Aadhar-Based Biometric Identification System for Tracking Missing Persons

KARUNA MIDDHA[1], MAAHIR SHARMA[2], DAKSH DUDEJA[3], AMRITA PANDEY[4]
[1] *Professor - Department of Computer Science, Maharaja Agrasen Institute of Technology*
[2, 3, 4] *Student, Department of Computer Science, Maharaja Agrasen Institute of Technology*

Abstract- ***This application is associated with the Aadhaar biometric information system developed to integrate information in order to combat human trafficking, especially child trafficking. It works through a dedicated portal with restricted access to authorized authorities only, all state and district-level police authorities may share the Aadhar-enabled data on missing persons. (Such as their photo, age, and fingerprints). The approved authorities may use this information to look for missing people in areas under their control. An integrated and validated information system would assist local police departments in locating missing people. This study shows that machine learning has a high potential to curb this problem due to ML's ability to mine, search and analyze big sets of data, performing matching tasks more quickly and reliably than the conventional methods.***

Indexed Terms- DAM, Aadhar Based Biometrics Identification, Biometrics, Database, Machine Learning, KNN Algorithm

I. INTRODUCTION

The high rate of abduction and trafficking cases in various locations is one of the main causes of missing individuals in India.

State/UT police get a large number of complaints of missing individuals each year. People can disappear for a variety of causes. Following their absence, some of them come back soon after without suffering any harm. Some of them, though, may have met a horrible end due to violence, suicide, or an accident. They can have also fallen prey to criminal activity like human trafficking. It might be challenging to determine whether a disappearance is purposeful or accidental. Both adults and children, as well as both genders, may be victims of human trafficking. Both domestic and international criminality may include people trafficking. It involves victim exploitation for labour and sexual purposes. The analysis of the literature demonstrates that women and girls are trafficked for sexual exploitation, and boys are just as exploited as camel jockeys. According to the literature study, men and women who leave their homes in quest of better living conditions do so for a variety of reasons, including low educational attainment, poor work prospects, and a lack of chances.

Many investigations have shown that a significant number of people vanish each year. Every missing person incident carries certain inherent risks, but certain groups of people are more vulnerable to violence than others. There are several reasons why someone could go missing, including mental illness, poor communication, miscommunication, misadventure, domestic violence, and being a victim of crime. The vast majority of people who have been recognized as victims of trafficking for sexual exploitation and 35% of those who have been identified as victims of forced labour are women and girls, according to the UNODC Global Report 2018 on human trafficking. Men account for more than half of the victims of forced labour trafficking at the same time.

Fortunately, a lot of people who are reported missing to the police are found quickly. However, some people are never discovered or end up being recognised as victims of crime or unfortunate circumstances. Additionally, there are those who are missing but have not been reported to the police or looked into. Some young kids flee their homes because of intolerable abuse and neglect. They become more susceptible to hazards of exploitation and criminal activity, such as prostitution, drug addiction, trafficking, and violence. Numerous

missing people have perished tragically in acts of violence, suicide, or accidents. Many missing people become victims of trafficking. It is not always simple to determine if a person's disappearance was done intentionally or not. Since it might be difficult to match a John Doe to their original identity just based on their things, biometrics are typically used in cases when missing persons are frequently discovered unconscious or in appalling conditions. The fingerprints of an unidentified victim are frequently used to identify them. Because there are millions of records, this procedure might take many days because the fingerprints are first scanned and transmitted to the Aadhar Team for matching. For several Law Enforcement Agencies to locate missing people in a timely manner, a centralised search platform is a great necessity.

Biometrics, often known as the identification of a person based on physical or behavioral traits, The identification of the iris, and the representation of hands or fingers are examples of physiological traits. The features that make up behavior are learned or acquired. dynamic voice verification, dynamic keystroke dynamics, and dynamic signature verification. There isn't one "ideal" biometric that can satisfy all requirements. Each biometric system has benefits and drawbacks of its own. However, for a biometric system to be useful, certain traits must be present. The biometric must first be founded on a distinct characteristic. Numerous scientific studies back up the claim that "no two fingerprints are the same."

In Fingerprint Recognition, the user coupled to a reader gently touches his finger against a small reader surface (optical or silicon) during the time of verification for less than 5 seconds; the reader is approximately 2 inches square. The reader is a computer that receives data from the scanner, transfers it to the database, and then compares it to the data there. The Automated Fingerprint Identification System (AFIS), which collects and stores fingerprints in the United States as well as other nations including Canada and the United Kingdom, has a database of fingerprinting methods. Fingerprints are individual to each person. Due to its high level of dependability, this approach is crucial.

