UIDAI

Unique Identification Authority of India
Planning Commission, Govt. of India (GoI),
3rd Floor, Tower II,
Jeevan Bharati Building,
Connaught Circus,
New Delhi 110001

Role of Biometric Technology in Aadhaar Enrollment

This report focuses on the biometric technology of the UID project for the purposes of UID enrollment. It goes into the proof of concept studies conducted in India, analysis of the study results, design decisions on biometric modes necessary in the Indian context, implementation of client and server side systems for enrollment and finally concludes with the accuracy and performance achieved by the UID biometric system using 8.4crore real enrolments.

Goal: The goal of the UID project is to assign a unique Identification number to each resident of India. The uniqueness constraint implies that during enrolment stage (creation of Aadhaar) each person will get one and only one Aadhaar number. To ensure that each person gets one and only one Aadhaar number it is necessary that the resident's identity information is captured and matched against every other resident (1:N check) who have previously enrolled - This process is called de-duplication.

Uniqueness & Biometrics: It is not possible to de-duplicate 1.2 billion residents by using demographic fields only (like name, address, age, gender etc) and moreover identity documents that rely only on demographic fields and personal reference checks are surrogates of identity and are vulnerable to forgery, falsification, theft, loss and other corruptions. In Indian context, biometrics were determined to be the most suitable factors for carrying out de-duplication. Hence it is necessary to enrol all residents along with their biometrics and build a clean database for the purposes of a National Identity system.

Biometric Standards Committee: The UIDAI's Biometric Standards Committee headed by Director General NIC (National Informatics Centre), published a report in December 2009 and advised that a biometric system based only on fingerprint might present challenges in India due to a large number of people engaged in manual labour and urged the UIDAI to consider the use of Iris in addition to fingerprints in order to improve inclusiveness and accuracy of the system.

Proof-of-Concept Study (PoC): The UIDAI conducted a Proof-of-Concept study during Mar-June 2010 in predominantly rural areas of Andhra Pradesh, Karnataka and Bihar published a report in December 2010 whose key findings included:

- Iris enrollment took less than a minute to capture and could be captured effectively from people, even from those who were blind.
- Children between 4-15 years could be biometrically enrolled correctly, and could be de-duplicated as accurately as adults.
- The accuracy levels achieved with a combination of fingerprint and iris were more than an order of magnitude (10x) better than using one or the other.

The PoC report concluded that "*The biometric matching analysis of 40,000 people showed that the accuracy levels achieved using both iris and ten fingerprints were more than an order of magnitude better compared to using either of the two individually. The*

multi-modal enrolment was adequate to carry out de-duplication on a much larger scale, with reasonable expectations of extending it to all residents of India"[1].

System Design: Based on the biometrics standards committee report, the PoC report, global learning's and expert opinions the UIDAI made the following design choices:

- Selected three biometric modalities of –10 fingerprints, 2 irises and face.
- Created standard client enrollment software - with quality checks for biometric and demographic data, consistency of capture process and encryption of enrolment data for security/data-protection.
- Built an enrollment server to perform demographic de-duplication, biometric de-duplication and manual adjudication of matches found by the system.
- Used commodity hardware, devices standards, open source software wherever possible, and defined standards and APIs (application programming interface) for interoperability and to avoid vendor-lock-in.

The Enrollment status of the UIDAI project as of Dec 31st 2011

- *36,000 Active enrolment stations*
- *59 Registrars, 83 Enrolment agencies in 32 states and union territories*
- *87,000 Certified enrolment operators*
- *11 - Models of certified biometric devices that are deployed in the field*
- *15 crore – Number of residents enrolled in the field*
- *10.25 crore – Number of Aadhaars generated by UIDAI*
- *10 lakhs/day – Peak enrolment processing rate*
- *100 trillion - Biometric person matches conducted every day to issue 10 lakh Aadhaars*

Biometric Accuracy and Effectiveness: In the last few months there have been media reports with misconceptions about the accuracy and efficacy of the UIDAI's biometric system. **The Parliamentary Standing Committee on Finance** (2011-12) that reviewed the National Identification Authority of India bill has referred to an "expert" who has stated "*it has been proven again and again that in the Indian environment, the failure to enrol with fingerprints is as high as 15% due to the prevalence of a huge population dependent on manual labour*". *These misconceptions have been addressed in the* box below:

[1] http://uidai.gov.in/images/FrontPageUpdates/uid_enrolment_poc_report.pdf

Biometric Accuracy

As of December 31st 2011, the UIDAI has true and tested statistics computed from real operational large-scale UIDAI system with the resident enrollment database size of 8.4 crore (84 million). It is unnecessary and inaccurate to attempt to infer UIDAI system performance from other systems which are ten to thousand times smaller. Specifically,

- ***Failure to Enroll(FTE) Rate: Zero**. As a policy, every unique resident, regardless of their biometrics can be enrolled and issued Aadhaar number.*
- ***Biometric Failure to Enrol Rate: 0.14%.** This implies that 99.86% of the population can be uniquely identified by the biometric system. The exceptions (0.14%) however are de-duplicated using demographic data and checked manually for fraud. The legitimate cases among these are issued Aadhaar number.*
- ***False Positive Identification Rate (FPIR): 0.057%.** In practical terms, it means that at a run rate of 10 lakh enrolments a day, only about 570 cases need to be manually reviewed daily to ensure that no resident is erroneously denied an Aadhaar number. The UIDAI currently has a manual adjudication team that reviews and resolves these cases. After manual adjudication, there is a negligible number of legitimate residents who are wrongly denied an Aadhaar number*
- ***False Negative Identification Rate (FNIR): 0.035%.** This implies that 99.965% of all duplicates submitted to the biometric de-duplication system are correctly caught by the system as duplicates. Given that currently approximately 0.5% of enrolments are duplicate submissions, only a few thousand duplicate Aadhaars would possibly be issued when the entire country of 120 crores is enrolled.*

The analysis resulting from such a large data set (8.4 crore records) is empirically repeatable and statistically accurate. There is no longer a need to rely on small sample size tests or hearsay from other projects. The UIDAI is now capable of measuring the accuracy, performance and scalability of the actual production system, which is already among the largest in the world. The results lay to rest unfounded claims that the underlying technology is untested, unreliable and based on unproven assumptions.

