Extremely-Large-Scale Biometric Authentication System - Its Practical Implementation

SU Leiming, SAKAMOTO Shizuo

Abstract

The Indian Unique ID (UID) is an extremely-large-scale system that attempts to identify India's 1.2 billion people, which is about 1/6th of the world population, using biometric authentication. As there are also other countries studying the introduction of national-scale authentication systems, NEC is currently conducting related R&D for the implementation of such systems. This paper describes a suitable system for use in processing of extremely large-scale biometric information.

Keywords

biometric authentication, multimodal, Indian unique ID system
large-scale database, scalability

1. Introduction

The process of confirming or identifying individuals is an essential component of many social scenarios. In particular, the series of terrorist attacks in the USA on September 11, 2001 significantly changed the level of importance of identifying individuals. However, personal identification based on possession of a card or use of an ID number or password is essentially incapable of preventing leaks, counterfeiting or impersonation. As a result, the usefulness of biometric authentication technology using the biometric information of individuals is emphasized, and its use in passports and immigration control is being adopted at an accelerating pace.

Citizen ID is one of the fields in which the application of biometrics is spreading rapidly. At NEC, we have already achieved certain results in this field, such as in the citizen ID system for the Republic of South Africa. This system uses fingerprint authentication that has been demonstrated to have the world's top recognition accuracy in benchmark tests held by the U.S. National Institute of Standards and Technology (NIST) [1] . Meanwhile, the Indian unique ID system (Indian UID) project has recently been started. This is an unprecedented extremely-large-scale system aiming at authenticating the 1.2 billion people of India, or 1/6th of the world population that uses biometrics to enable ID authentications by the Indian electronic government. Currently, we are tackling numerous new technical issues for the practical implementation of this system.

In the present paper, we introduce our efforts being made to support an extremely-large-scale biometric authentication system as represented by the Indian UID.

2. Indian UID

The Indian UID [2] forms the basis of the authentication program conducted by the Indian electronic government and it is expected that it will be applied in a wide range of authentication operations, such as tax payment, welfare and banking. The operation and issuing of the IDs are the responsibility of the Indian governmental agency called the UIDAI (Unique identification Authority of India), which will issue a unique ID for each Indian citizen in the near future.

To use IDs in various authentication operations, it is important that a unique ID is allocated to each citizen without duplication. This is guaranteed by collecting biometric information including the face, the fingerprints of ten fingers and the irises of both eyes and checking if they are duplicated in previously recorded data. The project was begun in 2009 and the registration/matching of 200 million citizens is scheduled to be completed by the end of 2012. Thereafter, the ID issuing speed will be accelerated in order to complete registration of 600 million citizens by the end of 2014 [3] .

De-duplication check of biometrics can be divided into the management of biometric information collection and the registration/matching based on biometric authentication. The overall scale of this biometric authentication system is unpre-

cedented globally and NEC is handling the registration/matching operations.

3. Issues with Extremely–Large–Scale Biometric Authentication Systems

Implementation of an extremely-large-scale biometric authentication system such as the Indian UID system requires solutions for various issues. We have identified the following four issues through analysis and study of the system.

Issue 1: Multimodal authentication accuracy
Issue 2: Optimization of multimodal authentication processing
Issue 3: Scalability
Issue 4: High-availability system designing

These issues are described separately in the following sections.

3.1 Issue 1: Multimodal Authentication Accuracy

The key to achieving high authentication accuracy is to collect biometric information of a quality, suitable for authentication. However, with an extremely-large-scale biometric authentication system such as the Indian UID, the collection itself can become an important issue. The Indian UID has to collect three modalities, which includes the face, the irises of both eyes and all fingerprints of both hands. To collect all of these with highest possible quality, considerations regarding the following matters are required.

Firstly, India has a vast land area and the second largest population in the world. A large data collection team is therefore deployed in order to reduce the time required to launch the system. In fact, the training and experience of each team vary widely. Therefore, quality of information depends on team members who collect biometric data. Another factor that contributes to quality variations is the large number of different devices provided by various manufacturers, used in collecting biometric information.

