SAND2014-0031C

Sensor Fusion for Intrusion Detection Under False Alarm Constraints

Matthew Pugh, Jacques Kvam and Jerry Brewer
Sandia National Laboratories
Livermore, CA 94550, USA
Email: {mopugh,jkvam,jbrewe}@sandia.gov

***Abstract*—Sensor fusion algorithms allow the combination of many heterogeneous data types to make sophisticated decisions. In many situations, these algorithms give increased performance such as better detectability and/or reduced false alarm rates. To achieve these benefits, typically some system or signal model is given. This work focuses on the situation where the event signal is unknown and a false alarm criterion must be met. Specifically, the case where data from multiple passive infrared (PIR) sensors are processed to detect intrusion into a room while satisfying a false alarm constraint is analyzed. The central challenge is the space of intrusion signals is unknown and we want to quantify analytically the probability of false alarm. It is shown that this quantification is possible by estimating the background noise statistics and computing the Mahalanobis distance in the frequency domain. Using the Mahalanobis distance as the decision metric, a threshold is computed to satisfy the false alarm constraint.**

I. Introduction

A. The Problem

Sensor fusion algorithms have been used in a myriad of applications to improve performance compared to single sensor systems by increasing detectability and resolution, decreasing false alarm rates, etc. See for example [1], [2]. The ultimate goal of sensor fusion algorithms is to make more informed and reliable decisions. Along these lines, the objective of this work is to develop a framework to characterize the false alarm rate of a multi-sensor intrusion detection system when the event signals are a priori unknown. False alarm constraints can be analytically or computationally characterized in many signal processing applications where the signal and noise models are known. For example, in the canonical detection theory problem of finding a sinusoid of known frequency, phase, and magnitude in an additive white Gaussian noise (AWGN) environment, a threshold placed after a matched filter traces out a theoretically computable receiver operating characteristic (ROC) curve (see [3]). In this example, the threshold is selected to meet the false alarm constraint (with a few extra steps to account for the constraint having a temporal aspect, i.e. false alarms per unit time).

SAND Number xxxx-xxxx. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

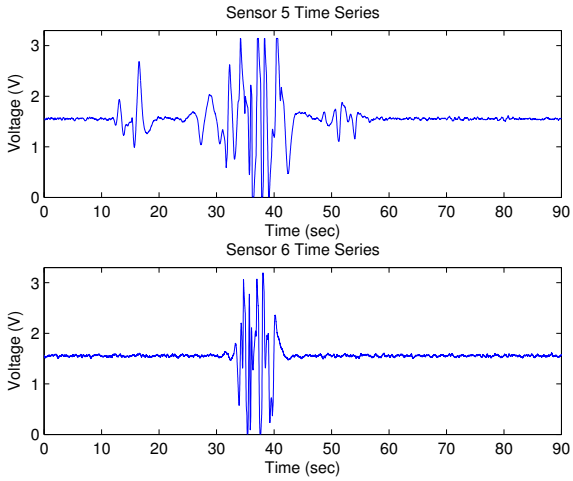

Fig. 1. Intruder Event as Observed By Two Different Sensors.

The signal model is unknown in the situations analyzed in this work. Figure 1 shows an example of data collected on a testbed from two PIR sensors as someone walked into the room being monitored (data collection will be covered in Section II). There are several things to notice from these time series. First, the response of the sensors can take on many forms. Looking specifically at sensor 5, the signal from 10 to 20 seconds looks very different from the signal between 25 and 30 seconds, and 30 to 42 seconds. This is an example of what is meant by not knowing the signal model. Secondly, comparing the signals from sensor 5 and 6, sensor 6 only observers the intrusion event from 30 to 42 seconds. One could use a threshold and logic architecture to detect the events, e.g. whenever any sensor's voltage deviates from the DC offset by more than 0.5 volts, declare an intrusion event. This could theoretically solve the problem if the background noise statistics are known, but this leads to the third observation, to be discussed later, that the background noise is non-independent and non-Gaussian. Thus using a voltage threshold and combinatorial logic, while possibly successful for detecting intrusions, cannot provide the required false alarm guarantees.

The central problem addressed by this contribution is to develop a sensor fusion algorithm to combine and process

signals from multiple PIR sensors to detect intrusion into a room under the constraint that the probability of false alarm not exceed some design parameter. This work only considers a "normal" operating environment, i.e. false alarms caused by abnormal environments such as rapid heating or cooling due to HVAC operations are not considered. The probability of false alarm constraint will be denoted α. An example of such a constraint could be the probability of a single false alarm in a single year is 0.01, or stated another way the algorithm does not have a single false alarm in a year with probability 0.99. In systems where false alarms are prohibitively expensive, a premium is spent on being able to characterize and meet a false alarm constraint at the possible expense of detectability.