Based on the analysis, it can be stated with confidence that UIDAI enrollment system has proven to be reliable, accurate and scalable to meet the nation's need of providing unique Aadhaar numbers to the entire population. It is now safe to conclude that the system will be able to scale to handle the entire population.

Abbreviations

ABIS	Automatic Biometric Identification System
API	Application Programming Interface
CIDR	Central Information Data Repository
FPIR	False Positive Identification Rate
FNIR	False Negative Identification Rate
FTE	Failure to Enroll
NFIQ	NIST Fingerprint Image Quality
POC	Proof of Concept
SDK	Software Development Kit
STQC	Standardisation Testing and Quality Certification Directorate
UIDAI	Unique Identification Authority of India

Contents

1 Introduction

UIDAI has adopted use of biometrics technology as part of its core strategy[2] in meeting its goal of preventing issuance of duplicate identity number to a resident. There is no method or technology, other than biometrics, that can catch a person who is disclaiming his real identity. Biometrics consists of methods for uniquely recognizing human beings based on one or more of their intrinsic physical or behavioural traits. By matching a person's biometric characteristics with everyone else's (known as de-duplication), the technology helps prevent issuance of duplicate identity (Aadhaar number) to a single person.

Identity documents that rely only on demographic fields and personal reference checks are surrogates of identity and are vulnerable to forgery, falsification, theft, loss, and other corruptions. In western countries such as the United States and the United Kingdom, documents such as driver's license, and passports are used as identity proofs but only because of the reliability of the birth certificates. A birth certificate acts as a breeder document (in conjunction with identity documents of the child's parents) in obtaining identity document for the child. Even in countries with reliable birth certificates, the issuance of identity documents in a way that assures a 'one person/one identity' policy has been problematic. This model does not work in India, so UIDAI's strategy has been to minimize dependence on unreliable breeder and identity documentation and to not depend upon the trustworthiness of the operator, but rather to leverage automation and technology in a way that reduces the total dependence on error-prone documents and people based processes.

[2] http://uidai.gov.in/UID_PDF/Front_Page_Articles/Documents/Strategy_Overveiw-001.pdf

2 Biometric Design Methodology

The UIDAI biometric system design has followed global best practices. In designing UIDAI's biometric system, UIDAI reviewed existing state-of-the-art biometric systems, consulted with the world's top biometric experts, conducted a proof of concept study and built a biometric system that is currently considered to be world's best and widely acknowledged to be so in numerous international biometric forums and conferences. UIDAI technical experts visited two of the world's largest biometric implementations: US-VISIT program and US Visa/Consular system. They had meetings with a large number of experts from several countries including Mexico, Bangladesh, UK, the US, Singapore and Australia. Two of the world's most renowned biometrics experts – Prof. Anil Jain[3] and Prof. James Wayman worked with the UIDAI team and helped with the design. Prof. Anil Jain is pre-eminent biometric expert and advisor to many national and international governments. Prof. Wayman has served as an expert to numerous national ID system programs including UK, Philippines and the US. Several other biometric experts including Prof. Arun Ross[4], Prof. John Daugman[5] and Prof. Venu Govindaraju[6] also contributed to UIDAI's design.

UIDAI technical staff visited, reviewed and analyzed existing biometric programs in India including E-shakti NREGA scheme in Bihar, Coastal ID card of RGI, Orissa's UNWFR program, AP's Iris based ration card enrollment, Employees State Insurance Scheme of India (ESIC) and RSBY. These learnings were incorporated into a report published in December of 2009 by the UIDAI's Biometrics Standards Committee[7]. Analysis of some of the programs is referenced in the committee's report. Based on other programs' results, the report acknowledged that fingerprint-only system might present challenge in India due to majority of population being engaged in manual labour and advised that UIDAI to consider using iris to complement fingerprints in order to improve inclusiveness and accuracy of the system.

In December of 2010, The UIDAI published a Proof-of-Concept (PoC) study[8] of biometric enrolments that were conducted between March 2010 and June 2010 in the predominantly rural areas of Andhra Pradesh, Karnataka, and Bihar. The UIDAI also carried out biometric enrolment of school children in the vicinity of Bangalore. About seventy five thousand people in all were enrolled during the first phase of the PoC study

[3] http://www.cse.msu.edu/~jain/
[4] http://www.csee.wvu.edu/~ross/
[5] http://www.cl.cam.ac.uk/~jgd1000/
[6] http://www.cubs.buffalo.edu/govind/
[7] http://uidai.gov.in/UID_PDF/Committees/Biometrics_Standards_Committee_report.pdf
[8] http://uidai.gov.in/images/FrontPageUpdates/uid_enrolment_poc_report.pdf

including people over 90 years of age,, and sixty thousand of the same people were re-enrolled during the second phase after a gap of three weeks, in order to test the biometric matching efficiency using known duplicates. While the biometrics committee report based its recommendations upon learnings from other programs and experts, the proof-of-concept study aided the UIDAI team in getting first hand field experience and in measuring the various process and accuracy parameters. It also confirmed empirically, the earlier recommendation of the 'Biometrics Standards Committee' that using iris in conjunction with fingerprints was a prudent decision.

The PoC was conducted to evaluate technical, operational, and behavioral hypotheses related to both the use of biometric devices and the overall enrolment process itself. It was also conducted to establish a baseline for the quality of biometric data that could be collected in rural India.

