Radioactive Source Localization in Urban Environments with Sensor Networks and the Internet of Things*

Clair J. Sullivan[1]

Abstract— The use of radiation detectors as an element in the so-called "Internet of Things" has recently become viable with the available of low-cost, mobile radiation sensors capable of streaming geo-referenced data. New methods for fusing the data from multiple sensors on such a network is presented. The traditional simple and ordinary Kriging methods present a challenge for such a network since the assumption of a constant mean is not valid in this application. A variety of Kalman filters are introduced in an attempt to solve the problem associated with this variable and unknown mean. Results are presented on a deployed sensor network.

I. INTRODUCTION

Since September 11, 2001 a great deal of research has been placed into the area of nuclear counterterrorism. A great problem exists detecting the illicit movement of special nuclear material (SNM) that could be used in a nuclear weapon or radioactive materials that could be used in a radiological disperal device or dirty bomb. This problem is not dissimilar to a variety of other types of detection problems in that there are signals emitted from the source (in this case, ionizing radiation in the form of gamma rays or neutrons) that may be detected with some probability. This signal exists within a background noise signal that determines both the probability of detection of a signal while also providing a source of potential false alarms. Since some level of adjudication is required for every alarm, it is clearly desirable to increase the probability of detection of the signal while also reducing the probability of false alarms. The probability of detection and false alarm can be determined either statistically or measured directly or indirectly.

Early generations of radiation sensor networks that were explored circa 2005 employed tens of detectors that were stationary and separated by distances that were much larger than the source detection range of even substantial masses of SNM. [1], [2], [3], [4] This was done to minimize cost while maximizing the coverage area. Further, the computational complexity requied to handle the data generated by even this small of a network was impractical at the time.

New generations of radiation sensor networks are coming online that will eventually employ thousands of mobile nodes. [5], [6], [7], [8] These small, mobile detectors can also be coupled to larger and more efficient stationary detectors for enhanced detection probability over the entire network. This paper describes the benefits and complications of working with large sensor networks for radiation detection, the added benefits sensor fusion can provide, and suggests opportunities presented by treating such networks as a data stream within the so-called Internet of Things.

II. GENERAL CONCEPTS IN RADIATION DETECTION

A variety of detector types exist for measuring ionizing radiation. In order to select a radiation detector, the following characteristics need to be considered in relationship to the desired application:

1) Type of radiation to be detected: Most radiation detectors are limited to the ability to detect one type of radiation (e.g. alpha particles, beta particles, gamma rays, or neutrons). However, some sources of SNM such as ^{239}Pu emit multiple different types of radiation. Being able to simultaneously detect and distinguish among these types offers an excellent advantage in terms of reducing false alarms. However, the data fusion problem presented by the use of data from multiple types of detectors must be considered.

2) Efficiency of the detector: As in other disciplines, it is desirable to have a detector with a high efficiency. However, in radiation detection, efficiency is determined by the detecting material, its density, and its geometry. Most notably, the larger the detector the larger its geometric efficiency. Based on the desired application, for example if the device must be human-portable, there can be significant limitations to how large a detector can be.

3) Anticipated distance to the source: If possible and controllable, it is important to get as close to the source of radiation as possible. This is because the flux of radiation decreases based on $\frac{1}{r^2}$. So if the distance to the source is double, the flux decreases by a factor of four.

4) Presence of intervening shielding materials: The flux of ionizing radiation is attenuated through well-understood physical processes. Typically the decrease in flux follows an exponential relationship is a function of the thickness of the absorber and the total cross section of the absorbing material to the incident radiation.

5) Available integration time: Radioactive decay is a Poissonian process whose distribution is known. As more

[1]Clair J. Sullivan is faculty member of the Department of Nuclear, Plasma, and Radiological Engineering, University of Illinois at Urbana-Champaign, 104 S. Wright St., Urbana, IL 61801, USA cjsulli@illinois.edu

events are collected by the detector, the uncertainty in the measured quantity decreases. Hence, the longer that a detector can count a given source, the better the statistics will be.

In order to create a practical sensor network to detect the illicit movement of sources, there are some clear requirements and implications on the above list. First, SNM like ^{235}U and ^{239}Pu and their associated daughter isotopes emit many different types of radiation. However, only gamma ray and neutrons have any appreciable probability to travel any great distance and are less impacted than the other types by the presence of shielding. Second, if the nodes in the network are limited to human-portable devices, it should be expected that the detectors will be small with limited efficiencies. We also must assume that the distance to the source is unknown as is the potential presence and type of intervening shielding materials. Lastly, the available integration time it typically short – on the order of seconds to one minute. To integrate for longer than that is not feasible since it would impact the movement of people, vehicles, and goods greatly.