Since the Indian UID deals with the citizen IDs of all people aged 15 or more, the quality dependency on differences in the living environments of people and their occupations also poses a problem. It is known that the biometric information tends to vary depending on the ethnicity and occupations of the people of such a vast country. For example, the fingerprints of blue-collar workers, particularly those in the agricultural sector are often damaged. The tradition of wearing a beard is also problematic for biometric authentication. In other words, any of the collection conditions including the devices used, the collection staff and the collection targets pose the potential of affecting authentication accuracy.

In assuming the role of provider of the biometric authentication system, we have to consider all of the conceivable quality degradation factors as described above and to develop matching algorithms so that the collected biometric information can be as acceptable as possible for authentication.

3.2 Issue 2: Multimodal Authentication Processing Optimization

Usually, in the field of information processing, the search index is developed in advance from the registered population in order to increase the speed of searching targeted individuals. Specially in biometrics, information taken at different times may not strictly match with each other due to aging and conditions prevailing at the time of collection. As a result, it is generally difficult to develop and utilize the index effectively. For example, in order to confirm that a person requesting registration is not duplicated, all that can be done is to match the data with all of the registered data. Especially in the case of the Indian UID that executes strict duplication checking, it is forbidden to narrow down the check targets using personal demographic information such as name, gender and age, which means that the amount of required computation is enormous.

For example, to check duplication among all of the 1.2 billion people in India, about 7.2×10^{17} matching operations are required. The number is as large as about 1.8×10^{17} times even for checking the 600 million people scheduled to be registered by the end of 2014. To enable duplication checks on 600 million people in three years, we have to implement a system capable of more than 1.9 billion matches per second without interruption.

It is therefore not desirable to match all of the individual irises or fingerprints, but use them optimally is necessary. The optimization of such multiple modalities requires deep knowledge of each biometric technology. There are actually only a few companies in operation that are capable of implementing such a system beyond the experimental level.

3.3 Issue 3: Scalability

Even when individual biometric authentication operations can be optimized, a scalability issue will be posed if the system is incapable of an efficient response to matching requests

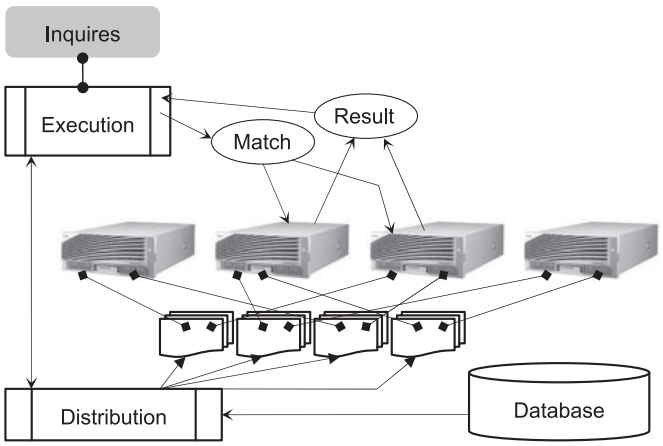

Fig. 1 Image of distributed processing configuration.

from the extremely-large-scale biometric database.

Fig. 1 shows the scheme of biometric authentication transaction. When a matching request is input to the system, it creates a plan for the registered biometric information distribution and for its execution on servers according to the internal load/processing situation. It also distributes the matching request to the required servers, collects the matching results from the servers and outputs the final response. Even slightest delays in the overhead, idling of matching servers will occur, which results in reduced matching efficiency or attainment of the scalability limit at which the transaction throughput cannot be increased even when matching servers are added.

3.4 Issue 4: Building a High–availability Environment

In general, the fault occurrence rate of this system is high because the system is composed of an extremely large number of servers and particularly that the servers in charge of matching are overused with CPU utilization rates of nearly 100%. However, as this system is used as a government authentication platform, it is subjected to severe fault tolerant performance requirements. It therefore encounters the following major issues before practical implementation.

1) No biometric information should be lost, even in case of a fault.
2) All operations should be able to continue, if the fault is in a single component only.
3) Multiple data centers should be operated as a countermeasure against disasters as well as for load distribution.

4. Implementation of Extremely–Large–Scale Biometric Authentication System

Out of the four issues mentioned above, this section introduces our solutions for issues 1 to 3, but excludes issue 4 which is related to actual implementation.

Solution 1: Multimodal authentication accuracy

In general, biometric information such as fingerprints, irises and faces vary on human ethnicity, customs and age. It can be understood intuitively that individuals have characteristic faces, hand sizes and finger lengths.