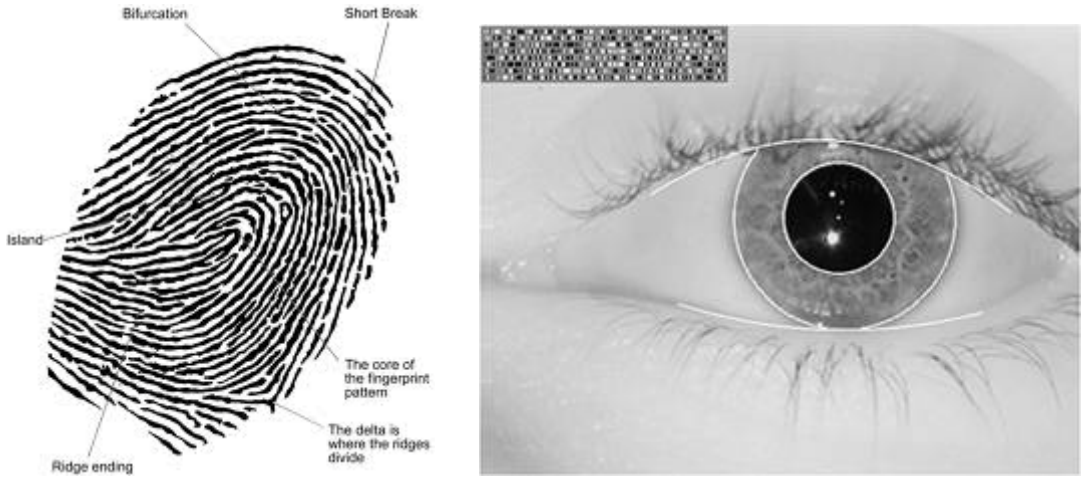


Figure: Fingerprint and Iris biometric

The key findings of the POC report, which have been presented internationally at scientific conferences and received extensive peer review, are listed here:

1. The PoC was successfully conducted over 1,35,000 biometric enrolments[9]. The relative ease of conducting the operation confirmed that biometric enrolment conforming to UID standards of quality and process was indeed possible on a large scale in rural India. The total biometric enrolment time for each individual, on average, was a little over three minutes. Of this, iris enrolment took a little under a minute, and was not perceived to be excessively difficult either by the resident or the enrolling operator. Specifically, many blind people also had their iris images captured successfully.
2. Multiple fingerprint scanners as well as iris capture devices were used in the PoC, and they performed according to expectations. The PoC was dispersed geographically and included many rural, often remote locations across three states. The enrolment was typically conducted with minimal infrastructure and sometimes in extreme

[9] Sample size of the PoC is considered large and statistically adequate for biometric technology assessment.

weather conditions. Residents varied in age all the way from four years to about ninety years.

3. In general older people took longer to enroll than younger people, and residents whose employment involved manual work took longer to enroll than the rest of the PoC population. Older people needed more assistance from operators to capture their biometrics. However, the range of enrolment times observed was well within expectations implying that the enrolment exercise for the population was indeed practical.
4. The enrolment variations tested in the process led to the conclusion that the best process was one where the resident remained stationary during enrolment and the operator did the positioning of the devices.
5. The enrolment of children at the school showed that children in the age range of four to fifteen could be biometrically enrolled using the same process as that used for adults and with no additional difficulty. The match analysis also showed that their iris images and fingerprints could be de-duplicated as accurately as those of adults.
6. The quality of the biometric capture was sensitive to the setup of the enrolment station and the process itself. Most importantly, the enrolment operator's instructions made a significant difference in the efficiency of the biometric capture.
7. The quality check process built into the enrolment software was very important and provided helpful feedback to the operator in capturing high quality images. The biometric matching analysis of 40,000 people showed that the accuracy levels achieved using both iris and ten fingerprints were more than an order of magnitude better compared to using either of the two individually. The multi-modal enrolment was adequate to carry out de-duplication on a much larger scale, with reasonable expectations of extending it to all residents of India.

The complete report is available on UIDAI's website. The final UIDAI design incorporated learnings from this PoC.

As of December of 2011, UIDAI has documented measurements taken from the real large-scale operational UIDAI system that has already issued over 10.25 crore (102.5 million) Aadhaar numbers. These measurements will be discussed later in this report under the heading of project status and system performance after we review the high level design of the UIDAI biometric system.