One key problem associated with detecting any source is the presence of natural background. A variety naturally-occuring radioactive materials (NORM, such as ^{40}K, ^{238}U, and $^{232/228}$Th) are present everywhere in varying quantities. The presence of NORM isotopes varies with position on the planet and also with time. It has been well-documented that fluctuations in weather can increase or decrease the local radiation background levels. [9], [10]

Therefore, it is necessary to establish a formalism for the measurement of background. If we assume the measurement of background radiation occurs at position (x, y) and time t, then

$$\mu(x, y, t) = \mu_1(x, y) + \mu_2(t) \tag{1}$$

where $\mu(x, y, t)$ is the Poissonian parameter representing the mean value of background radiation, $\mu_1(x, y)$ represents the fluctuation in background as a function of position, and $\mu_2(t)$ contains the time-varying parameter associated with the weather. Therefore, μ represents the mean value and, according to Poissonian statistics, is equal to the variance and must be greater than zero. Further, there are no requirements that μ be large. Hence the simplification of the Poissonian distribution to the Gaussian distribution cannot always be assumed, although frequently is to provide numeric simplification. Additionally, μ_1 and μ_2 are independent variables.

With this understanding of the nature of radiation background we next assume a measurement to be taken whose value is given by $d(x, y, t)$. It is then clear that the probability this measurement comes strictly from the background radiation distribution is given by

$$P[d(x, y, t)] = \frac{\mu^{d(x,y,t)} e^{-\mu}}{d(x, y, t)!}. \tag{2}$$

In this application, it is necessary to extract the above to set of D measurements where $d(x, y, t) \in D$. Then the probability that all the measurements come from background

is derived from Equation 2 as follows:

$$P(D) = \prod_{d(x,y,t) \in D} P[d(x, y, t)] \tag{3}$$

$$= \prod_{x \in X} \prod_{y \in Y} \prod_{t \in T} P[d(x, y, t)] \tag{4}$$

$$= \prod_{x \in X} \prod_{y \in Y} \prod_{t \in T} \frac{\mu^{d(x,y,t)} e^{-\mu}}{d(x, y, t)!}. \tag{5}$$

If the log-likelihood of dataset D is given by

$$l(D) = log(P(D)), \tag{6}$$

then it can be shown that

$$l(D) = \sum_{x \in X} \sum_{y \in Y} \sum_{t \in T} \{d(x, y, t) \log(\mu_1(x, y) + \mu_2(t))$$
$$- \mu_1(x, y) - \mu_2(t) = log(d(x, y, t)!).\} \tag{7}$$

Thus it is evident that it is possible to seek a maximum to the expression of 7, called the Maximum Likelihood Estimation. [11]

III. RADIATION SENSOR FUSION

The above formalism holds for a single detector making a series of measurements at different positions and times. However, when a network of sensors is used, the problem evolves into a data fusion problem. Even if each sensor in the network is completely identical (which is usually not a valid assumption), there is still a problem when it comes to combining the measurements of each sensor as they move through space. While the coordinate system of such a problem can be established through GPS, those coordinates themselves have error that can be up to several meters.

This sensor fusion problem like most common fusion problems can be treated in three distinct steps:

1) Align the data spatially in a common coordinate reference frame
2) Align the data temporally in a common time system
3) Align the sensor measurement values into a common unit (frequently normalized).

In this section, we shall focus our attention on the first and third items of the above list, assuming that the temporal alignment is without error.

A. Geospatial Alignment

One key goal of using a geotagged radiation sensor network is to establish the easiest way to align the coordinates of each node in the network, whose individual measurement of position has a potentially large error, into a continuous map. A common method of doing so is Kriging. [12] Suppose there are N detectors in the network such that $d_i(x, y, t)$, $i \in \{1, 2, ..., N\}$ is measured. Kriging takes the equation:

$$d(x, y, t) = \mu(x, y, t) + \sum_{i=1}^{N} w_i(d_i - \mu(x, y, t)) \tag{8}$$

and attempts to solve for the weights, w_i to provide the best unbiased solution. Typically, a variety of simplifications can

be explored for $\mu(x, y, t)$ in Equation 8, such as assuming it to be a known, constant value (simple Kriging) or even just an unknown, constant value (ordinary Kriging). In the case of radiation detection, it is not appropriate to assume that the mean is constant. As such, universal Kriging likely presents the most optimal solution to this problem.

