Volume 3, Issue 12, December 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Bipartite RankBoost Approach for Score Level Fusion of Face and Palmprint Biometrics

B.Sateesh Kumar*
Assistant Professor, CSE Dept
JNTUH CEJ
JNT University Hyderabad
AP, INDIA

Dr.A.Govardhan
Professor of CSE
School of Information Technology
JNT University Hyderabad
AP, INDIA

Abstract: Boosting is a learning method used extensively in the field of machine learning, that finds a way to combine "weak" classifiers and build up a "strong" classifier. One approach to boosting based score level fusion is to exploit classifiers for finding the best decision boundary between genuine and imposter instances in biometric systems. This paper explores the idea of using RankBoost and its bipartite version which is designed to minimize misranking error explicitly in multimodal biometrics of Face and Palmprint. Multimodal biometric system utilizes two or more individual modalities, to improve the recognition accuracy of conventional unimodal methods. Face and palmprint biometrics are being increasingly used as alternatives to passwords, PINs and visual verification. Palmprint is one of the most stable biometric characteristics and its implementation for security purpose over the last decade has been increased all over the world.

Keywords: Multimodal Biometrics, AUC, Boosting, Adaboost, RankBoost, Score Level Fusion.

I. INTRODUCTION

There is increasing need in the development of reliable, rapid, efficient and non-intrusive security control systems. A wide variety of systems requires reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services [15].Amongst these systems is the biometric system, which recognizes the individual based on their physical characteristics or the behavioural characteristics.

Identity authentication is a general task that has many real-life applications such as access control, transaction authentication (in telephone banking or remote credit card purchases for instance), voice mail, or secure teleworking. The goal of an automatic identity authentication system is to either accept or reject the identity claim made by a given person. Biometric identity authentication systems are based on the characteristics of a person, such as its face, Fingerprint, palm or signature. An identity authentication system has to deal with two kinds of events: either the person claiming a given identity is the one who he claims to be (in which case, he is called a *client*), or he is not (in which case, he is called an *impostor*). Moreover, the system may generally take two decisions: either *accept* the client or *reject* him and decide he is an *impostor*.

Nowadays, as a result of advancement in computation power and storage capacity of computers, fusion of many source of information became more accessible. Also verification systems are influenced by this progression and tend to use multimodal biometrics instead of unimodal biometrics. Multimodal systems acquire information from more than one source. Unibiometric identifiers use single source biometric evidence and often are affected by problems like lack of invariant representation, non-universality, noisy sensor data and lack of individuality of the biometric trait and susceptibility to circumvention. These problems can be minimized by using multibiometric systems that consolidate evidences obtained from multiple biometric sources. Multimodal biometric fusion combines measurements from different biometric traits to enhance the strengths and diminish the weaknesses of the individual measurements[14]. Multimodal Biometrics with various levels of fusion: sensor level, feature level, matching score level and decision level.

The performance of face authentication systems has steadily improved over the last few years. State-of-the-art methods use the projection of the gray-scale face image into a Linear Discriminant subspace as input of a classifier such as Support Vector Machines or Multi-layer Perceptrons. Unfortunately, these classifiers involve thousands of parameters that are difficult to store on a smart-card for instance. Recently, boosting algorithms has emerged to boost the performance of simple (weak) classifiers by combining them iteratively. The famous AdaBoost algorithm have been proposed for object detection and applied successfully to face detection.

In the physical characteristics of an individual, palmprint has the most stable and unique characteristics and is rich in the feature information. [16] Even two identical twins have different palmprints. The features of a palm are palm geometry, flexion creases (principle lines), secondary creases (wrinkles), ridge and valley features, delta features, minutiae features, singularity point and texture. By combining all these features, it is possible to build a highly accurate biometrics system [17].

