FINAL REPORT

Demonstration and Validation of Statistical Analysis Techniques for TOI Discrimination using Advanced EMI Sensor Systems

ESTCP Project MR-201156

DECEMBER 2014

Larry Carin
Levi Kennedy
Todd Jobe
Xianyang Zhu
Signal Innovations Group, Inc.

Distribution Statement A

This document has been cleared for public release

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
26-10-2012

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
09/2011-05/2012

4. TITLE AND SUBTITLE
2011 UXO Discrimination Study Camp Beale, CA: Final Report

5a. CONTRACT NUMBER
MR-201156

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Larry Carin (PI), Levi Kennedy, Todd Jobe, Xianyang Zhu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Signal Innovations Group, Inc.
4721 Emperor Blvd., Suite 330
Durham, NC 27703

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The main technical objective of the Camp Beale demonstration is to validate and substantially automate the SIG learning process using next-generation electromagnetic induction (EMI) sensor data for discriminating targets-of-interest. This process includes three major components, feature extraction, site learning and excavation. **Results by sensor: TEMTADS 2x2, BealeTrees -** All UXO were captured in the dig list with ~160 false alarms. **BUD, BealeTrees -** The two methods for selecting a single sounding at an anomaly, 'symmetric' and 'closest', capture all the UXO (Figure 16). The 'closest' method performed better than the 'symmetric' method in the number of unnecessary digs required to reach the last UXO: 180 vs. 200. **CH2MHILL MetalMapper, BealeOpen -** The final dig list missed two 37mm, and one ISO. An additional 150 digs would have been required to capture all these UXO.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Levi Kennedy
U	U	U	SAR		**19b. TELEPHONE NUMBER** *(include area code)* 919-323-3456

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

List of Figures ii
List of Tables iii
List of Acronyms iii
Keywords iv
Abstract 1
Objective 2
Background 3
SIG Isolate Process Overview 3
Materials and Methods 5
EMI Multi-Dipole Model and Feature Extraction 5
Subspace Denoising and Joint Diagonalization for Robust Feature Extraction 7
Feature Selection with BENet 7
Semi-Supervised Classification 8
Non-Myopic Active Learning 9
Results and Discussion 10
TEMTADS 2x2 – Beale TREES 10
Training Steps 10
Dig Lists 11
Performance 12
Retrospective 12
BUD – Beale Trees 13
Training Steps 13
Performance 16
Retrospective 17
MetalMapper – Beale Open 19
Training Steps 19
Dig Lists 20
Performance 20
Retrospective 21
Conclusions and Implications for Future Research/Implementation 24

Literature Cited 24
Appendix A: Points of Contact.....25

List of Figures

Figure 5. Flow diagram of the SIG Isolate process. 5

Figure 1: Example plot of the polarizability of a UXO as a function of time (right). The left image is the UXO with the polarizability axes shown. Because of the distinctive shape of UXO, at least of two of the polarizability axes will be coincident; in this case M2 (red) and M3 (blue) are coincident. 6

Figure 2: Denoising results from the TEMTADS sensor taken at Camp Butner. The polarizability curves for the denoised data are more consistent with this axial symmetric object. 7

Figure 3.. Receiver operating characteristic (ROC) curves for UXO classifier at SLO site with features selection using the BENet algorithm (red line) and without feature selection (blue line). The number of false alarms is lower for the classifier where feature selection was used. 7

Figure 4: A comparison between supervised and semi-supervised classifiers for a 2. Labeled data from both classes (red and green circles) are shown, along with unlabeled data (black dots). The supervised classifier is trained on only the labeled data and the decision boundary is shown (dotted line). The semi-supervised classifier is trained on both the labeled and unlabeled data and the decision boundary (solid line) makes the two classes linearly separable. 8

Figure 6. Fisher Information gain from training rounds 2-7..... 10

Figure 7. The BENet weights from round 7 with 90 labeled anomalies..... 10

Figure 8. Predicted probabilities for the remaining anomalies at the end of training. The predicted probabilities of the labeled data are shown as shapes. Reference lines for predicted probability 0.5 (black) and 70th quantile of the anomalies (gray) are shown. 11

Figure 9. Predicted probabilities for the remaining anomalies at the end of digging. The predicted probabilities of the labeled data are shown as shapes. 11

Figure 10. Final ROC curve for TEMTADS 2x2 in Beale Trees where Fuzes were considered UXO (left) and as clutter (right). 12

Figure 11. ROC curves for the generative model applied to the TEMTADS 2x2 sensor data from the BealeTrees site. The left graph shows the performance when the data are censored to only those anomalies with low error inversion models. The right graph shows the performance on all data. 13

Figure 12. The Fisher Information gain through all training rounds for the 'symmetric' method (left) and the ‘closest’ method (right). 14

Figure 13. Histogram of the selected template position for the BUD sensor across all anomalies in BealeTrees for the ‘symmetric’ method (left) and the ‘closest’ method (right). Template position 2 is the central template position..... 14

Figure 14. Distribution of posteriors for the 'symmetric' method (left) and the ‘closest’ method (right) at the end of training. 15

Figure 15. Posteriors distributions at the end of digging for the 'symmetric' method (left) and the 'closest' method (right). 16

Figure 16. ROC curves for the ‘symmetric’ method (top) and the ‘closest’ method (bottom). Fuzes are treated as UXO (left) and clutter (right). 17

Figure 17. Feature weights from BENet on all labeled anomalies for all UXO (blue) and fuzes only (red). 17

Figure 18. MAP estimates along two dimensions of a non-linear ordination of the anomaly features for the ‘closest’ method (top) and the ‘symmetric’ method (bottom). The anomalies that were marked as ‘dig’ in the dig list are shown (UXO in white, clutter in black). The missed fuzes for each method are also shown (red circles). 18

Figure 19. Polarizabilities for flag BE-0671 which is an ISO. The 'closest' method selected template position 4 (black box). when all other responses are more UXO-like. 19

Figure 21. BENET weights at the end of training for the MetalMapper sensor in the BealeOpen dataset. The selected feature is shown in red. 20

Figure 21. Scaled Fisher Information gain for 9 rounds of aquiring training labels via active learning for the MetalMapper sensor at the BealeOpen dataset. 20

Figure 22. Histogram of predicted probabilities of being UXO unlabeled MetalMapper data at the end of digging. 21

Figure 23. Histogram of predicted probabilities of being UXO for unlabled MetalMapper data at the end of training. Resubstituted predicted probabilities for the training data are shown as symbols: clutter (blue) and UXO (red). 21

Figure 25. ROC Curve for the MetalMapper retrospective discrimination of the Belae Open dataset. Only 40 training samples were used. 22

Figure 24. ROC Curve for MetalMapper sensor, BealeOpen dataset where fuzes are counted among the UXO (right) and where they are counted as clutter (left). 22

Figure 26. Archetypal responses for the missed UXO type (left column) along with the responses for the UXO that were missed in the dig lists (center and right columns). The responses are significantly different from the standard response and are unlikely to be capture by an automated discrimination method. 23

List of Tables

Table 1. Program Office Performance Objectives for Discrimination Analysis 2

List of Acronyms

BENet	Bayesian Elastic Net
BUD	Berkeley UXO Discriminator

EMI	Electromagnetic Induction
ESTCP	Environmental Security Technology Certification Program
MAP	Maximum *a posteriori*
MTL	Multi-task learning
NMAL	Non-Myopic Active Learning
PNBC	Parameterized Neighborhood-Based Classifier
ROC	Receiver Operating Characteristic
Rx	Receiver
RVM	Relevance Vector Machine
SIG	Signal Innovations Group, Inc.
SLO	San Luis Obispo Demonstration Site
SVD	Singular Value Decomposition
TEMTADS	Time-domain Electromagnetic Multi-sensor Tower Array Detection System
Tx	Transmitter
UXO	Unexploded Ordnance

Keywords

UXO, Bayesian classifier, active learning

Abstract

Objective: The main technical objective of the Camp Beale demonstration is to validate and substantially automate the SIG learning process using next-generation electromagnetic induction (EMI) sensor data for discriminating targets-of-interest. This process includes three major components: feature extraction, site learning and excavation. The end result of the discrimination is a list of classifications for all anomalies at a site. Discrimination performance is increased by maximizing the number of unexploded ordnance (UXO) correctly identified, maximizing the number of non-UXO anomalies correctly identified, specifying an appropriate dig no-dig boundary, and minimizing the number of anomalies that cannot be analyzed.

Technical Approach: SIG performed discriminations for three sensors across two datasets. The first two sensors were the Berkeley UXO Discriminator (BUD) and the Time-domain Electromagnetic Multi-sensor Tower Array Detection System (TEMTADS) 2x2 at the 'Beale Trees' site. The third was the CH2MHILL MetalMapper sensor for the 'Beale Open' site. SIG developed a sensor/target model using a multi-anomaly dipole inversion. Feature selection was performed using the Bayesian Elastic Net which has the benefit of retaining correlated and informative features for classification. Classification was performed using a semi-supervised, parametric Bayesian classifier. An additional retrospective analysis was undertaken for the TEMTADS 2x2 sensor that included a 'generative' approach: where the classifier modeled UXO responses directly rather than discriminating between UXO and clutter.

Results: Summarized by sensor and dataset

TEMTADS 2x2, BealeTrees - All UXO except were captured in the dig list ~160 false alarms. Approximately 20% of the UXO were captured during training

BUD, BealeTrees - The two methods for selecting a single sounding at an anomaly, 'symmetric' and 'closest', captured all the UXO (Figure 16). The 'closest' method performed better than the 'symmetric' method in the number of unnecessary digs required to reach the last UXO: 180 vs. 200..

CH2MHILL MetalMapper, BealeOpen - The final dig list missed two 37mm, and one ISO. An additional 150 digs would have been required to capture all these UXO.

Benefits:

A key improvement to the SIG Isolate Process was highlighted in this discrimination. A generative, rather than a discriminative approach, could improve performance in special, easily discriminated sites.

Objective

The main technical objective of the Camp Beale demonstration is to validate and substantially automate the SIG learning process using next-generation electromagnetic induction (EMI) sensor data for discriminating targets-of-interest. All elements of human interpretation and intuition are being incrementally constrained or removed from the process, resulting in an automated process, where all algorithm parameters and thresholds will either be determined by specified site parameters (i.e., expected or inferred munitions types) or by data-driven inferences (i.e., cross-validated operating threshold). SIG applied and matured each of the three key process phases that constitutes the SIG statistical learning approach to UXO discrimination - called the "SIG Isolate" process. The three phases of Isolate include: Phase I - feature extraction, Phase II – site learning, and Phase III – excavation. Each of the phases is described in detail below. Validation of Isolate entails meeting all of the discrimination performance objectives defined by the program office for each of the sites considered (Table 1).

Table 1. Program Office Performance Objectives for Discrimination Analysis

Performance Objective	Metric	Data Required	Success Criteria
Analysis and Classification Objectives			
Maximize correct classification of targets of interest	Number of targets-of-interest retained.	• Prioritized anomaly lists • Scoring reports from the Institute for Defense Analyses (IDA)	Approach correctly classifies all targets-of-interest
Maximize correct classification of non-target of interest (UXO)	Number of false alarms eliminated.	• Prioritized anomaly lists • Scoring reports from IDA	Reduction of false alarms by > 65% while retaining all targets of interest
Specification of no-dig threshold	Probability of correct classification and number of false alarms at demonstrator operating point.	• Demonstrator - specified threshold • Scoring reports from IDA	Threshold specified by the demonstrator to achieve criteria above
Minimize number of anomalies that cannot be analyzed	Number of anomalies that must be classified as "Unable to Analyze."	• Demonstrator target parameters	Reliable target parameters can be estimated for > 98% of anomalies on each sensor's detection list.