If we let $\vec{d}$ represent all measurement locations where $\vec{d} \in \{d_1, ..., d_N\}$ then we can treat any measured value $y(\vec{d})$ as:

$$y(\vec{d}) = \mu(\vec{d}) + Z(\vec{d}) \tag{9}$$

where $\mu(\vec{d})$ is the mean measured value at $\vec{d}$ and $Z(\vec{d})$ represents a stationary process with constant mean. [13] It should be noted that the first time is deterministic while the second term is probibalistic. In fact, Z can be thought to represent the Poissonian fluctuation in our measurement that has zero mean while $\mu(\vec{d})$ is the mean in radiation background as a function of position neglecting any time variation.

In universal Kriging, it can be shown that the expectation value of any measurement can be treated as the linear combination of a known function $f_l(\vec{d})$ with unknown coefficients, a_l. In this way, the expectation value of any measurement $y(\vec{d})$ can be found as:

$$E[y(\vec{d})] = \sum_{l=1}^{k} a_l f_l(\vec{d}). \tag{10}$$

Then for two different points d_1 and d_2 the expectation value can be expanded to reveal:

$$E[(y(d_1) - \mu(d_1))(y(d_2) - \mu(d_2))] \equiv E[Z(d_1)Z(d_2)] \tag{11}$$

which is simply the covariance between points d_1 and d_2. However, if truly the expectation value between two measurements is reflected by the product of the expectation values of two probibalistic functions with zero mean, then this covariance should be equal to zero. The validity of this statement will be evaluated in this paper.

B. Sensor Measurement Alignment

Every radiation detector, even of like models, has slight variability in its total absolute detection efficiency. These variations can be due to slight variations in detector sizes and physical properties achieved during the manufacturing process. These differences in efficiency impact the overall measured value for each sensor and are a function of energy. However, in this work we have focused on the overall count rate measured by each detector and leave the actual measured energy spectra for future work. Therefore, we will treat these differences as a constant value in energy from detector to detector.

It would be ideal to determine the detector-to-detector variation in efficiency a priori using a calibrated measure, but in practice for a large sensor network this is not practical. Therefore, a method must be derived to evaluate the measured detector response from the collected data itself.
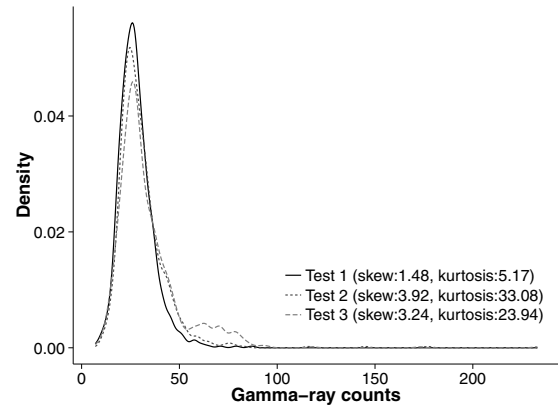

Fig. 1. Measured count rate distributions for three different detectors collected at the same time and location.

One possible way to do this is through the alignment of the measured mean in background. However, in order to do such an alignment, each detector must at the same location at the same time so that fluctations in position and weather are common among them. Depending on the sensor network density, this may not be practical either. But as the density of the network increases, the likelihood of two detectors being located within a similar region and a similar time also increases. These spatial and temporal overlaps should be sought among the data.

More important is the need to treat this alignment with robust statistics. It is possible and tempting to attempt to set a mean based on the assumption of Gaussian statistics. However, in an unknown data set, this is ill-advised due to the possible presence of outliers in the measured value. For example, data is presented in Figure 1 illustrating the measured background on a given day by three different detectors. As can be seen from the figure, the skew and kurtosis illustrate distributions that are far from Gaussian. To avoid potential issus created by the impact of outlier measurements, when the skew and kurtosis are high the mean should be calculated using a Winsorized and/or trimmed approach. [14], [15], [16]

IV. EXPERIMENT AND RESULTS

A series of 23 Kromek D3s detectors [17] were deployed in an urban environment for a period spanning approximately 5 months. These detectors were operated nearly continuously during that time and were hand-carried by a team of volunteers everywhere they went during this time. This included measurements of normal background, enhanced background locations, and measurements in the presence of true radioactive sources. The detectors collected gamma-ray count rates and spectra as well as neutron count rates once per second. Utilizing a bluetooh connection to a smartphone and the phone's on-board GPS, the location and measurement was transmitted wirelessly to a server. The resulting data set comprises over 37 million measurements and was archived in an S3 bucket on Amazon Web Services (AWS) and is nearly 40 GB in size.
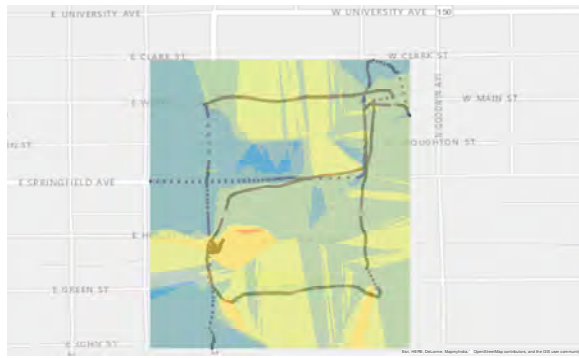
Fig. 2.   Detector 1, day 1

