Proceedings of the 60th Meeting of the Society For Machinery Failure Prevention Technology (MFPT), pp. 291-300, 6167, pp. 271-279, 2006

A MULTIPLE CLASSIFIER SYSTEM FOR AIRCRAFT ENGINE FAULT DIAGNOSIS

Weizhong Yan*, C. James Li[+] , Kai F. Goebel*

* Computing & Decision Sciences
GE Global Research Center
One Research Circle
Niskayuna, NY 12309
{yan, goebelk} @ crd.ge.com

[+] Department of MANE
Rensselaer Polytechnic Institute
Troy, NY 12181
lic3@rpi.edu

Abstract: Multiple classifier systems (MCS) are considered as one of the most significant advances in pattern classification in recent years. Numerous studies (both theoretical and empirical) have proved that MCS are effective in achieving improved classification performance for various application problems. Aircraft engine fault diagnosis plays a crucial rule in cost-effective operation of aircraft engines. By accurately detecting and reliably diagnosing impending engine faults, aircraft engine fault diagnosis can help to increase engine on-wing time, reduce maintenance turnaround time, reduce aircraft life-cycle costs, and increase flight safety. However, designing a reliable aircraft engine fault diagnostic system is a challenging task, due to a number of characteristics of aircraft engines. These characteristics include the wide range of flight regime that aircraft engines are operated over and that engines experience normal wear that needs to be differentiated from faults. Motivated by a goal of achieving the highest possible performance of fault diagnosis, we introduce MCS to aircraft engine fault diagnosis. By designing a real-world MCS-based aircraft fault diagnostic system, we demonstrate that MCS is effective in improving the performance of aircraft engine fault diagnostic systems.

Keywords: Aircraft engines; Classification; Classifier fusion; Diagnostics; Multiple classifier systems; Neural networks; Support vector machine

1. Introduction: Aircraft engine fault diagnosis (AEFD) is to detect, isolate, and assess malfunctions/faults and failures of engine system and its major components. AEFD is the enabler of the modern condition-based maintenance strategy for aircraft engines. The benefits of AEFD may include [1]: a) increasing flight safety by early detection of engine malfunctions; b) preventing costly component damage and/or catastrophic failure; c) reducing turnaround time by providing maintenance personnel with information on fault locations (by reducing time for manual fault isolation); d) reducing delays and cancellations by facilitating more on-wing maintenance; and e) increasing engine on-wing time by minimizing scheduled and unscheduled engine removal.

Reliably diagnosing engine faults is a difficult task due to the following intrinsic characteristics of aircraft engines:

a) There exist engine initial quality variations. Engine initial quality variation is the result of the variation of fabricating and assembling. Engine initial quality varies from engine to engine even within the same engine models.
b) Engine quality deteriorates over time. Engine deterioration is caused by many effects, such as, tip clearance changes in the rotating components, seal wear, blade fouling, blade

erosion, blade warping, foreign object damage, actuator wear, and blocked fuel nozzles [2]. Engine deterioration results in engine performance parameter changes over time (time-varying).

c) Aircraft engines are operated at different points in flight regime. When an aircraft travels from one point to another in flight regime, the engine performance parameters change following the principles of thermodynamics and aerodynamics.

The aforementioned importance and the challenges associated with aircraft engine fault diagnosis have motivated tremendous amount of research efforts. With advances in modern aircraft engines, designing a reliable and cost-effective AEFD system continues to be the most interesting research topic in fault diagnosis.

Researchers have applied a number of different techniques for engine fault diagnosis. These techniques range from the traditional gas path analysis (GPA) to modern soft computing technologies, such as, neural networks, fuzzy logic, and evolutional computation. Li's work [3] provides a good review of different techniques adopted for engine fault diagnosis. For complex classification problems like AEFD, it is generally accepted that single classifier design is often incapable in achieving the desired performance. For such complex problems, studies (both theoretical and empirical) have proved that multiple classifier systems (MCS), one of the most significant advances in pattern classification in recent years, are more effective in achieving improved classification performance. Multiple classifier systems, as the name implies, are a classification system that consists of an ensemble of classifiers. Outputs of these individual classifiers are combined in certain way to arrive at the final classification decision. MCS achieves higher classification performance by strategically utilizing the information provided by these individual classifiers. MCS has been widely used for various applications [4]. In the field of aircraft engine fault diagnosis, MCS has not been actively used for solving a real-world problem. We have seen a few studies on exploring the concept of using multiple models for diagnosis, e.g., [5-7]. In these studies, potentials of using information fusion to combine information from multiple sources (sensors, models, etc) are investigated. Ogaji and Singh [8] also proposed a nested neural system to dividing a complex problem into a number of smaller (and hence simpler) problems. It is our belief that MCS is a good technique for AEFD. In this paper, we apply MCS to aircraft engine fault diagnosis. Our objective here is to use a real AEFD as an example to illustrate design details of MCS and to demonstrate the effectiveness of MCS in improving classification performance of AEFD.

