Smart Card Identification Management Over A Distributed Database Model

Olatubosun Olabode
Department of Business Information System
University of Botswana, Olabode

Abstract: Problem statement: An effective national identification system is a necessity in any national government for the proper implementation and execution of its governmental policies and duties. Approach: Such data can be held in a database relation in a distributed database environment. Till date, The Nigerian government is yet to have an effective and efficient National Identification Management System despite the huge among of money expended on the project. Results: This article presents a Smart Card Identification Management System over a Distributed Database Model. The model was implemented using a client/server architecture between a server and multiple clients. The programmable smart card to store identification detail, including the biometric feature was proposed. Among many other variables stored in the smart card includes individual information on personal identification number, gender, date of birth, place of birth, place of residence, citizenship, continuously updated information on vital status and the identity of parents and spouses. Conclusion/Recommendations: A conceptualization of the database structures and architecture of the distributed database model is presented. The designed distributed database model was intended to solve the lingering problems associated with multiple identification in a society.

Key words: Distributed database, national identity, biometrics, identity card, fragmentation, smart card and fingerprint, Boston’s Health care for the Homeless Program (BHCHP), primary-foreign key, Data Communication (DC), Universal Serial Bus (USB)

INTRODUCTION

Setting up an explicit national ID card in the Nigerian States is an athema for government of Nigeria for a very long time. The Federal government has made concessive effort toward increased verifiability and standardization in the area of personal identification by establishing the National Identification Commission. Nigeria have expended so much money at ensuring every individual in the country is in the possession of an ID card. But this effort left so much to ask.

Bali (2009) notes that, the objective of a national ID card is to provide a universal government-issued document that validates an individual’s identity and that can then be used for purposes of identification in various functions. In the contribution of (Ojaide, 2010), the National Identification system is needed in many governments sectors for an effective application and execution of their governmental policies and programs. In sectors of the economy like tax administration, immigration system, healthcare system, driving licence administration, educational system, financial and banking system, crime control and policing. This identification is usually a unique identifier of some sort that is assigned to each individual in a country and is usually printed in identification document of some kind for easy identification. This unique identifier is usually stored together with personal data such as name, date of birth, place of birth, gender, address and possibly more.

Beynon-Davies (2007) notes, in the Information Society an individual may take on a number of different identities. One for each electronic service in the public, private and voluntary sectors with which the individual engages. As a consequence, an individual may accumulate a vast array of personal identifiers for such “services” and is also likely to accrue a range of physical representations or tokens of such multiple identification as in credit card, debit card, driving licence, passport, library card, parking permit. The need to have several identity can suggest the possibility of identification fraud of individuals and organizations in contemporary life.

A critical analysis of Nigerian situation as opinioned by Akinyemi (2011) suggest, owing to lack of national identification card, Nigerians walk about the country (and outside the country) with nothing to identify them thereby making it easier to get roped in crime or even to commit crime and get away. A lot of people considered to be missing by their loved ones are actually dead and found by people who cannot identify them for lack of ID
or DNA analysis. Similarly, a criminal may escape justice if he relocates from his Abeokuta home where he has committed a crime to nearby Ibadan to reside. He may decide to change his name from Kola to Musa. Since there is no ID card to identify him and proof his identity if he relocates from his Abeokuta home where he has committed a crime to nearby Ibadan to reside. He may decide to change his name from Kola to Musa. Since there is no ID card to identify him and proof his identity, the police officer will be looking for Kola but since he lives with Musa, he may not suspect that he is actually living with a wanted man.

A wide variety of biometric systems have been developed (Jain et al., 2004) for automatic recognition of individuals based on their physiological/behavioral characteristics. These systems make use of a single or a combination of traits like fingerprint, face, hand-geometry, iris, retina, palm-print, ear, voice, gait, signature, keystroke dynamics, for recognizing a person. Biometric recognition systems have been widely deployed in forensic, government and commercial applications. Sequel to this, the Nation needs a reliable National Identity Management System.

This research takes a look at the issue of National Identity card management in Nigeria and consequently presents a model for the conceptualization and realization of a Smart Card Identification Management over a Distributed Database Model. Instead of having multiple identity cards like driving licenses, work permits, employee cards, passports, voters cards, a single national identity card with in-built smart features including biometrics and advanced identification features that will eventually eliminate the need to carry multiple card will be okay.

Background study of national identity registration in Nigeria: The Nigerian National Identification project was originally conceived in 1977 but was executed in 2003. Registration exercise took place at 600,000 centers across Nigeria registering about 60,000,000 adults from the age of 18 years. The information collected in this exercise includes: name, gender, address, occupation, state/local government, height, photographs were collected and fingerprints taken.

