Signal Processing 132 (2017) 177–182

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

# Rao–Blackwell dimension reduction applied to hazardous source parameter estimation

Branko Ristic[a,*], Ajith Gunatilaka[b], Yan Wang[c]

[a] RMIT University, School of Engineering, Melbourne, VIC3000, Australia
[b] Defence Science and Technology Group, Land Division, Melbourne, Australia
[c] Tsinghua University, Institute of Public Safety Research, Department of Engineering Physics, Beijing, China

ARTICLE INFO

*Keywords:*
Bayesian estimation
Chemical–biological–radiological defence
Monte Carlo estimation
Plume dispersion

ABSTRACT

Parameter estimation of a source of chemical, biological or radiological emissions is a problem of great importance for public safety. The key parameters of interest are the source intensity and its location. This paper applies the concept of Rao–Blackwell dimension reduction to solve the posterior probability distribution function of source intensity, conditioned on source location, analytically. The paper is cast in the context of a source of a hazardous release of particles or gas and its turbulent transport through the medium. Numerical results, obtained by simulations and using an experimental dataset, demonstrate the statistical efficiency of the proposed method.

## 1. Introduction

The threat of a hazardous chemical–biological–radiological (CBR) attack, either in a form of a release of toxic biochemical aerosols into the atmosphere or an improvised nuclear device, has been well documented [1,2]. For the sake of public safety, it is very important to rapidly detect and localise the CBR source so that the mitigation actions can be carried out promptly. The problem of CBR source localisation has been studied for quite some time. The standard solutions are based on optimisation techniques, such as nonlinear least squares [3,4] and maximum likelihood estimation [5]. These approaches can fail due to local minima or poor convergence and in addition provide only point estimates without uncertainty attached to it. The alternatives are the Bayesian techniques [6–12], which estimate the posterior probability density function (PDF) of a source, thereby providing an uncertainty measure to any point estimate derived from it. Among the Bayesian techniques, Markov chain Monte Carlo (MCMC) is the dominant method for estimation of the posterior PDF, applied in [6,8–10]. Other approaches to posterior estimation are numerical integration [7], importance sampling [11], and approximate Bayesian computation [12].

In formulating the problem, one needs to specify the parameter vector, the dispersion model and the measurement model. The parameter vector $\boldsymbol{\theta}$ typically includes the source release rate or intensity, $Q_0 \in \mathbb{R}^+$, the position of the source $\mathbf{r}_0 \in \mathbb{R}^d$ ($d$=2 or 3), and possibly environmental parameters (e.g. the mean wind speed and canopy characteristics). The dispersion or propagation model describes via mathematical equations the mean concentration of an emitted biochemical substance, or the mean radiation field, at a given sensor location, as a function of the parameter vector $\boldsymbol{\theta}$. All dispersion/propagation models used in practice are nonlinear with respect to the source location $\mathbf{r}_0$, but linear with respect to intensity $Q_0$. This linear relationship is the key feature exploited in the paper. Finally, the measurement model relates the mean concentration to sensor measurements that are affected by stochastic fluctuations. The most adequate model for this purpose has been experimentally found to be the Poisson distribution – either in the context of measuring the count of the radiated photons [13] or the count of dispersed biochemical particles [14].

The source parameter estimation problem is approached in the Bayesian framework. The main novelty of the paper, compared to the earlier Bayesian approaches, is that it applies the concept of Rao–Blackwell dimension reduction [15,16] to the problem. In doing so, we derive the analytic expressions for: (a) the posterior PDF of source intensity $Q_0$ (conditioned on the source location) and (b) the likelihood of the source location parameter (only). Numerical analysis, using both simulated and experimental data, is carried out to demonstrate the gains of the proposed, dimension reduced, estimation method. The presentation of the paper, without loss of generality, will focus on a hazardous biochemical substance release into the atmosphere, using an open-field two-dimensional ($d$=2) dispersion model.

* Corresponding author.
*E-mail address:* branko.ristic@rmit.edu.au (B. Ristic).

http://dx.doi.org/10.1016/j.sigpro.2016.10.005
Received 3 March 2016; Received in revised form 22 July 2016; Accepted 10 October 2016
Available online 12 October 2016

