Radiation Measurements 46 (2011) 1138–1146

Contents lists available at SciVerse ScienceDirect

Radiation Measurements

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

Detection of radionuclides from weak and poorly resolved spectra using Lasso and subsampling techniques

Er-Wei Bai[a,*], Kung-sik Chan[b], William Eichinger[c], Paul Kump[a]

[a] Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242, USA
[b] Department of Statistical and Actuarial Science, University of Iowa, Iowa City, IA 52242, USA
[c] Department of Civil and Environmental Engineering, University of Iowa, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 29 April 2011
Received in revised form
21 August 2011
Accepted 22 August 2011

Keywords:
Lasso
Subsampling
Radionuclide
Detection
Identification

ABSTRACT

We consider a problem of identification of nuclides from weak and poorly resolved spectra. A two stage algorithm is proposed and tested based on the principle of majority voting. The idea is to model gamma-ray counts as Poisson processes. Then, the average part is taken to be the model and the difference between the observed gamma-ray counts and the average is considered as random noise. In the linear part, the unknown coefficients correspond to if isotopes of interest are present or absent. Lasso types of algorithms are applied to find non-vanishing coefficients. Since Lasso or any prediction error based algorithm is inconsistent with variable selection for finite data length, an estimate of parameter distribution based on subsampling techniques is added in addition to Lasso. Simulation examples are provided in which the traditional peak detection algorithms fail to work and the proposed two stage algorithm performs well in terms of both the False Negative and False Positive errors.

1. Introduction

There has long been an interest in detecting the covert movement of radioactive materials. With the collapse of the Soviet Union, the poor security associated with its nuclear stockpile and the War on Terror, there is an increased interest in uncovering smuggled nuclear materials. This has resulted in reassessment of the potential nuclear signatures of these materials which could be used to detect them, particularly with respect to the emission of gamma-rays. While a number of active detection methods have been proposed (Fetter et al., 1990; Moss et al., 2002; Weller et al., 2006) which can detect relatively small amounts of material, they are ultimately not suitable for use in public areas because they require potentially lethal doses of some type of radiation to be delivered to the object to be examined. This leaves passive sensing as the primary candidate method for use in areas where human is present. A number of both portable and non-portable gamma-ray spectrometers are commercially available; all suitable for such applications as checkpoint monitoring and nuclear searches where the detector and potential source are close and the expected signal is relatively large so that the gamma-ray signature peaks are well-resolved (August and Whitlock, 2005). Gamma-ray signature recognition is straightforward if signals are strong and high-resolution detectors can be used. Such detectors allow unambiguous identification of radioactive nuclide using photo peak search algorithms (Murray, 1998). However, there exist a large number of applications in which the detectors cannot reasonably be expected to be close to the sources, the exposure time may be short and there may have a wide range of materials between the source and the detector. In these cases, the detectors generally used have large volume (to increase sensitivity) and poor resolution. The signals from these detectors will be weak and difficult to separate from the background radiation or from the signatures of commonly used radioactive materials. This paper addresses the problem of identification of nuclides from weak and poorly resolved spectra. Despite the fact that we know in detail the gamma-ray emission spectra of the individual radioactive nuclide of interest, the exact nature of the gamma-ray signatures from the actual objects of interest are unknown. Signals from commonly encountered natural and man-made gamma-ray sources can compete with the signals of interest and cause false alarms. The emissions from nuclear weapons materials Pu239 for example are quite weak, with many of the lines at relatively low energies so that they are easily attenuated. In spite of these difficulties, nuclear weapons materials spectra still remain unique, but may be subtle and difficult to recognize by traditional peak searching detection algorithms.

* Corresponding author. Tel.: +1 319 335 5949.
E-mail addresses: er-wei-bai@uiowa.edu (E.-W. Bai), kung-sik-chan@uiowa.edu (K.-s. Chan), william-eichinger@uiowa.edu (W. Eichinger).

1350-4487/$ – see front matter
doi:10.1016/j.radmeas.2011.08.020

Another existing approach is transforming the measurements into a vector of background-adjusted difference referred to as the spectral comparison ratio (SCR) method (Wei et al., 2010; Anderson et al., 2008). Properties of the ratios can be used to determine whether the unknown measurement is similar to that of the benign measurement.

The contribution of the paper is the development of a new two stage detection algorithm that combines recent regression methodologies with subsampling techniques based on the idea of majority voting. The algorithm is built on a Poisson model of gamma-ray emission that leads to a linear regression model. Then, an algorithm referred to as the Lasso together with subsampling techniques are applied. Theoretical and simulation studies reported in this paper demonstrate promising results for the proposed algorithm.