3 Biometric System Design

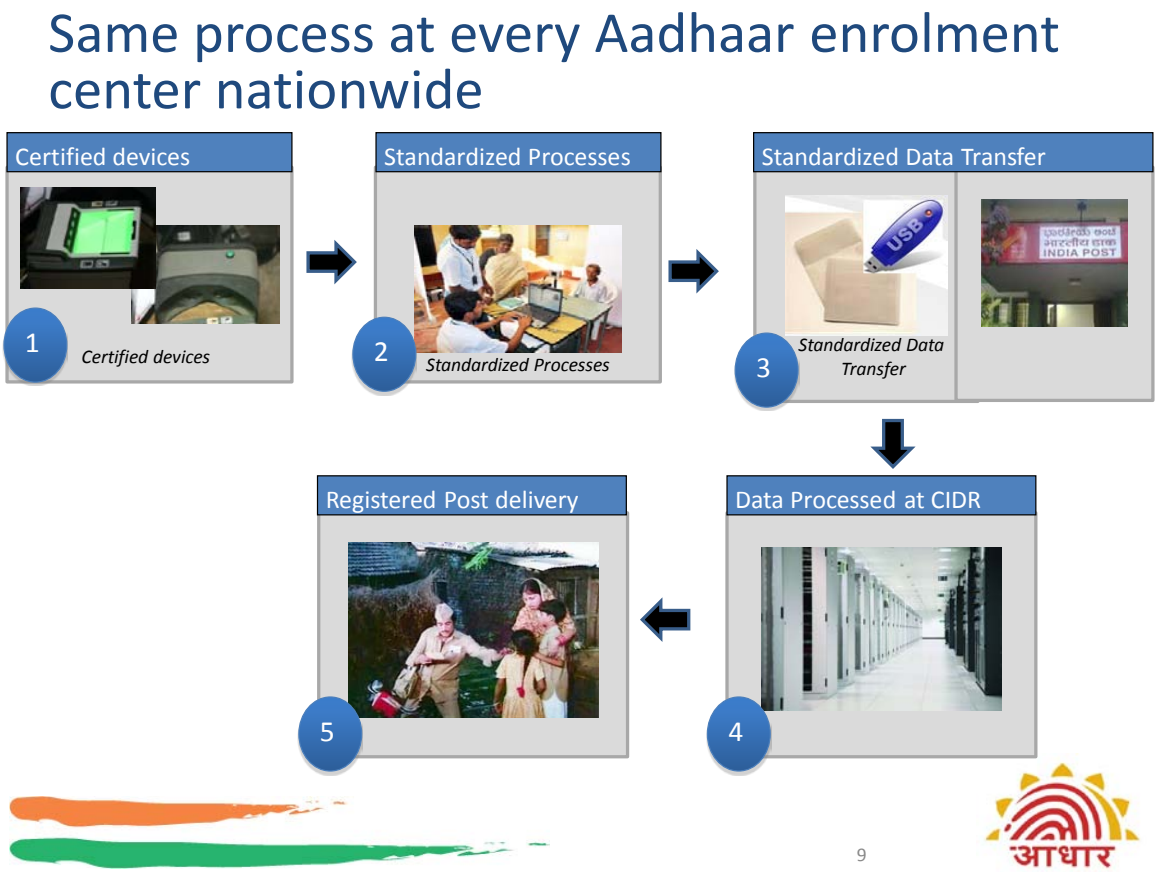
Following Biometric Standards Committee report, expert opinions, and learnings from the PoC, UIDAI selected three biometric modalities: face, all ten fingerprints and two irises. The decision to include iris in the UID initiative was a considered one, and took into account the critical needs of the project in ensuring the uniqueness of the Aadhaar number, and to also ensure that residents, particularly children and the elderly, are not excluded from enrolling for the UID. The PoC empirically demonstrated that iris is easy to capture, highly accurate, and not too expensive. By guaranteeing the universality and uniqueness of the UID, the initiative can have a substantial, transformational impact in the lives of residents.[10]

While the UIDAI Biometrics Standards Committee had already recommended the inclusion of iris, the PoC clearly demonstrated that iris capture was indeed necessary, and along with fingerprint, it was sufficient to de-duplicate and uniquely identify the entire population. The accuracy of the combined system is an order of magnitude better than fingerprints alone or iris alone, an important factor to consider for a population of 120 crore, and if the unique number is to be usable in high-security applications. Another reason for adding iris was inclusion. The use of iris also enables us to ensure the inclusion of the very poor, many of who work in physically intensive jobs, as well as children and the elderly. People working in jobs that require repeated use of fingers– for example, in fireworks factories or in areca nut plantations – often find their fingerprints degraded, which makes iris useful in ensuring uniqueness. The challenge with both fingerprint and face biometrics is that these have limitations when it comes to providing a unique number to children. Iris biometrics however, is reasonably stable in most persons, and can be collected from children as young as five years of age. This is an important factor considering the multiple programs aimed at child welfare.

The enrollment system is designed in two major parts: i) client-side and ii) server-side. The client-side is responsible for operator-assisted collection of relevant data from the resident in the field. The data is collected by client software provided by UIDAI which immediately encrypts and applies a digital signature to the data so that no one other than UIDAI's server can decrypt it, not even the operator, enrolling agency or even the registrar. Since data is encrypted, UIDAI's multi-registrar approach improves scalability and provides choice to residents without any negative effects on the data security. The encrypted data is transmitted to UIDAI Central Information Data Repository (CIDR) where it is fed to the server-side system. The backend server-side system uses multiple automatic biometric identification systems (ABISs) to determine whether the resident is unique (that is, the resident has never received another Aadhaar number before). The Aadhaar letter containing the UID number (assuming that the server found the resident

[10] http://uidai.gov.in/UID_PDF/Working_Papers/UID_and_iris_paper_final.pdf

to be unique) is sent from the server-side system back to the resident through a letter delivered by the department of post.



3.1 Enrollment software: Client side

The client-side system is used by trained and *certified* enrollment operators in the field to collect relevant data from residents. The data collected includes demographic and biometric data. As long as the resident has met the requirements under one of document/introducer/NPR methods, the resident will not be denied enrollment. *Biometric failure to enroll is not a reason to deny enrollment in Aadhaar*. Key features of the client-side system are:

1. **Standardization**. UIDAI has standardized demographic data and biometric data formats. Client software is provided by UIDAI to achieve consistency across the nation. . NPR has also adopted the same standards for their enrollment.
2. **Open source and avoidance of vendor lock-in**. Client software is supported on both open source Linux and Microsoft Windows platforms. To promote competition and avoid vendor lock-in, UIDAI has standardized the Application Programming Interface (API) between the devices and client software to allow use of any certified device. A total of eleven different devices (five fingerprint devices and six iris devices) have been tested and certified by STQC.

3. **Quality**. Strong and sophisticated quality control measures and checks are built intricately into the client software. The operators of the system are required to go through rigorous training and certification process for high quality of data and consistency across the country.
4. **Security**. The data collected by the UIDAI client software is immediately encrypted and signed by the software such that it can only be decrypted by the UIDAI server. No other party can access data at source or in transit.

