Criminal Check Program with an ID, Face, and Fingerprint Recognition

E. Niyonsaba
3962056
Robert Sobukwe Rd, Bellville
Cape Town, 7535
3962056@myuwc.ac.za

ABSTRACT

Public and private companies need an online way to check the criminal record of an applicant. There must be an online application, that connects to Department of Home Affairs (DHA) to authenticate the application using fingerprint, South African Identification number/ passport number and picture of the applicant's face using facial recognition. The application must also connect to the South African Police Service (SAPS) to check whether the applicant has a criminal record. Our project deliverable is to develop an application VeriCrimCheck that will allow prospective employers to check an applicant's criminal status. The application will use the following criteria: biographical information, biometric information (fingerprint) and facial recognition to build a profile of an application. To improve this application, the use of cloud computing will be essential.

1 INTRODUCTION ABOUT A BIOMETRIC SYSTEM

Over the centuries, many organizations put a lot of attention on expanding employment [1]. In order to get a job, the applicant must present an extract from his or her criminal record and this will determine whether the person in question can be hired [1]. In South Africa, as in other countries that have not yet benefited from the use of computer technology, the process of presenting the criminal record to the company in which employment is sought is done manually [2].

Our system will verify the identification and criminal status of individuals in using online databases (Department of Home Affairs and South African Police Services) relying on fingerprints, facial images and South African identification number (ID). The latter will meet the requirements of uniqueness of the identified person. Even though the criteria used are not the basis for the system's performance. As recognizing a face by a computer system is fast and there is no reliability [3], the use of fingerprints will provide this reliability. When necessary to retrieve data from the databases both parameters will be completed by using only the ID which is the traditional type of verification. The development of the prototype contains the use of faces and fingerprints as well as biographical details including the ID to check the criminal record of the applicant for a job.

In order for the transactions to be reliable, automatic identification must be used. Traditionally, a personal identification number (PIN) was used to ensure data reliability. Currently, this method of identification is no longer used in most businesses because it can greatly facilitate fraud [4]. Biometrics is used in reliable identification [4][5] with condition that cannot be changed : there must be no replacement of the following characteristics universality, uniqueness, permanence, and collectability [6][4]. There are a lot of questions we will need to ask ourselves when implementing this system. These questions are listed as follows [3][4]:

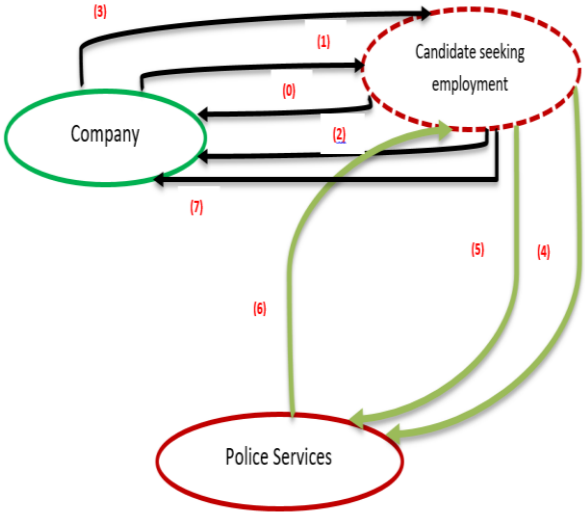
1) performance, referring to specifying a person's identification in terms of speed and robustness,
2) acceptability, referring to the measure individuals must follow to agree with biometrics in their daily activities and experiences, and
3) circumvention, showing how the system can be deceived using fraudulent methods.

This practical biometric system we are developing could accomplish the following:

1) accuracy and assurance of identification that must be accepted with reasonable resources,
2) avoid harm to subjects and ensure acceptance by the target population, and
3) sufficiently resist various fraudulent methods. Nine biometric techniques are to be used in current biometric systems. These are: face, facial thermogram, fingerprint, hand geometry, hand vein, iris, retinal pattern, signature, and voice print [5].

