TOPSIS Based Multi-Criteria Decision Making of Feature Selection Techniques for Network Traffic Dataset

Article *in* International Journal of Engineering and Technology · December 2013

3 authors, including:

Raman Singh
University of the West of Scotland

Ravinder Kumar Singla
Panjab University

TOPSIS Based Multi-Criteria Decision Making of Feature Selection Techniques for Network Traffic Dataset

Raman Singh [#1], Harish Kumar [#2], R.K. Singla [*3]

[#] UIET, Panjab University
Chandigarh (India)
[1] raman.singh@ieee.org
[2] harishk@pu.ac.in
[*] DCSA, Panjab University
Chandigarh (India)
[3] rksingla@pu.ac.in

***Abstract*— Intrusion detection systems (IDS) have to process millions of packets with many features, which delay the detection of anomalies. Sampling and feature selection may be used to reduce computation time and hence minimizing intrusion detection time. This paper aims to suggest some feature selection algorithm on the basis of The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). TOPSIS is used to suggest one or more choice(s) among some alternatives, having many attributes. Total ten feature selection techniques have been used for the analysis of KDD network dataset. Three classifiers namely Naïve Bayes, J48 and PART have been considered for this experiment using Weka data mining tool. Ranking of the techniques using TOPSIS have been calculated by using MATLAB as a tool. Out of these techniques Filtered Subset Evaluation has been found suitable for intrusion detection in terms of very less computational time with acceptable accuracy.**

Keyword- Feature selection, Multi criteria decision making, TOPSIS, Intrusion Detection System and Network Traffic Classification

I. INTRODUCTION

Computer network is growing day by day and on each passing moment billions of packets travel across any point on the Internet which is a large network of networks. These networks became backbone of economy and hence any attack on them may financially harm any company, organization or even countries. Misuse/Signature based Intrusion Detection Systems (IDS) may fail to detect zero day attack and hence networks are slowly moving towards anomaly based IDS. These systems need training by using traffic traces along with their features. In these systems accurate training is crucial as it learns normal behaviour of network so, traffic traces with good features are very important. After the training, IDS processes millions of packets with hundreds of features to detect the intrusions. Large number of feature requires more time to process this traffic. But detection of intrusion should be time bound to avoid any loss to network. Sampling process is used to reduce the size of training dataset used for IDS. Timely detection of intrusion can reduce losses due to attacks on the networks. In order to train IDS, training dataset consisting of network packets, are fed into this system. Hundreds of features of this dataset increases the overall detection time due to more computations. Feature selection may be used to reduce the feature set by maintaining accuracy in acceptable limits. There are many algorithms available for feature selection. Algorithms may behave differently for different types of dataset. So analysis is required to find out the suitable algorithm for IDS.

In this paper various features selection algorithms are compared on different parameters like accuracy, number of features, root mean square error (rms), true positive rate, false positive rate, precision, recall and receiver operating characteristics(roc) area. In certain circumstances it may become difficult to take decision on single parameter. For example, if one feature selection algorithm to be chosen out of a set of available techniques, then the decision cannot be taken only on the basis of single parameter like accuracy as it may increase computational time. The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) may be used in such situations where there are options to choose various feature selection algorithms along with different criteria like accuracy, number of features etc. Rest of the paper is organized as follows. Section I introduce topic, Section II describes the related work about feature selection and multi criteria decision making, section III describes experiment setup, methodology and dataset used, section IV discuss results and section V concludes the paper.

II. RELATED WORK

A. Feature Selection Techniques

Effectiveness of IDS is considerably depends on computational time required to process data for information extraction. Some features of packets may be redundant or non-productive. These may be discarded to reduce time taken by IDS to detect threats [1]. Feature selection plays crucial role in order to avoid over fitting and improve performance. However it introduces complexity and may result in lesser accuracy of machine learning algorithm [2]. There are many feature selection algorithms suggested by researchers and are widely used in many applications areas. Inter correlation among features is used to select features from different subsets in [3], [4]. In [5] correlation based feature subset to reduce the dimensionality and features of hand shape and palm texture is used. Chi-square test is performed to achieve discretization of continuous features in order to discard redundant and un-productive features [6]. In chi-square feature selection algorithm features are ranked individually by filter using squared statistics [7]. Researchers also used weighted chi-squared tests for dimension regression combined with sliced inverse regression which gives better performance [8].