Background

Using next-generation cued sensors, discrimination performance on real sites has shown the feasibility of advanced statistical analyses for distinguishing unexploded ordnance (UXO) from clutter. Signal Innovations Group, Inc. (SIG) has demonstrated the effectiveness of site-specific statistical learning for smartly selecting labeled training data to maximize target discrimination. This technology has been developed and validated under previous SERDP efforts by SIG and Duke University and was ready for application to ESTCP site demonstration at Camp Beale.

Many current analysis approaches rely on expert scientists to make educated decisions at multiple points in the discrimination analysis process. This situation is not scalable, transferable, or cost effective. The SIG approach standardizes the options and creates a documented process flow that can be explicitly followed. Key statistical technologies that were validated and automated during this effort include: physics-based target/sensor models, subspace denoising, automated and efficient feature extraction, data selection for classifier training. These techniques represent the state of the art in digital geophysics.

SIG Isolate Process Overview

The SIG Isolate process laid out in [5] can be summarized in the following 'recipe' (Figure 5):

- **Data Conditioning** - First, raw, unlabeled anomaly data are received. Then, quality control checks are performed. These include ensuring that the background subtractions are complete and determining if the raw sensor inputs need any scaling in order to be appropriate for the feature extraction software.
- **Subspace Denoising** - The anomaly data is denoised to ensure robust performance for discriminating late time-gate features.
- **Feature Extraction** - A robust multi-anomaly dipole model is fitted to the data. The polarizability parameters from this fitting become the set from which features are drawn for classifier training. In addition to the time-domain polarizabilities, a set of 9 'rate' features were calculated. These features were the calculated by fitting the time-domain polarizabilities of each axis to an exponential-decay model:

  $$p_i = r_{1i} + r_{2i} e^{\frac{-t}{r_{3i}}}$$

  where $i \in \{x, y, z\}$ is the current axis, p is the polarizability, t is time and $\{r_1, r_2, r_3\}$ are the fitted rate parameters. Though r_{1i} is unphysical, it is useful for adjusting for noise at late time gates and where odd responses would make the optimization difficult. The optimized values of the rate parameters were found using non-linear least squares.
- **Basis Selection** - A few of the many possible features are selected based on their physical interpretation as they relate to the anomaly, and, using these features, the most informative set of anomalies are selected via an information metric to begin classifier training.
- **Feature Set Augmentation** - The feature set is then augmented by adding early, mid and late time polarizabilities values.

- **Automated Features Selection** - For the now larger feature set, the most relevant set of features is selected using BENet.
- **Semi-supervised PNBC Training (STL or MTL)** - When the PNBC is trained only using data from the current site of interest, it is called Single Task Learning (STL). When the PNBC is trained for multiple sites simultaneously it is called multi-task learning (MTL).
- **Non-myopic Active Learning** - Based on the estimates made with the PNBC classifier, a new set of anomalies will be selected for labeling using NMAL. The goal at this step is to maximize the information gain from new labels requested from the set of unlabeled anomalies. The process is repeated as the PNBC classifier adequately learns data manifold. The stopping criteria for the learning process is apparent when the remaining unlabeled data points have approximately equal information for improving the classifier. At which point, labeling any one anomaly is no better than any other.
- **Excavation Adapted Threshold Selection** - At this point, the highest probability UXO are selected for excavation and labels. The classier continues to be retrained when new labels are revealed. This process continues until the highest probability UXO items excavated are all found to be clutter at which point digging stops.

The process outlined above falls into 3 broad phases: Feature Extraction, Site Learning, and Excavation. Details on each phase are given in the next subsections. The SIG Isolate process is relatively linear save for two feedback steps. The first feedback is in training the semi-supervised classifier, where additional anomaly labels are requested until the classifier reaches sufficient stability. The second feedback is during the excavation of anomalies, where the classifier is retrained with additional labeled anomalies until either the UXO/clutter predictions become highly separable or until high probability anomalies are substantially revealed to be clutter upon excavation.

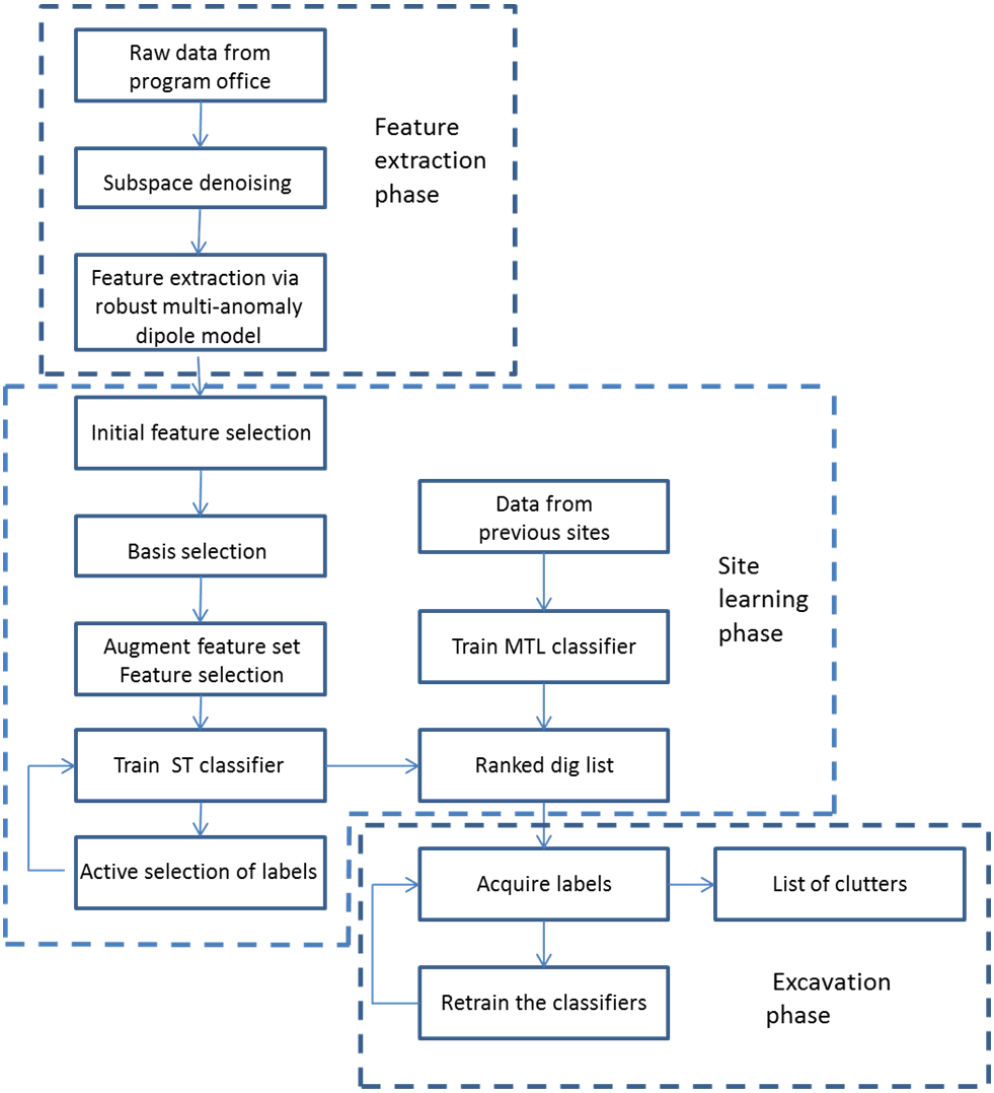

Figure 5. Flow diagram of the SIG Isolate process.

Materials and Methods

SIG applied the Isolate discrimination process at Camp Beale for three sensors across two sites. The first were the Berkeley UXO Discriminator (BUD) and the Time-domain Electromagnetic Multi-sensor Tower Array Detection System (TEMTADS) 2x2 sensors in the 'Beale Trees' site. The third was the CH2MHILL MetalMapper sensor for the 'Beale Open' site. The process for each sensor involved the following key technologies, including: parametric target/sensor modeling, robust feature extraction and selection, semi-supervised classifier training using active selection of labeled data and multi-task learning incorporating past demonstration site data. Each of these technologies is described briefly in the following subsections.

EMI Multi-Dipole Model and Feature Extraction

SIG extracted features by fitting raw sensor data to a physics-based parametric model [1], [2] that was developed under SERDP support and has been successfully demonstrated and validated in Camp Sibert, Camp San Luis Obispo (SLO), and Camp Butner analyses. It has been shown [3],

[4] that the induction response of simple targets can be efficiently represented in terms of one or more time/frequency-dependent magnetic dipoles. In particular, the magnetic dipole moment m of a target is represented as $m = M \cdot H^{inc}$, where H^{inc} represents the incident (excitation) magnetic field and M is a tensor that relates H^{inc} to m. Using reciprocity in wave propagation, the total magnetic field observed at the receiver coil, H^{rec}, can represented as

$$H^{rec}(\omega) \propto r_{st} \cdot U^t \cdot M \cdot U \cdot r_{st}$$

where r_{st} is a unit vector directed from the source to the target and the 3×3 unitary matrix U (that contains information about the target orientation) rotates the fields from the coordinate system of the sensor to the coordinate system of the target. To simplify the above expression, an assumption is made that the source responsible for the incident field H^{inc} can be characterized by a dipole. This assumption is valid if the dimension of the transmitter coil is much smaller than its distance from any buried anomaly. While this may be appropriate for many sensors, it is not appropriate for the multi-coil multi-axis sensors (TEMTADS, MetalMapper, or BUD). SIG has developed a model to synthesize H^{inc} for these systems; wherein the physical sizes/shapes of the transmitter and receiver coils are accounted for explicitly (this is done with a rigorous Biot-Savart analysis). This model has already been validated on both TEMTADS (supported by SERDP funding for Camp SLO analysis) and MetalMapper data (under SERDP contract MM-1708).

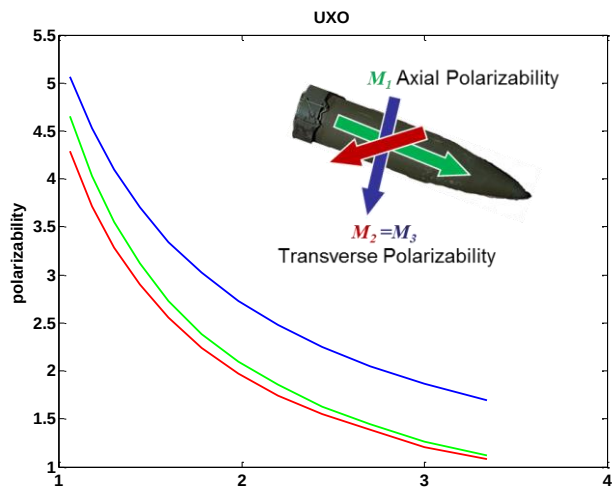

Figure 1: Example plot of the polarizability of a UXO as a function of time (right). The left image is the UXO with the polarizability axes shown. Because of the distinctive shape of UXO, at least of two of the polarizability axes will be coincident; in this case M2 (red) and M3 (blue) are coincident.