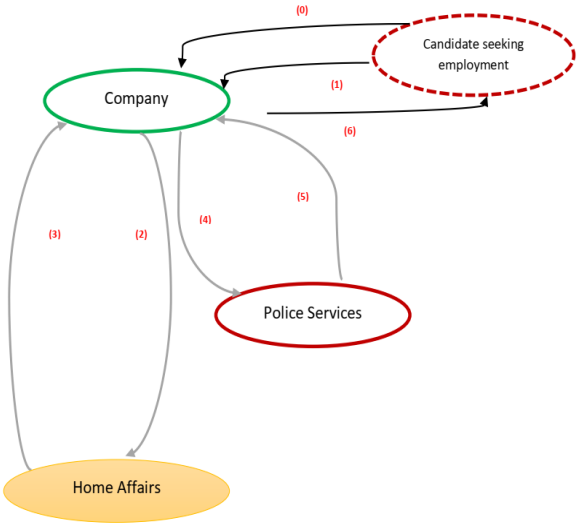
For our project, only the two will be our concern: the face and the fingerprint which will be accompanied by the traditional

identification, the identification number.



Legend:
(0): receipt of the form to be completed;
(1): complete the form;
(2): submit the form for signature by the company's agent;
(3): Repeat of the full form and signed by the candidate;
(4): bring the complete and signed form to the police for verification of the criminal status;
(5): Receipt of the form by the police and verification of the person's criminal status;
(6): reception of the result;
(7): return to the company for the filing of the certificate of non-sufficiency and decision making.

New Data Flow Diagram for Searching candidate's details.



Legend:
(0): receipt of the ID by the company agent;
(1): complete the for and take the picture and the fingerprint;
(2): sending the request to the home affaires database;
(3): display of the details from the home affaires database;
(4): send the request to the police to find out the details of the crime;
(5): display of the crime details from the police services database;
(6): proclamation of the decision taken by the company;

Conceptual Data Model

The Conceptual Data Model (CDM) is a representation of all the data in the domain, which does not take into account the technical and economic aspects of storage and access and without referring to the conditions of use by a particular processing operation. Its purpose is to formally write down the data that will be used by the information system.
Before building this model, it is necessary to make an inventory of the data from which redundancies, synonyms and polysemy are eliminated.

Rules of Criminal State Management by the police
The specification of management rules, carried out in parallel with data collection, allows the establishment of relationships between information. The rules for managing a person's criminal status are represented in the following table:

Rules for the Management of Criminal Records Extract Records

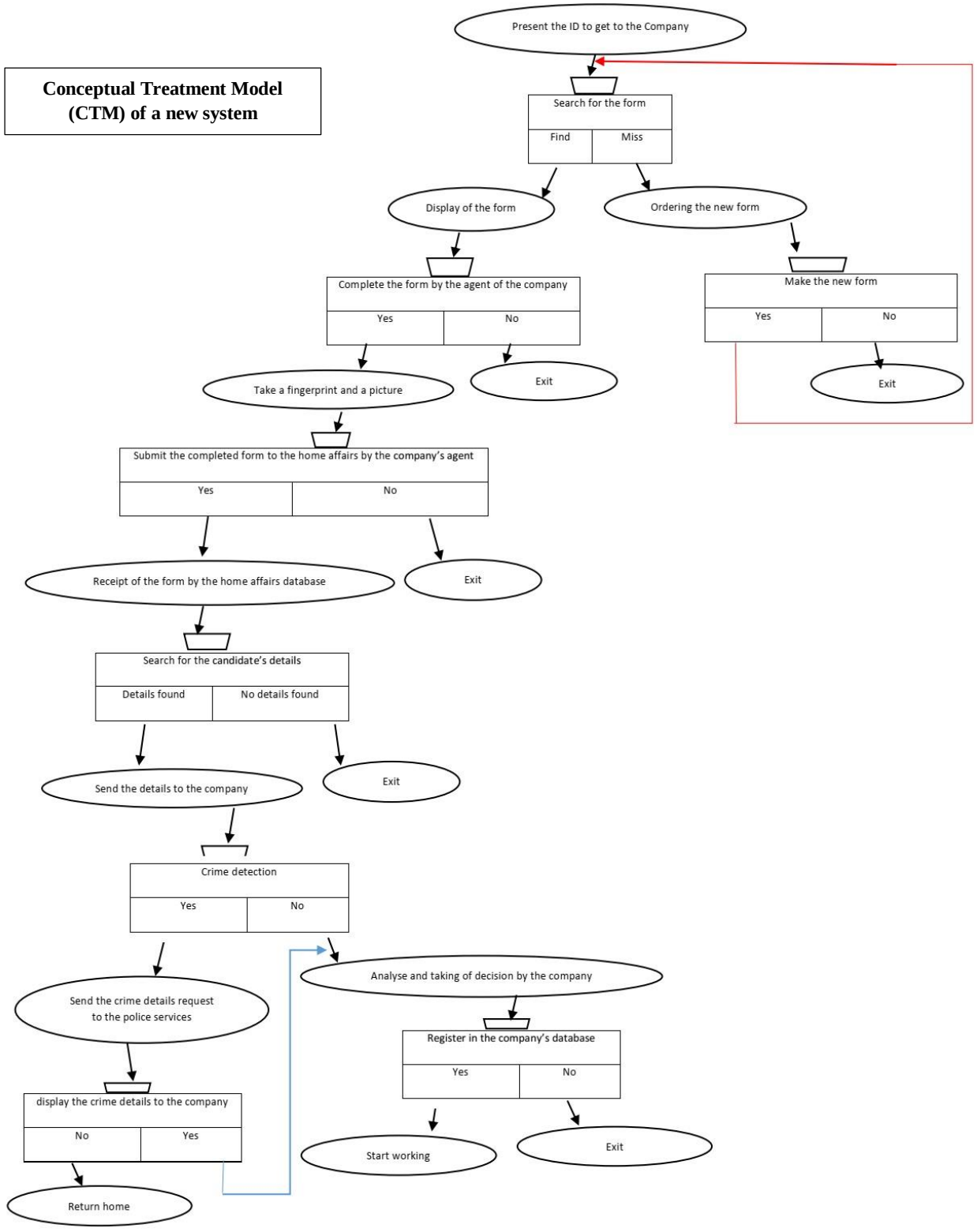
Rule number	Label
1	The South African Police Services database which deals with the crime only keeps fingerprints of people with unclear criminal records.
2	If a service requests the criminal status of the person in the database, a query is made in the database through the person's fingerprint; this query searches and returns the person's details if they exist in the database, otherwise it returns nothing.
3	Once the details are displayed, the certificate is printed and the session is closed.
4	A person's details are only searched when they want to make changes to them to get a criminal record or if they need to consult them in order to get some information.