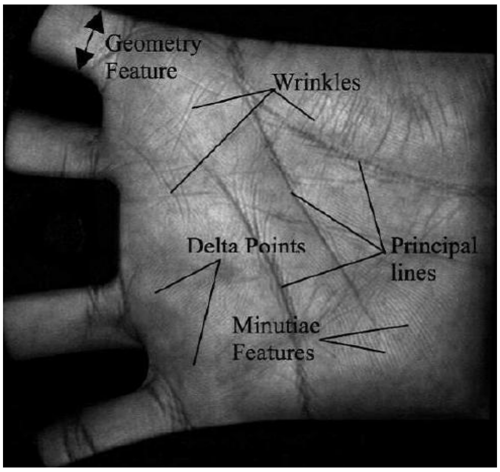

Figure 1: Different features of the palm

Usually acquisition of palmprint is done with the help of different types of cameras and scanners. Contactless image acquisition can be done with the help of digital cameras and video cameras. But quality of such images may not be good as it is contactless.

The score level fusion is the best in sense of simplicity and amount of information which supposed to be combined. Generally, there are three approaches to score fusion: 1) transformation based score fusion, 2) density based score fusion, 3) classifier based score fusion. Transformation based methods usually are applied after score normalization step. Sum rule, Product rule, Min rule and Max rule belong to this category, amongst them, Sum rule shows the best experimental results [1]. Density based score fusion methods are based on score distribution estimation. Well-known density estimation models like Naive Bayesian [2] and Gaussian Mixture Model (GMM) [3] have been used for fusion. Classifier based score fusion treat scores as features and try to find the best decision boundary like the case of binary classification problem [4, 5]. For instance, in [6] Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) have been used for classifier based score fusion. Multivariate polynomials of hyperbolic functions [7] and its combination with GradientBoost [8] have been applied for score level fusion.

Besides this taxonomy, some algorithms are introduced which try to minimize ranking error and therefore improving Receiver Operating Characteristic (ROC) curve, which is based on maximizing Area under the Curve (AUC) of ROC. In [9], Toh et al. Developed least square error based framework to do this, and examined their algorithm for score level fusion. Optimizing AUC in kernel based model was presented in [10], and Freund et al. introduced RankBoost [11] for this purpose. Continuing development of RankBoost, Rudin et al. [12] introduced margin based and coordinate descent of RankBoost. Also, they have proved that AdaBoost not only minimize classification error, but also under certain condition, can optimize AUC. Latest discoveries on AdaBoost capability is inspiring to exploit AdaBoost as a credible algorithm for score fusion. As a consequence, in this paper, boosting based method is used not only as a classifier, but also as an algorithm to optimize AUC in multimodal biometrics.

A. AUC of ROC curve

In many cases, in a binary classification problem cost of errors for two classes are not equal, thus, global measures like minimum total error rate does not reflect the real error. As a consequence, changing decision threshold to satisfy imposed cost is inevitable. ROC is a very comprehensive measure, to show how classifiers could deal with this problem. In addition, ROC curve helps to validate classifier performance for undetermined threshold and variable error cost situations [13]. ROC can be understood as a plot of True acceptance rate (TAR) versus false acceptance rate (FAR). In fact, for each possible value of decision threshold, it shows a pair of TAR and FAR values, thus, ROC curve can be determined completely by varying the decision threshold.

To reduce ROC curve to a single scalar value, AUC is used as a measure which is a performance metric that is invariant to unequal error cost and unbalanced class sample size. For example, AUC of base classifier is 0.5 and that of ideal classifier is 1 independent of inequality in error cost and sample size between two classes. AUC can be calculated as follows:

$$\mathrm{AUC} = \frac{\sum_{X0}\sum_{X1} I(h(X1)>h(X0))}{|X0| \quad |X1|} \qquad (1)$$

where |X0| and |X1| denote the number of instances for each class in binary classification problem, and I(u) denotes the indicator function. According to the Equation 1, it can be inferred that misranking error is an affine transform of AUC [12] and therefore minimizing misranking error will increase AUC. Hence minimizing misranking error is the key concept of AUC optimization and until now, wide variety of methods have been developed to train classifiers in order to optimize AUC. This paper focuses mainly on boosting based algorithms
which optimize AUC.