The rest of this paper is organized as follows. Section 2 gives the problem description about aircraft engine fault diagnosis. Section 3 describes design details of MCS for aircraft engine fault diagnosis. Classification results and discussions are given in Section 4. Section 5 concludes the paper.

2. Problem description: Aircraft engine fault diagnosis (AEFD) is to detect and isolate engine faults/malfunctions based on information such as sensed performance measurements and flight regime conditions. AEFD is a classical classification problem. Let's assume the sensed performance measurements and flight regime information at time t are represented by a vector $\mathbf{x}_t = [m_1, m_2, ... m_k]$. Also assume engine has $1+p$ possible conditions/health states (a normal condition plus p different faulty conditions), i.e., $\mathbf{C} = [C_1, C_2, ... C_{p+1}]$. AEFD is then to infer engine instantaneous condition by assigning $\mathbf{x}_t$ to one of the $1+p$ health states, $\mathbf{C}$. Mathematically, AEFD performs a nonlinear mapping from an input vector $\mathbf{x}_t$ to a condition space $\mathbf{C}$, i.e., $F: \mathbf{x}_t \in \Re^k \to \mathbf{C} \in \Re^{p+1}$.

Aircraft engines are a complex system and there are potentially a number of possible faults occurring in different sub-systems or components. In this paper, we limit our study to those faults related engine gas path (engine gas path faults). More specifically, we design an engine diagnostic system to detect and diagnose the following 6 engine gas path faults:

1. Fan fault (FAN) – Fan blade damage, typically occurring due to bird strikes or other foreign object damage (FOD) during takeoff.
2. Compressor fault (CMP) – Compressor blade damage, compressor contamination, or abnormal operation
3. High Pressure Turbine fault (HPT)– Typically a partial loss of one or more blades, most commonly during high power conditions.
4. Low Pressure Turbine fault (LPT) – Typically a partial loss of one or more blades, most commonly during high power conditions.
5. Customer Discharge Pressure fault (CDP) – Leakage in excess of the desired bleed level commanded by the aircraft and communicated to the Full Authority Digital Electronic Control (FADEC).
6. Variable Bleed Valve fault (VBV) – Variable bleed valve doors not closing according to FADEC issued command.

The 10 sensed performance measurements that are used as inputs to the diagnostic system are listed in Table 1. These are the standard engine performance measurements that are popularly used for engine gas path fault diagnosis. In addition, we construct 2 extra parameters based on the domain experts' experience. They are the ratio of compressor exit and inlet pressures and the ratio of fuel flow rate and compressor exit pressure. Including the two ratio parameters, we have a total of 12 variables that are used as inputs to the diagnostic system.

Table 1: 12 variables used as inputs for AEFD

No.	Variable Description
1	Fuel flow rate
2	Fan speed
3	Core speed
4	Compressor inlet pressure
5	Compressor Exit pressure
6	Fan tip exit pressure
7	Compressor inlet temperature
8	Compressor exit temperature
9	HP turbine exit temperature
10	LP turbine exit temperature
11	5/4
12	1/5

Engine performance measurements change as aircraft operating at different points in flight regime even though the engine condition is kept unchanged. Engine performance measurement changes as the result of flight regime mask/distort the changes induced by engine faults, thus make the engine fault diagnosis even more difficult. Addressing flight regime issue has always been an important part in designing engine fault diagnostic systems. In this paper, we take care of the flight regime issue by flight regime mapping as proposed by Yan [1]. It essentially compensates for flight regime induced changes, thus accentuates the engine condition related changes of engine performance measurements, by mapping the engine performance values from the actual flight regime to sea level static equivalent.