In this work a sensor fusion algorithm is developed that can quickly detect intrusions yet satisfies the specified false alarm constraints. To our knowledge, very little work has been done to create detection algorithms merging multiple sensors where the signal model is unknown while still meeting a false alarm requirement. One possibility for the lack of research in this area is that a premium is typically placed on characterizing the detectability of an algorithm. False alarms, when they occur, are viewed as a nuisance rather than a possibly catastrophic system failure. The approach taken in this paper is to place a premium on characterizing the false alarm rate of an algorithm which results in a loss detectability.

A brief description of our approach to this problem is now given. Because the signal model is unknown, the ability to quantify and meet a false alarm constraint rests in being able to characterize the background noise. Along these lines, days worth of background data were collected. Unfortunately, this data reveals that the background noise behaves poorly as it is non-Gaussian and non-independent. To combat this the data is transformed into the frequency domain, where it is observed that nearly all the real and imaginary frequency components are marginally Gaussian distributed, but the joint distribution over all the frequency coefficients is not jointly Gaussian. By selecting a small enough subset of the frequency components, the joint distribution looks nearly Gaussian. The subset of frequency components is selected using principal component analysis (PCA). Once the subset of coefficients is selected, taking the Mahalanobis distance of the PCA coefficients yields a chi-squared background noise distribution. The Mahalanobis distance has a history of use in outlier detection [4], [5], [6] and classification [7]. The combination of PCA followed by a Mahalanobis distance has also been used to find anomalous flights in [8] and detect network intrusions in [9]. Having established a metric with a known distribution on the background data set, a threshold can be computed on the Mahalanobis distance that achieves the false alarm constraint. It is then shown that applying this threshold to intrusion data yields a decision algorithm that has very good performance.

II. Data Collection

The testbed used to collect data consisted of a shielded room with eight custom designed sensor modules mounted along the walls. The sensor module is shown in the top photograph of Figure 2. Each module contains a Kionix KXRB5-2050 tri-axis accelerometer, a Marktech 5052TD photo diode and a Panasonic AMN24112 PIR sensor. In this work, we focus only on the PIR sensor. The eight sensor modules are mounted around the shielded room, as illustrated in the bottom photograph of Figure 2. The sensor modules are connected via a CAN bus. The PIR signal at each sensor module has a possible range of 0 to 3.3 volts and is uniformly scalar quantized with a 12 bit quantizer at a rate of 100 Hz. The quantization level was selected to be 12 bits because experiments with less resolution showed that the noise random variable looked too discrete and could not be modeled with high confidence using continuous random variables. Because the quantization level was set to 12 bits, this limited the sampling rate of the system to be 100 Hz given the bandwidth of the hardware used in the experiment.

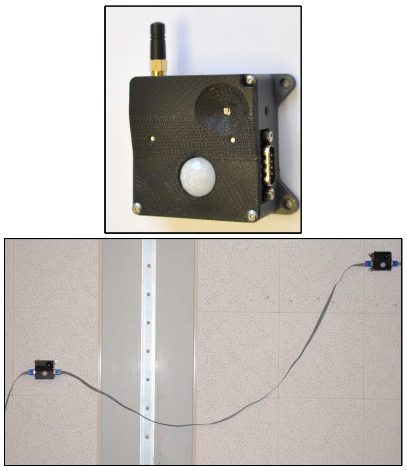

Fig. 2. Top: Sensor Module with PIR sensor, photodetector and three-axis accelerometer. Bottom: Sensor Modules placed along wall of room via CAN bus.

Days worth of data were collected with the shielded room empty and closed to gather as much background data as possible. Simple intrusion event data were collected and consisted of a person opening the door to the shielded room, walking about the room, exiting and closing the door. The PIR signal of such an intrusion event is illustrated in Figure 1. Because the objective is quantifying the false alarm constraint, data gathering focused on the normal operating environment, whereas data gathering would focus more on collecting event data when characterizing the detectability is the primary goal.

