PHM Decision Support under Uncertainty

Murat Yasar[1] and Teems E. Lovett[2]

[1,2] *United Technologies Research Center, East Hartford, CT, 06118, USA*

yasarm@utrc.utc.com
lovette@utrc.utc.com

ABSTRACT

Decision support systems aim to improve the quality of services and help operators perform their duties faster, more accurately and more efficiently by providing an immense amount of knowledge. As human operators cannot convey their complete understanding of the situation to the system, decision support systems face the challenge of interpreting human intent based on operator inputs, which introduces a high level of uncertainty into the system. In this paper, a decision support system is used for determining system health status as a decision aid to the operator. The goal of this work is to supplement sensor data with human inputs in a prognostic health management environment while minimizing the effects of uncertainty and to provide situational awareness for the operator.

1. INTRODUCTION

Advanced data mining and machine learning techniques are evolving rapidly. With the advent of “Big Data” technology and highly sophisticated data indexing, retrieval, clustering and classification processes, automated decision support has become an integral part of industrial, commercial and medical applications. Yet, the operator is still a significant source of both information and uncertainty in these applications such as given in Yasar, et al. (2009). For example, diagnostic/prognostic decision support systems play an increasingly important role in medical practice. These computer systems are designed to assist physicians or other healthcare professionals in making clinical decisions. They can provide decision support for a particular diagnostic (e.g. decision tree) or recommend therapeutic options by interpreting pulmonary function tests, analyzing electrocardiograms (e.g. pattern classification), etc. In any given decision process, the patient provides the critical inputs and the physician interprets the decision outputs based on the domain knowledge and patient’s history. Therefore, while interacting with the decision support system, the patient and the healthcare professional provide possibly subjective information with high uncertainty.

In a perfect world, a decision support system would be ubiquitous and omniscience. However, it is impossible to know or foresee all possible outcomes during the design of these systems. Therefore, passive decision support systems that depend on pre-defined strategies and models lead to sub-optimal solutions in most cases. Due to this apparent drawback, the decision support system turns to human operator to provide the context-specific domain knowledge. This is at the heart of “active learning.” Leveraging human inputs can enable achieving significant performance improvements as well as providing enough flexibility for contextual adaptation. However, such active learning schemes introduce a type of uncertainty that most machine learning techniques are not equipped for. As opposed to sensor data, human-generated information can become rather unstructured in general. Previous works in this area offer solutions based on constraining the interaction modality to limit the type of human inputs (Dasgupta & Hsu D, 2008). The idea is to extract implicit human intent to help make better decisions.

The key word here is “implicit”. Since intent, by definition, is a function of state of mind, it is too complex to be captured explicitly, even if a free-text input can be provided by human operator. (Although text mining may provide contextually relevant information, it would still be incomplete.) By making intent extraction implicit, the goal is essentially to have the human operator self-classify the intent into a set of possible inputs. This somewhat artificial, but necessary, discretization helps the decision support system to enumerate the outcomes and compute the most likely one. Going back to the diagnostic/prognostic decision support system example, it is possible to understand why the diagnostic tool is built on binary or multiple choice answers rather than free-text. At each decision junction, the decision support system uses heuristics to identify the most-likely outcome. By discretizing the possible inputs, the system tries to elucidate the human intent, in this case the perceived health condition. In fact, chaining together these human inputs creates a mathematical framework for modeling the decision making where outcomes become partially random and partially under the control of the human operator.

In this paper, the issues that arise due to interaction of human operator with the decision support system are addressed. Specifically, the aim of this paper is to provide a reasonable solution for interpretation of self-classified operator inputs that carry high uncertainty while making sure that the sensor data corroborated those inputs. This paper proposes a multi-tiered approach, where on one hand decision support system tries to auto-classify the sensor data (i.e. clustering) without considering the operator inputs; on the other hand, computes the sensor data associations with the operator inputs using a mixed-data metric (i.e. correlation).

This paper is organized in five sections including the present one. Section two elaborates the approach problem formulation for the proposed decision support environment and states the background definitions and assumptions. Section three describes the type of analysis that can be performed on mixed data modalities that enable automated decision making under uncertainty. In section four, the results for the developed analysis techniques on an anonymized dataset are presented. The paper is summarized and concluded in section five.

