Proc. of 3rd IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), (Washington DC, USA), September 2009.

Exploiting the "Doddington Zoo" Effect in Biometric Fusion

Arun Ross, Ajita Rattani, Massimo Tistarelli

Abstract— **Recent research in biometrics has suggested the existence of the "Biometric Menagerie" in which weak users contribute disproportionately to the error rate (FAR and FRR) of a biometric system. The aim of this work is to utilize this observation to design a multibiometric system where information is consolidated on a user-specific basis. To facilitate this, the users in a database are characterized into multiple categories and only users belonging to weak categories are required to provide additional biometric information. The contribution of this work lies in (a) the design of a selective fusion scheme where fusion is invoked only for a subset of users, and (b) evaluating the performance of such a scheme on two public datasets. Experiments on the multi-unit CASIA V3 iris database and multi-unit WVU fingerprint database indicate that selective fusion, as defined in this work, improves overall matching accuracy while potentially reducing overall computational time. This has positive implications in a large-scale system where the throughput can be substantially increased without compromising the verification accuracy of the system.**

I. INTRODUCTION

Multibiometric systems mitigate some of the limitations of uni-modal biometric systems by consolidating evidences from multiple sources of information [1]. These multiple sources of information can be in the form of multiple modalities (e.g., face and fingerprints), multiple impressions (e.g., face images captured under different poses), multiple classifiers (e.g., LDA and EBGM for face verification), multiple sensors (e.g., 2D and 3D sensors for face recognition) or multiple units (e.g., left and right index fingers). The fusion of these sources can be performed at various levels, viz., sensor level, feature level, score level, rank level and decision level [2].

Among the various levels of fusion, score level fusion offers the best tradeoff in terms of information availability and ease-of-fusion. While simple techniques such as the sum, product, min or max rules may be used to consolidate the multiple scores into a single fused score [2], recent research has resulted in the development of other types of fusion schemes that are data- and quality-driven [3], [4]. A variant of the sum rule - the weighted sum rule - may also be used to account for differences in matching performance of individual sources [5]. Efforts have also been made to perform user-specific fusion in which user-specific weights and matching thresholds are learned for each user [6], [7]. Such a scheme may be warranted due to the existence of different categories of users in the biometric database [8].

A. Ross is with the Lane Department of Computer Science and Engineering, West Virginia University, Morgantown, USA arun.ross@mail.wvu.edu

A. Rattani is with the Department of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi, Cagliari 09123, Italy ajita.rattani@diee.unica.it

M. Tistarelli is with the Department of Architecture and Planning, University of Sassari, Alghero 07041, Italy tista@uniss.it

Doddington et al [9] classify users, in the context of a speaker recognition system, into different groups based on their propensity to contribute to the False Reject and False Accept Rates (FAR and FRR) of a biometric system. The resulting classification, known as "Doddington's Zoo", assigns users into several categories labeled as Sheep, Lambs, Goats and Wolves. Sheep are well behaved users exhibiting low FRRs and whose feature sets are well separated from other users in the database. Goats are users who are intrinsically difficult to recognize and they tend to adversely degrade the performance by increasing FRR. Lambs are users whose biometric feature set overlaps significantly with other users in the database thereby contributing to a high FAR (zero-effort attack); according to Doddington et al [9], these users can be easily mimicked. Wolves are individuals who have the capability to spoof the biometric characteristics of other users: they too contribute significantly to the FAR of the system (spoof attack). Sheep (well behaved users) usually dominate the population. While Goats, Lambs and Wolves represent only a very small fraction of the database, their contribution in terms of the overall error rate of the biometric system can be large. Users in these categories are therefore referred to as *weak users* due to their negative impact on the error rate of the system.

Apart from the aforementioned categorization of users, Poh and Kittler [10] compute the "sheepishness index" for each user based on criterion such as the FRatio, and cluster the users into multiple groups based on this criterion. This is referred to as the "FRatio based approach" in this paper. Recently, Poh and Kittler [11] also propose a biometric menagerie index as the ratio of the intra-user variance to the expectation of the total variance, for ranking users in the database.

Existing work [9][10][11][12] confirm that weak users constitute only a small fraction of the database population; however, their contribution to error rates can be disproportionately high. Work reported in [13] proposes the use of user-specific score-normalization technique which has the effect of reducing the user-induced variability thereby improving the overall matching performance of the system. Thus far, efforts have been made toward classifying or ranking these weak users [9][10][11] or for reducing their overall impact [13]; however, no study has been reported as to how this categorization can be exploited in the context of a multibiometric system.