Nigerian is a fairly large country with about 150 million people. In the presentation of the Federal Government of Nigeria entered into a Turkey Agreement with SAGEM in 2001 in the sum of $215,911,000 (off-shore component) and N2,089,600,000 (on shore component) to deliver the Nigerian National ID Card project which is an integrated ID Card production facility. The contract was supervised by Federal Ministry of Interior through the then department for National civil Registration (DNCR). According to SAGEM records, SAGEM registered and processed 52,640,908 made up of total female 18,062,443 and total male 19,307,165 and total reject was 37,369,608. The project concluded in December 2006 and entire system subsequently shut down since then. Nobody could run or continue the project since the contractor left.

The Nigerian national identity card is a synthetic card, measuring approximately 3.4 by 2.1 inches. The card has no expiry date. The following cardholder information is found on the front of the card: surname; first name; middle name; maiden name; date of birth; height; sex; occupation; blood group; applicant's residence, including address, town/village, Local Government Area (LGA), state, “Ward” and “PU number). The back of the national identity card contains “particulars of next of kin”, including surname, first name, middle name and address. It also holds information on the applicant's origin (place and state), as well as the date and place of issue of the ID card. Three numbers also appear on the back of the card: the “Seq Number” (sequence number) and the “ID Form Number”, which are numerical and the “Personal Number”, which is alphanumerical. A bar code is found at the bottom of the back of the card, (IRBC, 2008).

The Commission put in place measures to prevent fraud, double registration and foreigners from obtaining the card, wrong documentation, registration of applications who were under the age of eighteen and errors on the cards. However more than seven million national identity cards registered had been cancelled due to ineffectiveness of the measures.

According to (IRBC, 2008), Nigeria plans to introduce a new national identity card to replace the existing card which will be compulsory. The purportedly, new card may be used for multiple purposes, including registration for voting, payment of taxes, credit transactions, access to health insurance, access to pension and social security, purchase of land and for the issuance of passports and the opening of bank accounts among others.

Theoretical Frameworks: Date (2004) and Tucker, (1997) note a distributed database system consist of a collection of sites, connected together via some kind of communications network, in which, each site is a full database system site in its own right, but the sites have agreed to work together so that a user at any site can access data anywhere in the network exactly as if the data were all stored at the user’s own site. This follows that a distributed database is really a kind of virtual database, whose component parts are physically stored in a number of distinct real databases at a number of distinct sites and each site has its own local real databases, its own local users, its own local DBMS and transaction management software including its own local locking, logging, recovery, software and its own local data communication manager. In particular, a given user can perform operations on data at that user’s
own local site exactly as if that site did not participate in the distributed system at all.

The distributing data across sites within state and local government will allows those data to reside where they are generated or most needed, but still to be accessible from other sites in the state and local government areas. Keeping multiple copies of the database across different sites will allows continuous database operations even when one site is affected by a natural disaster, such as flood, fire, or earthquake or manmade incidences. Distributed database systems is structured geographically or administratively distributed data spread across multiple database systems. Ekenstam et al. (2001) opined that the central function of a distributed database system is to provide access to data while maintaining the integrity and consistency of that data. The system must have the ability to support large numbers of users without sacrificing performance. Higher reliability and availability in the presence of equipment and network failures are requirements for mission critical enterprise data systems. These requirements are often at odds with each other, leading to solutions that compromise between availability, consistency, scalability and performance.

Much data pertaining to the identity of individual and their behavior is presented within contemporary information and communications technology systems as personal records, typically as data element in the data structures of some database system. One of the most popular data models for representing data structure elements and items is the relational data model. In this data model data structures constitute relation which in turn constitute aggregation of attributes. Each row of a table must be unique and for this purpose a primary key is declared on the table. Data from diverse tables are selected, projected and joined over relationships defined in terms of primary-foreign key matches (Watson, 2006).

The data which frequently resides on multiple sites inside an organization might be managed by several database management systems for multiple reasons such as scalability, performance, access and management, (Date, 2004; Ramakrishnan and Gehrke, 2003). Thus, the information requirements for executing transactions and answering questions might not reside in a single site. distributed database management Systems deal with distributed database as a single logical database and the principles and techniques of Database Management Systems are still applicable to the distributed one; although the distributed one has special characteristics. A distributed database management system is a software that support the transparent creation, access and manipulation of interrelated data located at the different sites of a computer network (Lightstone et al., 2007), according to Morris et al. (2010). A distributed Database Management System (DDBMS) governs the storage and storage of logically related data over interconnected computer system in which both data and processing are distributed among several sites. Each site of the network has autonomous processing capability and can perform local applications. Each site also has the potential to participate in the execution of global application, which is to improve the accessibility, compatibility and performance of a distributed database while preserving the appearance of a centralized database management system, (Lightstone et al., 2007). Moreover, Distributed database system are very complex systems that have many interrelated objectives of transparency, heterogeneity, autonomy, high degree of function, extensibility and openness and optimized performance. It should be noted however, that, data allocation is done largely at the discretion of the database designer or database administrator (Date, 2004; Lightstone et al., 2007; Connolly and Begg, 2010).