Fig. 3.   Detector 1, day 2

While this data set is modest in size, it exceeds the memory capabilities of the current average PC and will grow as the network grows and more data is collected. To address the anticipated growth in network size (a 1000-node network would be expected to generate approximately 200 GB of data per day) and need for real-time analyis of such data, algorithms were written using Apache Spark 1.6 and SparkSQL and deployed on a small Elastic Map-Reduce (EMR) cluster on AWS.

Prior to any analysis and, in particular, studies on spatial alignment, it was necessary to apply a Kalman filter to the geocoordinate data. [12] It was observed that the GPS data had many positional discontinuities ranging up to nearly 150 m in 1 second. Additionally, occasionally the GPS coordinates failed to be recorded. Thus a Kalman filter was applied to smooth out these discontinuities prior to any analysis performed.

It was first necessary in the spatial alignment to evaluate the assumption that the covariance in position of the measured background was zero. Using the aforementioned data set, it was necessary to select data in a geographic region known to have no enhanced background where more than one detector measured the same region at reasonably the same time. It was assumed that the measurements associated within a single day would be sufficient so long as there was no precipitation on that day.

Appropriate data was selected from within the data set and the covariance in measured gamma-ray counts as a function of position was calculated. It was found that the covariance was very small, ranging from 10^{-4} to 10^{-5}, which supports the assertion surrounding Equation 11.

Once this was assertained, ArcGIS was used on selected data from the above for universal Kriging. Data was selected by multiple detectors being located in th same geographic bounding box within the same day without a source beyond background being present. Examples of the results of these calculations are presented in Figures 2 – 5 for two different days and two different detectors. As can be seen, there are visibly noticable differences between the Kriged data of identical background measurements. This is likely due to the detector-to-detector efficiency variations since there was no normalization applied to this data.
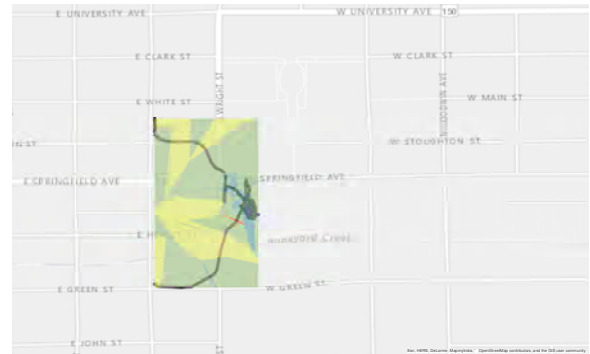

Fig. 4.   Detector 2, day 1

Fig. 5.   Detector 2, day 2

As an example of the impact of sensor measurement alignment, the difference in measured data within the sample shown in Figures 2 – 5 is illustrated through their measured mean count rate and standard deviation in Table I. Given a difference in average count rate of nearly 40%, it is clear that some method of aligning the sensor measurements is required.

The data were rescaled by their differences in mean background count rate for the day and then passed through the universal Kriging algorithm again. Despite the multiplicative factor difference in the count rate data, the overall structure of the resulting map was observed to be the same, as shown in Figure 6.

398

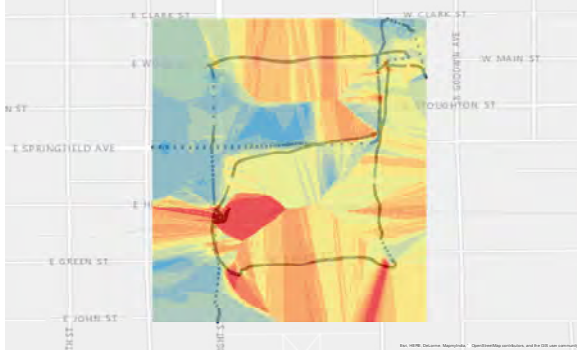|	μ	σ	No. of Points
Detector 1	26.394	5.382	14209
Detector 2	36.505	7.685	1180

Fig. 6. Kriged map showing data of Figure 2 scaled by a constant. The shading is in arbitrary units.