Based on the generalized forward model described above, the three principal polarization terms in the magnetic polarization dyadic, the orientation angles in the rotation matrix, as well as the position of the object can be extracted directly from the field data by using a nonlinear least-square solver (Figure 1). It is well known that the trigonometric functions in the rotation matrix are nonlinear multi-valued functions that result in many local solutions. However, the tensor is a symmetric matrix, having only six independent elements (parameters), and the general magnetic polarization dyadic is a linear function of the measured field data. This fact is exploited by extracting the six parameters directly from the measured data (using least-square inversion), rather than extracting the three principal polarization components and three rotation angles directly. The number of non-linear parameters in the model is greatly reduced in this implementation, and the problem of local solutions is significantly relieved, resulting in reliable convergence of the feature inversion. This procedure is also capable of performing simultaneous

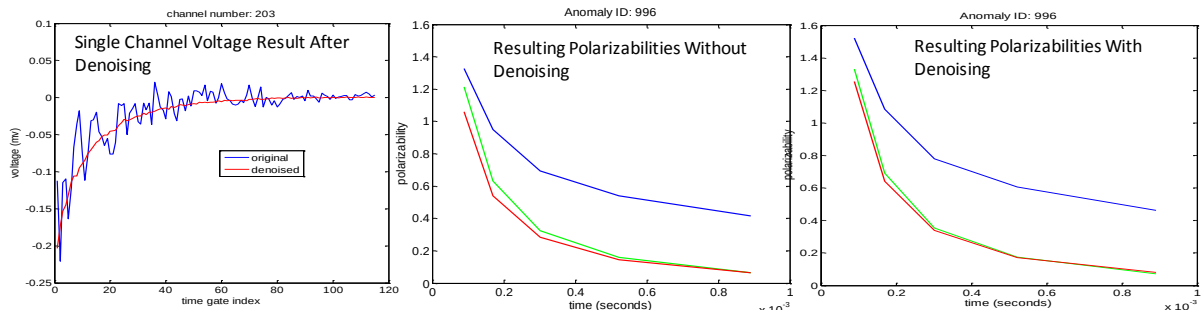

Figure 2: Denoising results from the TEMTADS sensor taken at Camp Butner. The polarizability curves for the denoised data are more consistent with this axial symmetric object.

feature inversion for multiple co-located (or nearby) anomalies, which was validated during the SIG SLO demonstration.

Subspace Denoising and Joint Diagonalization for Robust Feature Extraction

SIG has developed a denoising technique for the TEMTADS data that exploits the co-located arrangement of transmitters and receivers (Figure 2). The technique employs the Singular Value Decomposition (SVD) to exploit the redundancy in the Tx/Rx pairs of multi-coil sensors to identify the principal contributors of the signal due to the buried anomaly and suppresses the contributions due to noise. This technique has been successfully implemented and validated on the Butner demonstration for the TEMTADS sensor. A similar subspace denoising technique has been developed and applied for the Metalmapper sensor using the Matrix-Pencil decomposition which is a generalization of SVD. The benefit of this technique is particularly important for low SNR anomalies (i.e. deep or small objects) where much of the late time information is affected by noise but is important for accurate discrimination.

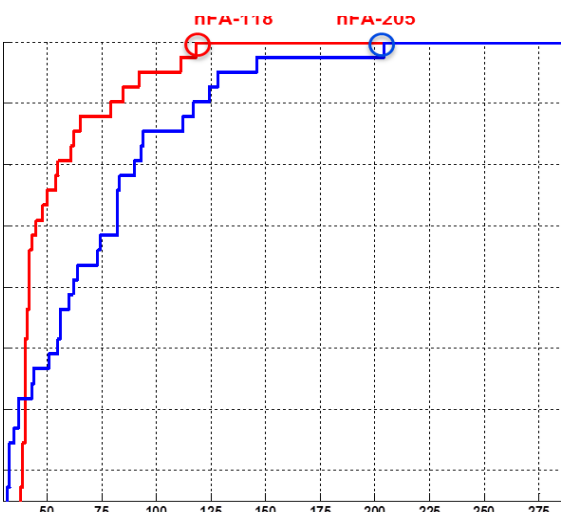

Figure 3.. Receiver operating characteristic (ROC) curves for UXO classifier at SLO site with features selection using the BENet algorithm (red line) and without feature selection (blue line). The number of false alarms is lower for the classifier where feature selection was used.

Feature Selection with BENet

Adaptive learning of a classifier *in situ* benefits from refining the appropriate set of extracted features for the targets under test. This occurs because of the 'curse of dimensionality' where the number of data points required to cover the breadth of a features space grows exponentially with the number of features considered. If the amount of training data does not sufficiently sample the feature space, then the learned classifier will lack statistical support and class estimate uncertainty is large. At SLO in particular, feature selection played a key role in classifier performance (Figure 3). Bayesian classification models perform feature selection by placing a sparseness prior on the inferred feature weights. The Bayesian elastic net (BENet) regression model used for feature selection employs a sparseness prior equivalent to a convex combination of L1-norm and L2-norm penalties in a least squares optimization

formulation [5], [2]. The sparseness prior of the BENet model jointly infers the essential subset of relevant features, including correlated features, for a given classification task. Rather than encouraging the selection of a single feature in a set of correlated important features (like similar approaches such as RVM), the BENet model encourages the selection of all correlated important features. By performing sparse and grouped feature selection, the BENet algorithm provides a more robust approach to feature adaptability and the interpretation of important features, ultimately requiring fewer training data samples to achieve robust statistical support.

Semi-Supervised Classification

Semi-supervised learning is applicable to any sensing problem for which all of the labeled and unlabeled data are available at the same time, and therefore, particularly for the current demonstration study. In most practical applications (including the recent demonstration at Camp Butner), semi-supervised learning has been found to yield superior performance relative to widely applied supervised algorithms. Figure 4 depicts the advantage of a semi-supervised approach to classification over its supervised counterpart. A classifier trained purely on labeled data (depicted as red and green circles) is shown as a purple dashed line and generates classification errors. In contrast, a semi-supervised classifier trained on both labeled and unlabeled data will generate perfect classification (depicted by the blue line). Note that the context provided by the unlabeled data was crucial in improving the classification performance in this case, since the labeled data were not representative of the two class distributions. As the number of training samples increases, the supervised classifier should approximate the semi-supervised classifier. Semi-supervised formulation treats the dataset (labeled and unlabeled) as a set of connected nodes, where the affinity w_{ij} between any two feature vectors (nodes) $\boldsymbol{f_i}$ and $\boldsymbol{f_j}$ is defined in terms of a radial basis function [6]. Based on the above formulation, one can design a Markov transition matrix $\boldsymbol{A} = [a_{ij}]_{N\times N}$ that represents the probability of transitioning from node $\boldsymbol{f_i}$ to $\boldsymbol{f_j}$. Assuming $\mathcal{L} \subseteq \{1,2,\dots,N_L\}$ represents the set of labeled data indices, the likelihood functional can be written as

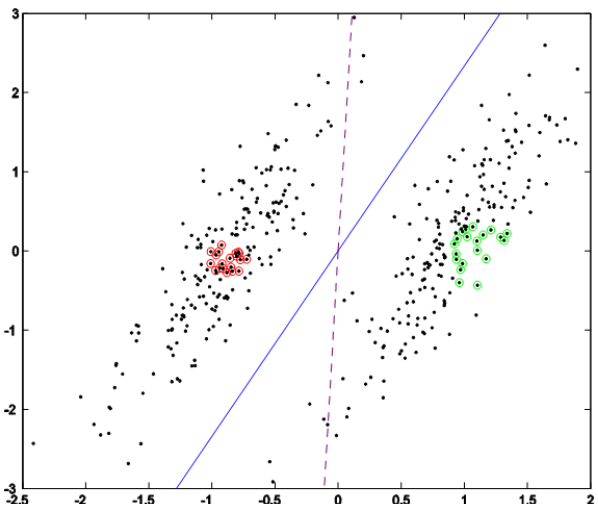

Figure 4: A comparison between supervised and semi-supervised classifiers for a 2. Labeled data from both classes (red and green circles) are shown, along with unlabeled data (black dots). The supervised classifier is trained on only the labeled data and the decision boundary is shown (dotted line). The semi-supervised classifier is trained on both the labeled and unlabeled data and the decision boundary (solid line) makes the two classes linearly separable.

$$(\{y_i, i \in \mathcal{L}\}|\mathcal{N}(\boldsymbol{f}_i), \boldsymbol{\theta}) = \prod_{i \in \mathcal{L}} p(y_i|\mathcal{N}(\boldsymbol{f}_i), \boldsymbol{\theta}) = \prod_{i \in \mathcal{L}} \sum_{j=1}^{N_i} a_{ij} p(y_i|\boldsymbol{f}_j, \boldsymbol{\theta})$$

where $\mathcal{N}(\boldsymbol{f})$ defines the neighborhood of $\boldsymbol{f}$. Estimation of classifier parameters $\boldsymbol{\theta}$ can be achieved by maximizing the log-likelihood via an Expectation-Maximization algorithm [7]. To enforce sparseness of $\boldsymbol{\theta}$ (enforcing most of the components of the parameter vector $\boldsymbol{\theta}$ to be zero), one may impose a zero- mean Gaussian prior on $\boldsymbol{\theta}$. A zero-mean Gaussian prior with appropriate variance can strongly bias the algorithm in choosing parameter weights that are most likely very small (close to zero). This technique is known as a parameterized neighborhood-based classifier (PNBC).

Non-Myopic Active Learning

Given that available training data labels at the beginning of a demonstration are not available and that excavations must be performed to reveal training data labels, one may ask in which order anomalies should be excavated to maximally improve the performance of the classifier algorithm. One useful criterion is to use the confidence on the estimated identity of the anomalies that are yet to be excavated. Specifically, one may ask which unlabeled anomaly label would be most informative to improve classifier performance if the associated label could be made available. It has been shown [8] that this question can be answered in a quantitative information-theoretic manner.

For active label selection, posterior distribution of the classifier is approximated as a Gaussian distribution centered on the maximum *a posteriori* (MAP) estimate. The uncertainty of the classifier is quantified in terms of the posterior precision matrix. The objective of AL is to choose a feature vector for labeling that maximizes the mutual information (I) between the classifier $\boldsymbol{\theta}$ and the new data point to be labeled. The mutual information can be quantified as the expected decrease of the entropy of $\boldsymbol{\theta}$ after new sample $\boldsymbol{f}_{i*}$ and its label y_{i*} are observed.

$$I = \frac{1}{2}\log\frac{|H'|}{|H|} = \frac{1}{2}\log\{1 + p(y_{i*}|\boldsymbol{f}_{i*}, \boldsymbol{\theta}) \times [1 - p(y_{i*}|\boldsymbol{f}_{i*}, \boldsymbol{\theta})]\boldsymbol{f}_{i*}^T H^{-1} \boldsymbol{f}_{i*}\}$$

It is important to note that the mutual information I is large when $p(y_{i*}|\boldsymbol{x}_{i*}, \boldsymbol{\theta}) \approx 0.5$. Hence, the AL prefers to acquire labels on those unlabeled samples for which the current classifier is most confused or uncertain. In this fashion the classifier learns quickly by not excavating anomalies that reveal redundant information. The process continues as new labels are revealed until the expected information gain for the remaining anomalies is approximately uniformly low. At that point the classifier is adequately trained and target inference on the remaining unlabeled anomalies can be reliably performed. By invoking the principle of submodularity in the algorithm optimization, the approach has been adapted to allow for the selection of multiple simultaneous labels at one time, making the technique operationally practical.

Results and Discussion

TEMTADS 2x2 – Beale TREES

Training Steps

There were 45 features in the TEMTADS 2x2 dataset. These included polarizabilities for 12 time-gates on the three axes (features 1-36), and 3 rate feature for each axis (features 37-45). 30 initial labels were requested from the program office. These features were selected based on their sequential Fisher Information gain. Nine of these 30 labels were UXO, and the type of UXO was quite varied. In fact, from this initial selection a label was acquired for each of the UXO types. This suggested two things. First, the UXO were not clustered in feature space relative to the distribution of clutter. Second, the UXO distribution was likely multi-modal. In other words, Beale Trees would require additional training data.