Hybrid approach with wrapper and filtered feature selection algorithm is used in classification along with memtic framework [9]. A feature selection algorithm should always select strongly relevant features by discarding weakly relevant. Consistency with other features can be calculated in order to find out strongly and weakly relevant features [10]. Feature subset evaluation is also used to reduce high dimension from such dataset which has more number of features than data elements itself [11]. In order to find suitability of features, researcher also used correlation based heuristic and then calculates gain ratio to build decision tree to find out the final subset [12]. In [13] researcher suggests two stage technique for feature selection. In first stage each feature of text document is ranked on basis of information gain and at second stage genetic algorithm and principal component analysis is applied to reduce dimensions. Relief attribute evaluation is also used with ranker in order to select best features [14]. In [15] symmetrical uncertainty coefficient is used to select qualitative features. Irrelevant and redundant features can be discarded by decision rule based heuristic.

B. Multi-Criteria Decision Making

The feature selection techniques are general in nature and can be applied for various kinds of dataset. One technique may give the best result for one dataset while under-perform for another dataset. Multi criteria decision making technique like TOPSIS is used to select the Computer Integrated Manufacturing (CIM) technologies. There are various CIM alternatives available and based on some features TOPSIS may be used to select one technology as per weights given to each feature [16]. Multi-attribute decision making (MADM) concept is also applied on cellular manufacturing system (CMS). In this similar parts are processed on the same machines and then grouped as a cell [17]. Authors also proposed a new TOPSIS inspired multi criteria case based reasoning (CBR) for binary business failure prediction (BFP) [18]. In [19] authors integrate fuzzy logic, survey questionnaires and MADM methods to propose a new disaster assessment model. Attributes weights of disaster indices are calculated by Delphi method and various MADM methods like TOPSIS, Preference ranking organization method for enrichment evaluation (PROMETHEE), Analytic hierarchy process (AHP) and Grey relational analysis (GRA). Comparison of different companies can be done by using multi criteria analysis modal based on some attributes. Modified TOPSIS is used to identify the relevance of financial ratio and then performance of various companies is evaluated for each financial ratio [20].

Some real life problem like group decision can also be solved by MADM. Each individual make efforts of judgments, comparisons, and rankings. Then collective decision is taken as a group, depending on various attributes or parameters [21]. TOPSIS is integrated with fuzzy concept to select e-service provider on the basis of some at-tributes. This can be used by any of the organization which wants to outsource its e-activities and have number of alternatives [22]. Fuzzy analytic hierarchy process is used to determine the weighting of subjective judgments in problem of selection of host country. MADM integrated with fuzzy concept is used to take decision like weather to take expatriate assignment for given host country [23]. Vendor selection problem is solved by hybrid model of multi criteria decision making (MCDM). Five step hybrid processes is integrated with analytical network process (ANP) and then modified TOPSIS is applied to rank the various products as per their performances [24]. Authors also suggested modified TOPSIS for multi criteria decision making problem in case of non-deterministic data like interval data [25].

MCDM is also used in maritime transportation industry in order to select most appropriate registry alternative for shipping fleet. AHP is used to assign relative importance of various internal and external attributes like environmental factors. TOPSIS is applied to rank the various shipping registry alternatives [26]. Location planning essential for urban distribution centres to save distribution cost and minimizing traffic congestion. Under such type of uncertainty fuzzy based TOPSIS is used to select the proper location [27]. In manufacturing industry, various different materials are avail-able with various characteristics and so selection of one material is difficult. TOPSIS may be used to select a material among available alternatives [28].

III. EXPERIMENT SETUP AND DATA SET

In 1998, MIT Lincoln Laboratory with DARPA sponsorship prepared intrusion detection dataset to use for performance analysis of IDS. The aim is to provide researcher a benchmark dataset to develop new techniques against network security threats. Various network traffic like e-mail, telnet, web, was generated and captured to create dataset. telnet, web, was generated and captured to create dataset. Different scenarios like managers, programmers, who uses various network services were also created. Various attack types like denial of service attack, remote to local, user to local and surveillance was simulated by using various operating system like Solaris, SunOS and Linux. In 1999, labelled dataset was made available for researcher. This is known as KDD 99 dataset for intrusion detection [29]. Later some of inherent problems of this dataset are solved and a new version is made available. This is called NSL KDD [30] dataset for intrusion detection. It has 41 different features and one class label. There are other network datasets available like PU-CAN [31] but this dataset is used as a benchmark to evaluate performance of various IDS. In this experiment the goodness of feature selection techniques is measured on the basis of nine parameters shown in table I.

