

Modeling Time-Varying Population for Biometric Authentication

Vandana Roy and C. V. Jawahar
Center for Visual Information Technology
International Institute of Information Technology
Hyderabad, India
jawahar@iiit.ac.in

Abstract

Population size plays a major role in determining the performance of any biometric authentication system, particularly when such systems are used for civilian applications. In this paper, we propose to improve the performance of a biometric authentication system by modeling the variation in the population participating in the process. We show that this technique is helpful when the number of users enrolled into the system is very large as compared to the number of users which actually participate in the process. We show results on aperiodically and periodically varying population using Markov models.

1. Introduction

Biometrics based recognition has traditionally been used for forensic applications, such as criminal and terrorist identification. However, with exponentially decreasing costs of hardware and computational costs, it has proliferated into applications of civilian domain. Such applications include physical access to buildings, time and attendance monitoring, disbursement of provisions etc. As the end user of these applications is common man, user psychology plays an important role in selecting the type of biometric trait. In [2], various factors that affect selection of biometric trait and a comparison of various biometric traits based on these factors are discussed. Various applications of biometrics have also been discussed. In [4], various challenges in biometrics namely, accuracy, scale, security and privacy are discussed in detail.

Strong biometrics (having the properties of uniqueness and permanence) such as fingerprints, iris, retina, etc. are not well suited for civilian applications mainly because of low user acceptability (fingerprints relate to criminal history while iris and retina require users to stare into the camera). Weak biometrics (having low degree of uniqueness and low degree of permanence) such as hand-geometry, face, speech, etc. are gaining popularity for civilian applications.

These applications are used over a very long period of time and the system keeps enrolling new users. This results in a huge population enrolled into the system over time. It is a well known fact that performance of any pattern-recognition system is inversely proportional to the number of classes. Hence over time, as the system enrolls new users, the performance of the system degrades. Hence, it is required to improve or maintain the performance of the system with the increase in number of users over time. In this paper, we throw light on various scenarios where number of users enrolled into the system is much larger than regularly participating users. This variation in participation of users can be modeled and the performance of the system can be optimized over the regularly participating population.

1.1. Related Work

In the field of biometrics, especially for face recognition, various discriminant analysis techniques have been used to select features to improve discriminating content. Linear Discriminant Analysis (2DoLDA) approach [10] employs LDA on face image matrices rather than applying it on vectorized images. 2DoLDA may be implemented in two different ways: Row-oriented LDA (RoLDA) and Column-oriented LDA (CoLDA). An iterative algorithm using a generalized bilinear projection based Fisher criterion was proposed to efficiently combine two complementary versions of 2DoLDA [11]. In [6], the authors proposed a kernel machine-based discriminant analysis method, which deals with nonlinearity of the face pattern distributions.

In order to improve the performance of a biometric system, Vatsa *et al* [9], proposed a technique by fusion of results of two biometrics with minimum training dataset. In [5], the author proposed to use multimodal biometrics by fusion of fingerprint, face and iris based biometrics to improve recognition accuracy for large user population. Another work [7] tries to solve the Large-scale Biometric Pattern (LBP) recognition problem by using Multi-agent system for LBP (MLBP). The authors used multi-agent system solution to this problem because of the need to distribute the

problem and parallelize it among multiple computational units.

Most of the existing works deal with the use of strong biometrics for civilian applications and the use of weak biometrics has not been explored in sufficient detail.

2. Large-Scale Weak Biometrics

There are three problems to be addressed while using weak biometrics for civilian applications. The first two deal with inherent problems of weak biometrics while the third problem arises for use in civilian applications.

Low Discriminating Content Weak biometrics can distinguish between individuals with low accuracy. This is because traits such as hand-geometry, face, etc. of individuals may be very similar to each other, which leads to low performance of the system. This can be addressed by transforming the features into a new space, where the features are better discriminating. Any appropriate discriminant analysis technique can be used to select features to improve discriminating content.