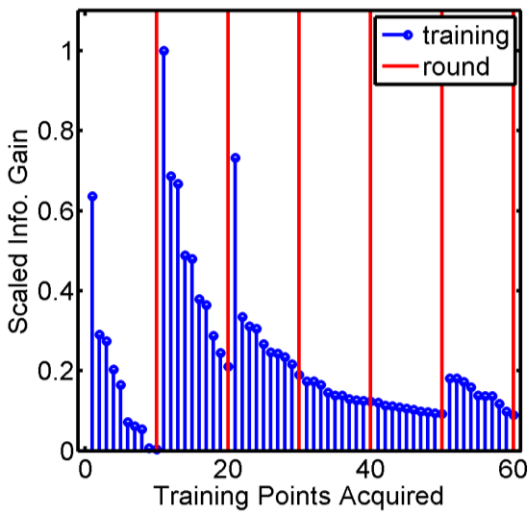

Figure 5. Fisher Information gain from training rounds 2-7.

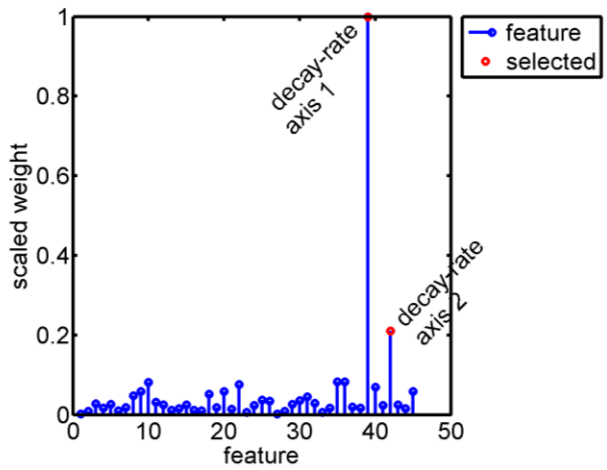

Figure 6. The BENet weights from round 7 with 90 labeled anomalies.

10 additional labels were requested for each round of training after the initial selection of 30 labels. These new labels were selected using NMAL. 5 rounds of training were performed in addition to the initial basis selection for a total of 80 labels acquired at the end of training. The decision to end training was based on a decrease in Fisher Information gain (Figure 6).

A new non-linear PNBC classifier was created for each round of training. The features used to create the non-linear kernel were selected based on BENet. As training proceeded the features selected by BENet became fixed. The dominant features at the end of training were the decay-rates of the x-axes and y-axes (Figure 7).

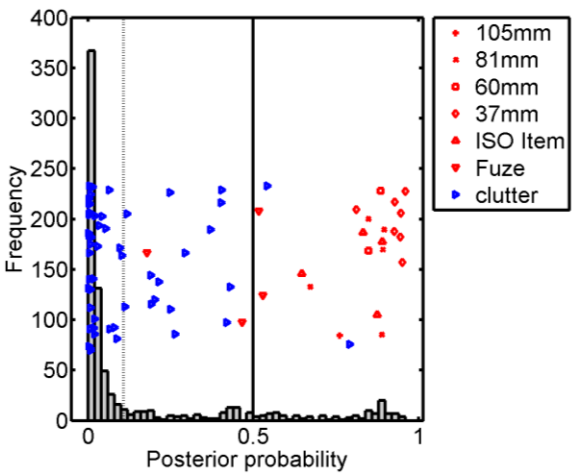

Figure 8. Predicted probabilities for the remaining anomalies at the end of training. The predicted probabilities of the labeled data are shown as shapes. Reference lines for predicted probability 0.5 (black) and 70th quantile of the anomalies (gray) are shown.

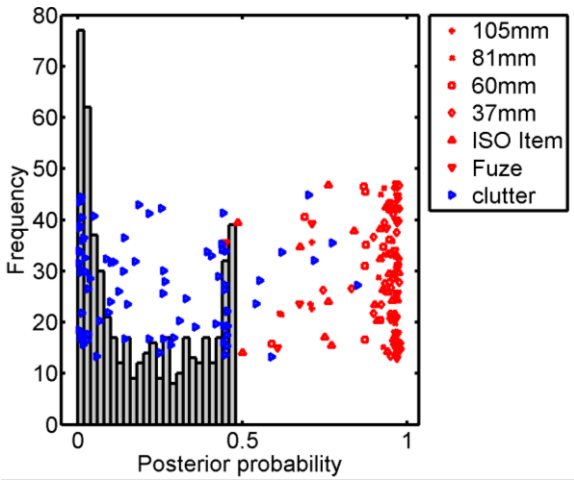

Figure 7. Predicted probabilities for the remaining anomalies at the end of digging. The predicted probabilities of the labeled data are shown as shapes.

Dig Lists

A PNBC classifier was training on all the labels acquired during training. Using this classifier, the predicted probability of being UXO was calculated for the remaining anomalies (Figure 8). These anomalies were sorted from most-likely UXO to most-likely clutter and the top 30% (including the training labels) were selected for digging. The total number of additional anomalies marked for digging was 188.

23 anomalies were deemed 'can't analyze' and marked for digging because their inversion fits were poor. Of the 23 bad fits, none were UXO, and 10 had no contact with an object when dug. In the retrospective analysis, no 'can't analyze' anomalies were separated. Instead difficult anomalies were simply placed in with the full dataset, and they were classified as clutter by the discrimination process. This improved discrimination performance.

A dig list was submitted to the program office and the labels for the dug anomalies were returned. No QC seeds were missing from this list. The PNBC classifier was retrained with the labels from anomalies that were marked as 'likely UXO' or 'can't decide'. Additional labels were requested for all anomalies whose predicted probability of being UXO was greater than 0.5. 15 new anomalies were requested in a 'stage 2' dig list. The procedure was repeated to generate a stage 3 list, where another 15 labels were requested. Digging stopped at this point because no anomalies existed whose predicted probability of being UXO was greater than 0.5.

Performance

The results from this final digging were that all the UXO were capture with ~160 false alarms (Figure 10). Approximately 20% of the UXO were captured during training. For final scoring the fuzes were considered to be inert and not counted as UXO.

Retrospective

There are three distinct approaches to classification: 1) the generative approach, 2) the discriminative approach, and 3) the functional approach. In the context of UXO, the generative

approach models the probability of being a target directly, without considering the distribution of clutter. The discriminative approach models the probability of being a target against the probability of being clutter. This is the technique we have used for classifying UXO in the performance assessment. The functional approach maps inversion values directly onto UXO based on training data and does not account for uncertainty in the sensor responses. This technique would be used in the case where sensor responses were perfect and anomalies could be classified with no uncertainty. At other sites, like the Pole Mountain Target Area, the functional approach may be sufficient, but this was not the case for Camp Beale. So, it receives no further consideration here.

One of the key benefits of using a generative approach is that digging can begin immediately from test pit data. In other words, no responses from clutter are necessary to train the model. The weakness of the generative approach is the possibility of missing hidden modes (in feature space) of UXO. Such modes would only be elucidated by exploring the clutter space as in the discriminative approach. In performance assessment the ROC curve for generative approaches will always be steeper initially than a discriminative approach. This is because training of generative models seeks only to find representative responses for UXO, whereas the discriminative approach seeks to find the boundary between UXO and clutter and explore areas of the feature space where little prior information is available. The generative approach is completely dependent on the library of UXO responses. So, if a hidden mode of UXO that do not exist in the library is present, then the number of false positives required to capture the last UXO will be greater than the discriminative approach. Given a generative model and a discriminative model with similar numbers of total clutter dug, the discriminative approach will tend to dig clutter during training and the generative approach will dig clutter during the 'dig phase'. This is obvious in ROC for the demonstration (Figure 10) where training accounts for half of the unnecessary digs.

The advantage of the discriminative approach is that the feature space can be explored in a structured way in order not to miss any UXO types that have odd responses. Taken together, the discriminative approach should be more effective in hard sites and the generative approach

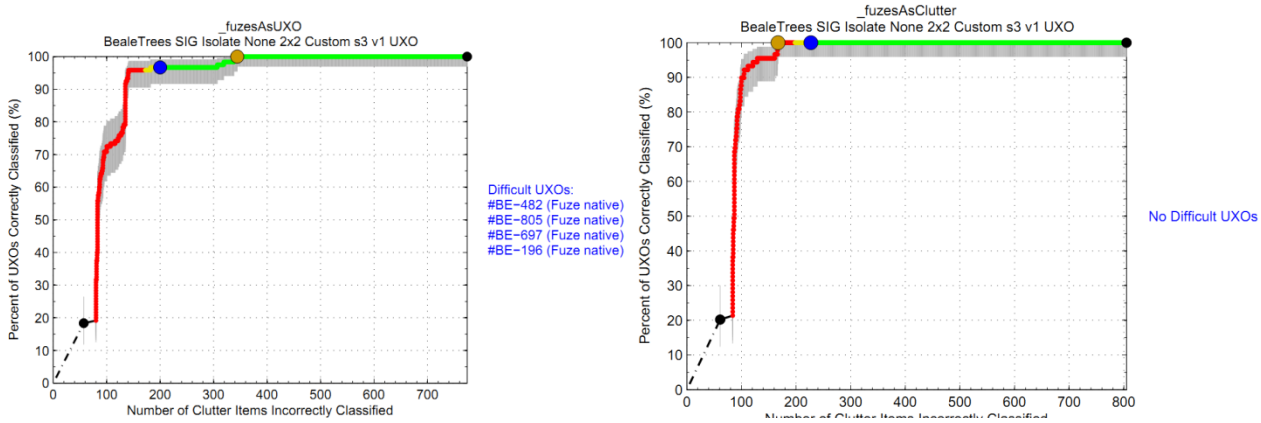

Figure 9. Final ROC curve for TEMTADS 2x2 in Beale Trees where Fuzes were considered UXO (left) and as clutter (right).

should be better in easy sites.

To compare the generative and discriminative approaches for the Beale Trees dataset the generative approach was applied to the TEMTADS 2x2 sensor data. Two generative scenarios were considered (Figure 11) and both treated fuzes as UXO. This represented the most difficult case for classification, even though in then end fuzes were deemed inert. In the first scenario the data were censored to only those anomalies whose inversion model had a low fit error (< 0.01). In the second, the data were uncensored. For the generative model on censored data, the number of unnecessary digs was drastically reduced over the original discrimination. The generative model applied to all the data, however, underperformed relative to the original discrimination. These results highlight the efficacy of the generative approach at 'easy' sites, where the inversion results are very good. It also demonstrates the inefficiency of the generative model on sites that have increased difficulty. In future demonstrations, SIG will develop an approach that is adaptively generative or discriminative based on the characteristics of the site.

BUD – Beale Trees

Training Steps

The SIG inversion model was not used for the BUD sensor data. Instead these inversions were provided by Erika Gasperikova (Lawrence Berkeley National Laboratory).