Conceptual Treatment Model (CTM)

2 APPLICATION CONCEPTUAL TREATMENT MODEL

The conceptual processing model leads to the determination of processes, means homogeneous units of concern. It represents events, results, operations and synchronizations. It describes the actions on the information system with their

triggers that make it possible to bring this mass of data to life.
It is a dynamic description of reality.



3 A GENERIC BIOMETRIC SYSTEM ARCHITECTURE

The system architecture to be implemented is made up of two modules: the enrolment module and the identification module, represented in Figure 1:

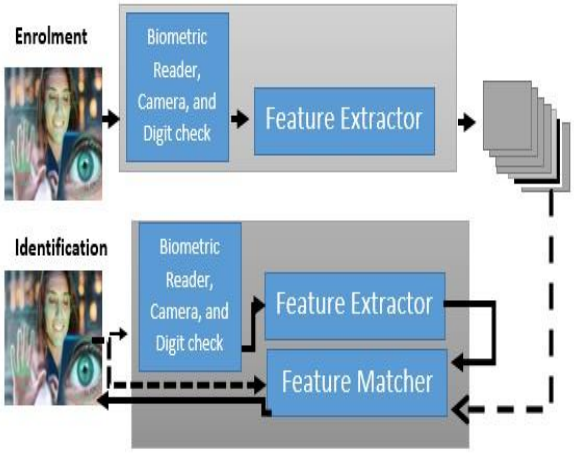


Figure 1:A generic biometric system architecture

The enrollment module is characterized as an information capturing component using fingerprint reader and the camera, as well as capturing the biological details of the person applying for the job. The application will store the collected individual information in a database if s/he is newly registered, followed by the identification component.

This system of checking the criminal status of a person seeking employment follows three techniques, face recognition, fingerprint recognition and identity number check.

4 FACE RECOGNITION AND FINGERPRINT VERFICATION

4.1 Face Recognition

The action of recognizing faces is classified in an active research area. Applications in this field start with the static and controlled verification of forensic identification photographs and end with the dynamic and uncontrolled identification of faces in a crowded environment [7]. For the identification of a person in a computer system, faces are recognized by referring to the static and controlled recognition of facial portraits of the person to be searched [7]. Research has shown that there has been the development of algorithms for recognizing faces and that these algorithms are capable of providing satisfactory solutions to the problem of checking a person's face in real time [8].