Low Degree of Stability Features of weak biometrics change over time due to physical changes (aging, person gaining or losing weight, etc). Thus, the performance of the system degrades with time. Most of these changes are not abrupt and an intelligent system can adapt to the changing features. The civilian applications need to operate in cooperative mode in which the users provide feedback to the authentication system. The system learns by updating the parameters of the feature selection framework based on the feedback. Jain *et al.* [3], proposed a method to learn the genuine and impostor probability distributions of each of the users and compute a different threshold as against a common threshold for all the users. This results in much higher classification accuracy with user dependent thresholds. A practical limitation of their technique is that large number of samples are needed to train the system. The proposed method overcomes this difficulty as learning takes place over time with new samples by online adaptation.

Varying Participating Population The participating population in a biometric system can vary periodically or aperiodically. For eg., users can go on leave for a particular duration (aperiodic) while a group can visit the company every year for some time-period (periodic). It is required to incorporate the prior information of participation of each individual so that authentication can be performed only on the participating population. We propose to incorporate the time-varying parameter, i.e., the prior probability of participation $P_t(\omega_i)$ of each user ω_i into the conventional framework for feature selection. Techniques which can be used to compute $P(\omega_i)$ are described in following sections.

An intelligent framework is needed, which can select and adapt to the changing features and optimize performance over the participating population.

3. Modeling Time-Varying Population

In the previous section, we discussed that the users can participate in the authentication process in different ways. The variation in participation can be periodic or aperiodic. In this section we propose to model the time-varying population using Markov methods.

HMM can be used to model the variation in participating population as it is based on finding the probability of observing a sequence of participation of each user over a long period of time. Periodically or quasi-periodically varying population causes a repetition of sequence of participation over a period of time. Hence, knowing the period of repetition, a HMM can be trained on the sequence and the prior probability of each individual can be determined.

In case of aperiodically varying population, there is no repetition of sequence involved and the variation in pattern of participation of each individual is for a short period of time. For such a sequence, Markov chain model is better suited because of its low computational complexity. Moreover, the use of HMM for modeling such aperiodically varying sequence will require training of the HMM for each user every day. Each of these situations and their solutions are discussed in detail in the following subsections.

3.1. Non-Periodically Varying Population

Consider an authentication system being used for physical access to an organization. Some employees may be absent for various reasons, like on-leave, sick, etc. This variation in pattern of participation of every employee is usually aperiodic. That is, the employee is absent for usually a short duration of time and then comes regularly later on. It is thus required to calculate the prior probability of participation of each employee using the information obtained from recent history.

Participation of every individual can be considered as a random variable whose probability on a particular day depends upon his history of participation. The simplest way to model this variation is by Markov chain [1] in which the probability of each random variable in the sequence depends upon the value of the previous element. Every individual can be in one of the two states: “0” (absent) and “1” (present) on any day. Assuming that participation of an individual depends upon his presence or absence on the previous day, we need to find the prior probability of his participation on the current day given the state he was in on the immediately preceding day.

This can be calculated based on frequency of transitions from last state to state “1” over frequency of transitions from last state to both the states.

$$P_t^j(\omega_i) = \frac{N_{j1}}{N_{j1} + N_{j0}}$$

Here, P_t^j is the probability of going to state “1” given that

the last state was j . N_{j1} is the number of transitions from state j to state “1”. This probability is estimated from a window of past M days.

3.2. Periodically Varying Population

Consider a situation in which the participating population varies periodically. For example, assume a huge batch of students in an institution being divided into multiple groups. Each group attends a particular class on fixed days of a week. Thus, students participate in authentication process periodically. In this case, population variation is periodic over a long period of time. Hence, a technique is needed which can make use of the prior information of participation of each user on each day of the period.

For this situation, a Hidden Markov Model (HMM) [1] can be used to predict the sequence of states user can be in over a period of time. There can be various reasons for absence of a user, eg. working in shifts, belonging to different batch, etc. The model for varying population should be able to account for these reasons while predicting his participation on any particular day. The hidden states in case of our problem are the status of the individual and the outcome or observation are symbols “1” for presence and “0” for absence. This is because only whether the person is participating or not is visible to the system but not the reason or the status of the person.