III. Intrusion Detection Algorithm

A. Time Series Issues

The signal model for both the background noise signal and possible intrusion event signals are unknown. The background noise statistics can be estimated accurately by collecting enough data. If the background noise statistics can be accurately measured and modeled with a known closed form probability distribution, then a threshold (possibly two sided threshold) can be selected such that when the signal exceeds the threshold, an intrusion is detected. Because it is assumed

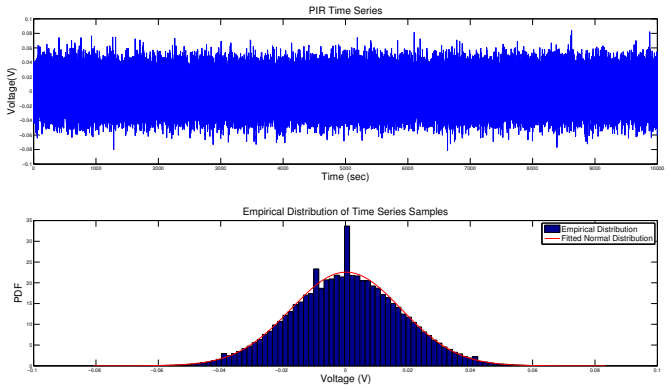

Fig. 3. Top: PIR time series. Bottom: Marginal distribution of PIR sensor data and a fitted normal distribution.

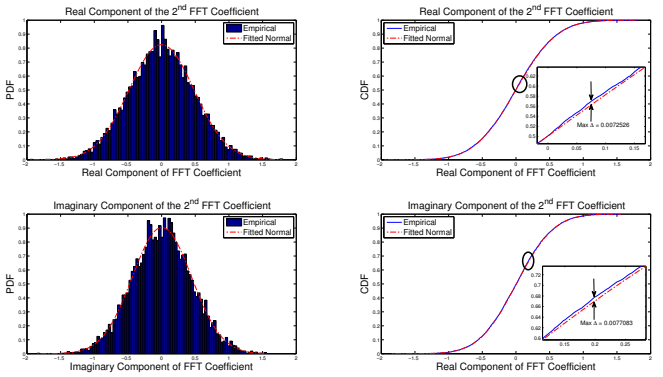

Fig. 4. Top: Empirical density and distribution for the real part of the 2^{nd} FFT coefficient and the fitted normal density. Bottom: Empirical density and distribution for the imaginary part of the 2^{nd} FFT coefficient and the fitted normal density.

the noise statistics are known, the threshold can be designed to meet the false alarm constraint by using the inverse cumulative distribution function. At this point it should be noted that having a closed form probability distribution that matches the noise distribution is essential in computing the threshold. If one wants to use the empirical distribution to compute the threshold, data must be collected on the time scale of the false alarm constraint, i.e. if one can only accept one false alarm per year, a year's worth of data must be collected, otherwise there is no way to extrapolate the results to meet the false alarm constraint. The collection of this volume of data is bypassed by having confidence in a closed form distribution for the noise. More data can provide more confidence in the estimates of the noise statistics.

A natural place to start is to estimate the background noise statistics in the time domain. The top plot of Figure 3 shows background data that was collected with the testbed for a single sensor over the course of 10^4 seconds. The bottom plot of the same figure shows the marginal empirical probability density (in blue) from the data and a normal distribution fitted to the data using maximum likelihood estimates of the mean and variance (in red). The empirical distribution looks fairly Gaussian except for some values which are over-represented, but it is imperative that we have high confidence in the fitting of the noise statistics to a closed form distribution. Using the Lilliefors' goodness-of-fit test ([10]) to test if the data appears to come from an normal distribution, the test rejects the null hypothesis (implying it does *not* comes from a normal distribution) at an α-level of 0.001, i.e. the hypothesis tests is 99.9% certain that the data does not come from a Gaussian distribution. Thus using the raw time series samples does not provide the desired closed form noise distribution.

B. Frequency Domain Approach

Analyzing the sensor data in the time domain did not provide a closed form distribution that matched the noise statistics. In this section the noise statistics will be analyzed in the frequency domain and the results are much more promising. To convert to the frequency domain, the FFT was computed using a window length of 128 with no overlap and no windowing function. A delay is introduced into the algorithm as any decision will have to wait for the FFT to be evaluated on the samples. In this case 128 samples was chosen to minimize the delay yet provide sufficient frequency resolution. An overlap can be introduced to reduce the delay, but this will increase the correlation between frequency coefficients between windows. For this reason, the overlap length is set to zero.