2. Modeling

The emission of gamma-ray from a radioactive source is a random variable. For a given range of energies (keV or spectrum channels), the number of gamma-ray counts registered by a detector per unit time and per unit amount of source material follows a Poisson distribution (Killian and Hartwell, 2000), $((\lambda)^k/k!)e^{-\lambda}$. The parameter λ primarily depends on the radioactive decay of the source, the characteristics of the detector material, with some contributions from environmental effects such as temperature effects on electronics and the position of the detector relative to the source. If β units of radioactive source material are present, the number of gamma-ray counts k follows

$$z \sim \frac{(\lambda\beta)^k}{k!}e^{-\lambda\beta}$$

where $\lambda\beta = Ez = E(z-\lambda\beta)^2$ is the mean value as well as the variance of the random variable. E stands for the expectation operator. $\beta = 0$ indicates that the source is absent. In the presence of multiple radioactive sources and the background radiation, the number of the gamma-ray counts registered at a detector is the sum of contributions from all sources and the background. For simplicity, the background may and will be considered as a radiation source.

In the presence of p radioactive sources, it is reasonable to assume that the contribution of individual source is statistically independent. Now, since the gamma-ray counts for each source is Poisson distributed,

$$\frac{(\lambda_i\beta_i)^k}{k!}e^{-\lambda_i\beta_i}, \quad i = 1,2,\ldots,p$$

the total gamma-ray counts by the properties of the Poisson distribution is also Poisson distributed (Ash, 2008).

$$z \sim \frac{\left(\sum_{i=1}^{p}\lambda_i\beta_i\right)^k}{k!}e^{-\sum_{i=1}^{p}\lambda_i\beta_i}. \tag{1}$$

The above analysis applies to each range of energy (keV) of the spectra. Suppose a spectrum has N channels. Then, the gamma-ray counts is a random vector $Z = (z_1, z_2,\ldots, z_N)'$ and each z_i is the gamma-ray counts in a given channel that follows a Poisson distribution,

$$z_i \sim \frac{(\lambda_{i1}\beta_1 + \lambda_{i2}\beta_2 + \ldots + \lambda_{ip}\beta_p)^k}{k!}e^{-(\lambda_{i1}\beta_1+\lambda_{i2}\beta_2+\ldots+\lambda_{ip}\beta_p)}, \quad i = 1,2,\ldots,N \tag{2}$$

where the vector $(\lambda_{1i}, \lambda_{2i},\ldots, \lambda_{Ni})^T$ is the spectrum of the ith radioactive source.

Define

$$\lambda = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1p} \\ \vdots & \ddots & \cdots \\ \lambda_{N1} & \cdots & \lambda_{Np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \tag{3}$$

where $(\lambda_{i1}\beta_1 + \lambda_{i2}\beta_2 + \ldots + \lambda_{ip}\beta_p)$ is the mean gamma-ray counts per unit time received by the detector at the ith spectrum channel from all radioactive sources. For a little abuse of notation, the random gamma-ray counts vector Z follows

$$Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{pmatrix} \sim \frac{(\lambda\beta)^k e^{-\lambda\beta}}{k!}$$

where $e^{-\lambda\beta}$ and $(\lambda\beta)^k$ are calculated component-wise. If we further assume that the number of gamma-ray counts at different channels are also independent, then the mean value and the variance of Z are given by

$$EZ = \lambda\beta = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1p} \\ \vdots & \ddots & \cdots \\ \lambda_{N1} & \cdots & \lambda_{Np} \end{pmatrix}\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$E(Z-\lambda\beta)(Z-\lambda\beta)^T = \begin{pmatrix} \lambda_{11}\beta_1+\ldots+\lambda_{1p}\beta_p & 0 & \cdots & 0 \\ 0 & \lambda_{21}\beta_1+\ldots+\lambda_{2p}\beta_p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{N1}\beta_1+\ldots+\lambda_{Np}\beta_p \end{pmatrix}$$

Let Z_i, $i = 1, 2,\ldots, l$, be l independent observations of Z. Define the empirical average of Z by

$$\overline{Z} = \frac{1}{l}\sum_{i=1}^{l} Z_i$$

that satisfies

$$\overline{Z} = \lambda\beta + \underbrace{\frac{1}{l}\sum_{i=1}^{l} Z_i - EZ}_{\overline{V}} = \lambda\beta + \overline{V} \tag{4}$$

which is cast in a linear regression form and can be used to estimate the strength coefficients β where $\overline{Z}$ is available from measurements, the relative magnitude matrix λ is known for given radioactive sources and $\overline{V}$ is a random variable which can be considered as noise.

It is easy to verify

$$E\overline{V} = \frac{1}{l}\sum_{i=1}^{l} EZ^i - EZ = 0,$$

$$E\overline{V}\,\overline{V}^T = \frac{1}{l}\sum_{i=1}^{l} E\left(Z^i-\lambda\beta\right)\left(Z^i-\lambda\beta\right)^T$$

$$= \frac{1}{l}\begin{pmatrix} \lambda_{11}\beta_1+\ldots+\lambda_{1p}\beta_p & 0 & \cdots & 0 \\ 0 & \lambda_{21}\beta_1+..+\lambda_{2p}\beta_p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{N1}\beta_1+\ldots+\lambda_{Np}\beta_p \end{pmatrix}$$

Notice that $\sum_{i=1}^{l} Z^i$ is also Poisson. If the gamma-ray counts are high, the Poisson $\sum_{i=1}^{l} Z^i$ is very close to a Gaussian distribution