V. CONCLUSIONS AND FUTURE WORK

The first results to fuse the measurements of a radiation sensor network have been presented. Geospatial alignment was demonstrated using univeral Kriging. However, the results of the fusion were limited due to problems in the alignment of the sensor measurements. The model that was used was a renormalization of the count rate data based on a direct scaling of the mean background measurements. It is clear from the results presented that a more sophisticated approach needs to be developed. Ideally such an approach would work utilizing the data collected by the detector in the field rather than a detailed measurement of the efficiency curves of the thousands of detectors expected to be deployed in the next few years. A series of methods that could potentially solve this problem and will be evaluated in the future include unsupervised or semi-supervised machine learning techniques to create an adaptive sensor normalization. A variety of machine learning methods exist that could classify a sensor's measurement as background versus background with a source, including stochastic gradient descent classification, support vector machines, and k-nearest neighbor analysis.

Additionally, as more data is collected, the overall understanding of the background fluctuations in an environment both with position and with time (as the result of weather fluctuations) can be quantified. Once an appropriate sensor normalization model has been identified, it is possible that the entire dataset could be Kriged to generate a normalized map of background for the network. This would have the benefit of opening the network more easily to Bayesian classification based on the use of a well-characterized prior distribution.

REFERENCES

[1] S. Brennan, A. Mielke, D. Torney, and A. Maccabe, "Radiation detection with distributed sensor networks," *Computer*, vol. 37, no. 8, pp. 57–59, 2004.

[2] R. Nemzek, J. Dreicor, D. Torney, and T. Warnock, "Distributed sensor networks for detection of mobile radioactive sources," *IEEE Transactions on Nuclear Science*, vol. 51, no. 4, pp. 1693–1700, 2004.

[3] D. Stephens and A. Peurrung, "Detection of moving radioactive sources using sensor networks," *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2273–2278, 2004.

[4] A. Mielke, S. Brennan, M. Smith, D. Torney, A. Maccabe, and J. Karlin, "Independent sensor networks," *IEEE Instrumentation and Measurement Magazine*, vol. 8, no. 2, pp. 33–37, 2005.

[5] J.-C. Chin, N. Rao, D. Yau, M. Shankar, Y. Yang, J. Hou, S. Srivathsan, and S. Iyengar, "Identification of low-level point radioactive sources using a sensor network," *ACM Transactions on Sensor Networks*, vol. 7, no. 3, pp. 1–35, 2010.

[6] A. Liu, J. Bunn, and K. Chandy, "An analysis of data fusion for radiation detection and localization," *Information Fusion (FUSION) Conference Record*, pp. 1–8, 2010.

[7] A. H. Liu, J. J. Bunn, and K. M. Chandy, "Sensor networks for the detection and tracking of radiation and other threats in cities," in *10th International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2011, pp. 1–12.

[8] D. S. Hochbaum and B. Fishbain, "Nuclear threat detection with mobile distributed sensor networks," *Annals of Operations Research*, vol. 187, no. 1, pp. 45–63, 2011.

[9] J.-F. Mercier, B. Tracy, R. d'Amours, F. Chagnon, I. Hoffman, E. Korpach, S. Johnson, and R. Ungar, "Increased environmental gamma-ray dose rate during precipitation: A strong correlation with contributing air mass," *Journal of Environmental Radioactivity*, vol. 100, no. 7, pp. 527–533, 2009.

[10] R. Livesay, C. Blessinger, T. Guzzardo, and P. Hausladen, "Rain-induced increase in background radiation detected by radiation portal monitors," *Journal of Environmental Radioactivity*, vol. 137, pp. 137–141, 2014.

[11] Z. Liu, "Mobile radiation sensor networks for source detection in a fluctuating background using geo-tagged count rate data," Master's thesis, The University of Illinois, 2016.

[12] D. F. Andrews, P. Bickel, F. R. Hampel, P. Huber, W. Rogers, and J. Tukey, *Robust Estimates of Location: Survey and Advances*. Princeton University Press, 2015.

[13] H. Mitchell, *Multi-Sensor Data Fusion, An Introduction*. Heidelberg, Germany: Springer-Verlag, 2007.

[14] K. Carling, "Resistant outlier rules and the non-gaussian case," *Computational Statistics & Data Analysis*, vol. 33, no. 3, pp. 249–258, 2000.

[15] R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press, 2012.

[16] ——, "Pairwise comparisons of trimmed means for two or more groups," *Psychometrika*, vol. 66, no. 3, pp. 343–356, 2001.

[17] "Kromek D3S," http://kromek.com, accessed: 2016-07-29.