Figure 4 shows the empirical distributions of the real and imaginary components of the 2^{nd} FFT coefficient and a Gaussian fit to the data. Visually, looking at the densities (the plots on the left hand side), it may appear worse than the time series distribution of Figure 3 from the previous section. However, looking at the empirical distribution function compared to the theoretical distribution (the right hand plots), they appear nearly identical. This is important as the Lilliefors' test is based on the supremum norm of the difference between the distribution functions. Indeed, the null hypothesis is *not* rejected even at a 99.9% confidence. This is not to say that we are 99.9% certain of the match, as this would require switching the role of the null and alternative hypotheses in the Lilliefors' test, which does not exist. The α-level of 0.001 is used to be consistent with our methodology for rejecting a distribution; the distribution is assumed to be Gaussian because the null hypothesis is *not* rejected. There is nothing special about the 2^{nd} FFT component and these results appear representative of most of the frequency components.

Since the Gaussian distribution matches the noise statistics in the frequency domain, a threshold can be adopted on the FFT coefficients to meet a false alarm constraint using a combinatorial logic architecture, e.g. design a threshold and declare an event when the first four FFT components exceed the threshold. See Chapter 8 of [1] for examples of this type of architecture. A different approach is proposed here which combines all the FFT information in a rigorous manner. Accepting that the real and imaginary components of the FFT coefficients of the background signal are normally distributed, the mean vector $\mu \in \mathbb{R}^{128}$ and covariance matrix

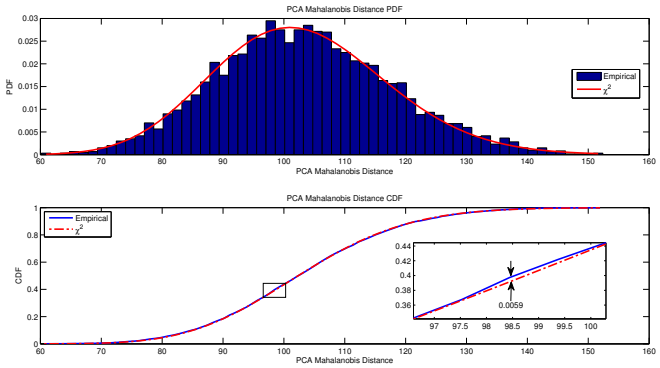

Fig. 5. Distribution of the Mahalanobis distance of the sensor fusion algorithm.

$\Sigma \in \mathbb{R}^{128\times 128}$ are estimated. The PIR signal is real, thus the 128 point FFT yields 65 unique complex coefficients, the first and middle (65^{th}) coefficients being purely real. Taking the real and imaginary components and stacking them in a vector leads to the vector in $\mathbb{R}^{128}$. After estimating the mean and covariance values, the Mahalanobis distance is computed

$$D_M(\hat{x}) = (\hat{x} - \mu)\, \Sigma^{-1}\, (\hat{x} - \mu) \tag{1}$$

where $\hat{x} \in \mathbb{R}^{128}$ are the unique FFT coefficients as previously described. When $\hat{x}$ is jointly Gaussian, the Mahalanobis distance $D_M(\hat{x})$ is known to be chi-square distributed with 128 degrees of freedom. Having a known distribution, a threshold can be chosen on the Mahalanobis distance to meet a false alarm constraint using the inverse CDF. It is beneficial to design the intrusion detection algorithm around the Mahalanobis distance because it has greater sensitivity than a logical architecture and combines all the FFT coefficient information in a consistent and probabilistic manner (see Figure 7 discussed later). To see this, consider any logical architecture where thresholds are placed on individual FFT coefficients. Now suppose that the all the FFT coefficients fall just below the all the thresholds. The logic architecture will not detect an intrusion, but the Mahalanobis distance of such an event will be large as the total distance as computed by $D_M(\hat{x})$ will be large. Here is a summary of a preliminary version of the detection algorithm based on a single sensor:

Algorithm 1: Compute estimates μ and Σ on background data (during a start of phase when conditions are controlled). Once estimated, compute a threshold x_{th} to meet the false alarm constraint α using the fact that $D_M(\hat{x})$ is chi-square distributed. For each new batch of 128 time series samples, compute the FFT $\hat{x}$ and then compute $D_M(\hat{x})$ and declare an intrusion if $D_M(\hat{x}) > x_{th}$.

Unfortunately Algorithm 1 will not work as stated. The Mahalanobis distance is computing a weighted energy in the frequency domain, but by Parseval's theorem a similar computation in the time domain should yield the same result. In the time domain however, there is little hope of achieving a closed form chi-squared distribution due to the non-Gaussianity of the samples.

