACT

We consider the problem of detecting and localizing a material release utilizing sparse sensor measurements. We formulate the problem as one of abrupt change detection. The problem is challenging because of the sparse sensor deployment and complex system dynamics. We restrict ourselves to propagation models consisting of diffusion plus transport according to a Gaussian puff model. We derive optimal inference algorithms, provided the model parametrization is known precisely, within a hybrid detection-localization hypothesis testing framework with linear growth in the hypothesis space. The primary assumptions are that the mean wind field is deterministically known and that the Gaussian puff model is valid. Under these assumptions, we characterize the change in performance of detection, time-to-detection and localization as a function of the number of sensors. We then examine some performance impacts when the underlying dynamical model deviates from the assumed model.

2. INTRODUCTION

The problem of detecting and localizing a material release utilizing sparse sensor measurements can be cast in a general framework applicable to many fields. In any distributed sensor network application, questions arise as to how the network should be configured to monitor the given environment as well as how the sensor measurements should be combined in an informative manner. In situations involving hazardous material, rapid detection is an important consideration while localization may be secondary.

It is frequently the case that only a sparse set of sensors can be deployed to monitor an area. In addition, these sensors are sensitive to local regions and as such yield only a myopic view of the environment when used in isolation. However, by defining a dynamical model it is possible to combine measurements from several sensors with non-overlapping fields of view so as to improve detection, time-to-detection, and localization beyond that obtained using a single sensor.

We restrict ourselves to scenarios where material propagation is reasonably well modeled by diffusion plus transport. Our goal is to characterize both the improvement of estimation performance attained by combining multiple sensor measurements as well as to characterize the sensitivity in performance to mismatched models.

We derive optimal inference algorithms for the case where the model is perfectly known, characterizing performance for a set of sensor configurations with regard to probability of detection P_d, time to detection T_d, and localization in time and space. This is accomplished via a series of controlled experiments using synthetic data.

Several researchers have considered Bayesian approaches to biological or chemical release detection. Release detection in hyperspectral imagery has been considered by [1] exploiting properties of isolated pixel intensities. Nofsinger uses ensemble Kalman filtering approaches for detecting chemical plumes [2]. Nehorai discusses plume modeling and moving release sources [3] without explicit consideration of the dynamics. Both O'Donnell [1] and Nehorai [3] utilize Generalized Likelihood Ratio formulations in their approaches. Our approach is novel in its use of abrupt change detection which allows for compact representation of hypotheses with only linear growth in complexity. Broad coverage of abrupt change detection is provided in [4].

3. PROBLEM FORMULATION

In a sparse sensor network, it will often be the case that there is a delay in observability of the event. The goal of the sensor network, then, is to intelligently combine sensor measurements distributed both temporally and spatially to first detect the release in a timely manner and secondarily to localize the release in space and time. The two goals are distinct in that reasonable detection may be obtained from early measurements while localization in both space and time generally relies on measurements aggregated over a longer time period. Bayesian filtering approaches provide a framework for this type of problem, presuming that the model parameters are available and inter-sensor communication is possible.

To develop a hybrid detection-localization hypothesis framework for abrupt changes, we modify the standard state space equations to include a term which represents a localized event in time and space (i.e. the release) whose effects persist over time.

Given a sparse sensor configuration and uncertainty regarding the underlying dynamics (e.g. rate of diffusion, knowledge of the wind field) it will generally be the case that the optimal inference approaches under an assumed parametrization will degrade in both detection and localization performance as compared to the matched case. Our results indicate that optimal inference on mismatched models yields reasonable detection performance at the cost of precise localization. Additionally, the methodology presented provides a framework within which to study these questions in more detail.

3.1. Bayesian approaches for localization

We divide the area of interest into a set of small regions or cells. The state of the system at time t is the mean particle concentration over one time step of the set of cells. The granularity of the cells is at a substantially finer resolution than the sensor placement.

We formulate a dynamical model for particle propagation by representing the wind velocities (u, v) as the sum of a deterministic and zero mean stochastic component, $(\bar{u}, \bar{v}) + (u', v')$, and assuming that the release is an instantaneous point source. This results in the following equation for computing the mean concentration of the material at a location (x, y) after a time t given an initial concentration

of c_o and diffusion coefficients (K_{xx}, K_{yy}) [5],

$$< c(x,y,t) > \quad = \quad \frac{c_o}{4\pi t(K_{xx}K_{yy})^{\frac{1}{2}}} e^{-\frac{(x-\bar{u}t)^2}{4K_{xx}t} - \frac{(y-\bar{v}t)^2}{4K_{yy}t}} \quad (1)$$

This Gaussian kernel is formed for every cell based on the given wind vector and distributes the material accordingly. The total mean particle concentration mapping, $A(t;\mathcal{V}_w)$, is the the sum of the new concentration maps from each cell.