The BUD sensor was placed in multiple locations at each anomaly flag following a template. This resulted in multiple inversion results for each anomaly. Sometimes there were three template positions, but for most anomalies there were five. Initially, the mean response of all template positions was used to represent the anomaly polarizabilities. The UXO responses were masked using this technique, though, because the variance in inversion results between template

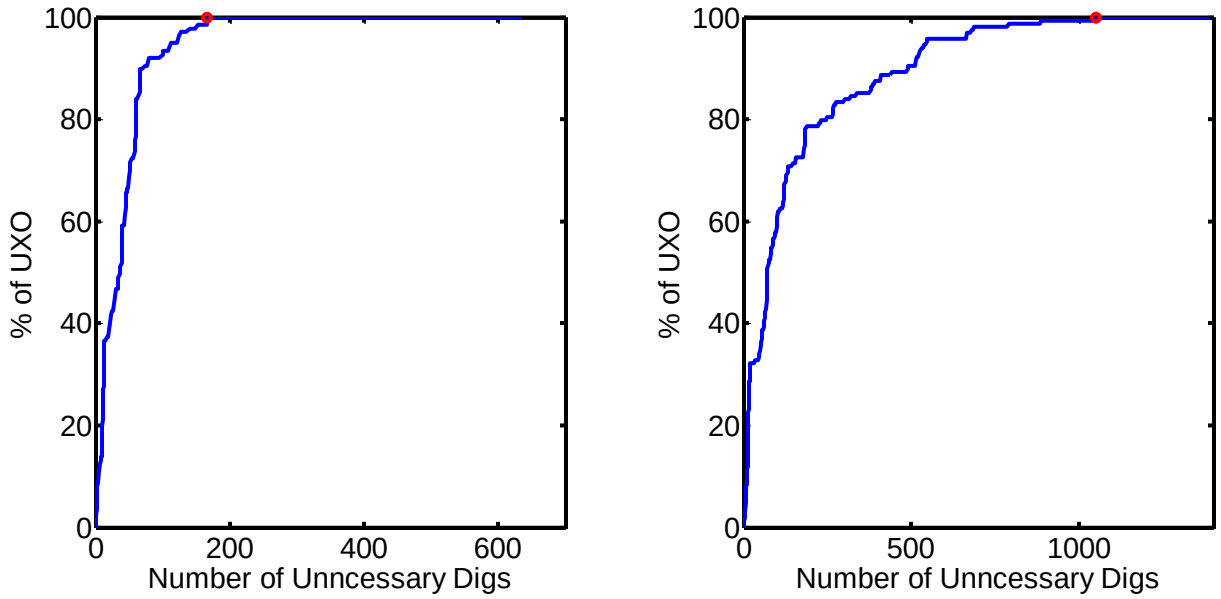

Figure 10. ROC curves for the generative model applied to the TEMTADS 2x2 sensor data from the BealeTrees site. The left graph shows the performance when the data are censored to only those anomalies with low error inversion models. The right graph shows the performance on all data.

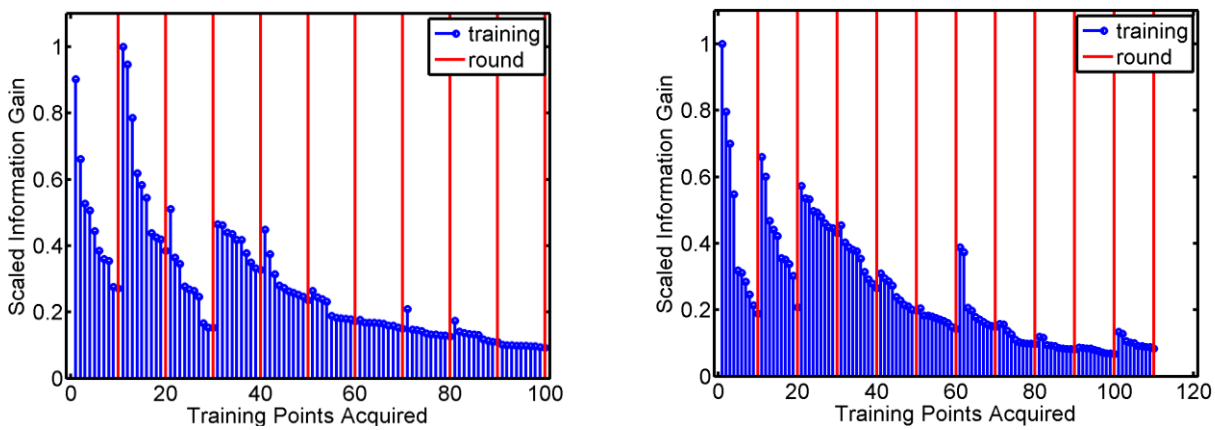

Figure 12. The Fisher Information gain through all training rounds for the 'symmetric' method (left) and the 'closest' method (right).

positions was high. This suggested that the sensor was focused on different anomalies at each template position. Or, one or more of the template positions was too far from the anomaly to provide reliable inversion results.

Instead of using some linear combination of all the template positions like the mean, two methods were explored for selecting a single template position to represent the anomaly. The first method selected the template position where the anomaly was estimated to be closest to the sensor, hereafter referred to as the 'closest' method. That is, the inversion estimate of the horizontal distance from the anomaly to the sensor was smallest. If the flag location were correct, template position 2, the central template position, should have been selected most frequently. However, template positions 1 and 4 were selected more frequently than template position 2 (Figure 12). This suggests that the flag locations relative to the anomaly were frequently beyond the distance that a single, centered sounding could illuminate.

The second method for selecting a single template position to represent the anomaly was to select the sounding that most symmetric along the transverse axes, hereafter referred to as the 'symmetric' method. Symmetry using the mean squared difference between log polarizabilities for the 2nd and 3rd axes:

$$S_i = \sqrt{\frac{\sum_{t=1}^{T}\left(p_{i,y,t} - p_{i,z,t}\right)^2}{T}}$$

where $i \in \{1, \ldots 5\}$ is the template position, p is the log polarizability at time $t \in \{1, \ldots T\}$ for the transverse axes y and z.

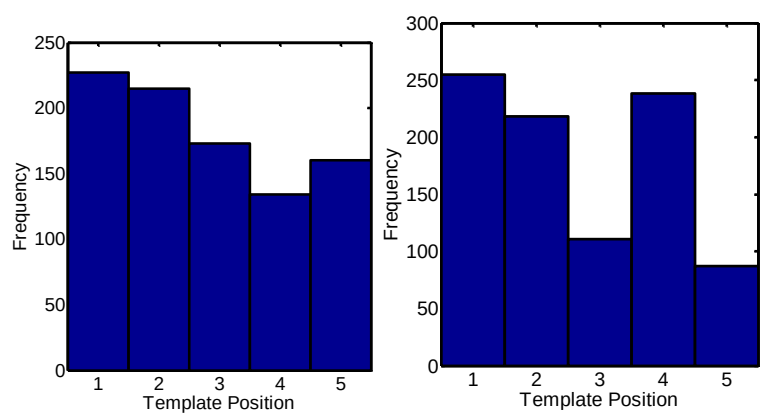

Figure 11. Histogram of the selected template position for the BUD sensor across all anomalies in BealeTrees for the 'symmetric' method (left) and the 'closest' method (right). Template position 2 is the central template position.

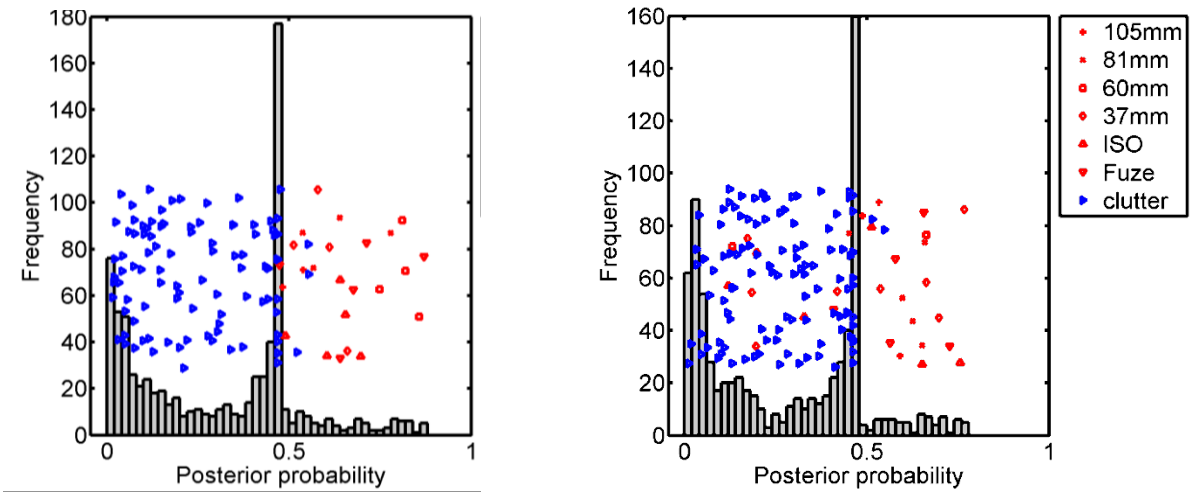

Figure 13. Distribution of posteriors for the 'symmetric' method (left) and the 'closest' method (right) at the end of training.

Having selected a single template position for each flag using these two methods, training proceeded according to the procedure outlined in the SIG Isolate Process Overview. An initial basis of 30 vectors was selected maximizing Fisher Information gain. This initial selection was the same for both the 'closest' and 'symmetric' methods. From this initial sample a set of relevant features was selected using BENet, and a non-linear PNBC classifier trained on these features. Then, a new set of 10 unlabeled anomalies were selected using NMAL. The labels for these anomalies were requested and the training process was repeated until new labels did not provide a substantial amount of new information. The total number of training rounds for the 'closest' was 11 and the total number for the 'symmetric' method was 10 (Figure 13). At the end of training there were many more features that were relevant for discrimination in the 'closest' method than the 'symmetric' method. This is likely due to the fact that the 'closest' method selected the wrong template position to represent an anomly more frequently than the 'symmetric' method did (See retrospective 4.2.3).

At the end of training the distribution of posterior probabilities of unlabeled samples was examined (Figure 14). The 'closest' method was not as separable as the 'symmetric' method at this stage; some predicted probabilities for the previously labeled UXO were less than 0.5. Unlabeled anomalies were selected for digging, beginning from the highest posterior probability and proceeding until 30% of the site was selected (including training data). There were three 'can't analyze' anomalies, whose positions were too far from the sensor to be reliable (BE-0609, BE-0879, and BE-0887).

These stage 1 dig lists, where 30% of the site was marked for digging, were sent to the program office and after receiving the labels back for this first round. The initial list missed 8 QC seeds in the 'closest' method and 3 seeds in the 'symmetric' method. The PNBC classifier was retrained based on these missed seeds. An additional 43 labels were requested for the 'closest' method and an additional 8 for the 'symmetric' method. These requests comprised the stage 2 dig lists. When labels were received the decision was made to stop digging the 'closest' method and to continue digging the 'symmetric' method . An additional 8 labels was requested in a

stage 3 list for the 'symmetric' method and then 22 labels. The final list was stage 5 and an additional 22 labels were request in this dig list. The total number of ranked anomalies (those dug but not counted in the training or 'can't analyze' sets) at the end of digging was 208 for the 'symmetric' method and 184 for the 'closest' method. The decision to stop digging the 'closest' method was based on the fact that the final 50 ranked anomalies were clutter, even though there remained 20 anomalies with a posterior probability greater than 0.5 (Figure 15). In retrospect, digging these remaining anomalies would not have revealed any UXO.

Performance

Both the 'symmetric' and 'closest' methods ended digging up all the UXO (Figure 16). The 'closest' method performed better than the 'symmetric' method in the number of unnecessary digs required to reach the last UXO: 180 vs. 200. Both methods are conservative given the fact that approximately 50 extra clutter were dug after all the non-fuze UXO had been revealed. Further, the features space had been thoroughly explored by NMAL. Over 100 training labels were acquired using NMAL in both methods. The 'closest' method, though, missed 3 fuzes and the 'symmetric' method missed 4 fuzes when these items were considered UXO. Both methods missed BE-0482, BE-558, and BE-0805, and the 'symmetric' method additionally missed BE-0697. The reasons why these fuzes were missed are described in the retrospective analysis.