TABLE I
Various Parameters and Desired Values

Sr. No.	Parameter name	Desired Values
1	Accuracy	Maximum
2	Number of features	Minimum
3	Root mean square error	Minimum
4	True positive rate	Maximum
5	False positive rate	Minimum
6	Precision	Maximum
7	Recall	Maximum
8	F-Measure	Maximum
9	Receiver operating characteristic	Maximum

In ideal situation, some parameters like accuracy, true positive rate should have maximum values while others like number of features, error, should have minimum value. All parameters are considered equally important and unit weight is assigned to each of them. However in special cases, some parameters may have more importance than the others, so weight has to be adjusted accordingly. The various available feature selection techniques can be compared on the basis of these parameters, however it is very difficult to point out a single technique to be used for dataset. In this experiment, there are 10 numbers of alternatives (1 full feature + 9 feature selection techniques) with each have 9 different features.

Network traffic dataset is analysed with full features and then different feature selection techniques are applied as mentioned in Table-II.

TABLE II
Feature Selection Techniques (FST) Used for Experimentation

Feature Selection Technique	FST No.	Feature Selection Technique	FST No.
Full Features	1	Filtered Subset Eval	6
CFS Subset Eval	2	Gain Ratio Attribute Eval	7
Chi Sqaured Attribute Eval	3	Info Gain Attribute Eval	8
Consistency Subset Eval	4	One Ra Attribute Eval	9
Filtered Attribute Eval	5	Symmetrical Uncert Attribute Eval	10

In Table II, FST No. corresponds to Feature Selection Technique Number. 10 experiments (1 full features + 9 various techniques) are conducted using each classifier resulting in total number of 30 experiments. These experiments are conducted by using Weka data mining tool [32]. Weka, developed in Java contains various machine learning algorithms which can be used in pre-processing, classification, clustering and other data mining tasks. In the next step, TOPSIS is implemented in Matlab technical computing tool [33] which is used for numerical computation, visualization and programming. It is used on the output obtained from each classifier. TOPSIS gives confidence value between 0 and 1. Feature selection technique can be suggested based on confidence value. Higher confidence value means preferred technique. In this experiment, various techniques are ranked based on the confidence values obtained after using each classifier.

Figure 1 shows the methodology adopted for this experiment. For each feature selection technique three different classifiers (Naïve Bayes [34], J48 [35] and PART [36]) are used. Results are collected and detailed analysis is carried out.