3. Design of MCS: Design of a successful MCS consists of two important parts: design of the individual classifiers that constitute the MCS and design of the fusion mechanism that integrates the outputs of the individual classifiers to arrive at the final classification decision [9]. Key to an effective MCS is the diversity of the individual classifiers. Strategies for boosting diversity include: 1) using different types of classifiers; 2) training individual classifiers with different data set (bagging and boosting); and 3) using different subsets of features. For AEFD concerned in this paper, we propose a 2-level MCS. At higher level is the multiple classifier fusion, where 3 different classifiers are designed and the outputs of the 3 classifiers are then fused to arrive at the final diagnosis decision. At lower level, each of the individual classifiers that originally have multiple target classes is decomposed to multiple binary classifiers. Figure 1 shows the overall architecture of the 2-level MCS. Design details of each of the 2 levels are given in the following two subsections.

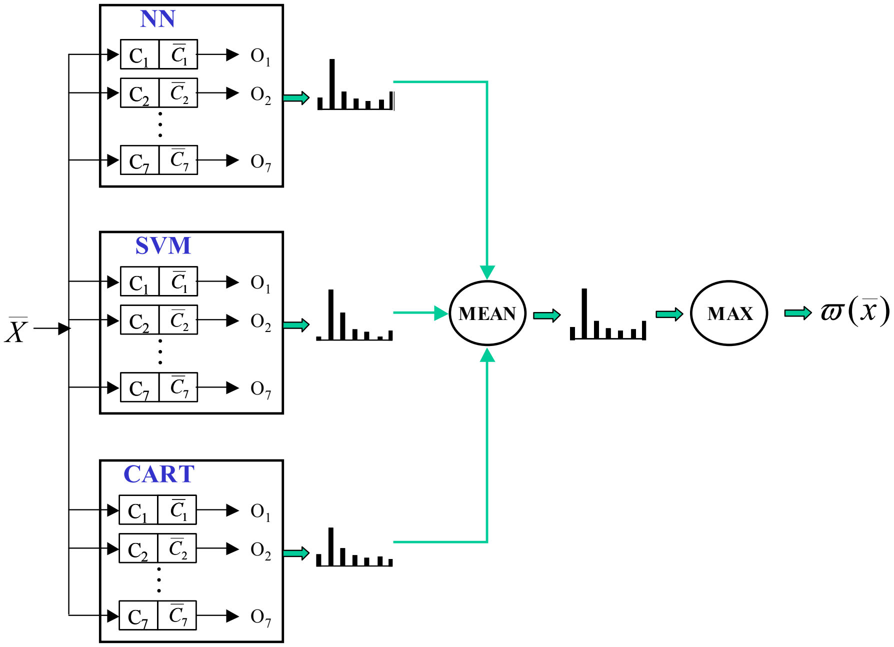


Figure 1: Overall structure of 2-level multiple classifier system

Multi-class decomposition: The AEFD system concerned in this paper involves diagnosing 7 different engine conditions (1 normal condition and 6 different types of faults). That is, the classifier to be designed has 7 different outputs, which is typically referred as a multi-class classification problem [10]. Multi-class classification can be handled directly, for example, for neural network classifier, we can structure network such that the number of output nodes equal to the number of classes. However, studies have proved that performance can be improved if the multi-class problem is decomposed into a series of binary classification ones [11]. There are several ways to decompose a multi-class classification problem into a series of binary classification ones in order to achieve better performance [12]. In this study, we take "one-vs-other" method since it is the simplest one. We decompose the AEFD classification problem that originally has 7 target classes into 7 binary classifiers. Each of the seven classifiers is trained to distinguish between patterns belonging to a class (C_i) and its complement ($\overline{C}_i$) (combining all data not belonging to class C_i). To classify an unknown input x, the outputs of the 7 binary classifiers form a 7-component vector, $O(x) = [O_1(x), O_2(x), ... O_7(x)]$ and the MAX-rule can be used to arrive at the final classification decision for the input x, i.e., $\varpi(x) = \arg\max_i O_i(x)$.

Multiple classifier fusion: The 3 individual classifiers used are the neural network (NN) classifier, the support vector machine (SVM) classifier, and the decision tree (CART) classifier. The 3 classifiers are of different types: NN classifier is weight-based, SVM classifier is distance or margin based, and CART is rule based. Using different types of classifiers as the constituent classifiers in classifier fusion is one of our design strategies in obtaining necessary diversity, thus achieving improved performance.

The neural network classifier is a 2-layer feed-forward network. It has one hidden layer with 25 hidden neurons and has one output nodes for binary classification. The activation functions for hidden and output nodes are logistic sigmoid function. The network is trained using the Levenberg-Marquadt learning algorithm for 500 epochs. To prevent saturation, the target values are scaled to 0.9 for positive cases and to 0.1 for negative cases. All other parameters are taken as the default values of MATLAB® Neural Network Toolbox.