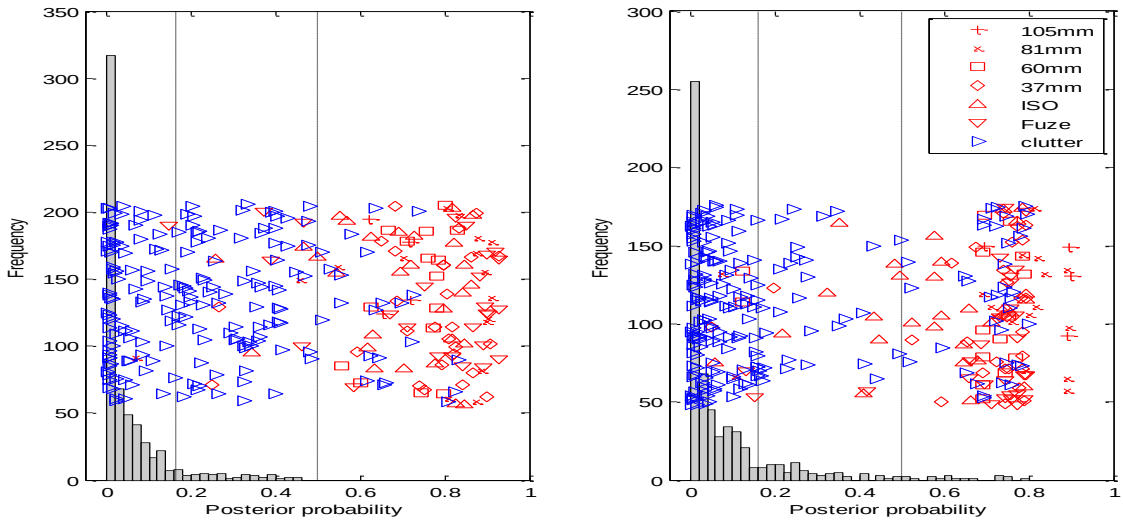

Figure 14. Posteriors distributions at the end of digging for the 'symmetric' method (left) and the 'closest' method (right).

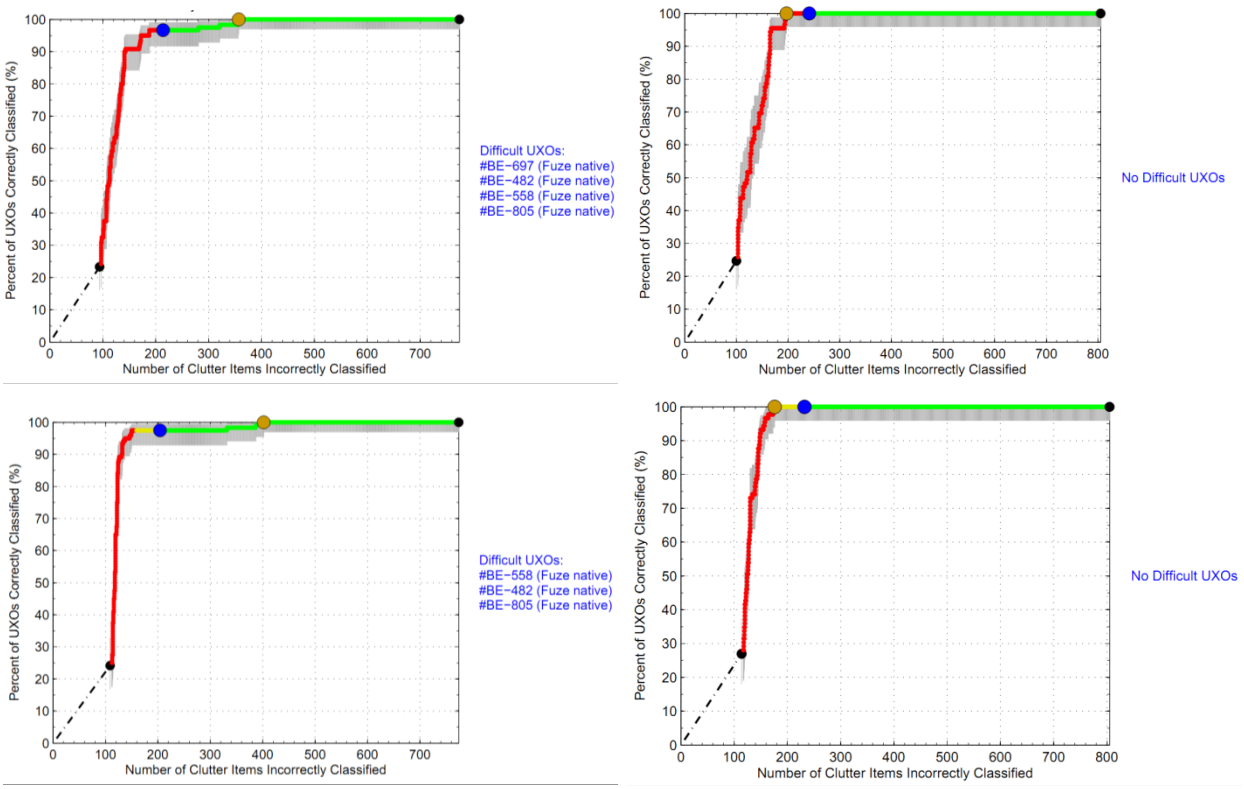

Figure 16. ROC curves for the 'symmetric' method (top) and the 'closest' method (bottom). Fuzes are treated as UXO (left) and clutter (right).

Retrospective

A group of fuzes were missed that, in retrospect, were fairly similar in feature space. Though, ultimately classified as inert, fuzes are representative of more 'difficult' UXO in that they don't have a typical munitions shape and are small. The primary reason for missing this mode of targets lies in the feature selection process. Each time a classifier was built, the feature space was pruned to only those features that were linearly discriminative of UXO and clutter. The set of features that was discriminative for most of the UXO, though, was not the same set of features that was discriminative for fuzes (Figure 17). A solution to this problem would be to create a separate classification for fuzes.

A secondary reason for missing these fuzes was that they were surrounded in feature space by clutter. Figure 18 shows that while the missed fuzes show clustering along the primary axes of feature variation, there are also many clutter close to them. During active learning labels were acquired via NMAL that were very close to this cluster of fuzes, and they were all clutter. So, any model fit to those labels would assign a low probability of being UXO to the fuzes.

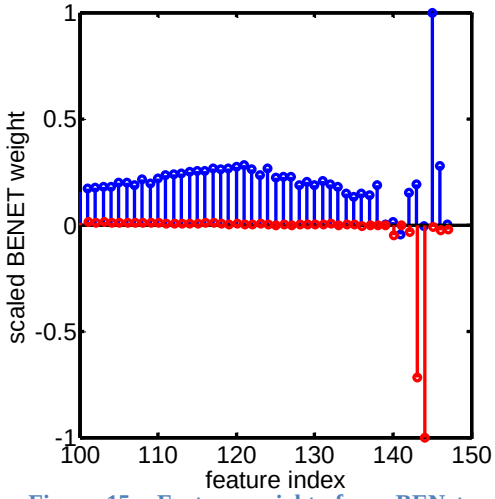

Figure 15. Feature weights from BENet on all labeled anomalies for all UXO (blue) and fuzes only (red).

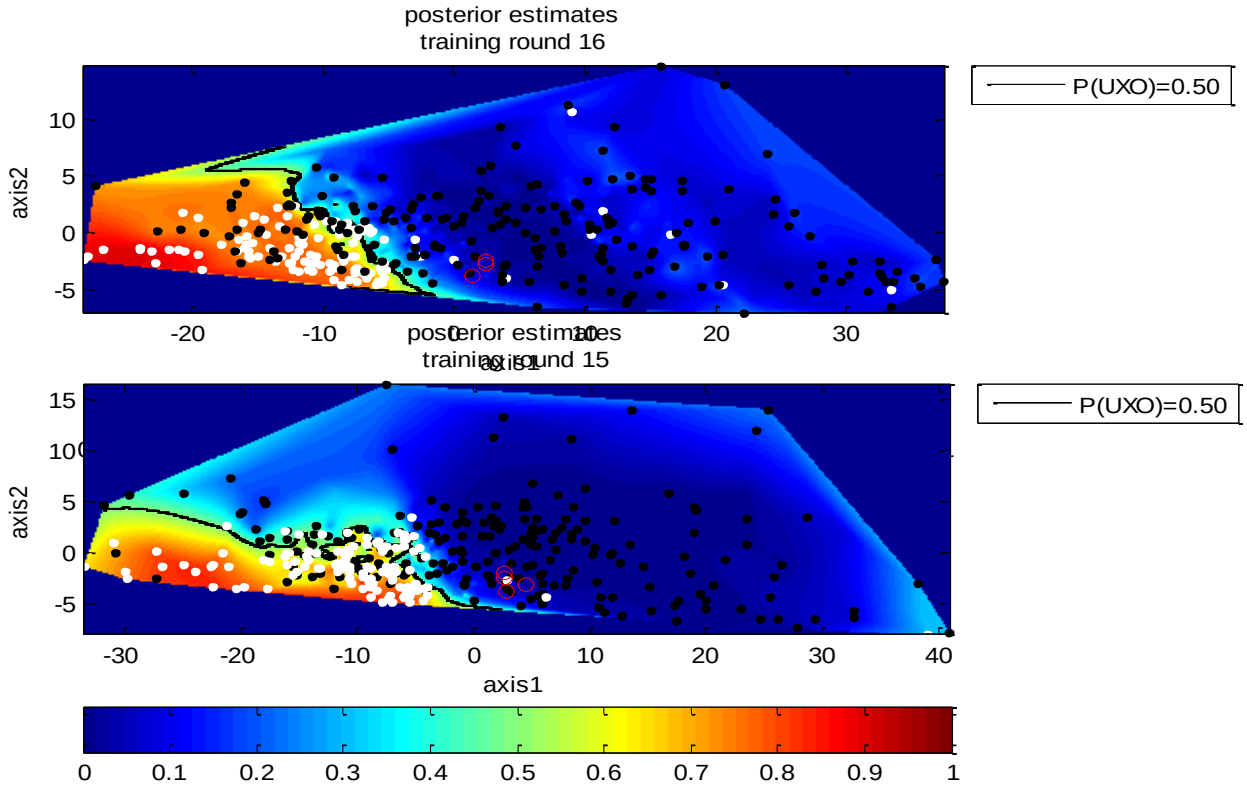

Figure 16. MAP estimates along two dimensions of a non-linear ordination of the anomaly features for the 'closest' method (top) and the 'symmetric' method (bottom). The anomalies that were marked as 'dig' in the dig list are shown (UXO in white, clutter in black). The missed fuzes for each method are also shown (red circles).

Again, this could be remedied by using features that were uniquely discriminative for fuzes rather than being discriminative of the entire dataset (as in Figure 18) or all UXO combined (as in the BENet feature selection during the demonstration)

In terms of selection algorithm performance, the 'symmetric' algorithm tended to have fewer incorrectly selected template positions than the 'closest' method. There are a few lines of evidence for this. The first is based on the variance of features for UXO which was lower for the 'symmetric' method than the 'closest' method. So, the feature space for the 'closest' method was compressed relative to the 'symmetric' method.