2. PROBLEM FORMULATION OF DECISION SUPPORT ENVIRONMENT

2.1. Definitions and Assumptions

First, we define the space for the decision support environment. Many different approaches can be found in the literature, among which the Markov Decision Processes seems to be a popular one. In this study, an industrial system is being considered where the overall decision support mechanism for diagnostics and prognostic health management is seen in Figure 1.

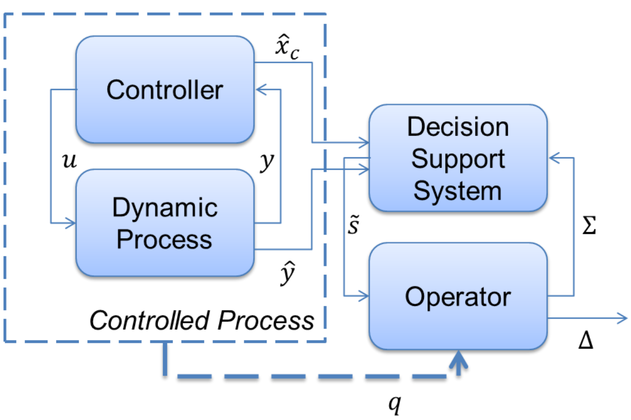

Figure 1. The information flow diagram for the decision support environment.

Assume that a controlled industrial process is being monitored by a human operator with the aid of a decision support system. Using the general system notation, $\{u, y\}$ represents the inputs and outputs of the process and x_c is the controller states. Only some (discrete) process state information, q, is available to the operator, whereas process outputs and controller states are partially observable by the decision support system, which is denoted by $(\hat{\bullet})$. Further assume that the decision support system interacts with the human operator to interpret operator decision state, $\Sigma \subseteq \Delta$, where $\sigma \in \Sigma$ is a set of possible decisions that may be taken by the operator that is known to the decision support system (which is, in general a subset of operator decisions, Δ.) This is where the uncertainty is introduced into the system. The decision support system produces a most likely process status, $\tilde{s}$. Note that $\tilde{s}$ could be an estimation of process state, i.e. $\tilde{s} = q|\max(Pr(q|\sigma))$ however, in general $\tilde{s}$ is a distribution over a superset of q. $(\tilde{\bullet})$ notation is used to explicitly specify estimates.

As mentioned before, the decision support system has two functionalities:

1. Clustering the process data to determine the possible process status that the data indicates
2. Correlating the human operator inputs and the process data to evaluate the operator inputs against the process status

Therefore, process status estimates of the decision support system evolve according to the equation:

$$\tilde{s} = f(\hat{x}_c, \hat{y}, g)$$

where f is the functional representation of the decision support system.

2.2. Data Types and Relationships

Beyond the assumptions stated earlier, the decision support problem is further complicated by the fact that different types of data are available for analysis. Although the terms "mixed data" and "heterogeneous data" have been used interchangeably in the literature, the mixed data term is preferred in this paper. At a very high level, what this means is that the data processed by the decision support system consist of continuous, discrete, transactional and categorical variables.

One of the biggest challenges for mixed data analysis is determining the structure of mixed data to perform semantic analysis (e.g. clustering, classification) to identify relevant patterns (Yasar et al. 2010 and Sarkar et al., 2014). Specifically, the associations between mixed data types require a quantitative measure of the strength of a phenomenon, which is called the effect size. Examples range from the correlation coefficient to the mean difference to information theoretic metrics such as neighborhood mutual information (Scheaffer, 1999 and Hu, 2011). In this work, rather than focusing on development of a new metric for the entire mixed data space, the aim was application of some of the better known clustering and classification methods in this new framework (Pages, 2014). It should be emphasized that the proposed decision support environment is flexible enough to accommodate applications both with and without the known ground truth.