## 2. Models and problem formulation

We adopt a dispersion model of turbulent transport through the medium from [14], used in a number of recent publications [17–21]. Consider a source of particle release into the environment (atmosphere and water), characterised by a constant emission rate $Q_0 > 0$ and located at $\mathbf{r}_0 = (x_0, y_0)^\intercal$. The particles propagate through the medium with the isotropic diffusivity $D$, but can also be advected by flow (due to wind or current), whose mean direction and average speed $V$ are known. Let us adopt the convention that the mean flow (wind) direction coincides with the direction of the $x$-axis. The average lifetime of a particle (before being absorbed) is $\tau$. A spherical sensor of small size $a$ at a location with coordinates $\mathbf{r} = (x, y)^\intercal$, non-coincidental with the source location $\mathbf{r}_0$, will experience a series of encounters with the released particles at the rate [14]:

$$R_{\boldsymbol{\theta}}(\mathbf{r}) = \frac{Q_0}{\ln\left(\frac{\lambda}{a}\right)} \exp\left[\frac{(x_0 - x)V}{2D}\right] \cdot K_0\left(\frac{\sqrt{(x - x_0)^2 + (y - y_0)^2}}{\lambda}\right) \tag{1}$$

where $K_0$ is the modified Bessel function of order zero and

$$\lambda = \sqrt{\frac{D\tau}{1 + \frac{V^2\tau}{4D}}}. \tag{2}$$

The rate of particle encounters $R_{\boldsymbol{\theta}}(\mathbf{r})$ is expressed in (1) as a function of the (unknown) parameter vector $\boldsymbol{\theta} = [Q_0, \mathbf{r}_0^\intercal]^\intercal$.

Suppose a network of $S$ spatially distributed sensors is measuring the concentration of emitted particles. The stochastic process of sensor encounters with emitted particles is modelled by a Poisson distribution: the probability that $i$th sensor at location $\mathbf{r}_i = (x_i, y_i)^\intercal$ encounters $z_i \in \mathbb{Z}^+$ particles during a time interval $t_0$ is then:

$$\mathcal{P}(z_i; \mu_i) = \frac{\mu_i^{z_i}}{z_i!} e^{-\mu_i} \tag{3}$$

where $\mu_i = R_{\boldsymbol{\theta}}(\mathbf{r}_i) t_0$ is the mean concentration at $\mathbf{r}_i$ and $i = 1, \ldots, S$. All sensors are assumed to be of the same (and known) size $a$. Because the environmental parameters $\tau$, $D$ and $V$, are also known, (3) represents the full specification of the likelihood function of $\boldsymbol{\theta}$, given a measurement $z_i$ (taken at location $\mathbf{r}_i$). Assuming the sensor measurements, conditioned on $\boldsymbol{\theta}$, are independent, the likelihood function of the measurement vector $\mathbf{z} = [z_1, z_2, \ldots, z_S]^\intercal$ can be written as a product $\ell(\mathbf{z}|\boldsymbol{\theta}) = \prod_{i=1}^{S} \mathcal{P}(z_i; t_0 R_{\boldsymbol{\theta}}(\mathbf{r}_i))$.

The parameter estimation problem is formulated in the Bayesian framework. The goal is to compute the posterior PDF $p(\boldsymbol{\theta}|\mathbf{z})$, which provides a complete probabilistic description of the information contained in $\mathbf{z}$ about $\boldsymbol{\theta}$. To compute the posterior distribution, in addition to $\ell(\mathbf{z}|\boldsymbol{\theta})$ one needs to specify the prior distribution of the parameter vector $\pi(\boldsymbol{\theta})$. Using Bayes' rule

$$p(\boldsymbol{\theta}|\mathbf{z}) = \frac{\ell(\mathbf{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int \ell(\mathbf{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}}. \tag{4}$$

Quantities of interest related to $\boldsymbol{\theta}$ (e.g., the posterior mean and variance) can be computed from $p(\boldsymbol{\theta}|\mathbf{z})$. Note that the prior $\pi(\boldsymbol{\theta})$ is typically non-Gaussian: the source position is often restricted to polygon regions, while $Q_0$ is strictly positive. Optimal Bayesian estimation is generally impossible because the posterior PDF cannot be found in closed-form. This is certainly the case for the signal model described above.

## 3. Rao–Blackwell dimension reduction

This section presents the key result: if the source intensity and its location are independent, the posterior PDF of source intensity, conditioned on the source location, can be calculated analytically. Consequently, Monte Carlo estimation can be applied to a reduced dimension of the parameter vector space.