With our system, the authentication algorithm for each modality can be tuned according to the circumstances of each case. In fact, the Indian UID is tuned optimally according to the actual circumstances of Indian people. For example, the recognition accuracy of faces with turbans and beards is improved by enhancing the algorithm conforming to the data obtained from Indian people and selecting optimum parameters.

Fingerprints are authenticated by considering young people who are still growing, old people with difficulties in fingerprint ridge recognition, and people with damaged fingerprints such as blue-collar workers particularly those in the agricultural sector. By also adding parameters considering Indian ethnic characteristics such as relatively long fingers, we achieved high authentication accuracy by repeating evaluations and verifications.

Solution 2: Multimodal authentication processing optimization

In order to achieve high-speed matching while maintaining high authentication accuracies, of the entire biometric data bank including all of the fingerprints, the irises of both eyes and the faces, only the minimum required number of optimum combinations is matched. This explanation adopts a simplified virtual model using only two items of biometric information, A and B, for ease of understanding.

The accuracy of biometric authentication is calculated based on the FAR (False Acceptance Rate) as a security index and the FRR (False Rejection Rate) as a usability index.

Let us assume that the authentication accuracy of biometric information A alone is $(FAR_a, FRR_a) \cdot S_a$, where S_a is the number of operations that can be processed by a single-CPU core.

Similarly the authentication accuracy of biometric information B is assumed to be $(FAR_b, FRR_b) \cdot S_b$. When it is additionally assumed that there is no correlation between the two items of biometric information, the authentication accura-

cy and processing capability can be predicted based on various combination methods.

- **Combination (1)**
  The matched person is determined to be a registered person when any of the biometric authentication results meets the decision threshold that the matched person is identical to a registered person (OR).
  Number of processing operations/core/sec.: $S_1 = 1 / (1/S_a + 1/S_b)$
  False rejection rate: $FRR_1 = FRR_a \times FRR_b$
  False acceptance rate: $FAR_1 = 1 - (1 - FAR_a) \times (1 - FAR_b)$
  With this combination, the number of processing operations/core/sec. and the FAR degrade but the FRR improves.
- **Combination (2)**
  The matched person is determined to be a registered person when both biometric authentication results meet the decision threshold that the matched person is identical to a registered person (AND).
  Number of processing operations/core/sec.: $S_2 = 1 / (1/S_a + 1/S_b)$
  False rejection rate: $FRR_2 = 1 - (1 - FRR_a) \times (1 - FRR_b)$
  False acceptance rate: $FAR_2 = FAR_a \times FAR_b$
  With this combination, the number of processing operations/core/sec. is identical to combination (1), but the FRR degrades and the FAR improves.
- **Combination (3)**
  Only when the authentication results using biometric information A meet the decision threshold that matched person is identical to a registered person, another authentication of the person matched is executed using biometric information B. Then when the authentication results using biometric information B meet the decision threshold, the person is determined as the registered person.
  Number of processing operations/core/sec.: $S_3 = 1 / (1/S_a + P_3/S_b)$
  where P_3 is the probability that the first authentication result meets the threshold.
  False rejection rate: $FRR_3 = FRR_a + FRR_b$
  False acceptance rate: $FAR_3 = FAR_a \times FAR_b$
  Unlike combinations (1) and (2), combination (3) adjusts the FAR and FRR to meet the authentication accuracy required by the system so that the number of processing operations/core/sec. can be improved.
  This combination can be effective by selecting optimum thresholds, particularly when biometric authentications with a large difference in processing capabilities are combined. For instance, if S_a = 1 million operations, S_b = 50,000 operations and P_3 = 10%, combination (3) can provide about 7 times higher processing speed than combination (1).
  The processing flow of the system we provide for the Indian UID project adopts the multi-stage matching method based on fingerprints, irises and faces, and optimizes the processing capability while maintaining high authentication accuracy. The multi-stage matching method partially links the AND/OR conditions. in various ways. In addition, it is provided with extensibility and availability in consideration of possible changes in the quality of the biometric data collected by the Indian UID in the future.

The optimization of this multimodal authentication was result of our advanced R&D of biometric authentication technologies on fingerprints, irises and faces [4)5)] as well as our work on the evaluation testing of the Indian UID.