Support Vector Machines (SVM) is a recently developed learning system originated from the statistical learning theory [13]. One distinction between SVM and many other learning systems is that its decision surface is an optimal hyperplane in a high dimensional feature space. The optimal hyperplane is defined as the one with the maximal margin of separation between positive and negative examples. Designing SVM classifiers includes selecting the proper kernel function and the corresponding kernel parameters and choosing proper C value. For aircraft engine fault diagnosis, we use the radial based function kernel for the SVM classifier. We set parameter γ that defines the spread of the radial function to be 5.0 and parameter C that defines the trade-off between the classifier accuracy and the margin (the generation) to be 3.0. SVM classifier design is performed using OSU SVM Classifier Matlab Toolbox (http://eewww.eng.ohio-state.edu/~maj/osu_svm/).

Decision trees are a popular non-parametric classifier that performs classification through a sequence of simple queries. Constructing a decision tree based on training data is a tree-growing process (also called tree induction process). CART (classification and regression tree) is one of generic tree-growing methods. Introduced by Breiman et al [14] on 1984, CART is a recursive and iterative procedure and has been considered as a powerful, fast, and flexible classification tool. To induce the decision tree classifier for the AEFD system, CART® for Windows by Salford Systems [15] is used. The Gini impurity is used as the splitting criterion. Pruning is performed based on the minimum cost-complexity principle, which results in a series of trees with different size. The final decision tree is selected based on the lowest estimated misclassification error under the test data.

While NN and CART classifiers output real numbers in the range of [0 1] (Note: these real numbers are not necessarily the posterior probabilities since they may not sum up to 1), outputs from SVM are not bounded to [0 1]. We scale SVM outputs using sigmoid function as follows [16]:

$$P(\varpi_i | x) = \frac{1}{1 + \exp(-\alpha \cdot o(x))} \tag{1}$$

where $O(\mathbf{x})$ is the output vector of SVM classifier for the input x and the coefficient α can be optimized using the data set. In this paper, we set α =1.0.

The outputs of the three individual classifiers are combined (fused) to arrive at the final classification decision. There are different methods for combining classifiers, ranging from simple majority voting to sophisticated Dempster-Shafer combination scheme. In this paper, we use the MEAN-rule to combine the classifiers due to its simplicity. Let $O(i,j)$ be the outputs (confidence value) of i^{th} classifier for j^{th} class. The MEAN-rule for N classifiers will give the probabilities of an input x belonging to each of the 7 classes as follows:

$$p_j(x) = \sum_{i=1}^{N} o(i,j) \cdot \tfrac{1}{N} \quad \text{for } j = 1,2,...7 \tag{2}$$

And the final class assignment is given by $\varpi(x) = \arg \underset{j}{MAX}(p_j(x))$.

4. Results:

The design data: The design data used in this paper is the simulated engine data. The simulation model is a component level model (CLM) of GEAE CFM56-7B engine used for commercial aircrafts. The simulated data contains all sensed engine performance parameters for both normal condition engines and engines with 6 gas path faults, totaling of 7 classes. Each class has 2805 data points, covering entire flight regime. The simulated data also includes engine-to-engine variation, sensor bias, and 5 levels of engine deterioration. Total number of data points is 19635.

Performance indices: Three performance indices (overall accuracy, false positive rate, and false negative rate) extracted from a confusion matrix are used for classifier performance comparison/evaluation. For multi-class classification, the three performance indices are defined as follows.

Let $CM(i,j)$, $\text{i, j} = 1,...\text{C}$ be the confusion matrix, where C is the number of classes. And assume Class 1 represents normal (fault-free) engine condition.

$$\text{Overall accuracy:} \quad OAC = \sum_{i=1}^{C} CM(i,i) \Big/ \sum_{i,j=1}^{C} CM(i,j) \tag{3}$$

$$\text{False positive rate:} \quad FPR = \sum_{j=2}^{C} CM(1,j) \Big/ \sum_{j=1}^{C} CM(1,j) \tag{4}$$

$$\text{False negative rate:} \quad FNR = \sum_{i=2}^{C} CM(i,1) \Big/ \sum_{i=2,j=1}^{C} CM(i,j) \tag{5}$$

The classifier is evaluated by 5-fold cross-validation [10]. Final confusion matrix for each of designs is the cell-by-cell sum of confusion matrices of the 5 validations.

Classification results: Confusion matrices for NN, SVM, and CART classifiers are shown in Tables 2 - 4, respectively. Table 5 shows the confusion matrix for fusion of the 3 individual classifiers. The 3 performance indices (OAC, FPR, FNR) are shown in Figure 2.