The second line of evidence for the ineffectiveness of the 'closest' method was the spread of UXO in the two dimensional ordination for the 'closest' method (Figure 18). There were noticeable UXO outliers in the dominant feature axes suggesting the polarizability responses for these UXO are substantially different from the mean responses of UXO. These outliers were not present in the 'symmetric' method, leading to the conclusion that these outliers were the result of selecting an incorrect template position in the 'closest' method. An example of inappropriate selection is shown in Figure 19. For this UXO (an ISO) every other template position had a 'UXO-like' response except the one chosen by the 'closest' method. The reason for such inappropriate selection is straight-forward. There was a piece of clutter underneath template position 4; one not associated with the flagged anomaly. This piece of clutter is closer to the sensor than the UXO was to any of the other template positions. So, the 'closest' method selected it.

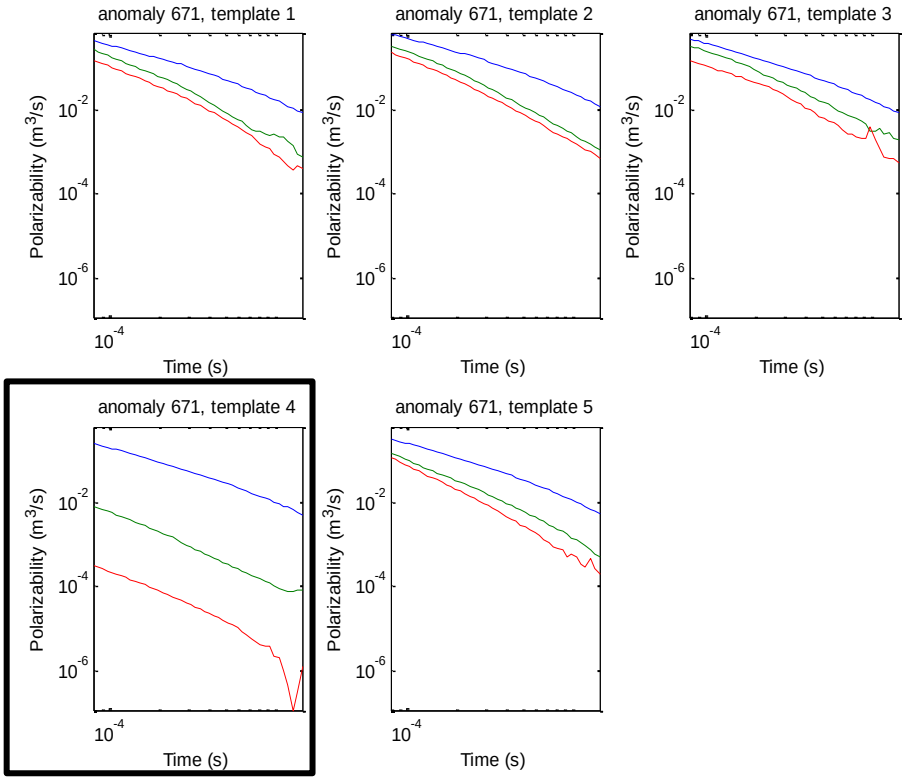

Figure 17. Polarizabilities for flag BE-0671 which is an ISO. The 'closest' method selected template position 4 (black box). when all other responses are more UXO-like.

MetalMapper – Beale Open

SIG performed discrimination analysis on the "Beale Open" site for the MetalMapper sensor operated by CH2MHill. There were 1470 flags to discriminate. Munitions included 3in, 75m,, 60mm, 57mm, and 37mm, mortars; ISOs. Fuzes, though included in the discrimination, were considered inert in the final scoring.

Training Steps

There were 45 features in the MetalMapper dataset used for discrimination. These included polarizabilities for 12 time-gates on the three axes (features 1-36), and 3 rate feature for each axis (features 37-45). Initially, 30 labels were requested from the program office. These features were selected based on their sequential Fisher Information gain. Four of these 30 labels were UXO. This included a 60mm mortar, two ISO, and one Fuze. As with the TEMTADS 2x2 data, this suggested that the different UXO likely were clustered along multi-modes in feature space and that those clusters of UXO were interspersed within the clutter. This dataset required additional training data.

10 new labels were requested at each round of training after the initial request for 30 labels. These new labels were selected using NMAL. 9 rounds of training in addition to the initial basis selection were performed for a total of 120 labels acquired at the end of training. The decision to end training was based on a decrease in Fisher Information gain (Figure 20).

A new classifier was created for each round of training. These classifiers were based on features that were selected specifically for that round using BENet. As training proceeded the features that were selected by BENet became relatively fixed. The dominant feature at the end of training

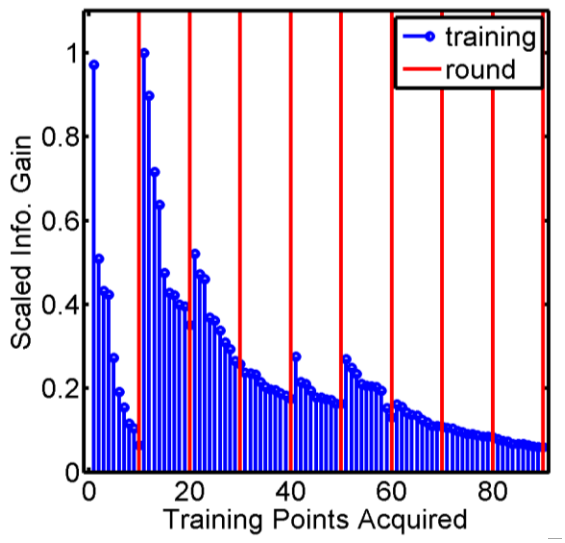

Figure 19. Scaled Fisher Information gain for 9 rounds of aquiring training labels via active learning for the MetalMapper sensor at the BealeOpen dataset.

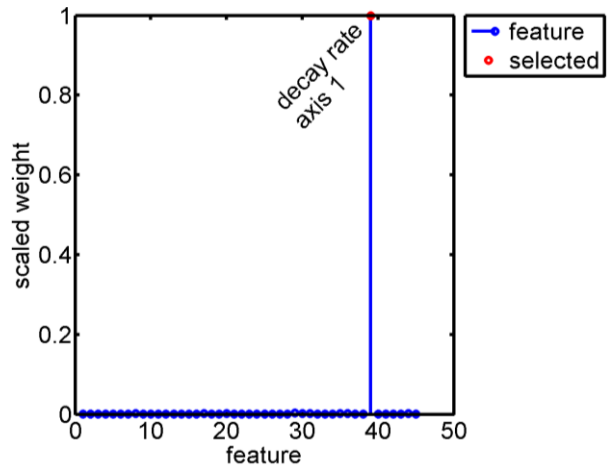

Figure 18. BENET weights at the end of training for the MetalMapper sensor in the BealeOpen dataset. The selected feature is shown in red.

was the decay-rate of the primary polarizability axis (Figure 21).

Dig Lists

A PNBC classifier was trained using all the labels that were acquired during training. The posterior predicted probability of being UXO was calculated for the remaining anomalies (Figure 22). These anomalies were sorted from most-likely UXO to most-likely clutter and the top 30% (including the training labels) were selected for digging. The total number of additional anomalies marked for digging was 285.

Additionally, 56 anomalies were deemed 'can't analyze' and marked for digging because their inversion fits were poor. Of the 56 'can't analyze' anomalies, two were UXO (ID 1951: 37mm and ID 2040: ISO). The initial dig list requested a total of 441 digs including all the training data and 'can't analyze' anomalies. This list missed 17 seeds. The missed seeds were of all sizes including 1, 81mm; 3, 60mm; 8, 37mm; and 4, ISOs.

With the labels for the missed seeds, the PNBC classifier was retrained and a second dig list was submitted. An additional 118 digs were requested for 559 total digs. A third and fourth stages were performed requesting an additional 90 and 37 digs. The final number of digs was 685. The decision to continue digging was based on the distribution of predicted probabilities for the remaining anomalies and the rate at which new UXO were found in the partial ROC curves. The distribution of predicted probabilities at the end of digging showed two things (Figure 23). First, the remaining anomalies were clustered around low probabilities. Second, the resubstituted probabilities of labeled anomalies were well-separated from the remaining anomalies.

Performance

The results for MetalMapper discrimination were not as good as the performance results for the Beale Trees dataset with the TEMTADS 2x2 and the BUD sensors (Figure 24). The final dig list missed two 37mm, and one ISO. An additional 150 digs would have been required to capture all these UXO.

Retrospective

The UXO missed in the dig lists showed marked differences from the standard responses for the UXO type (Figure 26). It is unlikely that any of these UXO would be captured by an automated discrimination method. The exception to this would be if these responses, though different from the standard UXO, were also different from the standard clutter response. In this case, the initial basis selection could capture them. During initial basis selection, the feature space is explored and outlier responses should be captured. To see if this were true, we examined how many basis selections were needed, (based on the gain in Fisher Information) to capture the anomalies missed in the original dig list. We selected 200 initial basis vectors instead of the original 30. This sample only captured one of the missed UXO (ID 2091, an ISO stats seed). So, the fuzes and the 37mm that were missed are not markedly different from clutter responses.

A retrospective discrimination was undertaken to see if performance could be improved. Five changes were made to the previous model. First, 35 time gates were used instead of the original 12. This resulted in better estimates of anomaly position. Second, the fuzes were classified separately from the clutter. This makes sense because the shapes of fuzes are different than the shape of other UXO. The third change that was made was in the number of training data that were acquired. Only 40 training labels were acquired. Fourth, there were no 'can't analyze' anomalies. For any anomaly that had poor fits, a two anomaly model was fit and the most-uxo-like of the two anomalies was used. two anomaly fits Finally, there was only a single dig stage, which means that there was no retraining and no 'missed QC seeds'.

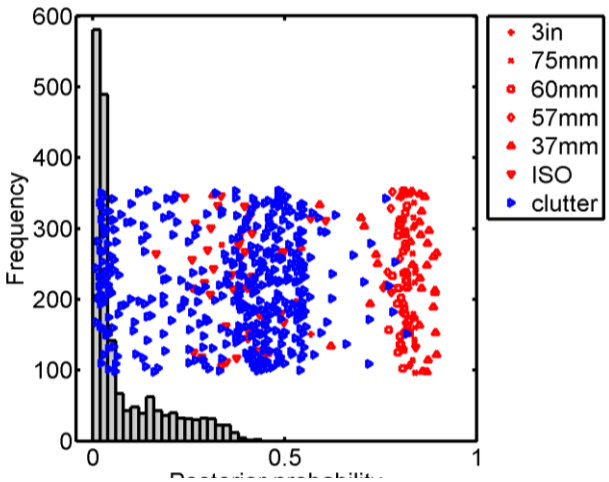

Figure 20. Histogram of predicted probabilities of being UXO for unlabled MetalMapper data at the end of training. Resubstituted predicted probabilities for the training data are shown as symbols: clutter (blue) and UXO (red).

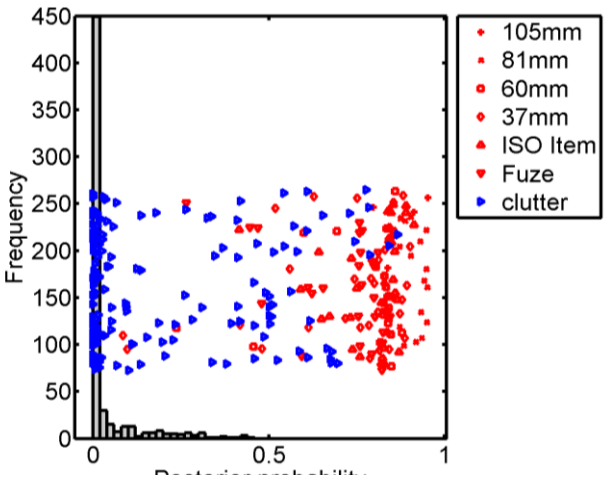

Figure 21. Histogram of predicted probabilities of being UXO unlabeled MetalMapper data at the end of digging.