In order to model localized abrupt changes, we add a term to the standard state space equations as in [6]. The advantage of this approach is that it allows one to model a localized release in time and space while capturing its persistent effects, which leads to an efficient hypothesis testing approach.

If we assume a release in cell *i* of strength β at time ϕ, the state space equations are given as follows:

$$x(t+1) \quad = \quad A(t;\mathcal{V}_w)x(t) + w(t) + \beta f_i(t,\phi) \quad (2)$$
$$y(t) \quad = \quad C(t)x(t) + v(t) \quad (3)$$

where $w(t)$ and $v(t)$ are independent Gaussian noise processes and $C(t)$ relates the state to the measurements $y(t)$. By linearity, we can decompose Equations 2 and 3 into $x_o(t)$ due to the background process and $\gamma_i(t,\phi)$ which captures the persistent effect of the release $f_i(t,\phi)$.

$$x(t) \quad = \quad x_0(t) + \beta\gamma_i(t,\phi) \quad (4)$$
$$y(t) \quad = \quad C(t)x_0(t) + v(t) + \beta C(t)\gamma_i(t,\phi) \quad (5)$$

The utility of the decomposition is that the problem reduces to one of detection in correlated noise (i.e. $x_o(t)$.) The detection formulations are greatly simplified by whitening the data using a Kalman filter and calculating the signature $\rho_i(t,\phi)$ of $\gamma_i(t,\phi)$ [6]:

$$\nu(t) = \nu_0(t) + \beta\rho_i(t,\phi) \quad (6)$$

3.2. Cell hypothesis testing framework

While the preceding framework has simplified the dynamical model we are still left with the task of testing various release hypotheses enumerated over space (cell index) and time. Additionally, the model is parameterized by the unknown release amount β which we estimate using a Generalized Likelihood Ratio (GLR).

Referring to Equation 6 and assuming that β is known for the moment, we can construct an indexed set of hypotheses. The null hypothesis, H_0, indicates no release while the alternative hypotheses, $H_{i,\phi}$, indicates an event in cell *i* occurring at time ϕ. Note that prior to time ϕ both the null and alternative hypotheses are identical.

$$H_0 \quad : \quad \nu(t) = \nu_0(t) \quad (7)$$
$$H_{i,\phi} \quad : \quad \nu(t) = \beta\rho_i(t,\phi) + \nu_0(t) \quad (8)$$

It is well known [6] that the sufficient statistics for this particular hypothesis test are the correlation between the signal $\rho_i(t,\phi)$ and the output $\nu(t)$,

$$d_i(t,\phi) \quad = \quad \sum_{\tau=\phi}^{t} \rho_i^T(\tau,\phi)V^{-1}(\tau)\nu(\tau) \quad (9)$$

and the energy of the signal

$$a_i(t,\phi) \quad = \quad \sum_{\tau=\phi}^{t} \rho_i^T(\tau,\phi)V^{-1}(\tau)\rho_i(\tau,\phi) \quad (10)$$

resulting in a log-likelihood of

$$l_i(t,\phi) = \beta d_i(t,\phi) - \frac{1}{2}\beta^2 a_i(t,\phi) = \frac{1}{2}\frac{d_i^2(t,\phi)}{a_i(t,\phi)} \quad (11)$$

for the hypothesis $H_{i,\phi}$ versus H_0 given sensor measurements up to time t. The last equality is a result of replacing β with its maximum likelihood estimate.