Solution 3: Scalability

To ensure scalability that can support an extremely-large-scale database, it is necessary to provide the control logic of the matching servers with a scale-out capability and scale transparency, which can be regarded as the biggest technological hurdle. Assuming that the processing capability of one server is equal to one unit, the overall processing capability by parallel processing of ten servers does not become ten units. This trend is more noticeable as the number of servers increase, and the improvement rate of the processing capability continues to degrade.

When the degradation rate becomes equal to the increase rate of the servers, the subsequent addition of servers leads to a stagnation of the processing capability. Assuming that the rate of effective performance degradation, due to factors other than the correlation effect between servers, is F and the number of servers is n, the effective output of the parallel processing of the servers can be expressed as $1/(1 - F + F/n)$. **Fig. 2** shows the graph in case F = 97%. When the correlation influence element is 3%, the execution output will never exceed 30 units, regardless of how many servers are added. If 1,000 matching servers are installed in parallel, the correlation influence element should be no more than 0.011%; thereby achieving an execution output of 900 units and an actual efficiency of 90%.

Our system manages the servers using the following techniques.

- Batch processing of internal job transports
- Multiphase transaction phase management

When processing transactions using matching servers of a

Extremely-Large-Scale Biometric Authentication System - Its Practical Implementation

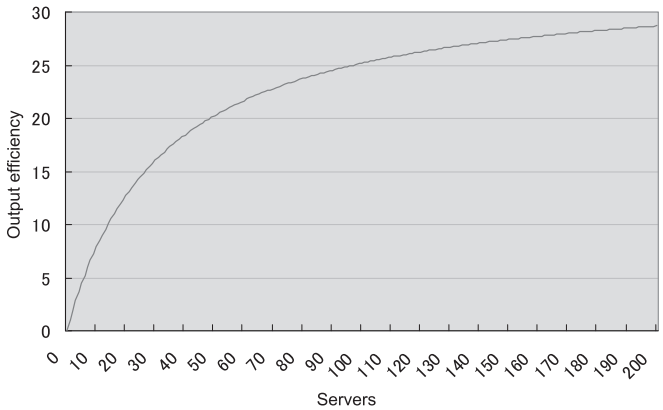

Fig. 2 Limit of scalability (when F = 97%).

scale of some hundreds of units, the overhead would be very high if the execution plans, placement plans and delivery/execution results are aggregated individually. To avoid this, we decided to manage transactions in a certain period with a single execution plan. We thereby succeeded in reducing the number of inter-server communications and the CPU power used in distributed management computations to several hundredths of the original numbers.

If transactions were managed individually, the status/progress monitoring cost would increase and a delay would tend to occur in the transition of internal processing. To prevent this occurrence and to improve the execution efficiency, minute management tasks are performed by internally dividing them into several phases.

5. Conclusion

The biometric authentication technology of faces and irises is constantly being improved and put to practical use following that of fingerprints. It is now customary for the biometric authentication technology to be applied to national-scale authentication systems such as for citizen ID systems. NEC aims to contribute to the creation of a safe and secure society in Japan as well as in other countries by making use of the technologies that have been developed over long years in the biometric authentication field, coupled with implementation knowhow of extremely-large-scale and high-reliability biometrics systems.

References

1) Shizuo Sakamoto, “Present Status and Prospects of Biometric Products and Solutions,” Vol.5, No.3, pp.14-17, Oct. 2010.
2) Unique Identification Authority of India (UIDA) http://uidai.gov.in/
3) The TIMES of India, Nov. 14th, 2009 / The Economist, Nov. 17th, 2011.
4) Masanori Mizoguchi and Masanori Hara, “Fingerprint / Palmprint Matching Identification Technology,” Vol.5, No.3 pp.18-22, Oct. 2010.
5) Hitoshi Imaoka, Akihiro Hayasaka, Yusuke Morishita, Atsushi Sato and Toshihiko Hiroaki, “NEC’s Face Recognition Technology and Its Applications,” Vol.5, No.3, pp.28-33, Oct. 2010.

Authors' Profiles

SU Leiming
Biometrics System Architect
2nd Government and Public Solutions Division
Government and Public Solutions Operations Unit

SAKAMOTO Shizuo
Executive Expert
2nd Government and Public Solutions Division
Government and Public Solutions Operations Unit