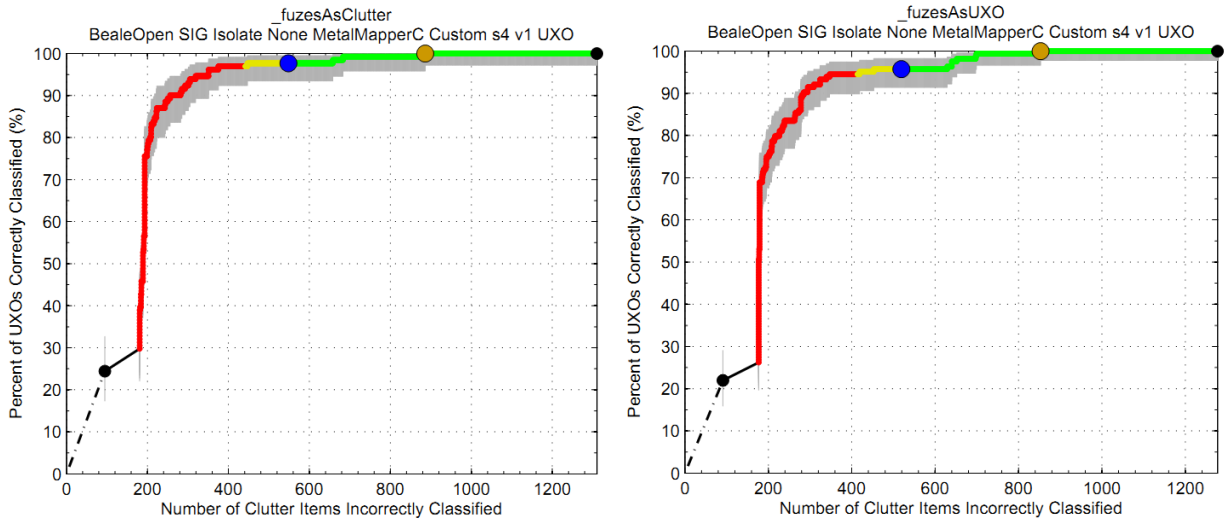

Figure 22. ROC Curve for MetalMapper sensor, BealeOpen dataset where fuzes are counted among the UXO (right) and where they are counted as clutter (left).

The performance of this classifier was better than that of the original dataset (Figure 25. However, two anomalies were difficult to discriminate. The first was ID 2091. This was the same ISO Stats seed that was missed by the original discrimination, and had an odd, clutter-like response (Figure 26). The second anomaly that was difficult to discriminate was 2277. This was an anomaly that we received as a Missed QC Seed in the original demonstration and should have been modeled as a multi-dipole anomaly. Barring these two, the retrospective analysis discriminated all the non-Fuze UXO with only 40 training labels, and 450 clutter digs, a significant improvement over the original performance of the classifier.

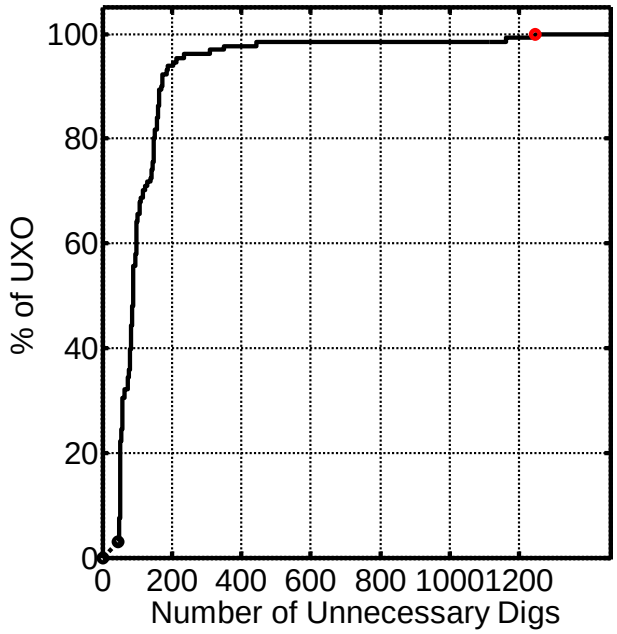

Figure 23. ROC Curve for the MetalMapper retrospective discrimination of the Belae Open dataset. Only 40 training samples were used.

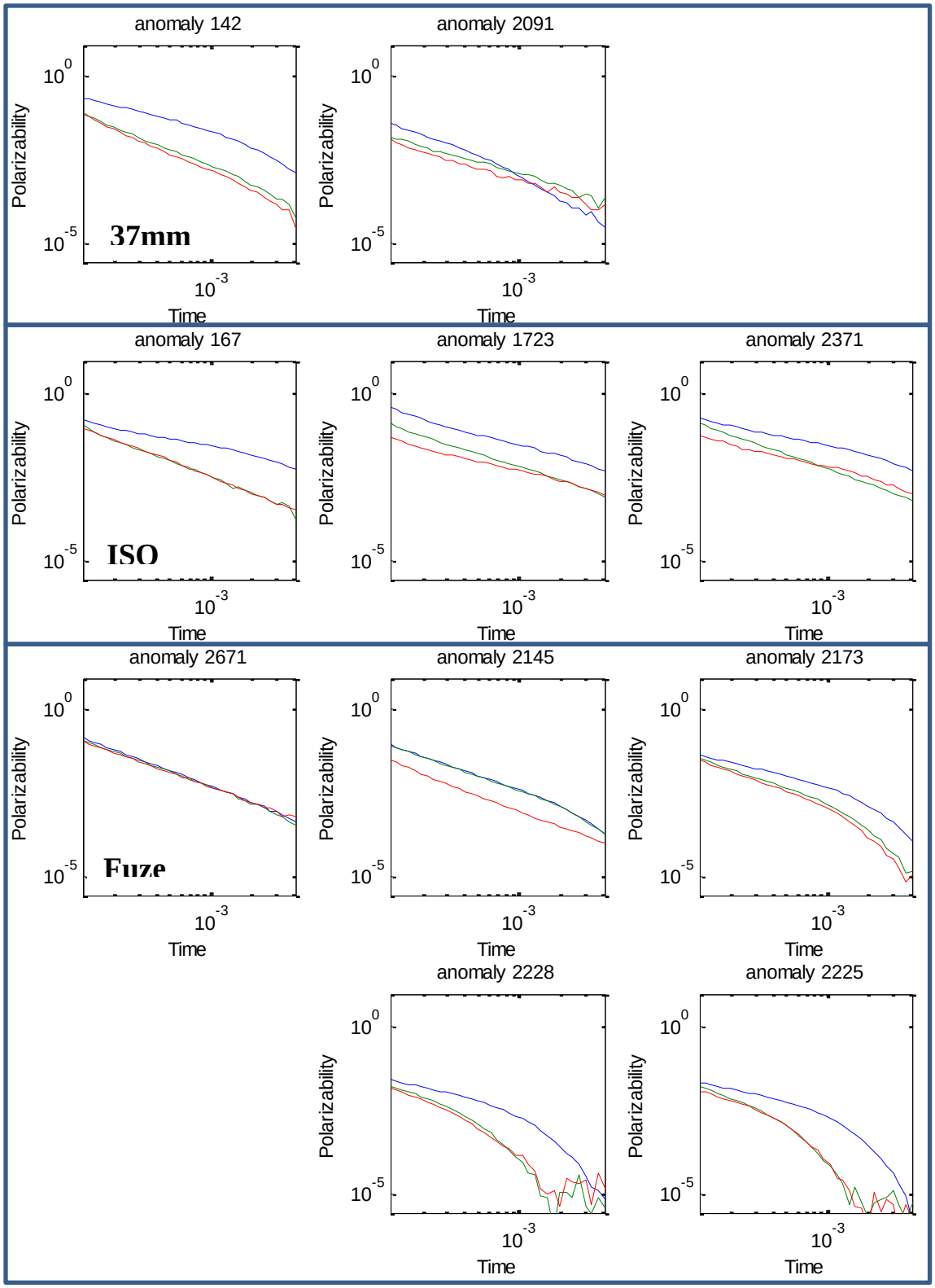

Figure 24. Archetypal responses for the missed UXO type (left column) along with the responses for the UXO that were missed in the dig lists (center and right columns). The responses are significantly different from the standard response and are unlikely to be capture by an automated discrimination method.

Conclusions and Implications for Future Research/Implementation

The SIG Isolate process has proven effective at providing efficient site discrimination with minimal training data. Camp Beale provides another example of this. The discrimination at Camp Beale also highlighted a key area where the SIG Isolate process can be improved. A generative, rather than a discriminative approach, could improve performance in special, easily-discriminated sites. There are two distinct advantages to the generative approach. The first is that it requires no training data *per se* and digging UXO can begin immediately. The second is that each UXO type can comprise its own model, relieving the aforementioned difficulty of classifying fuzes in a single class with all the other UXO types.

Literature Cited

[1] Y. Zhang, X. Liao, and L. Carin, "Detection of buried targets via active selection of labeled data: application to sensing subsurface UXO," IEEE Transactions on Geoscience and Remote Sensing, vol. 42, pp. 2535–2543, November 2004.

[2] Y. Zhang, H. Y. L. Collins, C. Baum, and L. Carin, "Sensing of unexploded ordnance with magnetometer and induction data: Theory and signal processing," IEEE Transactions on Geoscience and Remote Sensing, vol. 41, pp. 1005–1015, May 2003.

[3] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable target," IEEE Transactions on Geoscience and Remote Sensing, vol. 37, pp. 347–359, January 1999.

[4] L. Carin, H. Yu, Y. Dalichaouch, A. Perry, and P. Czipott, "On the wideband EMI response of a rotationally symmetric permeable and conducting target," IEEE Transactions on Geoscience and Remote Sensing, 2001.

[5] Y. Zhang, X. Liao, and L. Carin, "Detection of buried targets via active selection of labeled data: application to sensing subsurface UXO," IEEE Transactions on Geoscience and Remote Sensing, vol. 42, pp. 2535–2543, November 2004.

[6] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and other kernel-based learning methods. Cambridge University Press, 2000.

[7] X. Liao and L. Carin, "Migratory Logistic Regression for Learning Concept Drift Between Two Data Sets With Application to UXO Sensing," IEEE Transactions on Geoscience and Remote Sensing, vol. 46, no. 12, December 2008.

[8] D. MacKay, "Information-based objective functions for active data selection," Neural Computation, vol. 4, pp. 589–603, 1992.

Appendices

Appendix A: Points of Contact

POINT OF CONTACT Name	ORGANIZATION Name Address	Phone Fax E-mail	Role in Project
Lawrence Carin	Signal Innovations Group, Inc. 4721 Emperor Blvd., Suite 330 Durham, NC 27703	919 660-5270 919-323-4811 lcarin@ece.duke.edu	Principal Investigator
Levi Kennedy	Signal Innovations Group, Inc. 4721 Emperor Blvd., Suite 330 Durham, NC 27703	919-323-3456 919-287-2578 lkennedy@siginnovations.com	Project Management
Todd Jobe	Signal Innovations Group, Inc. 4721 Emperor Blvd., Suite 330 Durham, NC 27703	919-323-4811 919-287-2578 tjobe@siginnovations.com	Engineer
Xianyang Zhu	Signal Innovations Group, Inc. 4721 Emperor Blvd., Suite 330 Durham, NC 27703	919-323-4811 919-287-2578 xianyang@siginnovations.com	Engineer