Hidden Markov Model can be applied in following manner to our problem of finding prior probability of participation of each individual:

Hidden States: An individual can be in one of the states which determine if he will be able to participate. Let the hidden states be denoted by $H = h_j, j = 1 \dots M$, assuming M possible hidden states (eg. out of station, present, different batch, etc.).

Visible Symbols: There are two possible visible symbols: “0”(absence) or “1”(presence). Let $V = v_k$, where $k = 0, 1$

As the population is assumed to be varying periodically, we need to find the probability of participation of each individual on every day during the period. That is,

$$P_T(\omega_i) = p_t(\omega_i), t = 1 \dots T$$

$P_T(\omega_i)$ is the time varying series of probability of participation of the user ω_i and T is the periodicity (a week, a fortnight, or a month, etc.) Each user enrolled has a corresponding HMM for period T . It is required to train the HMM based on the sequence of participation observed over past T period of time. Training the HMM returns two probability matrices : $A = a_{ij}$ and $B = b_{jk}$, where a_{ij} denotes the probability of transition from h_i to h_j hidden state and b_{jk} is the probability of emitting v_k given that the system is in h_j state. Input for training is the sequence $V^T(\omega_i) = v_{kt}(\omega_i)$, where ω_i is the i^{th} user and $v_{kt} \in V$ on t^{th} day. It is required to find the probability that the user will be present

(i.e., $v_{kt} = 1$ on t^{th} day given that he is in state h_j . That is,

$$P_t(\omega_i) = P(v_{kt} = 1|h_{jt}) = b_{kj}.$$

Now, h_{jt} can be determined as a solution of the decoding problem of HMM which is defined as given the sequence of visible symbols, find the most probable sequence of states which produced the given symbols.

4. Results and Discussions

Various experiments were conducted to show that modeling the variation in population helps to improve performance of authentication. Both real and synthetic datasets were used for experiments. Incremental Biased Discriminant Analysis (IBDA) [8] was used to select the discriminating adaptive features. The prior probability of each user was incorporated into this framework. Markov chain was used to model the aperiodically varying population while HMM was used to model periodically varying population.

Generative Model A generative model to synthesize samples for a large number of individuals for experimental study in addition to the real data is proposed.

Characteristics of weak biometric samples:

1. Samples belonging to an individual obey Gaussian distribution. That is, $p(x/\omega_i, \theta_i) \sim N(\mu_{\omega_i}, \sigma_{\omega_i})$, where ω_i is the i^{th} user class and θ_i are the parameters of the samples of user i .
2. Weak biometric traits vary over time. Thus, the parameters of a user vary with time.
3. Clusters, belonging to different users, overlap with each other due to low discriminant content.
4. The population participating in authentication process changes with time. Let $P_t(\omega_i)$ be the prior probability of participation of the user ω_i at time instant t .

$$P^t(x/\omega_i) = \sum_{i=1} P_t(\omega_i) * p(x/\omega_i, \theta_i)$$

where $P^t(x/\omega_i)$ is the probability that the sample x of user ω_i will be generated on t^{th} time instant, say, day.

Datasets *Real Dataset (D1):* A hand geometry dataset collected from 12 users over a period of 60 days was used for experiments. The “raw” 24-dimensional features vectors extracted from hand images include lengths of each finger and their widths at five equidistant points.

Synthetic Dataset (D2): A synthetic dataset was created using the generative model as described before. This dataset consisted of samples from 150 users for 160 days.

Performance Over Time Experiments were conducted on datasets D1 and D2 to show the improvement in performance of the system for both the periodically and aperiodically varying population.