In Figure 1, the controller states and process observables can be discrete (usually binary), continuous or categorical (e.g. state of a finite machine used in supervisory control). On the other hand, the operator decision states are almost invariable categorical. This is due to the fact that human-machine interaction is preferred to be enumerable (based on the discussion in Section 1). What is reducing the complexity in this framework is that there is no need to make any further assumptions on the type and availability of other data points given in Figure 1, since only the relationships between the process and the decision support, and the operator and the decision support are of interest.

3. BACKGROUND FOR CLUSTERING AND CORRELATION ANALYSIS

3.1. Clustering

Data clustering is a data exploration technique that is used for features with similar characteristics to be grouped together, or partitioning data into dissimilar subsets. Therefore, clustering is usually associated with the idea of unsupervised learning. In the decision support system, clustering is used as a means to partition the mixed data generated by the controlled process into discrete partitions in order to further analyze against the discrete decisions made by the operator.

The most popular clustering algorithms include k-means and hierarchical clustering. For a review of clustering methods, the authors refer to Duda et al (2000) and Guyon & Elisseeff (2003). As clearly demonstrated, clustering methods are based on optimization of a numeric criterion defined from a distance or from a dissimilarity measure.

For categorical data, k-means algorithm has been used on binary data obtained after conversion of multiple category attributes into binary attributes. It is also possible to work on the categorical data directly with same optimization criterion. Several researchers developed k-means type approaches specifically for categorical data; these include k-modes and k-medoids algorithms (Jollois & Nadif, 2002).

The appeal of k-means algorithm is due to its simplicity. In essence, it partitions n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

$$D = \sum_{i=1}^{k} \sum_{j \in \mathbb{C}_i} d(x_j, \mu_i)$$

where d is a distance function and μ_i are the means of clusters $\mathbb{C}_i$. K-means algorithm tries to find μ_i by optimizing $\min_{\mathbb{C}} D$.

A key aspect of k-means algorithm is k being user defined. Although there are many methods of determining k, in this paper, an analytical approach to determine k is unnecessary. This is due to the fact that operator's decision states, Σ, are known by the decision support system and therefore simply $k = |\Sigma|$.

3.2. Correlation

Strength of association, also known as effect size, between mixed data types has been studied extensively in statistics. Some examples of effect sizes include the correlation between two variables, the regression coefficient in a regression, the mean difference, and the risk associated with an event (Wilson & Martinez, 1997).

The most basic idea of correlation is "as one variable increases, does the other variable increase (positive correlation), decrease (negative correlation), or stay the same (no correlation)". For example, Pearson correlation, also known as Pearson's r, uses population variances of two variables. Spearman correlation is similar, but uses the ranks to determine the association between ordinal variables. Yet another metric, Kendall's correlation, uses concordant (or discordant) ranks between all pairs of observations. Given two variables x and y, Pearson's r can be computed as

$$r = \frac{\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i - \mu_x)^2}\sqrt{\sum_{i=1}^{n}(y_i - \mu_y)^2}}$$

where n is the number of data points.

If the data has mixed types, (e.g. one variable is nominal and the other one is continuous) then it is meaningless to ask "what happens to the continuous variable as nominal variable increases (decreases)". However, there are measures of strength of association that can be used that are somewhat analogous. For example, by converting the nominal variable to binary format it is conceivable to perform a multiple regression. Part of the output of regression is the coefficient of determination, R^2. One interpretation of R^2 is that the model explains that much of the variance in the continuous variable. Another interpretation of R^2 is that by taking the square-root, it is possible to determine the multiple correlation coefficient, R. That is equivalent to the Pearson coefficient, r, between the observations and the predictions by the model. Although, it might be uncommon to use (multiple) regression for a categorical independent variable, it is in fact analogous to one-way ANOVA approach. The equivalent of R^2, the proportion of variance explained for multiple regression, in ANOVA is η^2. Therefore, η could be used as an analogous measure of association between a categorical variable and a continuous variable and can be computed as

$$\eta = \sqrt{\frac{\sum_{i=1}^{n}\left(y'_i - \mu_{y'}\right)^2}{\sum_{i=1}^{n}\left(y'_i - \mu_{y'}\right)^2 + \sum_{i=1}^{n}\left(y_i - \mu_y\right)}}$$

A good starting point to determine the association between categorical variables is Pearson's chi-squared (χ^2) test. It is a statistical test to evaluate the likelihood that any observed difference between the categories arose by chance. A commonly used measure of association for the χ^2 test is the Cramér's V. What is important in using Cramér's V for data association is that it can be used with variables having more than two categories. Cramér's V can be computed as

$$V = \sqrt{\frac{\chi^2/_n}{\min(c,r) - 1}}$$

where c and r are the number of columns and rows of a contingency table, respectively.