The client-side system, while geared towards enforcement of correct data collection process and policies, also includes a number of methods geared towards decreasing failure to capture and failure to enroll of biometric data and lowering the biometric de-duplication errors at the back-end. Few examples of checks and balances implemented at the client-side system are as follows:

1. Each biometric capture device is required to have a built-in auto-capture capability which ensures that biometric images are captured only when deemed to be valid fingerprints slap or iris images and are of sufficient quality.
2. Biometric data quality is measured using standardized automated algorithms and thresholds are utilized to decide whether a captured sample is insufficient quality to warrant immediate re-capture.
3. The enrollment client performs a number of consistency checks. For instance, it makes sure that each biometric capture attempt comes from the same resident (instead of coming from operator, family member or previously enrolling resident).
4. The client software confirms that all 10 captured fingerprints are distinct as well as the two irises are distinct. It ensures that no repeated biometrics is captured.
5. The captured biometric is checked against that of the operator and the residents who enrolled previously on the same computer to avoid any chance of mix-ups.
6. Any biometric exceptions such as missing fingerprint or iris are logged and supervisor verification is required. In extreme cases such as missing both hands and/or missing both eyes, additional photograph of the hand and face is taken for proof of disability.
7. Operator overrides of the policies set in the software are logged to facilitate further investigation of the capture process and operator actions.
8. The images from all attempts (up to four) are included in the resident data packet and sent to server for processing.

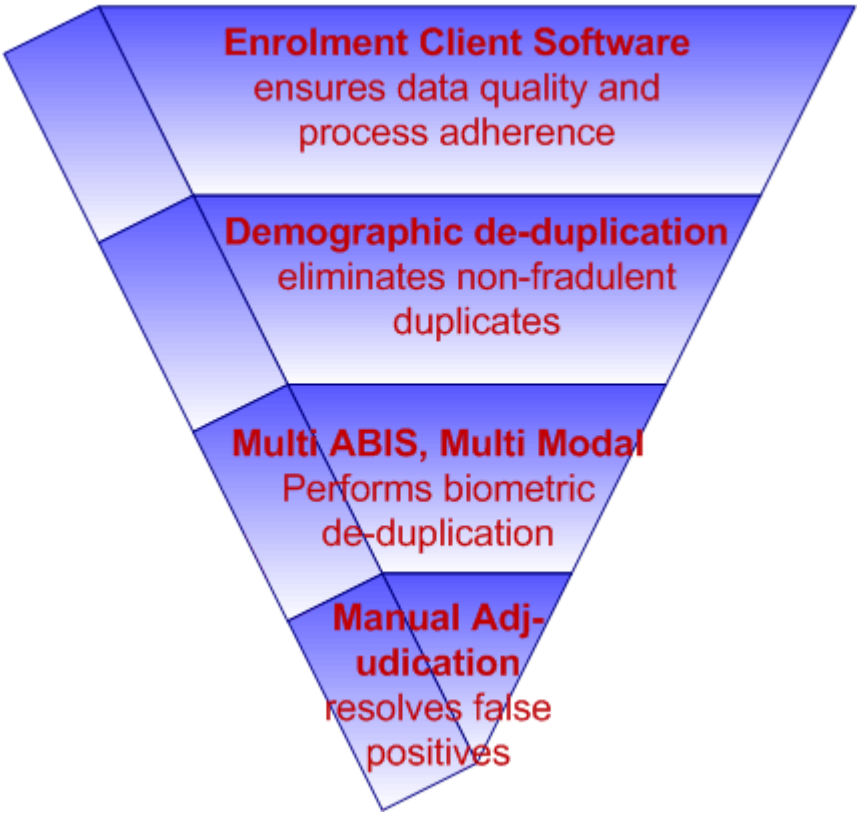
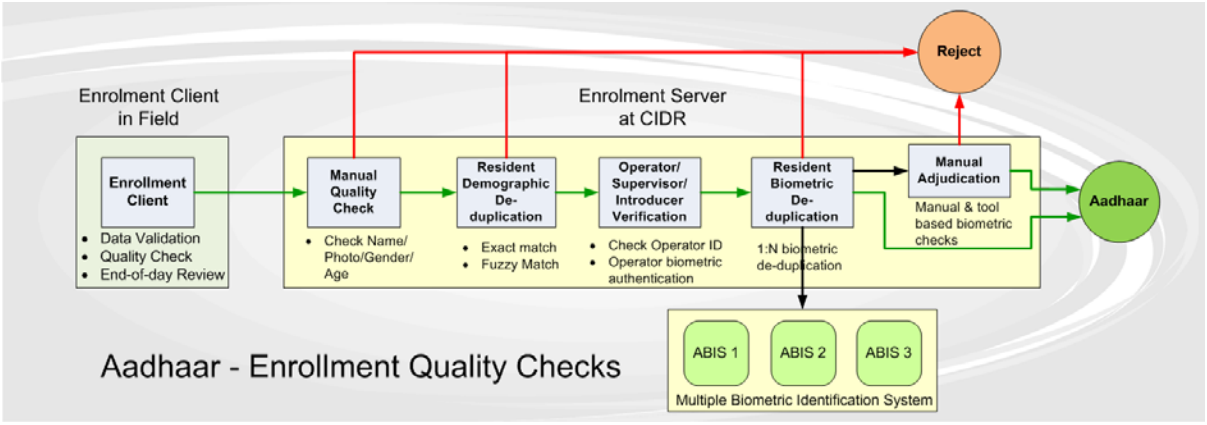


Figure: Multi-step process to maximize accuracy

3.2 Enrollment software: Server side

The server-side system is designed to scale to very high data and compute requirements as biometric de-duplication technology is computation and data intensive. Three main sub-systems on server side are utilized to provide highest accuracy while scaling to handle 120 crore population goal. Duplicates are identified at each level. Different algorithms also ensure that false rejection of Aadhaar number does not occur at any level.

- Demographic de-duplication
- Multi-ABIS Multi-modal Biometric de-duplication (using multiple ABISs and multi-modal biometrics of 10 fingerprints and 2 irises)
- Manual adjudication (primarily to resolve records identified as duplicates found by previous stages)



Demographic de-duplication is used primarily to catch trivial duplicates(non-fraudulent cases where all the demographic fields are identical) that are inadvertently submitted to the system, for example when a resident has not received Aadhaar number in a few days and decides to re-enrol at an enrollment station again. Secondarily, it is also used for children under the age of 5 year as biometrics is not captured for children that young. The UIDAI uses both exact-match and fuzzy-match strategies to improve the demographic de-duplication accuracy.