Non-periodically varying population: For these experiments, window of size 6 was chosen for computation of

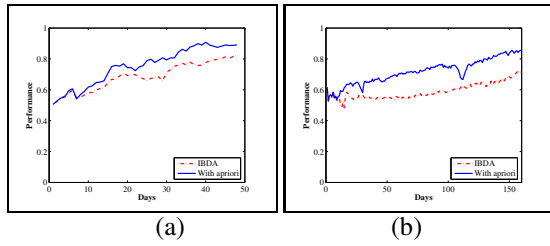


Figure 1. Effect on performance for Non-periodically varying population on datasets (a) D1 and (b) D2

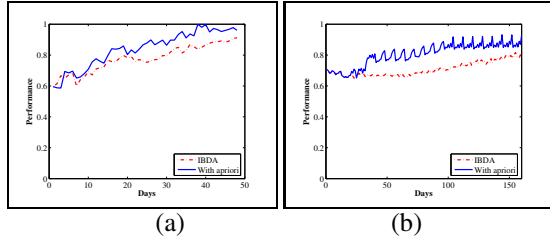


Figure 2. Effect on performance on periodically varying population on datasets (a) D1 and (b) D2

prior probabilities using Markov chain for both the datasets. Figure-1 shows the improvement in performance when apriori probability of participation of each individual was incorporated. However, on certain days (for eg. in Figure-1(a), on days 16, 21 and 32, and in Figure-1(b), on days 46, and 110), performance with apriori probability was observed to be almost same as that without incorporating apriori probability. This was because during those periods, prior probability of participation of each individual became almost equal to each other and hence, all individuals contributed equally toward the performance of the system. This is essentially same as using simple IBDA-transformed features as it assumes all users to participate equally-likely on each day.

Periodically varying population: Figure-2 shows the comparison of performance of the system when prior probability was incorporated over when it was not incorporated into the IBDA feature selection framework for periodically varying population. Figure-2(b) shows clearly that the performance changed periodically. This is because on each day of the 8-day period, different number of users participated and this pattern repeated after every 8-day period. This also shows that performance of the system was directly proportional to the number of users actually participating in the authentication process. This periodicity is not clearly visible with the dataset D1 in Figure-2(a) because same number of users participated on all the days during the period.

Effect of Varying Window Size Our method for estimating prior probability of each individual is based on his pat-

tern of participation for last M days. Table-1 shows the average performance of the system during periods of 10 days each on the dataset D2. Performance with window-sizes 1 and 3 were lower as compared to those with window-sizes 6, 8 and 12. Very small window size gives very less information based on which users' participation can be predicted in future. Optimal performance was observed with window-size 6. Larger windows resulted in degradation of performance because unwanted information used to predict the participation.

Effect of Entry and Exit of Users over Time We conducted experiments on dataset D2 for aperiodically varying population to show the pattern of change in performance when the users enrolled and exited from the system over time. Performance was recorded after a gap of 10 days between each consecutive observation. Table-2 shows that the system showed degradation in performance when new users entered into the system. However, when the users started exiting, performance of the system did not change much when apriori information was not incorporated and it showed an improvement in performance when apriori probabilities were included. This is because in the latter case, only the participating users contributed to performance of the system while in the previous case, samples of all the enrolled users were used for feature selection.

Users	30	50	100	120	140
Without apriori	0.936	0.852	0.813	0.789	0.708
With apriori	0.928	0.902	0.797	0.755	0.743
Users	110	90	60	40	20
Without apriori	0.684	0.677	0.653	0.667	0.652
With apriori	0.772	0.839	0.884	0.906	0.914

Table 2. Effect of users entering and exiting the system on performance

Effect of Varying Number of Users This experiment was conducted on dataset D2 with periodically-varying population. Table-3 shows that when apriori information was not incorporated, performance was observed to decrease with increasing number of users drastically while the degradation was observed to be at a much lower rate when apriori was included. This is because even though the number of users enrolled into the system increased, number of users participating was much lower. Hence the effective population size contributing to performance was low.