4. RESULTS OF CLUSTERING AND CORRELATION ANALYSIS FOR AN INDUSTRIAL PROCESS

Consider an industrial controlled process under the monitoring of a human operator who tries to determine the health status and a correct diagnosis for the system. The system consists of many components and subsystems and controlled by a digital controller. Various data streams are logged within the digital controller's memory and also via a data logger. An existing monitoring system is primarily used for transmitting the data to a central data server.

Now, consider a decision support system that computes $\tilde{s}_t$ which is an estimation of the system status, where t is an asynchronous event index. The decision support system uses clustering to determine an operator agnostic decision for each event $\tilde{s}_t = \mathbb{C}_t | \mu_i \ \forall i = 1 \ldots k$, where $\mathbb{C}_t$ is the determined cluster label at t given the static cluster centers μ_i.

The data used for clustering consists of discrete input/output of the digital controller, and discrete and continuous states of the system. There are total 76 variables. During operation the operator has four possible decision states (i.e. System inoperable, Normal operation, System in test, Unknown) to choose from. It is assumed that not all decision states are known by the decision support system, therefore $\Sigma \subseteq \Delta$; and "Unknown" state is a catch-all condition for all other operator decisions. There are 277 events that have the associated data available.

The data is clustered using k-means algorithm and Hamming distance into four clusters to match the available number of decision states. The four-class clustering results are shown in Figure 2. While the y-axis shows the individual events, x-axis illustrates the variables that constitute the data.

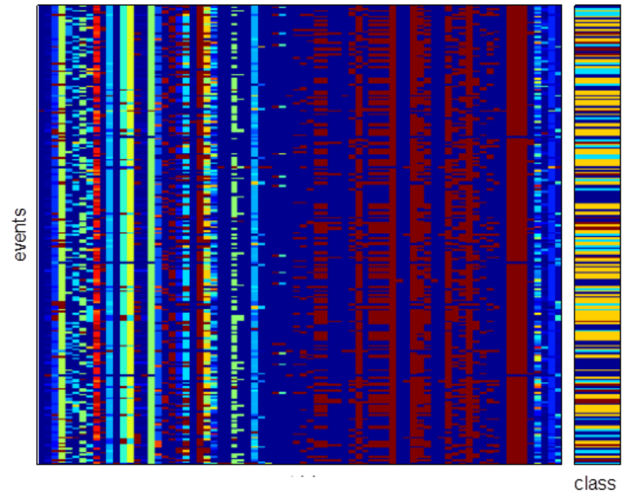

Figure 2. Four-class clustering results using Hamming distance on the process data.

It is possible to use the clustering results for feature selection. Features being the variables themselves, the goal is to find a subset of variables that would give the same clustering as the full dataset.

Consider Shannon entropy: $h = -\sum_{k=1}^{n} p_k \log(p_k)$.

Proposition: Given an initial cluster label $\mathbb{C}_i$ where $\mu_i \in \mathbb{R}^m$ are obtained by clustering m variables; there exists an ε such that $\hat{\mathbb{C}}_i = \mathbb{C}_i$ and $\hat{\mu}_i \in \mathbb{R}^l$ are obtained with l variables that satisfy $h > \varepsilon$ and where $l \leq m$.

For a chosen $\varepsilon = 0.1$, there are 34 variables that satisfied this criterion and these were used for obtaining the same clustering as the original 76 variables. The results are shown in Fig. 3.