4.2 Fingerprint Verification

By definition, a fingerprint is a pattern of ridges and grooves on the surface of a fingertip. The fingerprint is formed from accumulated dead cells having continuously flowing carnifications characterized by the scale-like shape of the exposed surface [9]. Fingerprint formation occurs when the child is still in the embryo in the early days of foetal development [10]. The use of the fingerprint by humans to make an identification a person has been used for many years. It is used for identification because it has biological properties that are easy to understand. The following list gives a summary of the properties of fingerprints:

1) Every person on planet earth has ridges and grooves in the skin that are unique to him/her, meaning there is no person with the same fingerprint characteristics as another person;
2) We change the configurations according to the individual. But these configurations can give a variety in less than the limits that allow a classification that is systematic;
3) There is a permanence of patterns and minute details of individual ridges and grooves and they remain unchangeable over time, but they may change if there are common injuries, scratches and scars [3] [10].

The local characteristics of the ridges and the determination of the uniqueness of a fingerprint is made by their relationships. Different local ridge features can be identified; these features are called minutiae [9]. The uneven distribution of these local features is noted. Many of them depend on the conditions under which they are and the quality of the fingerprints and are not observed most of the time in fingerprints. The following are the important features, and we mention two: the ridge termination and bifurcation of the ridge. These features are also called minutiae. A ridge termination is defined as the point where a ridge ends abruptly. A ridge bifurcation is the point where a ridge takes a point of divergence into branch ridges. A human fingerprint typically has about 40 to 100 minutiae [3]. The minutiae of a fingerprint can have the following characteristics: type, x and y coordinates and direction.

To verify fingerprints, the two main steps are unavoidable: extraction of minutiae and comparison of minutiae [9]: To identify automatically a fingerprint, the endings and bifurcations of the ridges are usually identical when one tries to compare them with each other [4].

The list of factors that explain the lack of well-defined ridge structures is as follows: abnormal skin ridge formations in fingerprints, post-natal marks, occupational marks and problems with acquisition devices.

So, an algorithm that reliably extracts minutiae would not have to assume perfect ridge structures and would have to degrade with the good quality of fingerprint images. A minutiae

extraction algorithm that we have proposed is made in three steps:

1) Estimate the orientation field. This is the step in which we observe the orientation of the fingerprint images taken at the entrance and there is the location of the region of interest,
2) extracting the ridges, consisting of extracting and thinning the ridges, and
3) detecting and final processing of characteristic points. These steps are also characterized by the extraction of thinned and refined ridge maps. The recording of the following parameters takes place at each detection of the characteristic point: x-coordinate, ycoordinate, the orientation, being defined as the local orientation of the associated ridge [3].

5 DECISION MERGING

The merging of decisions with many clues has been shown to improve the accuracy of a recognition system [8] . Generally, a multitude of clues can be merged using a next level found in the following [8]:

1) Level called abstract. The one that marks the output of each module. It is a set of possible labels with no confidence associated to the labels; in this level, one can use the simple majority to make the decision more reliable [11].
2) Named level of ranking; here we see the possible labels ranked in descending order of confidence, but there is no specification of confidence values;
3) Named level of measurement: at the output of each module, there is a set of possible labels. These labels have associated confidence values; here, more precise decisions can be made by integrating the different confidence measures into a confidence measure giving accurate information.

This prototype of the application we are designing and developing contains the merging of the decisions that will make the system work and provide the stand out level of the measure. There will be the verification of the first n possible identities that are established by the face recognition module and this verification will be done by the fingerprint verification module. In order for such a merging scheme to be possible, we need to define a measure indicating the confidence of the decision criterion and a criterion used to facilitate the merging of decisions.

In this application we need a matching algorithm to make the decision. This algorithm is called characteristic point matching. It is proposed to operate with the fingerprints taken by the fingerprint reader and a template. The model is based on comparing the normalized number of link sets of two minutiae that will match against a threshold. From the above an assumption will be made that the characteristic points in the finger region of the fingerprints of different persons must have a random distribution and there will be the probability that an impostor, represented by the letter I, is received by the system is {1 G(yI)} where yI represented the number of points taken two by two between the impostor and the person he intends to be.

6 DISTRIBUTION OF IMPOSTERS FOR FACE RECOGNITION AND FINGERPRINT VERIFICATION

It is very difficult to determine the distribution characteristics of the imposter to recognize the faces. The Face Verification module keeps the first numbers of matches to improve the likelihood of identifying an individual in the database. This is because of the ability to discriminate the weakness in recognizing faces.

The module increasingly classifies the first numbers (n) of the results of the Distance from Feature Space (DFFS) values. If the value the DFFS is large, the probability of the match is incorrect, so for it to be correct, the DFFS value must be small. The use of relative DFFS values is more preferred than the use of absolute DFFS values because the relative distances existing between successive DFFS values over time remain constant when compared to the average DFFS deviation.