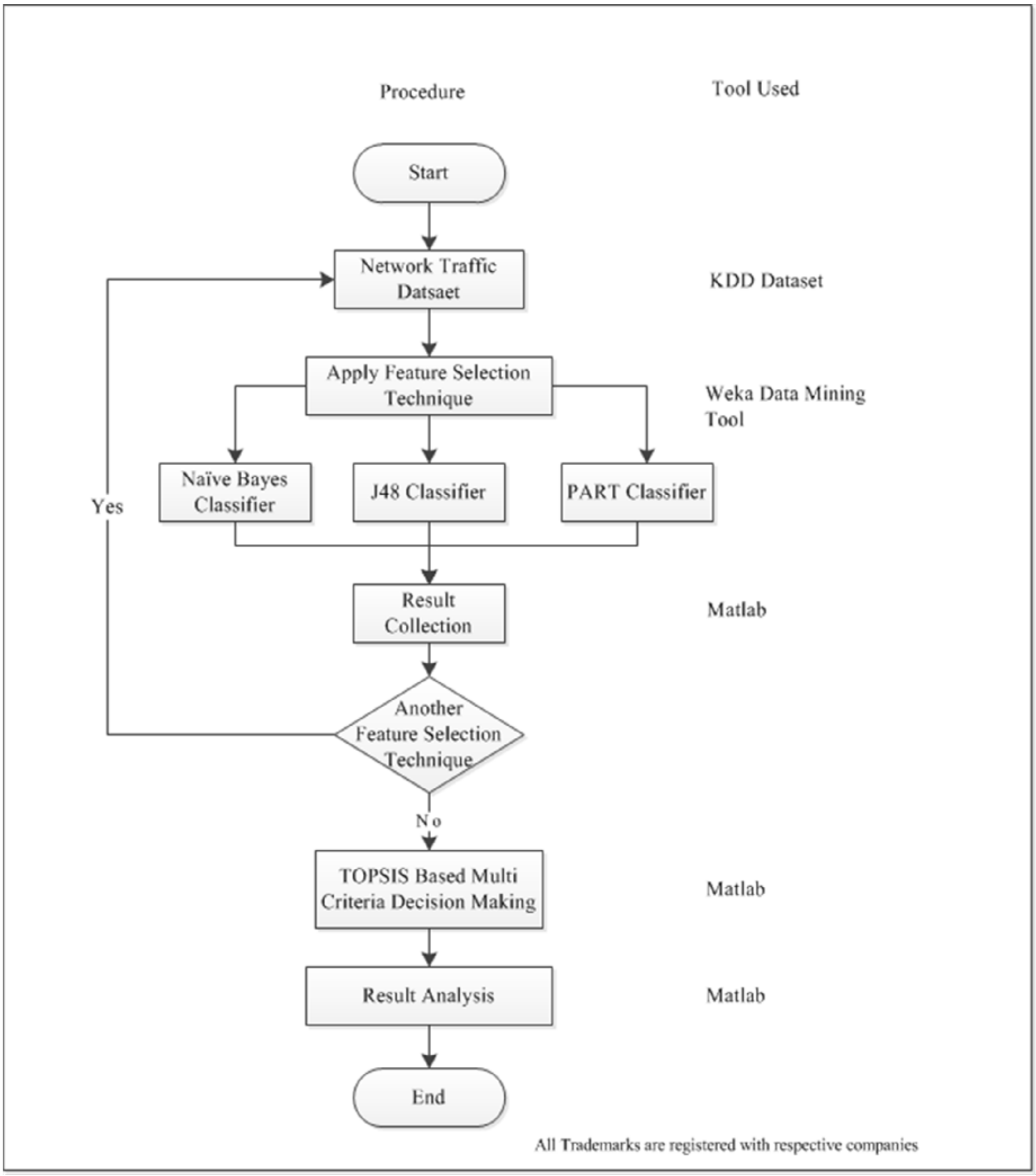

Fig. 1. Research Methodology used for experimentation

IV. RESULTS AND ANALYSIS

Performance of Naïve Bayes classifier is shown in table III. Values obtained of each feature selection techniques for each classifier after experiments are given in this table. These values are obtained for each parameters used in this study. Comparative analysis of various feature selection techniques with different classifier can be carried out on the basis of given parameters. Alternatives can be easily chosen on the basis of single parameters but as there are many parameters, choosing of alternative is difficult.

TABLE III
Performance of Naive Bayes Classifier

FST No.	Accuracy	No. of Features	RMS	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
1	93.565	41	0.0637	0.936	0.002	0.969	0.936	0.943	0.998
2	92.742	10	0.0647	0.927	0.001	0.965	0.927	0.937	0.998
3	93.209	30	0.0653	0.932	0.002	0.968	0.932	0.94	0.998
4	92.317	14	0.0692	0.923	0.002	0.968	0.923	0.934	0.998
5	93.492	30	0.0641	0.935	0.002	0.969	0.935	0.943	0.998
6	92.715	7	0.0648	0.927	0.001	0.964	0.927	0.936	0.998
7	89.037	30	0.0791	0.89	0.001	0.972	0.89	0.912	0.997
8	93.492	30	0.0641	0.935	0.002	0.969	0.935	0.943	0.998
9	93.492	30	0.0641	0.935	0.002	0.969	0.935	0.943	0.998
10	93.492	30	0.0641	0.935	0.002	0.969	0.935	0.943	0.998

Table IV shows the values of various parameters for each feature selection technique in case of J48 Classifier. Comparative analysis can be carried out from this table. There are various techniques which reduces number of feature while keeping accuracy comparable to full features dataset.