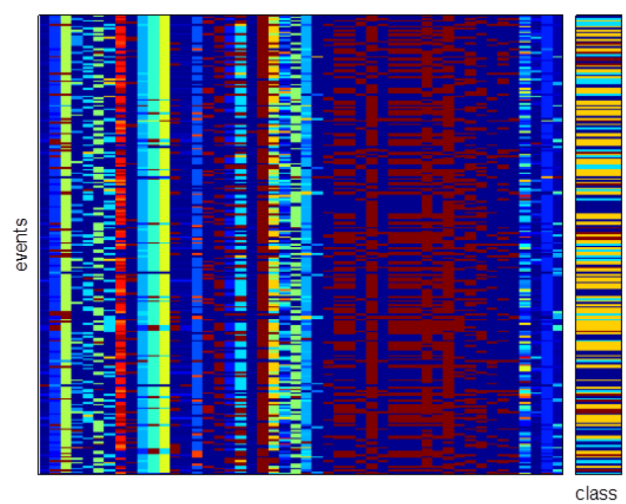

Figure 3. Four-class clustering results using Hamming distance on the reduced data set.

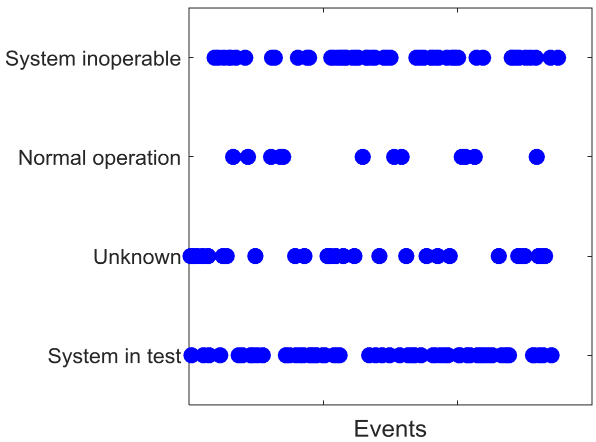

Figure 4. Operator decisions for the recorded events.

For the second stage of the decision support system, correlation and classification methods are utilized. The operator decisions, $\sigma \in \Sigma$, for each event constituted the class labels. The operator decisions are given in Figure 4 for each recorded event. It can be seen that most events correspond to conditions when the system is either inoperable or under testing, whereas a small sample of normal operating conditions are present. This is due to the fact that events are generated by the system controller to indicate potential failures of the system and the operator's role in this setting is to determine if there is an actual failure or not. Therefore, most of the events are expected to be either real failures or testing of the system.

The decision support system uses correlations between the process data and operator decisions to learn the strength of association, and then uses the learned model for classification of the process data to make a decision for each event $\tilde{s}_t = \mathbb{C}_t | m_i(\sigma_t)\ \forall i = 1 \ldots k$, where $\mathbb{C}_t$ is the determined class label at t and m_i is the association metric between the i^{th} variable and the operator decision state σ.

In this study, events with "Unknown" decision state were filtered out from the learning process (i.e. before performing correlation analysis). The question to answer here is "what are the correlations between the process data and decision states?" To answer that, first an appropriate metric for categorical data association and continuous data correlation was developed as described in Section 3. The results for the mixed-data correlations are presented in Figure 5. Note that out of 76 variables, only the ones that have an association metric greater than 0.4 are shown.

After the correlations were determined, top five indicators were chosen: Emergency operation input, Lower zone sensor, Upper zone sensor, Safe to operate, and Event condition. Using the model learned with these five indicators, classification was performed on the process data corresponding to events with "Unknown" decision state. The aim was to classify these events as one of "System inoperable", "Normal operation", "System in test" decision states. The re-classification results are given in Figure 6.

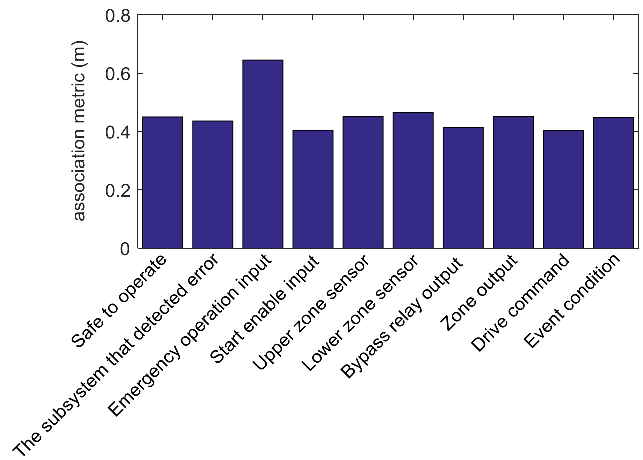

Figure 5. Correlations between process data and decision states.