II. BOOSTING METHOD

The Boosting method was first proposed in 1989 by Freund and Schapire which uses series of training data, with weights assigned to each training set. Series of classifiers are defined so that each of them is tested sequentially comparing the result of the previous classifier and using the results of previous classification to concentrate more on misclassified data[18]. All the classifiers used are voted according to accuracy. Final classifier, combines weight of the votes of each classifier from the test sequence. The Boosting algorithm is as follows

Boosting Algorithm for Classification

Input: $Z = \{z_1, z_2, \ldots, z_N\}$, with $z_i = (x_i, y_i)$ as training set.

Output: H(x), a classifier suited for the training set.

1. Randomly select, *without* replacement, $L_1 < N$ samples from Z to obtain Z_1; train weak learner H_1 on it.
2. Select $L_2 < N$ samples from Z with half of the samples misclassified by H_1 to obtain Z_2; train weak learner H_2 on it.
3. Select all samples from Z that H_1 and H_2 disagree on; train weak learner H_3.
4. Produce final classifier as a vote of weak learners $H(x) = \text{sign}\left(\sum_{n=1}^{3} \text{H}n(x) \right)$

Essential Boosting idea is combining together basic rules, creating an ensemble of rules with better overall performance than the individual performances of the ensemble components. Each rule can be treated as a hypothesis, a classifier. Moreover, each rule is weighted so that it is appreciated according to its performance and accuracy. Weighting coefficients are obtained during the boosting procedure which, therefore, involves learning. Mathematical roots of Boosting originate from probably approximately correct learning (PAC learning).

A. AdaBoost

The same researchers that proposed the Boosting algorithm, Freund and Schapire, also proposed in 1996, the AdaBoost (Adaptive Boosting) algorithm. The idea behind adaptive boosting is to weight the data instead of (randomly) sampling it and discarding it. The AdaBoost algorithm is a well-known method to build ensembles of classifiers with very good performance. It has been shown empirically that AdaBoost with decision trees has excellent performance , being considered the best off-the-shelf classification algorithm. This algorithm takes training data and defines weak classifier functions for each sample of training data. Classifier function takes the sample as argument and produces value -1 or 1 in case of a binary classification task and a constant value - weight factor for each classifier. Procedure trains the classifiers by giving higher weights to those training sets that were misclassified. Every classification stage contributes with its weight coefficients, making a collection of stage classifiers whose linear combination defines the final classifier[18]. Each training pattern receives a weight that determines its probability of being selected as a training set for an individual component. Inaccurately classified patterns are likely to be used again. The idea of accumulating weak classifiers means adding them so that each time the adding is done, they get multiplied with new weighting factors, according to distribution and relating to the accuracy of classification. *Discrete AdaBoost* or just *AdaBoost* was the first one that could change weak learners. Generally, AdaBoost has shown good performance at classification. The disadvantage of Adaptive Boosting is its sensitivity to *noisy data* and *outliers*. Boosting has a feature of reducing variance and bias, and a major cause of boosting success is variance reduction.

AdaBoost Algorithm for Classification

Input: $Z = \{z_1, z_2, \ldots, z_N\}$, with $z_i = (x_i, y_i)$ as training set . M, the maximum number of classifiers.

Output: H(x), a classifier suited for the training set.

1.Initialize the weights $w_i = 1/N$, $i \in \{1, \ldots, N\}$.

2.For m=1 to M

a) Fit a classifier $H_m(x)$ to the training data using weights w_i.

b) Let $err_m = \frac{\sum_{i=1}^{N} w_i I \quad (y_i \neq \text{H}_m(x_i))_1}{\sum_{i=1}^{N} w_i}$

c) Compute $\alpha_m = 0.5 \log(\frac{1-err_m}{err_m})$

d) Set $w_i \leftarrow w_i \exp(-\alpha_m I(yi \neq H_m(x_i)))$ and renormalize to $\sum_i w_i = 1$.