The only legitimate way to identify someone in India is via their AADHAR card. The physical and behavioral traits that can be kept in a database and compared during verification and identification are what biometrics and AADHAR are all about. Different aspects of the iris and fingerprints can be utilised to identify a person uniquely. The pin is nothing more than a biometric identification that is visible yet immovable. In this situation, seeding the AADHAR number with a bank account is required. The data is encrypted on several different layers. There are no remnants on the disc once it is decrypted in memory. Numerous security procedures are in place to secure the data center. All information will be kept in the Central Identities data repository (CIDR). In the event of the invalid access, a notice will be transmitted over GSM to higher authorities.

II. EXISTING SYSTEM

In case a police station finds an unidentified child, or a dead body that can’t be identified, the biometrics (Fingerprint and/or Iris) scans are sent over to the Cyber Cell which raises a request to search CIDR - the team then searches through the Database & tries to narrow the results down to records with the highest matching %. The UIDAI of possible records are extracted & the full face corresponding to them is matched. The final results can take over several days to arrive, this is a tedious process and needs to through multiple agencies & steps before even reaching the search stage.

III. PROPOSED SYSTEM

It is evident throughout our study that there is a need for a centralised searching system that can give access to the results of Biometrics in a fast & efficient way. We propose a web-based platform where authorised officials can fetch/feed lost and found person’s data, along with keeping track of that case using data linked with an Aadhar Card. The whole process is a multi-stage mechanism that involves:

A) Creation of CIDR Partitions:

The CIDR contains information about all the individuals who have registered for AADHAR. To facilitate faster searches in large databases, a

technique known as partitioning is employed. Partitioning refers to dividing data into chunks and storing them in such a way that only appropriate queries are directed to each partition. The major partition is going to be partitioning on the basis of Gender - Male & Female. Then creating partitions according to the regions with the highest rate of trafficking/murders / missing people reports.

The search queries over these partitions will be driven in an ordered way:

i) Active Number of Trafficking Cases in State/Region

ii) Active Number of Kidnapping Cases in State/Region

iii) Active Number of Missing Person Cases in State/Region

The data required to decide the order for these partitions will be feedback driven in accordance with the inputs received from Annual Surveys on Trafficking, Kidnapping & Missing People conducted by Law Enforcement Agencies. The data for the representation in the figures below have been taken from Crime in India (CII), an annual publication of the National Crime Records Bureau (NCRB). Crime in India contains data on police-recorded crime during the year. While recording missing persons in the system, the real motive is normally not known and becomes clear later during the investigation stage when the person is recovered. Motive is therefore not factored in the present analysis.

Representative Data in Figure (2) and Figure (3) can be utilized to decide the order in which partitions will be chosen from the directory and searched for. Scanning for States/Regions with a high number of missing/trafficking cases increases the probability of finding a match early on in the search stage.

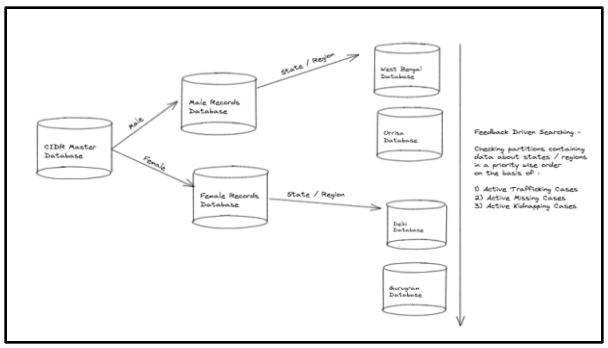

Fig 1: Multifactor Partitioning of Master Dataset

2016		2017		2018	
State/UT	Cases	State/UT	Cases	State/UT	Cases
Maharashtra	28316	Maharashtra	29279	Maharashtra	33964
West Bengal	24937	West Bengal	28133	West Bengal	31299
Madhya Pradesh	21435	Madhya Pradesh	26587	Madhya Pradesh	29761
Delhi UT	12067	Delhi UT	12202	Delhi UT	13272
Tamil Nadu	9596	Rajasthan	10230	Rajasthan	12525
Telangana	9238	Tamil Nadu	9564	Tamil Nadu	10403
Rajasthan	8414	Karnataka	8757	Odisha	10193
Karnataka	8092	Telangana	8405	Karnataka	9567
Gujarat	7105	Gujarat	7712	Chhattisgarh	9412
Chhattisgarh	6649	Chhattisgarh	7383	Gujarat	9246