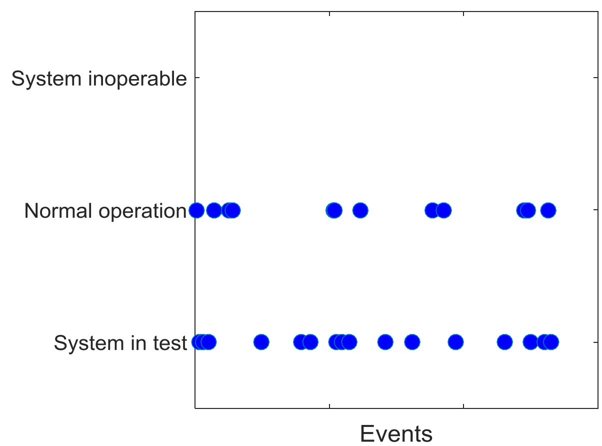

Figure 6. The classification of "Unknown" decisions to one of "System inoperable", "Normal operation", "System in test" decision states

As can be seen in Figure 6, the classification results indicate that those events with "Unknown" decision state were potentially either "Normal operation" or "System in test", and none were classified as an inoperable condition based on the observed patterns of the associated process data. Although it may seem counter-intuitive, the results are not totally unexpected since system failures are easier to be identified by the operators; therefore it is likely that they were not classified as "Unknown" by the operator when an event is generated.

Based on the correlation and classification analysis conducted, it can be concluded that it is very important to find appropriate metrics for correlation and classification analysis. That is the reason to develop a heterogeneous metric that enables automated identification of strong indicators in the process data. This method also provides the classification of events based on the learned model of operator decision state and re-classification of “Unknown” decisions. However, this technique is only applicable to univariate statistics, and does not account for combined effects (patterns) of process data on the decision state. Also only the linear relationships were captured in data association metrics of this study. As such, some features with significant nonlinear effect on the system status may be inaccurately excluded. It is possible to use mutual information based metrics to alleviate this problem which is under investigation.

5. CONCLUSION

This paper addressed how to use several well-established data analytics techniques to process system and controller data as a decision support tool. The decision support system is developed in the context of an industrial controlled process operating under the monitoring of a human operator. While the human operator only had access to high-level information about the system, the decision support was assumed to have more data sources available to it and therefore help reduce the uncertainty around the decision. Specifically, clustering and correlation techniques were established to provide better situational awareness for the human operator.

The main advantage of using a combination of clustering, correlation, and classification instead of performing a direct classification analysis with operator labeled event data is to address the sparsity of labels compared to the number of features. Both the classification and correlation has been used for feature selection to reduce the dimensionality of the feature space (from 76 to 5). A direct multi-class classification algorithm (even an ensemble method such as boosted decision trees) would be susceptible to overfitting if not preceded with feature selection. Another reason for not using a direct approach is that it would require balanced class labels or weighted priors. Since the labels are the primary source of uncertainty, introducing additional arbitrary weighing to class labels might have amplified the effects of uncertainty, therefore reducing the class separability.

Although, the main goal of this effort was to develop a decision aid for the human operator, the existing data and advanced machine learning methods can be used to create system performance and health condition metrics. Further, big data technology stacks can be used to manage larger data sets and execute parallel complex analytics.

In a future state, this work can also enable a few human operators to continually monitor thousands of systems with higher operational efficiency and situational awareness. A great example is air traffic controllers who help pilots keep aircraft safely separated from other aircraft or obstacles while in flight or on the ground. They monitor a large number of airplanes at any given time. This sheer number makes it prohibitive for the air traffic control to respond to unusual emergencies in a timely manner. For example, if an aircraft deviates from the planned route, it takes time to respond and the results can be catastrophic such as Malaysia Airlines flight 370 and Germanwings flight 9525.