3. Output $H(x) = \text{sign}\left(\sum_{m=1}^{m} \alpha_m \text{H}_m(x) \right)$.

The function I(c) used in steps 2.b and 2.d is an indicator function I(c) = 1, if c ="true" and I(c) = 0 if c ="false".
The algorithm stops when m = M or if $err_m > 0.5$; this last condition means that it is impossible to build a better ensemble using these weak classifiers, regardless of the increase of their number. This way, M represents the maximum number of classifiers to accommodate in the learning ensemble.

Table I: Comparison Of Boosting And AdaBoost Algorithms

Feature	Boosting	Adaptive Boosting
Data Processing	Random Sampling without replacement	Weighting (No Sampling)
No. of Classifiers	Three	Upto M
Decision	Majority Vote	Weighted Vote

B. Boosting Based Score Fusion Scheme

In Score Level Fusion, the matching scores of each subsystem are combined using techniques such as Weighted Sum rule, Weighted Product, Linear Discriminant, Decision Tree and the Bayesian Rule to find composite matching score which is then sent to the decision module. Currently, this appears to be the most useful fusion level because of its good performance and simplicity. This fusion level can be divided into two categories: combination and classification. In the former approach, a scalar fused score is obtained by normalising the input matching scores into the same range and then combining such normalised scores. In the latter approach, the input matching scores are considered as input features for a second level pattern classification problem between the two classes of client and the Impostor [14].

One approach to score level fusion is to exploit classifiers for finding the best decision boundary between genuine and imposter instances. Let $x(n) = [x_1^{(n)}, x_2^{(n)}, \ldots, x_N^{(n)}]$ denote the scores of N different biometric matchers for n'th instance, and $y^{(n)} \subseteq -1; 1$ denotes its corresponding label. Note that −1 and 1 refer to imposter and genuine classes, respectively. The goal is to find a decision function like $H : R^N \rightarrow -1; 1$ to recognize labels of unseen instances. Through boosting based algorithms, a function is sought that minimizes classification error.

C. Bipartite RankBoost

Freund et al. introduced RankBoost in detail and reported the experimental results of applying this algorithm for meta-searching and movie recommendation problems [11]. Similar to AdaBoost, this algorithm attempts to find a combination of "weak rankers" to create highly accurate single ranker. Furthermore, they described efficient implementation version of RankBoost, i.e., RankBoost.B, for bipartite feedback. Broadly speaking, the term of bipartite feedback is used when there are two sets of instances and the problem is to rank all instances of one set over another set. Since minimizing ranking error implies optimizing AUC of ROC curve and due to similarity of bipartite feedback and binary classification problems, which are cost-sensitive and has unbalanced classes, this special case of Rankboost algorithm can be applied to such classification problems in which optimization should be done over wide range of decision thresholds.

Algorithm RankBoost.B

Given: disjoint subset X0 and X1 of X.

Initialize:

$$\upsilon_1(x) = \begin{cases} 1 / |X_1| & \text{if } x \in X_1 \\ 1 / |X_0| & \text{if } x \in X_0 \end{cases}$$

For t=1, ..., T :

- Train weak learner using distribution
$$D_t = \upsilon_t(X_0)\, \upsilon_t(X_1)$$
- Get weak ranking $h_t : X \in R$
- Choose $\alpha_t \in R$
- Update:

$$\upsilon_{t+1}(x) = \begin{cases} \dfrac{\upsilon_t(x)\exp(-\alpha_t(x))}{Z_t^1} & \text{if } x \in X_1 \\ \dfrac{\upsilon_t(x)\exp(\alpha_t(x))}{Z_t^0} & \text{if } x \in X_0 \end{cases}$$

Where Z_t^1 and Z_t^0 normalize υ_t over X1 and X0:

$$Z_t^0 = \sum_{x \in X0} \upsilon_t(x)\exp(-\alpha_t h_t(x))$$

$$Z_t^1 = \sum_{x \in X1} \upsilon_t(x)\exp(\alpha_t h_t(x))$$

Output the final ranking: $H(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$

Figure 2.Pseudocode of RankBoost for bipartite feedback.