Since weak users contribute significantly to the error rates of the system, the performance of the biometric system can be significantly improved by detecting and strengthening these specific users using an additional source of biometric information. This type of *selective fusion* offers a good tradeoff between uni-modal and multibiometric system. The proposed concept utilizes the fact that less correlated sources of information for a user (e.g., face and iris, or left and right-index fingers) have a very high probability of belonging to different categories. Thus, for example, the first modality of a user may belong to the *well-behaved* category while the second modality may belong to the *weak* category. In such cases, the user may be required to provide only one of the modalities during authentication. However, for other users it may be necessary to obtain both the modalities and enhance verification accuracy.

This paper is organized as follows: section II discusses the methods for identifying weak users based on previous work in the literature; section III describes the selective fusion techniques designed in this work; a description of databases used and the experimental protocol is provided in section IV; experimental results are discussed in section V and conclusions are drawn in section VI.

II. CATEGORIZING USERS USING DODDINGTON ZOO AND FRATIO BASED APPROACHES

A. Doddington Zoo Based Approach

As reported in section 1, Doddington classifies the users in the biometric database into different groups, viz., sheep, goats, lambs and wolves. However, for the purpose of the current study, only sheep, goats and lambs have been assumed to be present in the database. Since concerted spoof attacks have not been considered in this work, the existence of wolves is not assumed. Goats and lambs are identified using the statistical framework based on the concept of percentiles of match scores as proposed in [9]. The p^{th} percentile of N ordered values is obtained by calculating the rank r as follows

$$r = \frac{p}{100} \times N + \frac{1}{2}. \tag{1}$$

For the detection of specific user categories, let $S(i,j,k)$ represent the score obtained by comparing the i^{th} sample of user j with the query sample of user k.

Goats: To identify goats, only the genuine scores of a user are required.

1) The mean genuine score of each user j, $SGEN_j$, is computed.
2) The set of mean genuine scores corresponding to all users is sorted.
3) The users whose $SGEN_j$'s are above the 70^{th} percentile (for distance scores) or lower than the 30^{th} percentile (for similarity scores), computed using (1), are labeled as goats. (See Figs. 1 and 2) .

Lambs: To identify lambs, the impostor scores corresponding to a user are necessary.

1) The maximum impostor similarity score between every pair of users is computed. Let $S_{max}(j,k)$ denote the maximum of all impostor scores between users j and k. For *distance* scores, the minimum of all impostor scores is computed; thus, accordingly $S_{min}(j,k)$ is calculated.
2) For every user j, the mean impostor score is computed as follows using (2):

$$SIMP_j = mean_k[S_{max}(j,k)] \tag{2}$$

3) The set of mean impostor scores corresponding to all users is sorted.
4) The users whose $SIMP_j's$ are above the 90^{th} percentile (for similarity scores) or lower than the 10^{th} percentile (for distance scores), computed using (1), are labeled as lambs (see Figs. 3 and 4).

Note that the percentile values used to detect goats and lambs are different from those specified in Doddington's original paper [9]. Further, these values can be altered based on the database and the nature of the application. Since the focus of this work is on demonstrating how the user categories generated using Doddington's method can be efficiently exploited, the percentile values have been arbitrarily selected.

As per the definitions above, it is possible for a user to be a goat as well as a lamb. To resolve this ambiguity, such users are automatically assigned to the category of goats.

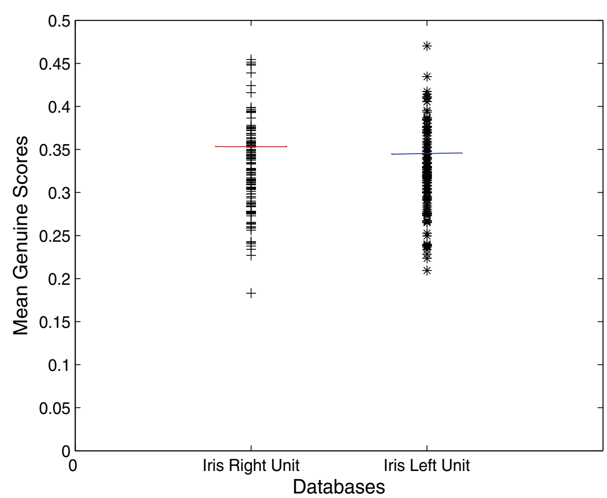

Fig. 1. The genuine mean distance score distribution of all the users in the database for the left and right irises in the CASIA v3 database. The users with scores above the red (Right Iris) and blue (Left Iris) lines are considered to be goats.