TABLE IV
Performance of J48 Classifier

FST No.	Accuracy	No. of Features	RMS	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
1	97.553	41	0.0332	0.976	0.003	0.975	0.976	0.975	0.998
2	97.036	10	0.036	0.97	0.002	0.972	0.97	0.969	0.998
3	97.55	30	0.0332	0.976	0.003	0.975	0.976	0.975	0.998
4	97.534	14	0.0332	0.975	0.003	0.976	0.975	0.976	0.998
5	97.552	30	0.0332	0.976	0.003	0.975	0.976	0.975	0.998
6	97.026	7	0.0362	0.97	0.002	0.972	0.97	0.969	0.998
7	97.478	30	0.0337	0.975	0.003	0.976	0.975	0.975	0.998
8	97.552	30	0.0332	0.976	0.003	0.975	0.976	0.975	0.998
9	97.552	30	0.0332	0.976	0.003	0.975	0.975	0.998	0.3112
10	97.552	30	0.0332	0.976	0.003	0.975	0.975	0.998	0.3112

Table V shows, values of different parameters obtained by PART classifier for each feature selection technique. These inputs are used by TOPSIS method to rank feature selection techniques.

TABLE V
Performance of PART Classifier

FST No.	Accuracy	No. of Features	RMS	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
1	97.552	41	0.0326	0.976	0.003	0.976	0.976	0.976	0.998
2	97.009	10	0.036	0.97	0.002	0.971	0.97	0.969	0.998
3	97.57	30	0.0324	0.976	0.003	0.976	0.976	0.976	0.998
4	97.574	14	0.0327	0.976	0.003	0.975	0.976	0.975	0.998
5	97.529	30	0.033	0.975	0.004	0.974	0.975	0.975	0.998
6	97.001	7	0.0359	0.97	0.002	0.971	0.97	0.969	0.998
7	97.514	30	0.0333	0.975	0.004	0.975	0.975	0.975	0.998
8	97.529	30	0.033	0.975	0.004	0.974	0.975	0.975	0.998
9	97.529	30	0.033	0.975	0.004	0.974	0.975	0.975	0.998
10	97.529	30	0.033	0.975	0.004	0.974	0.975	0.975	0.998

There are ten numbers of choices/alternatives with nine attributes for three classifiers. These values can be used to rank choices based on Multi-Criteria decision making methods. TOPSIS is implemented using MATLAB technical computing tool to obtain confidence value of each feature selection technique for each classifier separately. Table VI shows the confidence values of various feature selection techniques for each classifier.

TABLE VI
Confidence Level Based on TOPSIS Method

FST No.	Naive Bayes Classifier	J48 Classifier	PART Classifier
1	0.1538	0.8960	0.1958
2	0.9198	0.3112	0.8924
3	0.3120	0.6948	0.3639
4	0.6147	0.3112	0.7264
5	0.3165	0.9354	0.2866
6	0.9817	0.3097	0.9291
7	0.4419	0.3112	0.2856
8	0.3165	0.3112	0.2866
9	0.3165	0.3112	0.2866
10	0.3165	0.3112	0.2866

Based on these confidence values, ranks have been calculated for various techniques. Techniques with equal confidence values are assigned the same rank. Least value of rank indicates the most preferred technique which means Rank 1 is most preferred method. Rank calculated from confidence values are shown in Table VII. Final rank is obtained by taking average value of three different ranking values obtained using three classifiers.

TABLE VII
Final Ranking of Feature Selection Techniques

FST No.	Naive Bayes Classifier	J48 Classifier	PART Classifier	Final Rank
1	7	6	6	6
2	2	2	2	2
3	5	4	4	4
4	3	3	3	3
5	6	4	5	5
6	1	1	1	1
7	4	5	5	5
8	6	4	5	5
9	6	4	5	5
10	6	4	5	5

Feature Selection Technique no. 6 (i.e. Filtered Subset Eval) is ranked first because it obtains rank 1 for each classifier. All other techniques are assigned final ranks aggregating individual ranking of different classifier.

V. CONCLUSION

Choices can be made easily when there are only one attribute but selection among many available techniques sometimes become difficult job because they have multiple attributes. Effectiveness of feature selection techniques largely depends on the type of dataset. One technique may give better result on one type of dataset but may under-perform on other type of dataset. TOPSIS can be used to suggest one, among some available choices where each choice has various attributes. Some commonly available feature selection techniques are used on network traffic dataset and then TOPSIS is applied for ranking various techniques. From these experiments it can be concluded that Filtered subset Evaluation gives acceptable accuracy while keeping number of feature as low as seven. Consistency subset evaluation with only fourteen features gives even more accuracy than full featured dataset in case of PART classifier. However it takes slightly more computational time than Filtered subset Evaluation. According to this experiment Filtered subset Evaluation comes out to be the best feature selection technique for network traffic dataset. In future work some classifier may be suggested for this technique to increase the accuracy.