The pseudo code for Rankboost.B is shown in Figure 2. As it can be inferred from this figure, despite similarity of AdaBoost and RankBoost.B, there are some discrepancies between them. Although final output in two cases is a linear combination of weak learners, training of weak learners and computing of α_t's are accomplished in different ways.

III. PERFORMANCE EVALUATION

To evaluate the performance of the system a database containing palm and face samples was required. We collected the hand and face database ourselves using a scanners. The spatial resolution of the hand images is 180 dots per inch (dpi) / 256 grey levels. As the hand and the face databases contain samples belonging to different people, a "chimerical" multimodal database was created using pairs of artificially matched palm and face samples that were made for testing purposes. The database was divided into two sets: the training set and the testing set. The training set consisted of 200 image pairs of 50 people (4 image pairs per person) and was used as a training database for individual modalities, to get the weightings associated with different modalities. The testing dataset consisted of 240 image pairs of 60 people (8 image pairs per person) and was used exclusively for the evaluation of the system performance. Out of 8 image pairs for each person, 5 were used in the enrolment stage and 3 were used for testing. The tests involved trying to verify every test pair for every one of the 60 people enrolled in the database. The results of the experiments, expressed in the terms of FRR (false rejection rate) and FAR (false acceptance rate), vary depending on the selected verification threshold T (Figure 3).

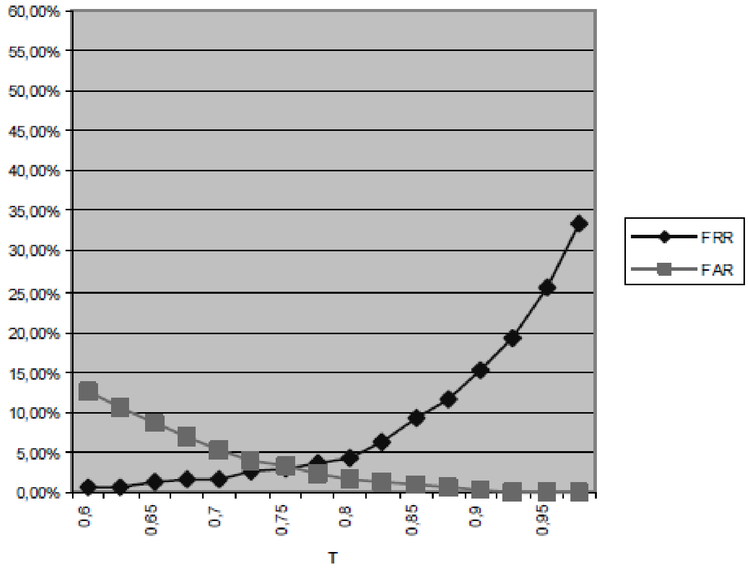

Figure 3: The verification results using the multimodal system depending on threshold