B. FRatio Based Approach

Poh and Kittler [10] cluster the users in a database such that there are significant differences in the recognition accuracy of users across clusters. The technique used for clustering in [10] is summarized below:

1) Compute the performance criterion value (see below) for each user

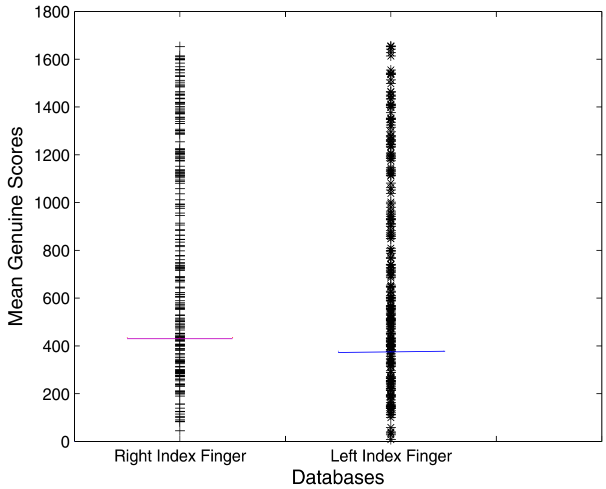

Fig. 2. The genuine mean similarity score distribution of all the users in the database for the left and right index fingerprints. The users with scores below the magenta (Right Index Finger) and blue (Left Index Finger) lines are considered to be goats.

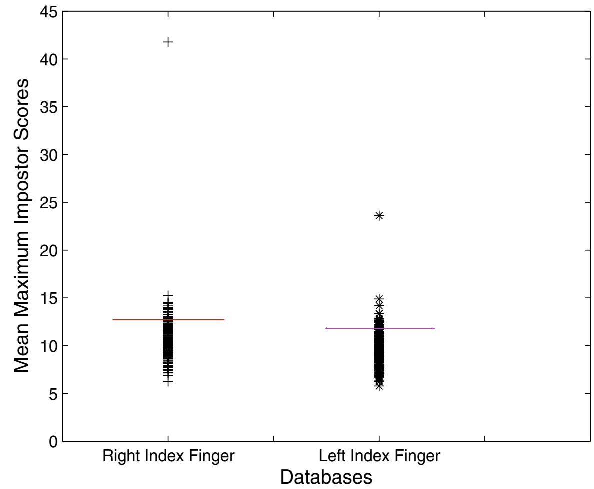

Fig. 4. The mean maximum impostor similarity score distribution of all the users in the database for the left and right index fingerprints. The users with scores above the red (Right Index Finger) and magenta (Left Index Finger) lines are considered to be lambs.

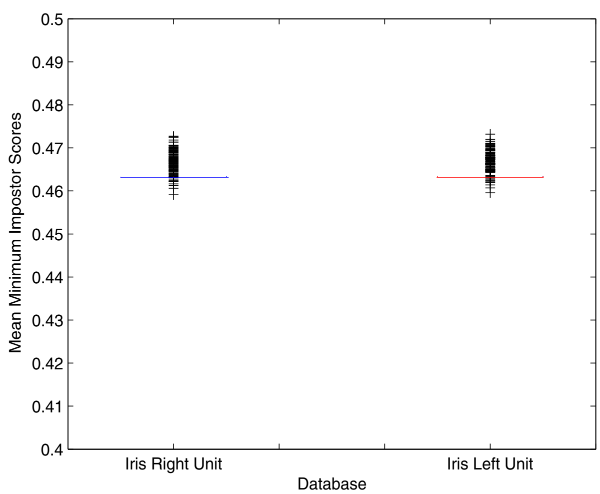

Fig. 3. The mean minimum impostor distance score distribution of all the users for the left and right irises in the CASIA v3 database. The users with scores below the blue (Right Iris) and red (Left Iris) lines are considered to be lambs.

2) Sort the users in increasing order on the basis of the selected criterion
3) Group the sorted users into multiple categories based on a pre-determined partitioning ratio
4) The first group comprises of the weakest users in the database while the last group contains the well behaved users

The authors in [10] investigated different criteria (viz., FRatio, Fisher Ratio and d-prime), and the FRatio (3) was eventually deemed to be useful for their purposes:

$$FRatio = (\mu_c - \mu_I)/(\sigma_c + \sigma_I) \quad (3)$$

where, μ_c and μ_I are the mean genuine and impostor scores, respectively, while σ_c and σ_I are the variance of the genuine and impostor scores, respectively, of a single user. Suppose there are 120 users in a database. For each user the FRatio is computed using (3). The higher the FRatio, the greater the class-separability between the genuine and impostor score distribution of the user. Next all the users are sorted in an ascending order on the basis of their FRatio. Based on a pre-determined clustering ratio obtained through learning on the training set, these users are partitioned into multiple clusters. For example, if the ratio is 1:2:3:4, then the individual clusters will have 12, 24, 36, and 48 users, respectively - the first 12 users being the weakest and the last 48 being the well-behaved users.