Under the assumption of there being one and only one release, the number of hypotheses grows linearly with time. The resulting hypothesis tree after three time steps is shown in Figure 1. In practice, a sliding window is used to implement the inference algorithm. This is justified by two related issues: it takes time for releases to propagate to the nearest sensor and after a period of time additional measurements do not contribute significantly to the probability of detection. Specifically, the window enforces that only hypotheses of a release occurring a maximum of M timesteps in the past are considered. The likelihood ratio test comparing H_1, the hypothesis that

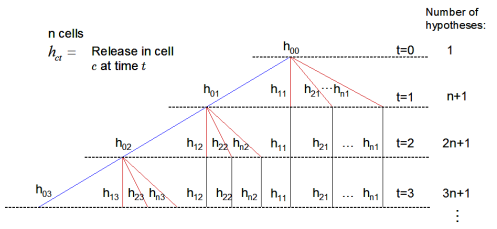


Fig. 1. Hypothesis tree for a region with n cells at time step 3. Hypothesis h_{ct} represents a release in cell c at time t.

a release has occurred at *any* time or location, to H_0, the hypothesis of benign conditions, is:

$$l_{\hat{i}}(t,\hat{\phi}) \quad \gtrless \quad \eta \quad (12)$$

where the value of η is set using Neyman-Pearson criterion to specify a fixed probability of false alarm, P_F and $l_{\hat{i}}(t,\hat{\phi}) \triangleq \max_{i,\phi} l_i(t,\phi)$.

4. EMPIRICAL RESULTS

We present a series of experiments in which we examine the utility of this inference framework in combining information from multiple sensors. The first set of experiments utilize a known mean wind field, specifically no transport term, to characterize probability of detection, time to detection, and localization in time and space as a function of the number of sensors used. In the second set of experiments, a sequence of increasing wind fields are incorporated into the dynamical model, but not into the inference procedure. In the third set of experiments, noise is generated on the mean wind field provided to the simulation, but the inference procedure only uses mean wind field. The purpose of the second and third set of experiments is to examine the robustness of this inference procedure to model deviations.

To simplify the analysis, we consider the case where the release is in the center of a room populated with an even grid of sensors. We hypothesis over a central 13x13 cell area of interest, but simulate over an area of 25x25 to reduce edge effects. There are four sensor configurations we examine: 1,2,4, and 16. The 16 sensor are spaced regularly throughout the 13x13 region and each other configuration is a inner subset of the 16 sensors. A sensor is assumed to observe the particle concentration throughout the entire cell in which it is

located with a measurement noise standard deviation of 10 material units per cell area. The diffusion coefficient of the Gaussian kernel is symmetrically 0.5. The process noise standard deviation is taken to be 100 material units per cell area.

For Sections 4.1 and 4.2, we restrict ourselves to a pure diffusion model in order to avoid the effects of wind which could blow the material towards or away from the nearest sensor. In Section 4.3, we examine robustness to increasing wind bias and temporally and spatially white noise on the mean wind field. For every release amount and sensor configuration scenario, 100 Monte Carlo simulations were performed.

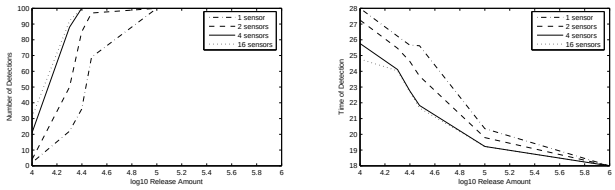
The size of the sliding window, M, was determined to be 14 time steps by analyzing the minimum time from the onset of the release by which the accrual of additional information was insignificant in the case of pure diffusion. In the four or 16 sensor configurations, pure diffusion provides the worst case scenario in time to nearest sensor.

The log-likelihood ratio threshhold was determined experimentally to achieve a false alarm rate of 0.01 on 1,500 Monte Carlo simulations under benign conditions (i.e. no release) for the four and 16 sensor configurations. The values are 9.55 and 10.85, respectively. These thresholds were used in all experiments conducted.

4.1. Detectability

Figure 2(a) shows the number of detections achieved out of 100 simulations as a function of release magnitude for various sensor configurations. The figure demonstrates that significant gains are made in detectability of small release amounts going from one to four sensors while there is only marginal improvement from four to 16. For a sufficient release amount, all sensor configurations are able to reliably detect. The reason for the insignificant increase in performance between four and 16 sensors in Figure 2 is that under pure diffusion with a release in the center, at least four sensors will see the release in both configurations. For large release amounts these four sensors provides significant evidence of a release while for small release amounts, the outer sensors do not accrue enough information to improve detection.

Similar reasoning holds in the corresponding time to first detection plots in Figure 2(b). All releases occurred at time step 16. The convergence for large release amounts is indicative of the time required for the release to propagate to the nearest sensor and is purely a function of the distance between the release and the sensor and thus sensor density.



(a) (b)

Fig. 2. (a) Number of detections per 100 simulation data sets and (b) time to first detection versus release amount. Actual release time was 16 seconds.