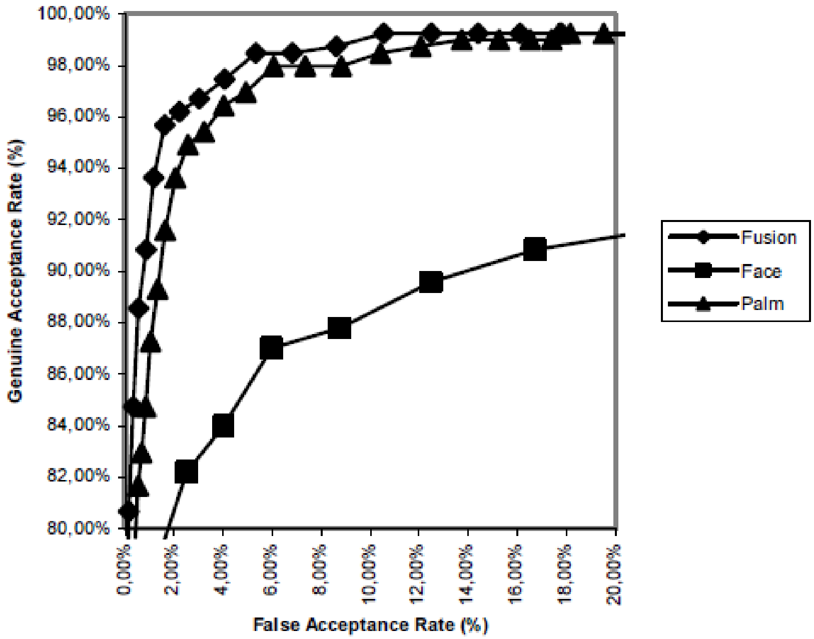

Figure 4: Comparison of unimodal and multimodal system verification results

Our multimodal system can achieve an EER (equal error rate) of 3.08% for T =0.748 and the minimum TER (total error rate) = 5.94% for T = 0.8. The comparison of both unimodal systems (palm and face modality) and a multimodal system is given in Figure 4. From the results it is clear that the verification based on the palmprint easily outperforms the verification based on the face. It can also be seen that the fusion of palmprint and facial features improves the verification score.

IV. CONCLUSIONS

In this paper, we applied Bipartite RankBoost in multimodal biometrics score level fusion with understanding its ability in AUC optimization. We reformulated weak learner used in Bipartite RankBoost for applications in which scores are meaningful. We explained that weighting instances inversely proportional to the number of instances in the corresponding classes Based on the results of our experiments we conclude that Classifier approach outperforms compared to transformation based score fusion and density based score fusion and AdaBoost achieves the same level of performance compared to Bipartite RankBoost.

REFERENCES

[1] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers, IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 3, pp. 226-239, 1998.

[2] N. Poh and S. Bengio, Eer of fixed and trainable fusion classifiers: A theoretical study with application to biometric authentication tasks, Proc. of Conference on Lecture Notes in Computer Science, pp. 74-85, 2005.

[3] K. Nandakumar, Y. Chen, S. C. Dass, and A. K. Jain, Likelihood ratio-based biometric score fusion, IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 2, pp. 342-347, 2008.

[4] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and J. Bigun, Discriminative multimodal biometric authentication based on quality measures, Pattern Recognition, vol. 38, no. 5, pp. 777-779, 2005.

[5] Y. Ma, B. Cukic, and H. Singh, A classification approach to multi-biometric score fusion. Proc. Of Conference on AVBPA , pp. 484-493, 2005.

[6] S. Bengio, C. Marcel, S. Marcel, and J. Mariethoz, Confidence measures for multimodal identity verification, Information Fusion, vol. 3, no. 4, pp. 267-276, 2002.

[7] K. A. Toh, and W. Y. Yau, Combination of hyperbolic functions for multimodal biometrics data fusion, IEEE Trans. Systems Man and Cybernetics, vol. 34, no. 2, pp. 1196-1209, 2004.

[8] M. Parviz, and M. S. Moin, Multivariate polynomials estimation based on gradientboost in multimodal biometrics, Proc. of Conference on ICIC, vol. 15, no. 3, pp. 471-477, 2008.

[9] K. A. Toh, J. Kim, and S. Lee. Maximizing area under roc curve for biometric scores fusion, Pattern Recognition, vol. 41, no. 11, pp. 3373-3392, 2008.

[10] S. Gao and Q. Sun, Improving semantic concept detection through optimizing ranking function, IEEE Trans. Multimedia, vol. 9, no. 7, pp. 1430-1442, 2007.

[11] Y. Freund, R. D. Iyer, R. E. Schapire, and Y. Singer, An efficient boosting algorithm for combining preferences, Journal of Machine Learning Research, vol. 4, pp. 933-969, 2003.