REFERENCES

[1] Isabelle Guyon, André Elisseeff, "An introduction to variable and feature selection", *The Journal of Machine Learning Research,* Vol. 3, pp 1157-1182, 2003.

[2] Walter Daelemans, Veronique Hoste, Fien De Meulder and Bart Naudts, "Combined optimization of feature selection and algorithm parameters in machine learning of language"*, in proc. 14th European Conference on Machine Learning,* Cavtat, Croatia September 22-26, 2003, pp 84-95.

[3] M.A.Hall, "Correlation – based feature subset selection for machine learning", *Ph.D. Dissertation, Department of Computer Science, University of Waikato, Hamilton, New Zealand* , 1999.

[4] Lei Yu and Huan Liu, "Feature selection for high-dimensional data : A fast correlation-based filter solution", *in proc. Twentieth International Conference on Machine Learning ICML-2003*, Washington DC, August 21-24, 2003, pp 856-863.
[5] Ajay Kumar, and David Zhang, "Personal Recognition Using Hand Shape and Texture", *IEEE Transactions On Image Processing*, Vol. 15, No. 8, pp 2454-2461 , 2006.
[6] H. Liu and R. Setiono, "Chi2: Feature selection and discretization of numeric attributes", *in proc. IEEE International Conference on Tools With Artificial Intelligence*, Herndon, Virginia, 5-8 November 1995, pp 388–391.
[7] Cantu-Paz, E., Newsam, S., and Kamath, C., "Feature selection in scientific applications", *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA: ACM Press, 2004.
[8] Zhishen Ye and JieYang, "Sliced inverse moment regression using weighted chi-squared tests for dimension reduction", *Journal of Statistical Planning and Inference*, Vol. 140, No. 11, pp 3121–3131, 2010.
[9] Zexuan Zhu, Yew-Soon Ong, and Manoranjan Dash, "Wrapper–Filter Feature Selection Algorithm Using a Memetic Framework", *IEEE Transactions On Systems, Man, And Cy-bernetics—Part B: Cybernetics*, Vol. 37, No. 1, pp 70-76 , 2007.
[10] Roberto Ruiza, José C. Riquelme, Jesús S. and Aguilar-Ruiz, "Incremental wrapper-based gene selection from microarray data for cancer classification", *Pattern Recognition*, Vol. 39, pp 2383 – 2392, 2006.
[11] Julia Handl and Joshua Knowles, "Feature Subset Selection in Unsupervised Learning via Multiobjective Optimization", *International Journal of Computational Intelligence Research*, Vol. 2, No. 3, pp 217–238, 2006.
[12] M.A. Hall and L.A. Smith, "Practical feature subset selection for machine learning", *in proc. 21st Australian Computer Science Conference*, pp 181–191, 1998.
[13] Harun Uguz, "A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm", *Knowledge-Based Systems*, Vol. 24, No. 7, pp 1024–1032 , 2011.
[14] Witten,I.H. and Frank,E., "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, San Francisco , 2005.
[15] Patricia E.N. Lutu and Andries P. Engelbrecht, "A decision rule-based method for feature selection in predictive data mining", *Expert Systems with Applications*, Vol. 37, No. 1, pp 602–609, 2010.
[16] Yusuf Tansel, "An experimental design approach using TOPSIS method for the selection of computer-integrated manufacturing technologies", *Robotics and Computer-Integrated Manufacturing*, Vol. 28, No. 2, pp 245–256, 2012.
[17] Arshia Ahi, Mir.B. Aryanezhad, Behzad Ashtiani and Ahmad Makui, "A novel approach to determine cell formation, intracellular machine layout and cell layout in the CMS problem based on TOPSIS method", *Computers & Operations Research*, Vol. 36, No. 5, pp 1478 – 1496, 2009.
[18] Hui Li, Hojjat Adeli, Jie Sun and Jian-Guang Han, "Hybridizing principles of TOPSIS with case-based reasoning for business failure prediction", *Computers & Operations Research*, Vol. 38, No. 2, pp 409–419, 2011.