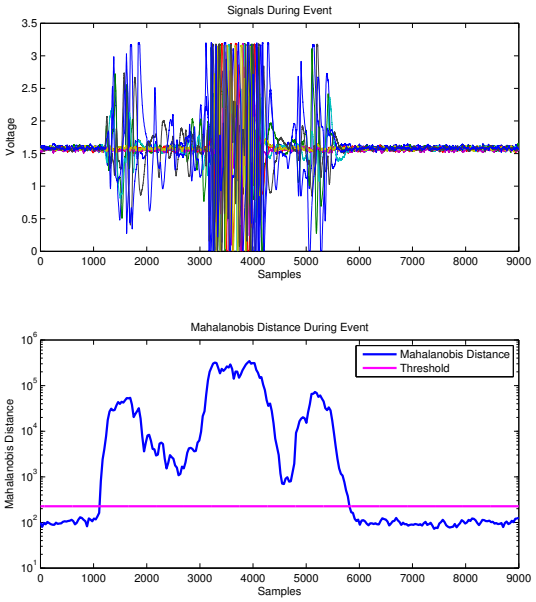

Fig. 6. Top: Event time series for eight testbed sensors. Bottom: The output metric (Mahalanobis distance) of the sensor fusion algorithm and decision threshold.

To address this issue, rather than using all of the FFT coefficients to compute the Mahalanobis distance, a subset of coefficients will be selected as it will be shown that this leads to a distribution that very closely matches a chi-squared distribution. This leads to the question of how to select this subset. Experiments have shown that almost any sufficiently small subset of FFT coefficients will work in terms of yielding a chi-squared distribution, but detectability of intrusion events could be compromised. A reasonable subset selection criterion would be to select the coefficients that have the most energy. Using this idea, principal component analysis (PCA) was used to find the subspace of FFT coefficients that describe the most energy. The principal components that accounted for 95% of the energy were used, which was 9 components for the sensor time series in the Figure 3. PCA can be performed because linear transformations of the Gaussian distributed FFT coefficients preserves Gaussianity. PCA also conveniently reduces the dimension of the feature space from 128 to 9 in this case. Theoretically the Mahalanobis distance computed on the 9 principal components should be distributed as a chi-squared random variable with 9 degrees of freedom. Using a two-sided Kolmogorov-Smirnov test ([11]), the null hypothesis of being a chi-squared distribution is *not* rejected at an α-level of 0.001 providing additional justification. The selection of the PCA dimension such that at least 95% of the variance is captured is a obviously a heuristic and future research will investigate optimizing the subset selection method.

The system can have N sensors, which are assumed to independently observe the environment. In the testbed, there are $N = 8$ sensors. The above procedure is performed on each of the individual sensors, and the PCA components of each sensor are added together at the central controller to yield the

final statistic which is compared to a threshold to decide if an event has occurred. This is possible because the distribution of the sum of the PCA components is known. Here is a summary of the sensor fusion algorithm:

Algorithm 2: On the background data set, compute the PCA subspace of each sensor in the frequency domain that accounts for 95% of the variance. Let β_i denote the dimension of the PCA subspace for sensor i. On the background data, compute $\mu_{i_{PCA}} \in \mathbb{R}^{\beta_i}$ and $\Sigma_{i_{PCA}} \in \mathbb{R}^{\beta i \times \beta_i}$, the PCA mean and covariance, for each sensor. For each batch of 128 time samples, compute $\hat{x}_i$ for each sensor, project it into the PCA domain to produce $\hat{x}_{i_{PCA}}$ and then compute $D_{M_i}(\hat{x}_{i_{PCA}})$ using $\mu_{i_{PCA}}$ and $\Sigma_{i_{PCA}}$ for each sensor. Sum the statistic from each sensor $D_{M_{total}} = \sum_{i=1}^{N} D_{M_i}$ and declare an intrusion if $D_{M_{total}} > x_{th}$, where x_{th} is compute from the the PCA statistics (see Equation 2).

Each D_{M_i} is distributed as a chi-squared random variable with β_i degrees of freedom, so $D_{M_{total}}$ has a chi-squared distribution with $\beta_{total} = \sum_{i=1}^{N} \beta_i$ degrees of freedom. Suppose we want to achieve a false alarm rate of α, e.g. probability of a false alarm in a year is 0.001. Let N be the number of 128 sample windows that occur in the time frame of interest (e.g. one year) and let F be chi-squared distribution function with β_{total} degrees of freedom. The threshold x_{th} can be computed with the following equation:

$$x_{th} = F^{-1}\left(\alpha^{\frac{1}{N}}\right) \tag{2}$$

where F^{-1} is the inverse distribution function. This equation follows from the theory of maximum order statistics ([12]). The largest of N computed Mahalanobis distances should exceed the threshold x_{th} with probability α. By assumption, the Mahalanobis distances are independent (and this essentially seems to be the case), so this probability can be given by the equation

$$F^N(x_{th}) = \alpha. \tag{3}$$

Inverting this equation yields the expression for the threshold x_{th} given in Equation 2.