Using the chain rule one can write the posterior PDF as follows:

$$p(\boldsymbol{\theta}|\mathbf{z}) = p(Q_0|\mathbf{r}_0, \mathbf{z}) p(\mathbf{r}_0|\mathbf{z}) \tag{5}$$

Suppose for a moment that we can calculate the posterior $p(Q_0|\mathbf{r}_0, \mathbf{z})$ analytically. Then we need to apply a Monte Carlo estimation method to compute only $p(\mathbf{r}_0|\mathbf{z})$, which according to Bayes' rule is given by:

$$p(\mathbf{r}_0|\mathbf{z}) = \frac{g(\mathbf{z}|\mathbf{r}_0)\pi(\mathbf{r}_0)}{\int g(\mathbf{z}|\mathbf{r}_0)\pi(\mathbf{r}_0) d\mathbf{r}_0}. \tag{6}$$

The problem with (6) is that the likelihood function $g(\mathbf{z}|\mathbf{r}_0)$ is also unknown – only $\ell(\mathbf{z}|\boldsymbol{\theta})$ is known. Hence, in order to apply the Rao–Blackwell dimension reduction, we need not only the analytic expression for $p(Q_0|\mathbf{r}_0, \mathbf{z})$, but also for $g(\mathbf{z}|\mathbf{r}_0)$.

Due to the independence assumption between the source intensity and its location, $\pi(\boldsymbol{\theta}) = \pi(Q_0)\pi(\mathbf{r}_0)$. Let us assume the prior $\pi(Q_0)$ is a Gamma distribution with shape parameter $\eta_0$ and scale parameter $\vartheta_0$, that is

$$\pi(Q_0) = \mathcal{G}(Q_0; \eta_0, \vartheta_0) = \frac{Q_0^{(\eta_0 - 1)} e^{-Q_0/\vartheta_0}}{\vartheta_0^{\eta_0} \Gamma(\eta_0)}. \tag{7}$$

For suitably chosen hyperparameters $\eta_0$ and $\vartheta_0$, this prior can be diffuse, with the support covering a large span of possible values of $Q_0$. Recall that the likelihood function $\ell(\mathbf{z}|\boldsymbol{\theta})$ is a product of Poisson distributions, which we can write in a slightly different form as:

$$\ell(\mathbf{z}|\boldsymbol{\theta}) = \prod_{i=1}^{S} \mathcal{P}(z_i; Q_0 \cdot \rho_{\mathbf{r}_0}(\mathbf{r}_i)) \tag{8}$$

where

$$\rho_{\mathbf{r}_0}(\mathbf{r}_i) = t_0 R_{\boldsymbol{\theta}}(\mathbf{r}_i)/Q_0 \tag{9}$$

is independent of $Q_0$. This is a consequence of the dispersion model being linear[1] with respect to $Q_0$.

Recall that the conjugate prior of the Poisson distribution is the Gamma distribution [22]. Therefore, we can expect the posterior $p(Q_0|\mathbf{r}_0, \mathbf{z})$ to be also a Gamma distribution, $p(Q_0|\mathbf{r}_0, \mathbf{z}) = \mathcal{G}(Q_0; \eta, \vartheta)$, with parameters $\eta$ and $\vartheta$ that can be calculated analytically as a function of $\mathbf{r}_0$ and $\mathbf{z}$. Note that the fact that the likelihood is a product of Poisson distributions (rather than a single Poisson) does not change the scheme, because one can think of this product as an arbitrary order sequence of updates of the Gamma distributed random variable with Poisson distributed measurements, which results in a sequence of Gamma distributed random variables.

**Proposition:** *The parameters $\eta$ and $\vartheta$ of the posterior $p(Q_0|\mathbf{r}_0, \mathbf{z}) = \mathcal{G}(Q_0; \eta, \vartheta)$ can be calculated as follows:*

$$\eta = \eta_0 + \sum_{i=1}^{S} z_i, \tag{10}$$

$$\vartheta = \frac{\vartheta_0}{1 + \vartheta_0 \sum_{i=1}^{S} \rho_{\mathbf{r}_0}(\mathbf{r}_i)}. \tag{11}$$

**Proof.** The proof is based on two properties of the Gamma distribution: (i) If $X \sim \mathcal{G}(\eta, \theta)$ then for any constant $c > 0$, $cX \sim \mathcal{G}(\eta, c\theta)$ [23]. (ii) If $\mathcal{G}(\mu; \eta, \theta)$ is a prior distribution and $n$ is a sample from the Poisson distributed likelihood function with parameter $\mu$, then the posterior is [22] $\mathcal{G}(\mu; \eta + n, \theta/(1 + \theta))$. Consider for simplicity only the first measurement, $z_1$, collected at position $\mathbf{r}_1$. The likelihood function $\ell(z_1|\boldsymbol{\theta})$ is Poisson distributed with mean $\mu_1 = Q_0 \rho_{\mathbf{r}_0}(\mathbf{r}_1)$, where $Q_0 \sim \mathcal{G}(Q_0; \eta_0, \vartheta_0)$. According to Properties (i) and (ii),

[1] All commonly used dispersion/propagation models are linear with respect to $Q_0$. While in the paper we adopt the model from [14], the method is universal.