Multi-ABIS Multi-modal Biometric de-duplication The biometric data de-duplication is at the heart of the system. The UID has procured 3 ABIS providers to perform biometric de-duplication, since they bring significant advantages:

1. The deployment of multiple ABISs improves the accuracy of de-duplication. If any ABIS identifies a potential duplicate, it is sent to the other ABIS for verification. By combining the results of all 3 ABIS's the overall biometric de-duplication accuracy goes up.
2. The utilization of three different de-duplication engines with different implementations and different fusion strategies also helps to detect various kinds of software or data collection errors. In certain enrolments (for example in suspected duplicates and enrolments with poor quality biometrics) the enrollment data is sent to more than one ABIS to minimize the chance of an identification error. Another fascinating aspect of the continuous improvement process, originates from the feedback that is provided from the server-side system to the enrollment agency about the quality of their data. When enrollment agencies receive frequent report on the quality of their enrolments, it leads them to improve their training and processes – since their payments are linked to successful Aadhaars generated and not number of enrolments conducted. This improvement clearly shows up in measurements of quality that is performed frequently at the back-end.
3. The three ABISs compete for work based on their throughput capacity. This competition allows for continuous improvement in throughput and accuracy.
4. By distributing and sharing the de-duplication load across the 3 ABIS vendors the multi-ABIS solution gives the system a threefold increase in throughput. This is the reason why the UID system is able to achieve 10 lakh Aadhaars/day.

5. It ensures that the there is no vendor lock-in, if one of the ABIS vendors needs to be replaced (for whatever reason - technical or contractual) it can be done without bringing the entire system to a grinding halt.

Manual adjudication (required for duplicates found by the ABIS) is implemented as a semi-automatic process. The duplicates found by ABIS are processed by biometric SDKs to check if a process related issue has caused this (for example, mix of operator and resident biometric or repeated biometric of the resident etc.). Finally, the duplicate is analyzed manually and the final decision is made by a human expert, leaving negligible chance that a legitimate resident is denied Aadhaar.

Security & Data Privacy: UIDAI system has been designed with utmost care to ensure security and privacy of data. Several features have been implemented to ensure that the resident's data remains completely private even within UIDAI partners and stakeholders, these are:

1. The data being sent to ABIS is completely anonymized meaning none of the ABIS systems have access to resident's demographic information (name, address, gender and date-of-birth), they are only sent biometric information of a resident with a reference-number and asked to de-duplicate. The de-duplication result that the ABIS returns with the reference number is mapped back to the correct enrolment-number by the UID's own enrolment server. This is akin to removing names from an examination answer paper so the examiner does not know whose paper he is evaluating.
2. The ABIS providers only provide their software and services. The data is stored in UID storage and it never leaves the UID's secure premises.
3. They ABIS providers also do not store the biometric images (source), they can only store template for the purpose of de-duplication.
4. The enrollment packet after it is encrypted by the enrolment client software in the field is sent to the UIDAI's CIDR (Central Identity Repository), the enrolment server decrypts the packet for de-duplication but never stores the decrypted packet in storage.
5. The original biometric images of fingerprints, irises and face are archived and stored offline and hence cannot be accessed through an online network.

4 Project Status & System Performance

There are over 36,000 active enrollment stations operated by 83 active enrollment agencies contracted through 59 active registrars across 32 states and union territories. Each station enrolls 50 residents per day, on an average. As many as eleven different models of fingerprint and iris devices are deployed in the field.

As of 31st December 2011, more than 15crore (150 million) enrolments have taken place in the field using UIDAI's client-side system, and over 10.25 crore (102.5 million) Aadhaar numbers have been generated by the server-side system. The overwhelming majority of the remaining nearly 5 crores is in transit from the field to the data center. The throughput has consistently increased since the start of the program both in the field and at the server back-end. In the month of December of 2011 alone, 2.23 crore (22.3 million) Aadhaar numbers were generated. The system is capable of processing at the rate of 10 lakhs (1 million) Aadhaar numbers per day and will continue to ramp up to meet the anticipated growth. The system is now the largest in the world based on its daily processing rate and one of three largest in terms of its database size. It is already the largest multi-modal biometric deployment in the world. The system produces daily measurements of - accuracy, throughput and quality. These measurements are based on industry wide accepted methods of calculating accuracy and quality. It should be emphasized that the system is now sufficiently large to clearly estimate the performance necessary to enrol the entire population. The system performance can no longer be a matter of speculation or extrapolation from small samples. It is unnecessary and inaccurate to attempt to infer UIDAI system performance from other systems which are ten to thousand times smaller.

Three key measures define the system effectiveness

1. **Biometric Failure to Enroll rate.** Per UIDAI policy, failure to enroll in Aadhaar is not allowed. That is, Aadhaar is a right of every Indian resident and cannot be denied. Therefore overall failure to enroll is set by policy to be *zero*. However, if certain residents are not able to provide their biometric (called B-FTE henceforth), biometric de-duplication cannot be carried out on their enrollment. Therefore the UIDAI measures B-FTE on an ongoing basis. It has been reported in a news item that this number in Indian context could be as high as 15%.
2. **False Rejection**. When new enrollment data from a resident is sent to CIDR, the system de-duplicates the resident packet to ensure that the resident has not previously been given a Aadhaar number. If the biometric de-duplication system rejects the new enrollment as being a duplicate, it is checked by manual adjudication process. If biometric system makes lots of errors, the volume of cases to adjudicate can go up significantly. This error by the biometric system is called false rejection of Aadhaar or False Positive Identification Rate (FPIR) of the biometric system. FPIR is a technical term and has very precise meaning in biometric literature. It has been reported in a news item that CIDR's FPIR could be so high to render the system useless.