Comparison with Re-training the System Traditionally in order to improve performance of the system with entry or exit of users, pattern recognition systems require the system to be re-trained with existing users. We used periodically-varying data for this experiment and it was conducted with all the 150 users enrolled into the system but different number of actual participating users. This was compared with

Period	1	2	3	4	5	6	7	8	9	10
1	0.516	0.497	0.527	0.523	0.484	0.536	0.517	0.504	0.496	0.509
3	0.643	0.659	0.654	0.663	0.675	0.663	0.690	0.721	0.753	0.793
6	0.686	0.717	0.718	0.733	0.742	0.750	0.776	0.782	0.825	0.8514
8	0.672	0.693	0.698	0.692	0.719	0.714	0.727	0.751	0.814	0.833
12	0.579	0.592	0.625	0.641	0.642	0.650	0.672	0.677	0.682	0.686

Table 1. Effect of varying window size on performance of the system

Users	10	30	40	60
Without apriori	0.937	0.929	0.913	0.876
With apriori	0.943	0.940	0.921	0.889
Users	80	110	130	150
Without apriori	0.859	0.818	0.782	0.742
With apriori	0.870	0.853	0.851	0.849

Table 3. Effect of Number of Users on Performance with periodically-varying population

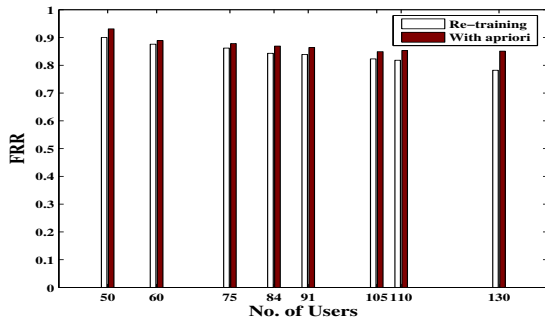


Figure 3. Comparison of performance with performance obtained after re-training the system

performance when the system was re-trained with the existing users. Figure-3 shows that our framework showed almost same performance as that obtained after re-training the system.

5. Conclusion

We proposed a strategy to improve performance of a weak biometric system for civilian applications. We argued that the effective population participating in the process is usually very less as compared to the total population enrolled into the system. We proposed to incorporate the prior probability of participation of each user into the adaptive feature selection framework. This results in optimized performance of the system over time over regularly participating population. Through various experiments we showed that incorporating prior probability helps to improve performance of the system under various scenarios.

References

- [1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley Interscience Publication, 2 edition, 2000.
- [2] A. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. In *IEEE Trans. on Circuits and Systems for Video Technology*, volume 14, pages 4–19, January 2004.
- [3] A. K. Jain and A. Ross. Learning user-specific parameters in a multibiometric system. In *International Conference on Image Processing (ICIP)*, volume 1, pages 57–60, September 2002.
- [4] A. K. Jain, A. Ross, S.Pankati, S. Prabhakar, L. Hong, and J. L. Wayman. Biometrics a grand challenge. In *Int'l Conference on Pattern Recognition (ICPR)*, volume 2, pages 935–942, August 2004.
- [5] T. Ko. Multimodal biometric identification for large user population using fingerprint, face and iris recognition. *Proceedings of 34th Applied Imagery and Pattern Recognition Workshop (AIPR05)*, pages 219–224, October 2005.
- [6] J. Lu, K. Plataniotis, and A. Venetsanopoulos. Face recognition using kernel direct discriminant analysis algorithms. In *IEEE Transactions on Neural Networks*, volume 14, pages 117–126, January 2003.
- [7] R. Meshulam, S. Reches, A. Yarden, and S. Kraus. Mlbpmas for large-scale biometric pattern recognition. *5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1095–1097, May 2006.
- [8] V. Roy and C. V. Jawahar. Hand-geometry based person authentication using incremental biased discriminant analysis. *Proceedings of the National Conference on Communications (NCC)*, pages 261–265, January 2006.
- [9] M. Vatsa, R. Singh, and P. Gupta. Multi biometric system for verification with minimum training data. In *Indian Conf. on Computer Vision, Graphics and Image Processing (ICVGIP)*, pages 569–574, December 2004.
- [10] M. Visani, C. Garcia, and J.-M. Jolion. Two-dimensional-oriented linear discriminant analysis for face recognition. In *Proc. of the International Conference on Computer Vision and Graphics (ICCVG)*, pages 1008–1017, September 2004.
- [11] M. Visani, C. Garcia, and J.-M. Jolion. Bilinear discriminant analysis for face recognition. In *3rd International Conference on Advances in Pattern Recognition (ICAPR)*, pages 247–256, August 2005.