In order for the first n results to have an incorrect probability of matching, there must be a change in position. When trying to represent this module by a function, the way to distribute imposters must be represented by a function that decreases the positon order.

6.1 Merging decisions

Distributing imposters for face recognition, distributing imposters for fingerprint verification and distributing imposters for the ID number check give assurance quantities for each of the first numbers of matches found by the face recognition module as shown on the Figure *2*. In order not to lose the general view, we take an assumption that at most one of the possible numbers of identities being established by the face recognition module for a person is the reality of a person's identity. Finally, the integration will decide either to reject all possibilities or to accept only one of these possibilities as an identity that is real. In practice, there is a specification that the False Acceptance Rate (FAR) that we will find in this application system must not be greater than the value taken [4]. As a result, the objective of merging decisions is to provide a criterion for deciding who ensures satisfaction for specifying the FAR [4].

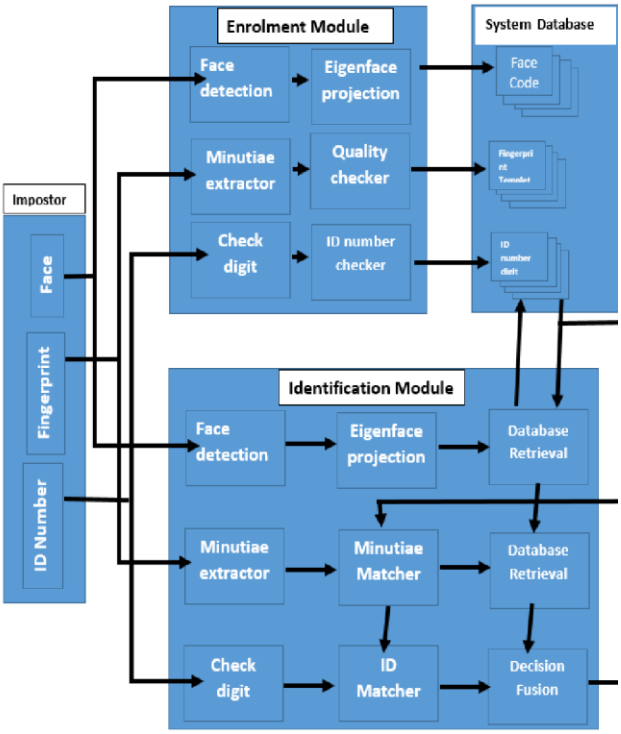


Figure 2: Architecture of the impostor's biometric identification.

The reason for taking an assumption that DFFS exists between two different individuals. The assumption that the two different individuals must have a fingerprint match score between them should not be overlooked. If two individuals have similar faces it does not mean that they have the similarity of fingerprints and vice versa. The above hypothesis should not be confused with the situation where impostors want to make tricks using counterfeit faces and/or counterfeit fingerprints.

7 Tools used

In the online criminal check project, tools used are as follow:
Programming language: C# (C/sharp) implemented in Visual studio. We are using this programming languages because of many reasons. The main reason is that we have a previous knowledge of the language in the bachelor projects. C# is an oriented language that has a big number of toolsets and frameworks to support it and all of them are stored by Microsoft. There many libraries to build websites, to implement security, to work with file system, … which are provided by .Net and is the open-source. Microsoft supports the C# language in a special way, spreading new features and improved syntax faster than other programming languages. C# is one of the languages that are popular, and is about the same as Java. C# is a very flexible language, which makes it a more advantageous language than others. This language with the .NET Framework and Visual Studio develops a multitude of applications that are used in everyday life. Some of the applications are: native Windows applications, REST APIS, mobile applications, web sites, ... This language ranks among the ten most appreciated by application developers. All the qualities mentioned above have pushed us to use this language as we considered it a good language to use.

For Fingerprint Enrollment, we are using Digital Personal U.are.U 4000B provided by the University, while in capturing the face we are using the webcam and biological information we use the .Net functions. We are calling required functions to register and check fingerprints and faces after the application will register, check and recognize the ID, Fingerprint, and Face. To register the fingerprint template is changed in a binary format which is easy to be stored in SQL Database and distributed. To identify a finger there is a comparison of templates of the fingers that are scanned from the application. For the face, the comparison is made between the new face and the old faces. Features used in our project are ID, Fingerprint, and Face capturing (for the enrollment module), ID, Fingerprint, and Face verification, and ID, Fingerprint, and Face Identification, for extract and match ID, Fingerprint, and Face. MySQL is used as a back-end data base.