3. **False Acceptance**. When new enrollment data from a resident is sent to CIDR, the system performs de-duplication to ensure that the resident has not previously been given an Aadhaar number. If in case the biometric system accepts the resident as new when in reality it was actually a duplicate, the resident will end up with two Aadhaar numbers. This error by the biometric system is called false acceptance of Aadhaar or False Negative Identification Rate (FNIR). FNIR is a technical term and has very precise meaning in biometric literature.

4.1 Measured System Performance

It is now possible to carry out full scale performance measurement of real UIDAI production system and not rely on small samples or hearsay about other projects. The measurements were thus carried out on the entire database (called gallery) using large number of identification records[11] (called probes) when the UID database size reached 8.4 crore records.

1. **Failure to Enroll (B-FTE).** The biometric failure to enroll rate is measured to be **0.14%. It means that 99.86% of the population has biometric that is usable for de-duplication purpose**. The exceptions (0.14%) of the population were not able to provide fingerprint and iris images and thus would be de-duplicated using demographic data and checked manually for fraud. The legitimate cases among these will be issued Aadhaar number. The UIDAI's actual B-FTE of 0.14% is more than 100 times lower than speculated in unfounded reports critical of the system.

2. **False Reject of Aadhaar (FPIR). FPIR is computed in the operational system by submitting a new record** to the system for de-duplication. If the system finds a duplicate (a HIT is said to have occurred), the pair is manually inspected and if the HIT is returned by the system in error, it is counted towards false positive identification errors. An **FPIR of 0.057%** was measured when the gallery size was 8.4 crore (84 million) and probe size was 40 lakhs (4 million). The false rejects (legitimate residents who are falsely rejected by the biometric system) were a count of 2309 out of the 40 lakh probes. These must go through adjudication process that involves manual review where the errors from the biometric system are corrected. In practical terms,it means that at a run rate 10 lakhs enrolments a day, only about 570 cases need to be manually reviewed daily to ensure that no resident is erroneously denied an Aadhaar number. Although this number is expected to grow as the database size increases, it is not expected to exceed manageable values even at full enrolment of 120 crores. The UIDAI currently has a manual adjudication team that reviews and resolves such cases.

3. **False accept (FNIR):** To compute FNIR, 31,399 known duplicates were used as probe against gallery of 8.4 crore (84M). The biometric system correctly caught 31,388 duplicates (in other words, it did not catch 11 duplicates). The computed FNIR rate is

[11] These tests are between 10 to 1,000 times larger than any test conducted on live data in the world.

0.0352%. Assuming current 0.5% rate of duplicate submissions continues, there would only be a very small number of duplicate Aadhaars issued when the entire country of 120 crores is enrolled. Aadhaar expects to be able to increase the quality of all collections as the system matures. Consequently, we expect the potential number of false acceptances to decrease further below this already operationally satisfactory number.

4.2 Scaling to 120 Crore

Performance

As the UIDAI enrols more people, more resources are required to perform biometric de-duplication. The CIDR will require more computing resources as the data base grows. However, this process is extremely scalable (parallelizable) and throughput can be maintained with the addition of hardware. UIDAI has done sufficient modeling as it grew from one crore to ten crore, it can be stated with high confidence that this throughput can be maintained till the entire nation is covered. In fact, the design of UIDAI system is such that throughput can be increased beyond 10 lakhs per day with reasonable addition of hardware.

Accuracy

Similarly, as the UIDAI enrol more residents, the internal parameters of the biometric system need to be adjusted to ensure that the biometric accuracy of the system does not degrade as the database grows.

Three well understood and accepted phenomena help us adjust these internal parameters.

1. False accept (FNIR) rate remains steady and does not increase with the increase in the database.

2. False reject (FPIR) rate grows linearly with the database size.

3. It is possible to trade off FNIR with FPIR. In other words, if we decrease FPIR, FNIR will increase. This relationship is modelled as "receiver operating characteristics (ROC)" and has very precise meaning in the biometric literature. UIDAI has modelled this relationship on the real production data.

Based on the model, the UIDAI expects the accuracy of the system to remain within the same order of magnitude as reported above. Hence it can be stated that system will be able to scale to handle the entire population without significant drop in accuracy.

5 Analysis

It is instructive to analyze the results to help us learn and make improvements in the system going forward.

5.1 Duplicates found correctly

In the previous section, when looking at results of FNIR, it was observed that a majority of duplicates found by the biometric de-duplication system are indeed found correctly. It is important to analyze these duplicates and understand why they are occurring. Some examples of duplicates being submitted to the system (and correctly caught by the system) are as follows:

1. Mixed biometrics: In this case, multiple attempts of the same modality (say fingerprints) belong to two different individuals. Among the duplicates, this issue was relatively frequent as it accounted for approximately 20% of the duplicates (correctly found by the ABISs). This issues is likely to have occurred due to operators not following due process. It has now been resolved with newer versions of the enrollment client software that includes checks to prevent this situation.
2. Anomalous biometrics: In this case, each modality is consistent, but different modalities have been captured from different individuals. Among the duplicates this issue was also relatively frequent as it accounted for approximately 20% of the duplicates (again, correctly found by the ABISs). Improvements in the enrollment client software will also reduce the incidence of such cases.

This analysis has resulted in UIDAI taking corrective action and building checks and balances in the enrollment client software so that the issue of inadvertent submission of duplicates can be eliminated in the field at the client-side system. The advantage of addressing such issues at the client side is that the operator and resident are still present at the enrollment location and can take corrective action immediately.

5.2 Impact of biometric sample quality

It is well known from the biometrics scientific literature that quality of biometric samples play an important role in the accuracy of biometric matching. For this reason, it was deemed important to measure and control the quality of biometric data in both the client-side system and the server-side system. The client-side system has checks and balances (for example, automated capture that determines presence of good quality biometric, quality-based re-capture, etc.) that enforces capture of good quality data from the field. The server-side system aggregates the quality information and analyzes it with respect to different geographic regions, enrollment agencies, individual operators, biometric devices, etc.