Furthermore, service technicians for the industrial systems can use this type of decision support systems to get an accurate and complete picture of system health problems, which may enable improved service efficiency and correctness.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Mark Thomson and our corporate research sponsors at UTRC.

REFERENCES

Dasgupta S. & Hsu D. (2008). Hierarchical sampling for active learning. *Intl. Conf. on Machine learning*, Helsinki, Finland.

Duda, R. O., Hart, P. E. & Stork, D. G. (2000). *Pattern Classification (2Nd Edition)*. Wiley-Interscience.

Guyon, I. & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, vol. 3, pp. 1157-1182.

Hu, Q. H., Zhang, L., Zhang, D., Pan, W., An, S., & Pedrycz, W. (2011). Measuring relevance between discrete and continuous features based on neighborhood mutual information. *Expert Systems with Applications*, vol. 38, no. 9, pp. 10737–10750.

Jollois, F.-X. & Nadif, M. (2002). Clustering Large Categorical Data. *Advances in Knowledge Discovery and Data Mining*, pp 257-263. Berlin: Springer.

Pages, J. (2014). *Multiple Factor Analysis by Example Using R*. Chapman and Hall/CRC.

Sarkar, S., Sarkar, S., Virani, N., Ray A. & Yasar, M. (2014). Sensor Fusion for Fault Detection & Classification in Distributed Physical Processes. *Frontiers in Robotics and AI*, vol. 1, article 16.

Scheaffer, R. L. (1999). *Categorical Data Analysis*. NCSSM Statistics Leadership Institute.

Wilson, D. R. & Martinez. T. R. (1997), Improved Heterogeneous Distance Functions. *Journal of Artificial Intelligence Research*, vol. 6, pp. 1-34.

Yasar, M., Beytin, A., Bajpai, G., & Kwatny, H. G. (2009). Integrated Electric Power System Supervision for Reconfiguration and Damage Mitigation. *IEEE Electric Ship Technologies Symposium*, Baltimore, MD.

Yasar, M., Ray, A., & Kwatny, H. G. (2010). Minimum rotation partitioning for data analysis and its application to fault detection. *American Control Conference*, Baltimore, MD.

BIOGRAPHIES

Murat Yasar, Ph.D. is with the United Technologies Research Center (UTRC) working on diagnostics, prognostics and PHM applications related to elevators, HVAC equipment, gas turbine engines and aircraft components. He holds Master’s degrees in Mechanical Engineering and Electrical Engineering and a Doctorate in Mechanical Engineering, all from the Pennsylvania State University, University Park, PA. Before joining UTRC, Dr. Yasar was involved with various programs ranging from aircraft health management to supervisory control design for shipboard systems at Techno-Sciences, Inc. and InnoVital Systems, Inc. for seven years. Dr. Yasar is an expert in control systems design, signal processing and predictive analytics and has worked on many projects as a principal investigator throughout his career. His recent focus has been on development and implementation of machine learning and predictive analytics for PHM applications. Dr. Yasar is a member of ASME and IEEE.

Teems Lovett, Ph.D., is principal engineer, Embedded Systems & Networks at United Technologies Research Center (UTRC). He earned Ph.D. and M.S. degrees in electrical engineering from the University of South Carolina in 2002 and 2004. He is a UTRC research engineer in the areas of simulation of complex systems, applied numerical methods, data analytics, cyber-physical security and software architecture. His prior work experience includes: Space Exploration Software Simulation Engineer, L3 Communications, contracted to NASA’s Johnson Space Center, Houston, Texas; in this role he was a simulation software architect for the next generation Astronaut training facilities and provided requirements oversight for the Orion capsule. Prior to working at JSC he spent 5 years at United Technologies Aerospace Systems (UTAS) in the Electric Systems division as a Staff Engineer working on system simulation, system integration, and verification and validation for aircraft electric systems (Boeing 787). His current research interests include the simulation of complex systems, simulation under uncertainty and PHM. He is a member of IEEE.