[12] C. Rudin and R. E. Schapire, Margin-based ranking and an equivalence between adaboost and rankboost, Accepted for Journal of Machine Learning Research, 2008.

[13] T. Fawcett, An introduction to roc analysis, Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006.

[14] Madhavi Gudavalli, Dr.S.Viswanadha Raju, Dr. A. Vinaya Babu And Dr.D.Srinivasa Kumar, *Multimodal Biometrics - Sources , Architecture & Fusion Techniques: An Overview*, 2012 IEEE International Symposium on Biometrics and Security Technologies.

[15] Anil K. Jain, ArunRoss and SalilPrabhakar “An Introduction to Biometric Recognition”, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 14, NO. 1, JANUARY 2004.

[16] K.Krishneswari, S.Arumugam , “A Review on Palm Print Verification System” International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM) ISSN: 2150-7988 Vol.2 (2010), pp.113-120 .

[17] Xiangqian Wu, David Zhang, Kuanquan Wang, Bo Huang,“Palmprint classification using principal lines” 2004 Pattern Recognition Society. Published by Elsevier Ltd.

[18] B.Sateesh Kumar, “ Boosting Techniques on Rarity Mining”, International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X Volume 2, Issue 10, October 2012 , pp.27-35.

AUTHORS BIOGRAPHY

B.Sateesh Kumar , is working as Assistant Professor of Information Technology, JNTUH college of Engineering JAGITYALA, JNT University, Hyderabad. He received B.Tech (CSE) from Kakatiya University, M.Tech(SE) from JNTU Hyderabad and pursuing Ph.D(CSE) from JNTU, Hyderabad. He received “BHARATH JYOTHI AWARD” from New Delhi in December 2011. He organized a five day workshop on “Computer Awareness for Police Department” Under TEQIP in Department of CSE at JNTU College of Engineering, Kakinada. He guided many projects for B.Tech, M.Tech and MCA students. His research interest is in the area of Data Mining. He is one of the coordinators for *UGC sponsored program for SC/ST students* conducted at JNTU Kakinada. He is nominated as MEMBER OF BOARD OF STUDIES in CSE department at P.R.Govt.College, Kakinada. He was the reviewer of International conference on intelligent systems and data processing (ICISD 2011) under Multiconf-2011 organized by

Department of IT, GH Patel college of engineering at Vallabh, Vidyanagar. His research articles are accepted in international Conferences and journals.

Dr.A.Govardhan is presently a Professor of Computer Science & Engineering at School of Information Technology and Executive Council Member, Jawaharlal Nehru Technological University Hyderabad (JNTUH), India. He served and held several Academic and Administrative positions including Director of Evaluation, Principal, Head of the Department and Students' Advisor. He did his B.E.(CSE) from Osmania University College of Engineering, Hyderabad in 1992, M.Tech from Jawaharlal Nehru University(JNU), New Delhi in 1994 and Ph.D from Jawaharlal Nehru Technological University, Hyderabad in 2003. He is the recipient of 21 International and National Awards including A.P. State Government Best Teacher Award, Bharat Seva Ratna Puraskar, CSI Chapter Patron Award, Bharat Jyoti Award, International Intellectual Development Award and Mother Teresa Award for Outstanding Services, Achievements, Contributions, Meritorious Services, Outstanding Performance and Remarkable Role in the field of Education and Service to the Nation. He is a Chairman and Member on several Boards of Studies of various Universities. He is the Chairman of CSI Hyderabad Chapter. He is a Member on the Editorial Boards for Eight International Journals. He is a Member of several Advisory Boards & Academic Boards, a Committee Member for several International and National Conferences. He has 2 Monographs by Lambert Academic Publishing, Germany and Published in USA. He has guided 18 Ph.D theses, 125 M.Tech projects and he has published 152 research papers at International/National Journals/Conferences including *IEEE, ACM, Springer and Elsevier*. His areas of research include Databases, Data Warehousing & Mining and Information Retrieval Systems.