At the time of writing of this report, the following measurements were obtained:

1. **Poor quality fingerprints**: 2.9% of residents were measured to have poor quality fingerprints as defined by their fingerprints yielding a score of 4 or 5 (on a scale of 1-5 with 5 being the worst quality) by the National Institute of Standards and Technology (NIST) NFIQ fingerprint image quality algorithm. By examining the records with low quality fingerprints, it is observed that majority of people who have poor quality fingerprints actually have good quality irises. It is important to note that for this group of people, the poor quality of their fingerprints alone does not degrade the de-duplication accuracy of the multi-modal biometric system.
2. **Poor quality fingerprint and poor quality irises**: 0.23% of residents have both poor quality fingerprint and poor quality irises. These are the residents who are susceptible to errors from the multi-modal biometric de-duplication system. Therefore, UIDAI team pays significant amount of effort in refining processes and building checks and balances in the system to measure, control, and reduce poor quality biometric data. It should be noted though that the state-of-the-art ABISs (such as those procured by UIDAI) are well versed in dealing with poor quality biometric data. Yet, by keenly measuring the quality and continuously improving the process that improves the quality of biometric data, UIDAI is making sure that the quality of collected biometric data stays high and does not degrade. This gives us the confidence that the system will scale to the entire population with the same quality of biometric as measured at the time of writing of this report.

6 Conclusion

In December 2009, UIDAI committee on biometrics published its report titled "Biometric Design Standards for UID Applications". The committee acknowledged that most other large-scale biometrics deployments were fingerprint-only and a fingerprint-based system may present challenge in India due to high manual labour practiced by majority of the population. The committee therefore held extensive meetings and discussions with international experts and technology suppliers. A technical sub-group analyzed fingerprint data collected from Delhi, UP, Bihar, and Orissa and found that the quality of the data was actually not substantially different from the western population. The committee said that it is possible to improve the accuracy of fingerprint system especially considering the need to scale to 120 crore population by additionally using iris. "*Iris can provide accuracy comparable to fingerprint. Therefore fused score of two uncorrelated modalities will provide better accuracy than any single modality and could achieve the target accuracy.*"[12]

In December of 2010, UIDAI published a report titled "UID Enrollment Proof-of-Concept Report". The report documents the findings of enrollment proof-of-concept study commissioned by UIDAI in three rural areas of Andhra Pradesh, Karnataka, and Bihar. Among many interesting findings on both process and technology, the report says "*The biometric matching analysis of 40,000 people showed that the accuracy levels achieved using both iris and ten fingerprints were more than an order of magnitude better compared to using either of the two individually. The multi-modal enrolment was adequate to carry out de-duplication on a much larger scale, with reasonable expectations of extending it to all residents of India*"[13].

As of December 2011, the UIDAI has true and tested statistics computed from real operational large-scale UIDAI system at a gallery size of 8.4 crore (84 million), which is more than 4,000 times the sample size that was available at the time of enrollment PoC. There is no longer a need to rely on small sample size tests or hearsay from other projects. The UIDAI is now capable of measuring the accuracy, performance and scalability of the actual production system, which is already among the largest in the world. The analysis resulting from such a large data set is empirically repeatable and statistically accurate. Based on the analysis, it can be said that the enrollment system has proven to be reliable, accurate and scalable to meet the nation's need of providing unique Aadhaar numbers to the entire population. Specifically, the following are observed:

1. **Failure to Enroll(FTE) Rate:** Zero. As a policy, every unique resident, regardless of their biometrics can be enrolled and issued Aadhaar number.
2. **Biometric Failure to Enrol (B-FTE) Rate: 0.14%.** This implies that 99.86% of the

[12] http://uidai.gov.in/UID_PDF/Committees/Biometrics_Standards_Committee_report.pdf
[13] http://uidai.gov.in/images/FrontPageUpdates/uid_enrolment_poc_report.pdf

population can be uniquely identified by the biometric system. The exceptions (0.14%) however can still be de-duplicated using demographic data and checked manually for fraud. The legitimate cases among these will be issued Aadhaar number.

3. **False Positive Identification Rate (FPIR): 0.057%**. In practical terms, it means that at a run rate of 10 lakh enrolments a day, only about 570 cases need to be manually reviewed daily to ensure that no resident is erroneously denied an Aadhaar number. Although this number is expected to grow as the database size increases, it is not expected to exceed manageable values even at full enrolment of 120 crores. The UIDAI currently have a manual adjudication team that reviews and resolves such cases.
4. **False Negative Identification Rate (FNIR): 0.035%**. This implies that 99.965% of all duplicates submitted to the biometric de-duplication system are correctly caught by the system as duplicates. Given that currently approximately 0.5% of enrolments are duplicate submissions, there would only be a very small number of duplicate Aadhaars issued when the entire country of 120 crores is enrolled.
5. **Scalability**. The system is currently processing 10 lakhs (1 million) enrolments a day with enrolment database (gallery) of 9.8 crore (98 million). It has scaled (grown) as expected. The additional computing power required to handle increasing number of enrolments will not grow at an abnormally high (non-linear) rate; it is well within the design and expectations of the UIDAI.

The key measures reported above have also been computed at different gallery sizes up to 8.4 crores. Based on the trend, the UIDAI expects the accuracy of the system to remain within the same order of magnitude as reported above. It is now safe to conclude that the system will be able to scale to handle the entire population. The results lay to rest unfounded claims that the underlying technology is untested, unreliable and based on unproven assumptions

It is the policy of Aadhaar to maintain continuous quality improvement. Consequently, UIDAI will continue to monitor performance, adjust parameters as needed and institute new processes and procedures to not only maintain the currently low error rates, but to even improve system performance as the system grows to the ultimate goal of 120 crore.