[19] Gang Kou, Daji Ergu and Yu Shi, "An integrated expert system for fast disaster assessment", Computers & Operations Research, In press – corrected proof , 2012.
[20] Hepu Deng, Chung-Hsing Yeh and Robert J. Willis, "Inter-company comparison using modified TOPSIS with objective weights", *Computers & Operations Research*, Vol. 27, No. 10, pp 963-973, 2000.
[21] Hsu-Shih Shih, "A Multiattribute GDSS for Aiding Problem-Solving", *Mathematical and Computer Modeling*, Vol. 39, No. 11-12, pp 1397-1412 , 2004.
[22] Cengiz Kahraman, Nüfer Yasin Ates, Sezi Çevik and Murat Gülbay, "Fuzzy Multi-Attribute Cost–Benefit Analysis of E-Services", *International journal of intelligent systems*, Vol. 22, No. 5, pp 547–565, 2007.
[23] Mei-Fang Chen and Gwo-Hshiung Tzeng, "Combining Grey Relation and TOPSIS Concepts for Selecting an Expatriate Host Country", *Mathematical and Computer Modeling* Vol. 40, No. 13, pp 1473-1490, 2004.
[24] Huan-Jyh Shyur and Hsu-Shih Shih, "A hybrid MCDM model for strategic vendor selection", *Mathematical and Computer Modeling*, Vol. 44, No. 7-8, pp 749–761, 2006.
[25] G.R. Jahanshahloo, F. Hosseinzadeh Lotfi and A.R. Davoodi, "Extension of TOPSIS for decision-making problems with interval data: Interval efficiency", *Mathematical and Computer Modeling*, Vol. 49, No. 5-6, pp 1137-1142, 2009.
[26] Ahmet Kandakoglua, Metin Celik and Ilker Akgun, "A multi-methodological approach for shipping registry selection in maritime transportation industry", *Mathematical and Computer Modeling*, Vol. 49, No. 3-4, pp 586-597, 2009.
[27] Anjali Awasthi, S.S. Chauhan and S.K. Goyal, "A multi-criteria decision making approach for location planning for urban distribution centers under uncertainty", *Mathematical and Computer Modeling*, Vol. 53, No. 1-2, pp 98–109, 2011.
[28] R. V. Rao and J. P. Davim, "A decision-making framework model for material selection using a combined multiple attribute decision-making method", *International journal of advanced manufacturing technology*, Vol. 35, No. 7-8, pp 751–760, 2008.
[29] Lippmann Richard P. , Fried David J. , Graf Isaac, Haines Joshua W. , Kendall Kristopher R., McClung David, Weber Dan, Webster Seth E. , Wyschogrod Dan, Cunningham Robert K., and Zissman Marc A., "Evaluating Intrusion Detection Systems: The 1998 DARPA Off-line Intrusion Detection Evaluation", *in proc. DARPA Information Survivability Conference and Exposition*, Hilton Head, South Carolina, January 25-27, 2000, pp 12 – 26, 2000.
[30] (1999)The NSL-KDD Data Set,[online].Available:http://nsl.cs.unb.ca/NSL-KDD/
[31] Raman Singh, Harish Kumar and R.K. Singla, "Traffic Analysis of Campus Network for Classification of Broadcast Data", *in proc. 47th Annual National Convention of Computer Society of India, International Conference on Intelligent Infrastructure*, Science City, Kolkata, December 1-2, 2012, pp 163-166, 2012.
[32] (2013) Weka: Data Mining software, [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka
[33] (2012) MATLAB and Simulink for Technical Computing, [Online]. Available: http://www.mathworks.in
[34] George H. John and Pat Langley, "Estimating Continuous Distributions in Bayesian Classifiers", *in proc. Eleventh Conference on Uncertainty in Artificial Intelligence*, Montreal QU, Canada, Aug 18-20 1995, pp. 338-345, 1995.
[35] Quinlan J. R., "*C4.5: Programs for Machine Learning*", Morgan Kaufmann Publishers, San Francisco, CA, USA , 1993.
[36] Eibe Frank and Ian H. Witten, "Generating Accurate Rule Sets Without Global Optimization", *in proc. ICML 98 : Fifteenth International Conference on Machine Learning, Madison*, Wisconsin, USA , July 24-27, 1998, pp 144-151 , 1998.