III. USER SPECIFIC FUSION AND SELECTION OF MODALITIES

In this section, the methods used to exploit the existence of these user categories are described. Two scenarios are envisioned, i.e, user specific fusion and user specific selection of modalities.

A. User Specific Fusion

In this scenario, it is assumed that all users are enrolled in a uni-modal system and only users deemed to be weak are expected to provide additional biometric information during authentication. Thus, the algorithm has two main steps:

1) Compute the weak users using the Doddington and FRatio based approaches;
2) Strengthen these detected users using additional source of information.

The proposed approach, apart from enhancing the overall matching accuracy of the system, also increases throughput time since only a small proportion of the users have to

TABLE I

SELECTING MODALITIES. THE COLUMNS LABELED *Modality 1* AND *Modality 2* INDICATE THE USER-GROUP TO WHICH THE MODALITY OF A USER BELONGS TO AND THE COLUMN LABELED *Selection* INDICATES WHICH MODALITY IS CHOSEN DURING ON-LINE OPERATION OF THE BIOMETRIC SYSTEM

Modality 1	Modality 2	Selection
Goat	Goat	Fusion
Lambs	Lambs	Fusion
Sheep	Sheep	Best performing modality
Goat	Lamb	Modality 1*
Goat	Sheep	Modality 2
Sheep	Lamb	Modality 1
Sheep	Goat	Modality 1
Lamb	Goat	Modality 2*
Lamb	Sheep	Modality 2

$*$ An alternative in this case would be to fuse the two modalities

provide multiple biometric information. The proposed approach will also be useful in situations where a uni-modal system is already deployed and the re-enrollment of all users for additional source of information may not be a viable solution (e.g., due to cost). Thus the system provides a good trade-off between uni-modal and multi-modal capabilities. The proposed technique will not encounter the problem of a small training set as the genuine match score data of a user will become incrementally available during the online operation of the system. This strategy is an incremental fusion technique because the second source of biometric information is only included after observing the system performance over a period of time.

B. User-Specific Selection of Biometric Modalities

In this second scenario, it is assumed that the multi-biometric information of all users are available to begin with; over a period of time, it is determined which fraction of users will continue to use multiple modalities.

A typical multi-biometric database stores multiple sources of information for all the users. For example, a multi-modal face and fingerprint database stores both face and fingerprint templates for all the users; during the online operation of the system each user is expected to present both the modalities (face and fingerprint) for authentication. In such instances, it may be possible to categorize individual modalities (for each subject) as being well-behaved or weak. For example, using Doddington's approach, the face modality of an individual may be categorized as "sheep" while the fingerprint may be categorized as "goat". In such a case, instead of retaining both sources of information for this individual, i.e., face and fingerprint, only the stronger source can be retained, i.e., face, without compromising the overall matching performance of the system.

1) User-Specific Selection of Modalities Using Doddington's Approach: Based on Doddington's approach, each modality of a user can be assigned to one of three categories. By observing the categories of both modalities (of a user), user-specific selection of modalities can be done using Table I

Thus, both sources of information are retained only when both are estimated to be lambs or goats. Although these rules (in Table I) are presented using the assumption that the number of available sources is two, the method can be scaled to a large number of possible sources.

2) User-Specific Selection of Modalities Using FRatio Based Approach:

1) Identify different clusters of users ranging from the weakest to strongest using the FRatio based criteria for each modality.
2) If both the modalities of a user are "well-behaved", then use only one of the modalities during on-line operation.
3) Fusion is performed only for those users whose modalities are deemed to be weak

Since well behaved' users always dominate the population, the user-specific selection of modalities will avoid using multiple modalities for most of the users. The scheme will substantially reduce the requirement of providing all biometric information during the online operation of the system. This will avoid the additional cost of processing, feature extraction and matching for each modality. The technique will also facilitate efficient storage management and updating of templates. Automated template updating can be conducted in an intelligent manner: templates belonging to goats may have to be updated frequently to account for intra-class variations while those corresponding to lambs may have to be updated cautiously due to their intrinsic overlap with other users.