Figure 6 shows the time series of all eight sensors and the corresponding Mahalanobis distance and threshold x^{th} for $\alpha = 0.001$ for a year. It is seen that the Mahalanobis distance exceeds the threshold during the event, indicating an intrusion. Figure 7 scales the sensor time series from Figure 6 so that the maximum value does not exceed the voltage threshold recommended by the manufacturer to declare an event. Thus, none of the sensors would have declared an event for the time series in Figure 7 and therefore no decision algorithm based on a combinatorial logic architecture would have declared an event. However, using Algorithm 2, an event is declared. This example shows the power in coherently combining all of the sensor time series via probabilistic models.

IV. Conclusion

A sensor fusion algorithm is developed to meets a false alarm requirement. The algorithm accomplishes this by characterizing the background noise statistics in the frequency domain and projecting the frequency domain coefficients into a lower dimension space via PCA. The PCA dimension is selected to capture at least 95% of the variance. The PCA coefficients are Gaussian distributed. The decision metric is the Mahalanobis distance computed on the PCA coefficients. The decision metric has a known distribution on the background noise, a chi-squared distribution, which allows a threshold to be computed to achieve the false alarm requirement. The threshold is derived in a straight forward manner from the theory of maximum order statistics.

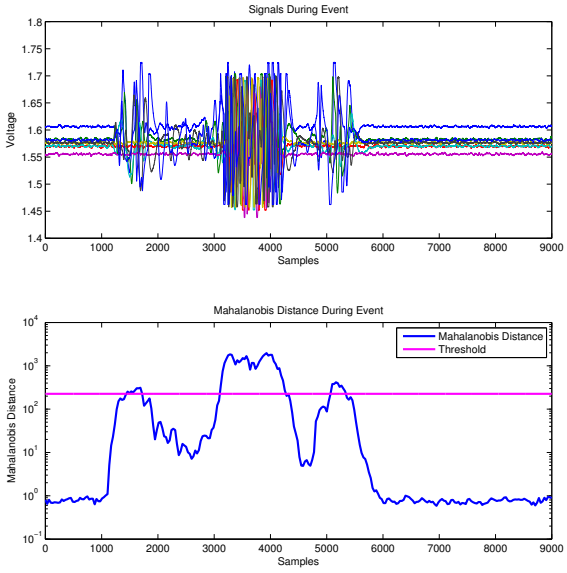

Fig. 7. Top: Scaled event time series. Bottom: Sensor fusion algorithm output metric on scaled time series and decision threshold

References

[1] L. A. Klein, *Sensor and data fusion: a tool for information assessment and decision making*, vol. 324. SPIE press Bellingham, WA, 2004.

[2] J. R. Raol, *Multi-sensor data fusion with MATLAB*. CRC Press, 2010.

[3] S. M. Kay, *Fundamentals of Statistical signal processing, Volume 2: Detection theory*. Prentice Hall PTR, 1998.

[4] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, pp. 85–126, 2004.

[5] H.-P. Kriegel, P. Kröger, and A. Zimek, "Outlier detection techniques," in *Tutorial at the 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2009.

[6] A. Bagherjeiran, E. Cantu-Paz, and C. Kamath, "Design and implementation of an anomaly detector," tech. rep., Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2005.

[7] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[8] B. G. Amidan and T. A. Ferryman, "Atypical event and typical pattern detection within complex systems," in *Aerospace Conference, 2005 IEEE*, pp. 3620–3631, IEEE, 2005.

[9] H. Shah, J. Undercoffer, and A. Joshi, "Fuzzy clustering for intrusion detection," in *Fuzzy Systems, 2003. FUZZ'03. The 12th IEEE International Conference on*, vol. 2, pp. 1274–1278, IEEE, 2003.

[10] H. W. Lilliefors, "On the kolmogorov-smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, vol. 62, no. 318, pp. 399–402, 1967.

[11] F. J. Massey Jr, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[12] H. A. David and H. Nagaraja, *Order Statistics*. Wiley. com, 2004.