4.2. Localization uncertainty

In Figure 3 we examine one time slice of the hypothesis space (t=30). At any given time, there are multiple hypotheses of releases in a given cell, each differentiated by various times of releases. The maximum likelihood ratio of all hypotheses for each cell is plotted for various sensor configurations. Each hypothesis has an associated time of release and maximum likelihood release amount. The real sensor measurements provided to the inference algorithm for the four sensor case is plotted in the lower right-hand figure. For each of the other sensor configurations, the measurements provided were a subset of those provided in the four sensor case.

The diagrams give an illustration of the degree of localizability achievable with different sensor configurations. In the one sensor case, the uncertainty is approximately circular indicating that the location of the release cannot be distinguished within a ring around the sensor. The two sensor and four sensor cases provide progressively better localization. As expected, two sensors have difficulty distinguishing along a line while four sensors are able to localize to a point. The degree of shading in the plots indicates the likelihood ratio value and hence the detection probability.

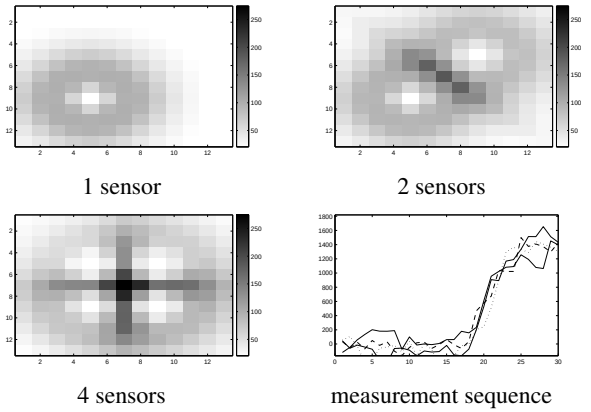


Fig. 3. Maximum likelihood ratio at every cell in region of interest for 1,2, and 4 sensor configuration. The measurement sequence for the 4 sensor configuration sequence is also shown.

4.3. Effect of unmodeled wind bias and variance

To analyze the effects of errors in the assumed known wind field, data produced under various wind biases and variances was provided to the inference algorithm which still assumed a pure diffusion model. The wind bias was always taken to be straight to the right such that it tends to transport the release between the two right-hand sensors in the four sensor configuration. The wind variance was an additive spatially and temporally white Gaussian noise on the known underlying mean wind field (pure diffusion, in this case). We examine the trends for a material release of size 1e5.

Probability of detection is fairly robust against wind variance, but drops off with increased wind bias because a strong enough bias will push the cloud of material between the sensors before it has dispersed enough to be registered. Specifically, for wind biases greater than 1.5 some of the 100 simulated releases were not detected. For all simulated wind variance cases, 100% of the releases were detected.

Figure 4 summarizes the localization performance for the four and 16 sensor configurations in three cases: nominal, unmodeled wind bias, and unmodeled wind variance. The plots indicate the x-direction localization mean and error variance. In the nominal and wind variance scenarios, this is nearly identical to the statistics in the

y-direction because we are operating under a zero mean wind field. In the wind bias scenario, this direction represents the principal axis of the covariance (same as the wind direction). The first two bars in each grouping indicate the estimates produced at the time of first detection while the second two bars are estimates after the system has accrued information until the detected hypothesis is "almost" outside the hypothesis window. We define "almost" to be a fixed lag parameter, n, which enforces that we stop accruing information when the window in which hypotheses of a release n time steps before the time corresponding to the initial best hypothesis are still considered. This is done because after the added time of gathering information, a hypothesis at an earlier time step may be better than the initially detected hypothesis. The main idea of this version of the algorithm is to alarm at the time of first detection and then wait an allotted period of time until producing localization results. The acceptable amount of time to wait is application dependent. Thus, the results in Figure 4 are indicative of the bounds on localization performance. Not