IV. DATABASE AND EXPERIMENTAL PROTOCOL

Experiments are conducted using the CASIA Version3 multi-unit iris and WVU multi-unit index finger databases.

A. Database

CASIA Database Version 3: 122 subjects having at least 10 samples each of the left and right iris units are selected from the CASIA version 3 database [14]. The iris images are segmented using the Geodesic Active Contour technique proposed by Ross and Shah [15]. Each segmented image is preprocessed using histogram equalization and then Gabor filter-based features are extracted from the preprocessed image. A binary iris code is generated by quantizing the phase information from the Gabor responses. The Hamming distance is used to generate the match score between two iris codes [16].

WVU Fingerprint Database: 239 users having five samples each of the left and right index fingerprint units are used from the WVU fingerprint database [17]. The commercial VeriFinger software based on minutiae is used to compute the match scores between two fingerprint images.

B. Training and Test Sets

CASIA Database Version 3: Five images per user are used for training and the remaining samples are used for testing. Since the hamming distance is a symmetric distance classifier, i.e., Score(A,B) = Score(B,A), where A and B are two iris codes, there are 10 genuine scores generated per user for training and testing. Impostor scores for the training set are generated by comparing each of the five images belonging to a user against the five training images of all the other users resulting in a total of $121 \times 5 \times 5 \ = 3025$ scores. The same number of impostor scores is generated for the test set also.

WVU Fingerprint Database: The fingerprint matcher is a non-symmetric similarity matcher, i.e., Score(A,B) may not be equal to Score(B,A), where A and B are two fingerprint templates. Genuine scores for a user are generated by comparing each of the five images to all the other images of the same user and the scores belonging to the first two samples are used in the training set $(2 \times 4 = 8)$ and the scores belonging to the other three images are used in the test set $(3 \times 4 = 12)$. Impostor scores for a user are generated by comparing each of the five images of that user to all the images of all the other users. Impostor scores belonging to the first two samples are used in the training set $(2 \times 238 \times 5 = 2380)$ and those belonging to the remaining three samples are used in the test set $(3 \times 238 \times 5 = 3570)$ for each user.

The training set scores for each database are used to identify the weak users using the Doddington and FRatio based approaches. Test set scores are used to evaluate the performance of the selective fusion techniques. The fusion is performed at the match score level using the simple sum rule [2].

V. EXPERIMENTAL RESULTS

A. Detection of Weak Users

Iris: Using the method described in Section 2, 37(30.32)% goats are identified at the 70^{th} percentile and $11(9.02\%)$ lambs are identified at the 10^{th} percentile. Thus, in total, 48 out of 122 users i.e., 39.34% of users, are identified as weak users using Doddington's approach for the left and right iris units.

Using the FRatio based approach, the users are partitioned into five non-overlapping groups in the ratio of $1:2:3:4:5$. For the left iris, the first group consisting of 24 users (19.67%) are identified as belonging to the category of weak users. For the right unit, 48 users (39.34%), pertaining to the first two groups are identified as weak users.

Fingerprint: For the fingerprint database, 72 users are identified as goats (30.12%) at the 30^{th} percentile and 24 users are identified as lambs (10.04%) at the 90^{th} percentile, resulting in a total of 96 weak users (40.16%). Using the FRatio based approach, users are partitioned into 4 clusters in the ratio $1:2:3:4$. 24 weak users (10.04%) are identified for the left finger and 48 weak users (20.08%) are identified for the right finger. Note that owing to the use of percentiles in Doddington's Zoo, the same numbers of weak users are identified for the left and right, fingerprint and iris units.

In the case of left and right iris units 3 and 5 users, respectively, had a designation of both goat and lamb. For the left fingerprint unit, no user overlapped between the two categories. However, for the right fingerprint unit, 3 users overlapped between the lamb and goat categories. To resolve this problem, they were assigned to the category of goats. The detected percentages of weak users based on Doddington's approach are not small in number due to the use of fixed percentile values; however a fewer percentages of weak users are reported using the FRatio based approach.

B. Experimental Validations

Experiment #1: The goal of this experiment is to study the efficacy of user-specific fusion (Section III(A)) only for weak users. Here, it is assumed that only one modality is initially available. So in the discussion below, two types of possibilities are considered: (a) the left-unit of the fingerprint/iris is first available and the right-unit is then added for some users; (b) the right-unit of the fingerprint/iris is first available and the left-unit is subsequently added for some users. The ROC curves show the performance enhancement obtained on the test set after fusing the second unit of information only for the weak users.