From Tables 2 - 5 and Figure 2, we can see that, individually, neural network classifier has the best performance in terms of the 3 performance indices while the CART classifier has the worst performance.

Classifier fusion yields classification performance that is better than that of the best individual classifier (NN classifier in this case). Comparing Table 4 to Table 1, fusion results in overall accuracy increase from 83.65% to 85.51%, while reducing both false positive rate from 23.74% to 21.21% and false negative rate from 5.19% to 4.63%. The overall improvement of fusion may seem not to be so significant. However, considering the fact that the diagnostic system we designed is model-free and the complexity of AEFD, we can see the effectiveness of fusion. Also, the fusion scheme we used is the simple mean-rule. Using more sophisticated fusion methods may lead to a greater improvement in classification performance, which will be investigated in our future work.

Table 2: Classification Results for NN classifier

		Predicted Classes							Per Class Accuracy	Performance Indices
		NF	**FAN**	**CMP**	**HPT**	**LPT**	**CDP**	**VBV**		
True Classes	**NF**	2139	7	106	150	197	70	136	76.26	
	FAN	0	2786	13	0	2	0	4	99.32	
	CMP	118	7	2454	62	94	57	13	87.49	**OAC=83.65**
	HPT	188	2	76	2210	291	16	22	78.79	**FPR=23.74**
	LPT	288	2	81	317	1893	177	47	67.49	**FNR=5.19**
	CDP	202	1	86	7	190	2292	27	81.71	
	VBV	77	1	4	8	37	28	2650	94.47	

Table 3: Classification Results for SVM Classifier

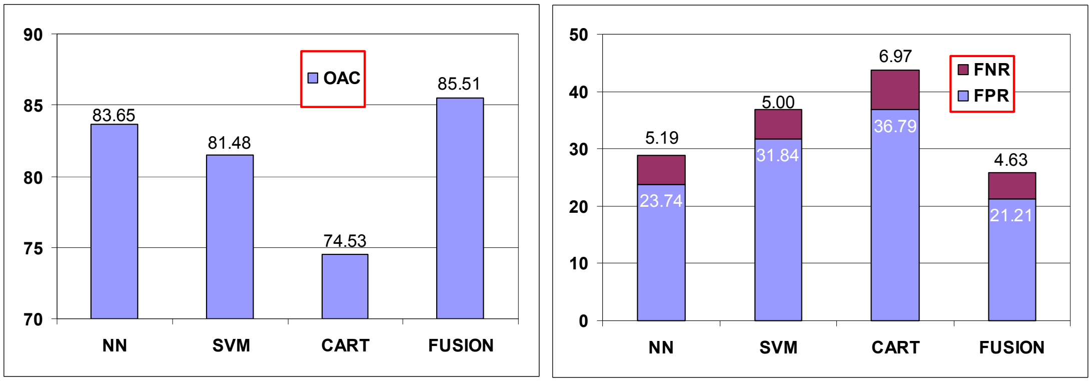
		Predicted Classes							Per Class Accuracy	Performance Indices
		NF	**FAN**	**CMP**	**HPT**	**LPT**	**CDP**	**VBV**		
True Classes	**NF**	1912	0	181	253	237	75	147	68.16	
	FAN	0	2803	2	0	0	0	0	99.93	
	CMP	98	1	2512	60	77	52	5	89.55	**OAC=81.48**
	HPT	183	0	107	2221	249	18	27	79.18	**FPR=31.84**
	LPT	302	0	115	547	1563	229	49	55.72	**FNR=5.00**
	CDP	197	0	105	4	184	2288	27	81.57	
	VBV	61	0	2	4	16	23	2699	96.22	

Table 4: Classification Results for CART Classifier

		Predicted Classes							Per Class Accuracy	Performance Indices
		NF	**FAN**	**CMP**	**HPT**	**LPT**	**CDP**	**VBV**		
True Classes	**NF**	1773	3	160	320	316	84	149	63.21	
	FAN	2	2797	5	1	0	0	0	99.71	
	CMP	148	2	2250	117	169	111	8	80.21	**OAC=74.53**
	HPT	319	0	181	1835	414	32	24	65.42	**FPR=36.79**
	LPT	366	0	182	449	1436	304	68	51.19	**FNR=6.97**
	CDP	187	0	106	35	302	2105	70	75.04	
	VBV	151	0	8	36	112	60	2438	86.92	