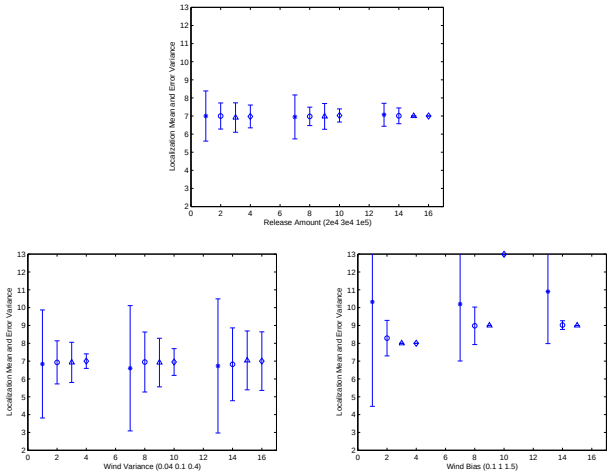


Fig. 4. Localization mean and error variance of 4 and 16 sensors at time of first detection ($*$ and o) and after accruing information ($\triangle$ and $\diamond$) for three values of release amount (top), unmodeled wind variance (bottom left), and unmodeled wind bias (bottom right).

surprisingly, the trend indicates that larger release amounts can be localized better and regardless of release amount, 16 sensors outperform four sensors. For smaller release amounts, 16 sensors do not gain much by waiting to accrue additional information. However, both sensor configurations achieve significantly better localization in the case of larger release amounts. These results can be explained by the fact that the signal to noise ratio at later time steps or further from the release location is lower. Thus, the added information accrued by the outlying sensors in the 16 sensor case for small releases will be negligible.

Localization estimates for the unmodeled wind bias case indicate that for smaller wind biases accruing information significantly helps localization. However, for larger wind biases the 16 sensor configuration inaccurately pinpoints the release to the edge of the room. This is because the inner four sensors only measure a small amount of the release as it passes through while the outer ring measure a much larger amount of material. The inference procedure assumes pure diffusion, so the most likely case is that the release was far from the inner four sensors. As a whole, these results highlight the point that localization performance suffers more quickly than detection performance when the wind field is mismatched to the inference procedure.

By comparing the wind variance case with the corresponding 1e5 release amount of the nominal case, one can analyze the degradation in localization performance caused by the model mismatch. Performance degrades with increased unmodeled wind uncertainty. In all cases, accruing more information provides better localization results. As with the matched model, the 16 sensor configuration outperforms the four sensor configuration. The disparity becomes less pronounced in the high variance with accrual situation. Comparing these results with those from an unmodeled wind bias, we see that the trends for randomness on the wind field better mirror those of the nominal matched case. This is because the disturbances on the propagation of particle concentrations caused by noise on the wind field are more similar to the type of errors modeled in the dynamic equation than the effects caused by a bias on the wind field.

5. DISCUSSION

We have presented a Bayesian state estimation approach to detecting and localizing a material release. Our approach allows for integration of measurements from multiple sensors over time. We have demonstrated the utility of this formulation and characterized the performance of a set of sensor configurations with regard to detection, time to detection and localization performance. We have also investigated aspects of model mismatch due to incorrect wind field assumptions. From the experiments conducted, we see that model mismatch impacts localization performance more than detectability.

The formulation presented provides a framework for answering questions about the interaction between release amount, release location, sensor density, and sensor placement. The appropriateness of this formulation for a given application depends on the validity of the assumptions we have made, such as a known mean wind field and abrupt releases of material. The linearity of a diffusion plus transport dynamic model makes this framework suitable in many scenarios.

Our approach integrates sensor measurements at a centralized processor. By arranging sensor nodes into groups, our approach provides the basic building blocks for a distributed processing configuration. The issues of how this arrangement should be conducted are a topic of future research. In addition, it is clear that knowing the model parameters is a critical factor in these approaches. Methods from machine learning may provide approaches to learning the underlying model parameters so as to reduce sensitivity to mismatch.

6. REFERENCES

[1] E. ODonnell et. al., "Identification and detection of gaseous effluents from hyperspectral imagery using invariant algorithms," *Proc of SPIE*, vol. 5425, pp. 573–582, April 2004.

[2] G.T. Nofsinger et. al., "Plume source detection using a process query system," in *SPIE Def&Security Symp Proc*, April 2004.

[3] T. Zhao and A. Nehorai, "Detection and localization of a moving biochemical source in a semi-infinite medium," in *Proc IEEE SAM Workshop*. 2004, IEEE.

[4] M. Basseville et. al., *Detection of Abrupt Changes: Theory and Applications*, Prentice Hall, 1993.

[5] J. Seinfeld and S. Pandis, *Atmospheric Chemistry and Physics*, John Wiley & Sons, 1997.

[6] A.S. Willsky et. al., "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems," *IEEE T Automat Contr*, vol. 21, no. 1, pp. 108–112, February 1976.