Fig 2: States with Highest Number of Missing Women

2017		2018	
District	Missing Women	District	Missing Women
Mumbai Commr.	4718	Mumbai Commr.	5201
Pune Commr.	2576	Pune Commr.	2504
Thane Commr.	1798	Thane Commr.	2352
Pune Rural	1559	Nagpur Commr.	1645

Fig 3: Districts of Maharashtra with Highest Number of Missing Women

B) Web Portal for Law Agencies:

The Web Based Portal gives the agencies a way to raise a search query on the CIDR Database. The Police official can upload a FingerPrint Scan and submit the query in a few steps. The internal work involves:

i) Gender Classification -

Step 1 -The fingerprint undergoes pre-processing i.e. noise removal, cropping, etc.

Step 2 - The fingerprint is converted into a grayscale image.

Step 3 - The greyscale image is normalised to 156x156 and defines matrix co-occurrence.

Step 4 - The local binary pattern statistical features were extracted.

Step 5 - Apply the K-NN classifier and find the class of the unknown fingerprint by using the database generated in LBP.

The K-nearest neighbor algorithm (K-N N) classifies the objects of training samples that are closest in feature space for the categorization of gender. An item is categorised by majority voting of its neighbors, with the objects allocated being the most prevalent class among its K closest neighbors (K is a positive integer which is typically small). When K=1, an algorithm is thought of that assigns the item to the class of its closest neighbor without performing an explicit training step. The fused feature vector of the fingerprint input is compared with the feature vector in the database during the final classification phase using the KNN classifier.

ii) Partition Matching -
Based on the predicted gender, the query is routed to the partition which matches the target's gender, furthermore, the Backend Service keeps the partition ID's in a separate local storage along with the priority. The priority for each partition is decided by the sum of active kidnapping/missing/trafficking cases which are associated with a state, similarly, each sub-partition also contains the priority of regions within a state. This targeted searching strategy increases the probability of finding a hit early on in the searching stage.

iii) Fingerprint Matching -
The target's fingerprint must be compared to each partition's fingerprint dataset in the absence of UIDAI. Utilising minutiae extraction and false minutiae removal is accomplished.

Minutiae Extraction: The minutiae point is extracted using a 3x3 pixel window. Pixels are the 0s and 1s values that make up a 3x3 window. It is referred to as a bifurcation point if the center pixel with a value of 1 matches its three neighbors with a value of 1, else it is a ridge terminating point. However, it is a straightforward procedure with a few false bifurcation sites. This approach is inaccurate.

False Minutiae Removal: False minutiae are the points that are caused by gaps, breakdowns, noise, etc. False minutiae points are occasionally eliminated from images during post-processing since the enhancement process sometimes makes it possible to do so. These actions erase them. 1) Determine the average separation (D) between two adjacent ridges. 2) Next, determine how far apart two-minute marks are from one another. They might be both points of bifurcation, one point of bifurcation and one termination, or both points of termination. The computed distance is designated as d. 3) Remove both of the minutiae if d is less than D and they are in the same ridge. 4) After marking true minutiae points and erasing false ones, the picture is identified using these.

iv) Report Generation -
Finally a report is generated by finding the UIDAI associated with the fingerprint, the UIDAI enables access to various key information about the target including Home Address, Father's Name, Associated Bank Accounts, Contact Number, Emergency Contact Number etcetera.

The PDF Report is generated & sent to all the officials working on the case via Email & SMS using Twilio Voice/SMS/Email API.