Table 5: Classification Results for Classifier Fusion

		Predicted Classes							Per Class Accuracy	Performance Indices
		NF	FAN	CMP	HPT	LPT	CDP	VBV		
True Classes	NF	2210	1	109	161	155	50	119	78.79	OAC=85.51 FPR=21.21 FNR=4.63
	FAN	0	2804	1	0	0	0	0	99.96	
	CMP	100	0	2522	48	77	52	6	89.91	
	HPT	165	0	83	2302	226	14	15	82.07	
	LPT	252	0	80	343	1916	174	40	68.31	
	CDP	196	0	78	6	148	2354	23	83.92	
	VBV	66	0	2	5	28	23	2681	95.58	



(a) **(b)**

Figure 2: (a) overall accuracies for all 3 classifiers and fusion; (b) False positive rates and false negative rates for all 3 classifiers and fusion.

5. Conclusions: Aircraft engine fault diagnosis is crucial in cost-effective condition-based maintenance in aircraft industry. Aircraft engine fault diagnosis is also a difficult classification problem due to the intrinsic characteristics of aircraft engines. The importance and the challenge draw tremendous amount of research interests on designing a reliable and accurate diagnostic system for aircraft engines. Motivated by the goal of achieving the highest possible performance of fault diagnosis, this paper introduces recently emerged MCS technique to aircraft engine fault diagnosis. By designing a 2-level MCS for a real world engine fault diagnostic system, we demonstrate that MCS can be an effective technique for obtaining improved classification performance for aircraft engine fault diagnosis.

REFERENCES:

[1] Yan, W.Z. (2003), “Diagnostic system configuration optimization and its application to aircraft engine fault diagnosis”, PhD Dissertation, Department of Mechanical Engineering, Rensselaer Polytechnic Institute

[2] Merrington, G., Kwon, O-K, Goodwin, G., and Carlsson, B. (1991), “Fault detection and diagnosis in gas turbines”, ASME Journal of Engineering for Gas Turbines and Power, Vol. 113, No. 4.

[3] Li, Y.G. (2002), “Performance-analysis-based gas turbine diagnostics: A review”, I. Mech.E. Proceedings of the Institute of Mechanical Engineering, Part A: Journal of Power and Energy, Vol. 216, No. A5.
[4] http://www.diee.unica.it/mcs/
[5] Goebel, K., Krok, M. and Sutherland, H. (2000), “Diagnostic information fusion: requirements flowdown and interface issues”, Proceedings of the IEEE 2000 Aerospace Conference – Advanced Reasoner and Information Fusion Techniques, vol. 6, pp.155-162.
[6] Volponi, A., Brotherton T., Luppold, R., and Simon, D.L., (2003), “Development of an information fusion system for engine diagnostics and health management”, NASA TM—2004-212924, ARL–TR–3127, *JANNAF Conference*, 2003.
[7] Roemer, M., Kacprzynski, G., and Schoeller, M. (2001), “Improved diagnostic and prognostic assessment using health management information fusion”, Proceedings of AUTOTESTCON, IEEE Systems Readiness Technology Conference, August 20-23, 2002, pp365-377.
[8] Ogaji, S. and Singh, R. (2003), “Gas path fault diagnosis framework for a three-shaft gas turbine”, Proceedings of the Institution of Mechanical Engineers, Vol. 217, Part A, pp. 149-157.
[9] Roli, F., Giacinto, G., and Vernazza, G. (2001), “Methods for Designing Multiple Classifier Systems”, *MCS 200, LNCS 2096, pp*78-87.
[10] Friedman, J. H. (1996), “Another approach to polychotomous classification”, Technical Report, Department of Statistics, Stanford University.
[11] Hastie, T. & Tibshirani, R. (1998), “Classification by pairwise coupling”, The Annals of Statistics, Vol. 26, Issue 2, pp451-471.
[12] Dietterich, T.G. & Bakiri, G. (1995), “Solving multiclass learning problems via error-correcting output codes”, Journal of Artificial Intelligence Research, 2, pp263-286.
[13] Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995
[14] Breiman, L, Friedman, J.H., Olshen, R.A., and Stone, C.J., (1984), *Classification and regression trees*, Chapman & Hall/CRC, Florida
[15] Steinberg, D. and Colla, P. (1997), “CART--Classification and Regression Trees”, San Diego, CA: Salford Systems, 1997
[16] Duin, R.P.W. and Tax, D. M. (1998), “Classifier conditional posterior probabilities”, Lecture notes in Computer Science, 1451:611-619.