Fig. 5, shows the performance of the iris recognition system based on the left unit (EER 1.7%) and the right unit (EER 3.0%). The legend *FRatio-left* shows the performance enhancement (35.29% improvement) obtained when the left unit is fused with the right unit for 19.67% users. Similarly, the legend *FRatio-Right* shows the performance enhancement (83.3% improvement) obtained when the right unit is fused with the left unit for 39.34% users. However, to demonstrate that this improvement is due to the concerted selection of users (as opposed to random selection), the fusion is performed by randomly selecting the same percentage of users in the database for each unit. It is observed that the performance enhancement is much less when random selection is employed. (For the sake of clarity, results are shown only for one unit using random selection; however similar results are obtained for the other unit too). The legend marked *Multibiometric System* indicates the performance when fusion is performed for all users.

Fig. 6 shows the improvement in performance when the Doddington approach is used to select users for fusion: the improvements for left (81.17%) and right (60.3%) units are marked as *Doddington Zoo-left* and *Doddington Zoo-Right*, respectively. The performance enhancement obtained by random selection based on the right iris unit is inferior by 45% compared to that obtained by the concerted selection of weak users.

In the case of fingerprint, the baseline accuracy of the right unit (EER 2.2%) is better than the left unit (EER 3.7%) (Fig. 7). Figs. 6 and 7 exhibit an interesting result, in which the performance of the fusion due to concerted selection is very similar to that of the Multi-biometric system when fusion is done for all subjects. Fig. 8 presents the results on the fingerprint database using the FRatio based approach.

Thus experimental results confirm that substantial increase in performance can be obtained by strengthening the weak users who represent only a portion of the entire database, thus lowering the throughput in comparison to a traditional multi-modal system.

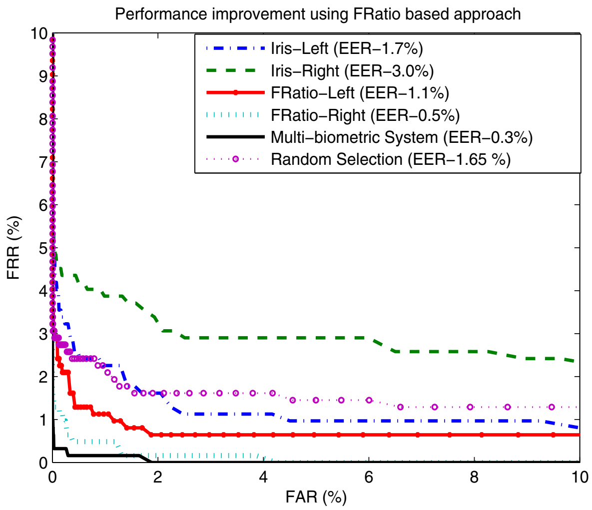

Fig. 5. The performance enhancement observed on the Left (35.29%) and Right iris units (83.3%) when user-specific fusion is performed for 19.67% users for the left iris and 39.34% users for the right iris. The weak users are identified using the FRatio statistic [10]

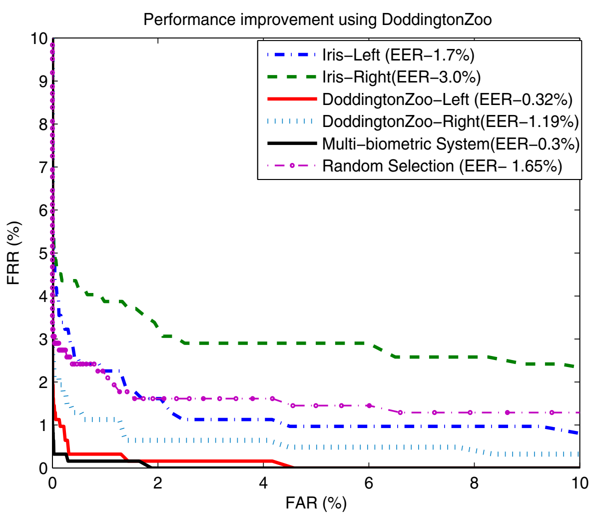

Fig. 6. The performance enhancement observed on the left (81.17%) and right iris units (60.3%) when user specific fusion is performed for 39.34% users identified using the Doddington statistic [9]

Experiment #2: The aim of this experiment is to show the performance improvement when user-specific selection of modalities (sources) is done using the procedure outlined in section III B(1). Table II shows that the combination involving sheep dominate the population on both the databases. Since fusion is performed when both the sources are categorized as weak, only 16.39% users in the CASIA database and 21.75% users in the WVU fingerprint database require fusion. The EER of the system after employing user-specific selection of modalities based on the Doddington approach is 2% (1.3% EER for multi-biometric system) for fingerprint and 1.1% (0.3% EER for multi-biometrics) for iris.