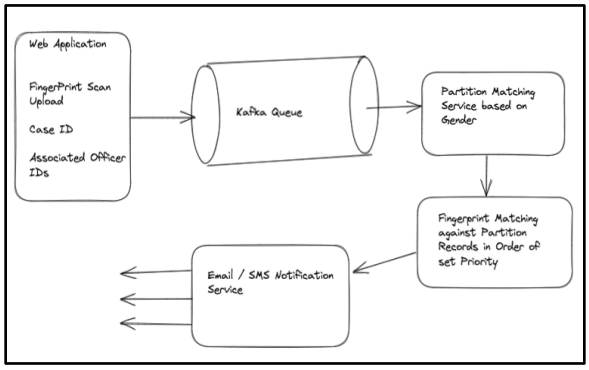

Fig 4 : High Level System Design

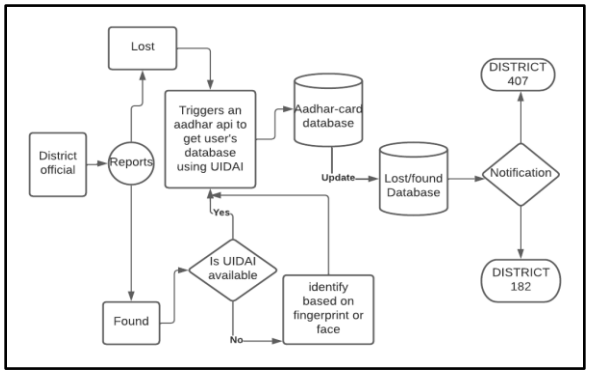

Fig 5: System Flow Diagram

IV. ADVANTAGES

1) This system creates a centralised application for Law Agencies to identify a Missing/Trafficked/Dead person/child.
2) By leveraging priority-based searching and partitioning the data based on State/Region, we can ensure that the data is found in the most accurate & efficient way possible.
3) This is a scalable & efficient distribution system that ensures that no time crucial query is dropped. A Messaging Queue like Kafka has been integrated into the system to ensure that all the requests are processed in an asynchronous way.
4) Multiple searches can occur based on the hosted system capabilities, whenever a Kafka Record is consumed, the service uses up a thread from a ThreadPool and submits the task of Partition Matching to the thread, by doing this we ensure that a multithreaded environment is being utilised for concurrent processing of queries.

V. FUTURE SCOPE

1) Heavy Machine Learning Models need to be deployed on the cloud, this is a cost extensive operation. There are several ways to counter this using Cache Systems for commonly queries clusters.
2) Some manual input from Law Enforcement Agencies is still required as the Gender Classification & Fingerprint Matching Models can never be 100% accurate - there might be still inaccuracies due to Image Quality, various Image Enhancement Algorithms can be implemented to make the model more accurate.
3) Data Mining models can be integrated to automatically mine annual/semi-annual crime reports from the publically available Government Websites & classify the cases, after that, an aggregated dataset can be built to find out the number of active kidnapping/missing/trafficking cases in a district/area at fixed periodic intervals. This will be used to set cluster priority while searching.

CONCLUSION

By creating a centralized application & adopting a priority-based biometrics search system, we can improve the performance of the existing system. Our application also brings down the manual effort required from the Law Agencies' perspective, maintenance of a previous record of searches is also helpful in generating reports - which can be used for future references & legal purposes.

The usage of distributed asynchronous system methodologies like Messaging Queues, Thread Pools & Rate Limiters within the system also ensures the availability of the service - which is essential in time-crucial queries.

REFERENCES

[1] A. Jain, L. Hong and S. Pankanti, "Biometric Identification", Communications of the ACM, vol. 43, no. 2, pp. 91-98, 2000. Show Context Access at ACM Google Scholar

[2] Anil K. Jain and Arun Ross, "Introduction to Biometrics" in Handbook of Biometrics, Springer, pp. 1-22, 2008, ISBN 978-0-387-71040-2. Show Context CrossRef Google Scholar

[3] Sahidullah and Md, Enhancement of Speaker Recognition Performance Using Block Level Relative and Temporal Information of Subband Energies, 2015. Show Context Google Scholar 4. "Biometrics for Secure Authentication", (PDF)Retrieved. Show Context Google Scholar

[4] Report on Missing Women & Children in India - National Crime Records Bureau

[5] S. Prabhakar, S. Pankanti, A. K. Jain ”Biometric recognition: security and privacy concerns” IEEE Security and Privacy, pp. 33-42, March/April 2003

[6] L. O’Gorman “Comparing passwords, tokens, and biometrics for user authentication”. Proceedings of the IEEE, Vol. 91, No. 12, pp.2021-2040, December 2003.

[7] M. Faundez-Zanuy “Biometric recognition: why not massively adopted yet?”. IEEE Aerospace and Electronic Systems Magazine. Vol.20 nº 8, pp.25-28, ISSN: 0885-8985. August 2005.

[8] M. Faundez-Zanuy “Privacy issues on biometric systems”. IEEE Aerospace and Electronic Systems Magazine. Vol. 20 nº 2, pp13-15, February 2005.

[9] "Biometrics for Secure Authentication", (PDF) Retrieved

[10] Website: http://www.bioelectronix.com/what_is_biometrics.html

[11] Website: http://www.biometricnewsportal.com/biometricsissues.asp.

[12] M. Faundez-Zanuy “Are Inkless fingerprint sensors suitable for mobile use? IEEE Aerospace and Electronic Systems Magazine, pp.17-21, April 2004