MySQL is one of the most widely used databases around the world. Its popularity can be explained by its open source status, reliability, compatibility with major hosting providers, cost effectiveness and easy management. Many organizations take advantage of the secure state of their data and the high transaction support MySQL provides to ensure the security of their data and improve customer interactions. Nevertheless, this database presents some small challenges.

The following are the major reasons why we decided to use MySQL. Finally, we will mention in passing some of the challenges commonly encountered by users of this database management system. :

1 Transactions are secured: For MySQL, transactions are made in a single unit, meaning that the authorization of each step of the operation is impossible. If there is a failure in any of the transaction steps, the whole operation is cancelled. MySQL integrates all data so that users can perform online transactions without fear of losing their data. The transaction is not performed until all the steps of the process are completed and if there is an error, you have to start from scratch.

2. Scalability on demand

One of the benefits MySQL offers is unparalleled flexibility, making it easy to manage deeply-anchored applications with ease. This is also true in large data centers where large amounts of information are stacked up in a critical state.

3. High availability

MySQL is available on a constant basis, so companies that use it benefit from it all the time they need it, and this availability is the main feature of MySQL. Wherever you can search on MySQL you will find that it handles millions and millions of requests and many transactions while providing the garrison of even unique caches, indexes with a lot of text, and optimizing transaction speed.

4. reliability at all times

All companies have a common concern, protecting commercially sensitive information. MySQL secures data using its unique data protection features. It has data encryption that is powerful enough to prevent data from being viewed by people who do not have authorization on it. It also has a mechanism that is very powerful limiting the number of people who can access the server and is able to prevent the use of the application by users in front of their machines.

5.Quick start capability

To download and install MySQL, you use the time not exceeding 15 minutes. MySQL is exceptionally fast. It can self-manage some operations like: automatic restart, give itself space and automatically change configurations to facilitate data management.

MySQL has some drawbacks as we pointed out at the beginning.

The development using a long time, high costs when saving data, query caches, connection that is not constant, if you use LAMP. To manage the connection, the solution is to use XAMP as it is possible to be started or stopped using a single command, is easy to transport, and facilitates debugging of scripts.

Progress to Date:

Application prototype developed:

1.able to capture biographical details.
2.camera functionality implemented and facial recognition implemented
3.database connection (local databases, DHA and SAPS will follow as a Proof of Concept and if time and processes permits)
4.documentation ongoing

What needs to be done:

1.facial recognition- implement ML or AI for recognition. In essence, a need to match two faces from two images, one captured and one in the DHA/SAPS databases.
2.biometrics - need to purchase a fingerprint scanner
3. populate databases with test data

REFERENCES

[1] P. Bari, "Modeling the Internet Congestion Control Using a Smith Controller with Input Shaping *."

[2] J. Reid, "CRIMINAL RECORD CHECKS AND POLICE CLEARANCE CERTIFICATES IN SOUTH AFRICA." [Online]. Available: http://www.ifacts.co.za/wp-content/uploads/2018/09/18. [Accessed: 03-Mar-2020].

[3] L. Hong and A. Jain, "Integrating Faces and Fingerprints for Personal Identification," vol. 20, no. 12, pp. 1295–1307, 1998.

[4] L. B. Tran and T. H. Le, "Multimodal Biometric Person Authentication Using Fingerprint , Face Features Multimodal Biometric Person Authentication using Fingerprint , Face Features," no. September, 2012.

[5] M. H. M. I. Kabir, "A Simple Approach to Recognize a Person Using Hand Geometry," no. November, 2010.

[6] A. N. Marana, J. R. Falguera, and F. S. Falguera, "Biometrics for Human Identification," no. January, 2006.

[7] A. S. Tolba, "Face Recognition : A Literature Review," pp. 88–103, 2019.

[8] L. Hong and A. Jain, "Integrating Faces and Fingerprints for Personal Identification," no. November, 2014.

[9] K. R. Moses, P. Higgins, M. Mccabe, S. Prabhakar, and S. Swann, "CHAPTER AUTOMATED FINGERPRINT IDENTIFICATION SYSTEM (AFIS)," 2005.

[10] D. Impedovo and G. Pirlo, "Automatic Signature Verification : The State of the Art," no. November 2018, 2008.

[11] P. Identification, "BIOMETRICS Personal Identification in Networked Society BIOMETRICS Personal Identification in Networked Society," 2006.

Project Plan