Table III shows the selection of modalities for the users based on the FRatio approach described in Section III B(2).

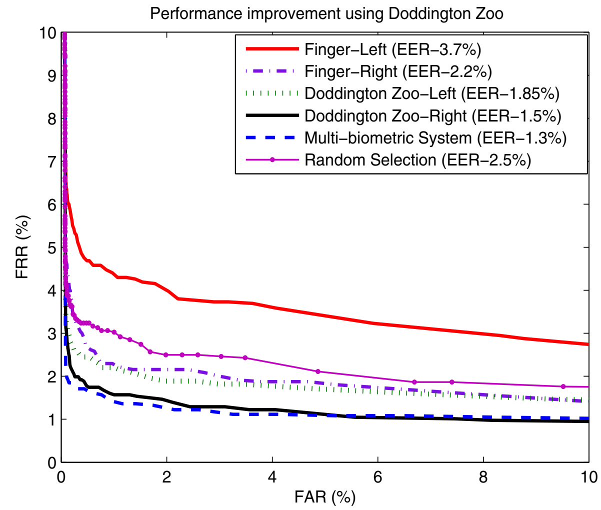

Fig. 7. The performance enhancement obtained on left (50%) and right fingerprint units (31.81%) when the additional source of information is used for 40.16% users identified using the Doddington approach [9]

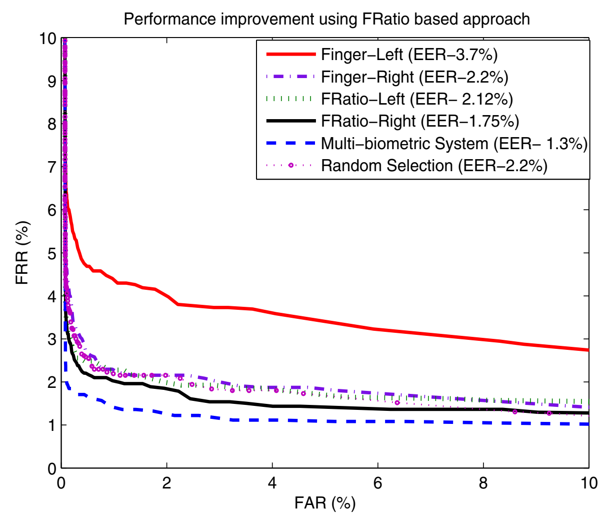

Fig. 8. The performance improvement obtained on the left (42.70%) and right fingerprint units (20.45%) when user-specific fusion is performed for 10.04% users for the left finger and 20.08% users for the right finger. The weak users are identified using the FRatio based approach [10]

TABLE II

THE DISTRIBUTION OF USERS BASED ON THE CATEGORIES ASSIGNED TO INDIVIDUAL MODALITIES FOR THE TWO DATABASES.

Right Unit	Left Unit	CASIA (%)	WVU (%)
Goats	Goats	14.75	20.50
Lambs	Lambs	1.64	1.25
Sheep	Sheep	50	46.02
Goat	Lamb	3.27	0
Goat	Sheep	12.29	9.62
Sheep	Lambs	1.63	7.11
Sheep	Goat	13.93	7.94
Lamb	Goat	1.64	1.67
Lamb	Sheep	0.82	5.85

TABLE III

THE SELECTION OF UNITS FOR USERS BASED ON THE FRATIO APPROACH

Dataset	Left Unit Only (%)	Right Unit Only (%)	Fusion (%)
Fingerprint(WVU)	40.16	40.16	19.66
Iris(CASIA)	33.60	33.60	32.78

Only 19.66% and 32.78% of users are required to use both the units of information in the CASIA and WVU databases, respectively. The EER of the system after performing user-specific modality selection using the FRatio based approach is 1.56% (1.3% EER for multibiometric system) for finger and 0.4% (0.3% EER for multibiometric system) for iris. Thus, there is minimal performance loss, especially in the case of FRatio based selection, accompanied by a substantial gain in throughput of the system as only a small fraction of users are needed to provide both the modalities during the online operation of the system. This user-specific selection has other advantages too as stated in section III.

VI. SUMMARY AND FUTURE WORK

This paper exploited the presence of different categories of users for user-specific fusion and modality selection. The following conclusions can be drawn:

1) Performing selective fusion for only weak users in the database enhances the performance of the system substantially. This enhancement cannot be achieved when the same percentage of users is randomly selected. Thus, this scheme offers a good tradeoff between unimodal and multibiometric systems.
2) Performing user-specific selection of modalities substantially increases the throughput of the system by fusing information for only a small percentage of users. Apart from this, the concerted selection of users has the added advantage of efficient management of user templates without compromising matching performance.
3) The Doddington and FRatio based methods for detection of different categories of users yield comparable results for both scenarios discussed in this paper.
4) The presence of weak users may also be due to the low and varying quality of biometric data obtained from them. Thus, incorporating quality information in the proposed framework will be necessary. This will result in the design of a dynamic fusion scheme based on intrinsic user characteristics as well as quality of the biometric data acquired from the user.

VII. ACKNOWLEDGMENTS

This work was done when Rattani was a Visiting Scholar at West Virginia University. Ross was funded by US NSF CAREER Award # IIS 0642554. Thanks to the anonymous reviewers for their suggestions and comments.

REFERENCES

[1] A. K. Jain and A. Ross, "Multibiometric Systems," *Communications of the ACM, Special Issue on Multimodal Interfaces*, vol. 47, no.1, 2004, pp. 34-40.

[2] A. Ross and A. K. Jain, "Information Fusion in Biometrics," *Pattern Recognition Letters*, vol. 24, issue 13, 2003, pp. 2115-2125.

[3] K. Nandakumar, Y. Chen, S. C. Dass and A. K. Jain, "Likelihood Ratio Based Biometric Score Fusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, 2008, pp. 342-347.

[4] K. Kryszczuk and A. Drygajlo, "Improving classification with class-independent quality measures: Q stack in face verification," *2nd International Conference on Biometrics (ICB)*, Seoul Korea, August 2007.

[5] A. K. Jain, K. Nandakumar and A. Ross, "Score Normalization in Multimodal Biometric Systems," *Pattern Recognition*, vol. 38, no. 12, 2005, pp. 2270-2285.

[6] A. K. Jain and A. Ross, "Learning User-specific Parameters in a Multibiometric System," *Proc. International Conference on Image Processing (ICIP)*, Rochester, New York, 2002, pp. 57-60.

[7] K.-A. Toh, X. Jiang and W.-Y. Yau, "Exploiting Global and Local Decisions for Multimodal Biometrics Verification," *IEEE Transactions on Signal Processing*, vol. 52, no. 10, October 2004, pp.3059-3072.

[8] M. Wittman, P. Davis, P. J. Flynn, "Empirical Studies of the Existence of the Biometric Menagerie in the FRGC 2.0 Color Image Corpus," *IEEE Computer Society Workshop on Biometrics*, 2006.

[9] G. Doddington, W. Liggett, A. Martin, M. Przybocki and D. Reynolds, "Sheep, Goats, Lambs and Wolves: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Recognition Evaluation," *Proc. of International Conference on Spoken Language Processing (ICSLP)*, vol. 4, 1998, pp. 1351-1354.

[10] N. Poh and J. Kittler, "A Methodology for Separating Sheep from Goats for Controlled Enrollment and Multimodal Fusion," *Proc. of Biometrics Symposium (BSYM)*, Tampa, Florida, 2008.

[11] N. Poh and J. Kittler, "A Biometric Menagerie Index for Characterising Template/ Model-Specific Variation," *3rd International Conference on Biometrics (ICB)*, Alghero, Italy, 2009.

[12] A. Hicklin, C. Watson, and B. Ulery, "The Myth of Goats: How Many People Have Fingerprints That are Hard to Match," *Technical Report NISTIR 7271*, NIST, 2005.

[13] N. Poh, S. Bengio and A. Ross, "Revisiting Doddington's Zoo: A Systematic Method to Assess User-dependent Variabilities", *Proc. of Second Workshop on Multimodal User Authentication (MMUA)*, Toulose, France, 2006.

[14] CASIA Iris Databases: http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp

[15] S. Shah and A. Ross, "Iris Segmentation Using Geodesic Active Contours," *IEEE Transactions on Information Forensics and Security (TIFS)*, 2009 (to appear).

[16] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, January 2004, pp. 2130.

[17] S. Crihalmeanu, A. Ross, S. Schuckers, L. Hornak, "A Protocol for Multibiometric Data Acquisition, Storage and Dissemination," *Technical Report*, West Virginia University, 2007.