Figure 1 present a functional component of a DDBMS a typical DDBMS consist of four major component (Connolly and Begg, 2010), of the Local DBMS component responsible for controlling the local data at each site that has a database and has its own local system catalog that store information about the data held at that site. It contains the Data Communication (DC) component which is a software that enables all sites to communicate with each other and the Global System Catalog (GSC) with functionality to hold information specific to the distributed nature of the system such as fragmentation and allocation schemas (Connolly and Begg, 2010; Morris et al., 2010) and the Distributed DBMS component is the controlling unit of the entire system. A distributed system requires functional characteristics that can be grouped and described as transparency features. This are discussed in Morris et al. (2010) as distributed, transaction, failure, performance and heterogeneity transparency.

The database is physically distributed across the data sites by fragmenting and replicating the data (Ceri et al., 1987). Given a relational database schema, fragmentation subdivides each relation into horizontal or vertical partitions. Horizontal fragmentation of a relation is accomplished by a selection operation which places each tuple of the relation in a different partition based on a fragmentation predicate. Vertical fragmentation, divides a relation into a number of fragments by projecting over its attributes.
Fragmentation is desirable because it enables the placement of data in close proximity to its place of use, thus potentially reducing transmission cost and it reduces the size of relations that are involved in user queries. Based on the user access patterns, each of the fragments may also be replicated. This is preferable when the same data are accessed from applications that run at a number of sites.

In this case, it may be more cost-effective to duplicate the data at a number of sites rather than continuously moving it between them Tucker (1997); Hevner and Yao (1979) and Ozsu and Valduriez (1991) consider a network of interconnected computers. Figure 1 presents a model of the distributed database environment. Each computer, known as a node in the network, contains a Distributed Database Management System (DDBMS) and a possibly redundant portion of the database. Data are logically viewed in the relational data model. The unit of data distribution is a relation. The DDBMS will maintain system directories so that each query will receive a no redundant consistent mapping of its required data.

A relation r is fragmented into fragments \( r_1, r_2, \ldots, r_n \), either horizontally or vertically. According to (Codd, 1970; Date, 2004; Ramakrishnan and Gehrke, 2003), horizontal fragmentation involves a relation r is divided into a number of subsets, \( r_1, r_2, \ldots, r_n \). Each tuple of relation r must belong to at least one of the fragments, so that the original relation can be reconstructed. Canonically, a horizontal fragment can be defined as a selection operation on the global relation r. That is, a predicate \( p \) to construct fragment \( r_i \), \( r_i = \sigma_{p_i}(r) \) and to reconstruct the relation r, the union of all the fragment is taken, thus: \( r = r_1 \cup r_2 \cup \ldots \cup r_n \).

In turn, vertical fragmentation of \( r(R) \) involves the definition of several subsets of attributes \( R_1, R_2, \ldots, R_n \) of the schema R so that: \( R = R_1 \cup R_2 \cup \ldots \cup R_n \) each fragment \( r_i \) of R is defined then as \( r_i = r_i (\prod_{k=1}^{n} R_k) \).

And to reconstruct r, the natural join is taken as:

\[
    r = r_1 \sigma_{R_1} r_2 \sigma_{R_2} \ldots \sigma_{R_n} \]

One way to ensure a successful relation reconstruction is to include the primary-key attributes of R in each \( R_i \).

Connolly and Begg (2010) included the mixed fragmentation of a relation consisting of a horizontal fragment that is subsequently vertically fragmented or a vertically fragmented that is then horizontally fragmented. This approach is defined using the selection and projection operations of relational algebra. Given a relation R, a mixed fragment is defined as:

\[
    \sigma_p(\pi_{a_1,\ldots,a_n}(R)) \]

\[
    \pi_p(\sigma_{a_1,\ldots,a_n}(R))
\]

Date, (2004) presented this scenario in form fundamental fragmentation rules as:

Rule 1: Completeness. If a relation instance R is decomposed into fragments \( r_1, r_2, \ldots, r_n \), each data item that can be found in R must appear in at least one fragment. This rule is necessary to ensure that there is no loss of data during fragmentation.

Rule 2: Reconstruction. It must be possible to define a relational operation that will reconstruct the relation R from the fragments. This rule ensure that functional dependencies are preserved.

Rule 3: Disjointness. If data item \( d_i \) appears in fragment \( R_i \), then it should not appear in any other fragment. Vertical fragmentation is the exception to this rule, where primary key attributes must be repeated to allow reconstruction. This rule ensures minimal data redundancy.

Access to various data item in a distributed system is usually accomplished through transaction, which must preserve the ACID properties (Connolly and Begg, 2010). The transaction can either be local or global transaction.

MATERIALS AND METHODS

National identification card over a distributed database model: National identity register databases system can be used to manage subject data. When
photographs, signatures or finger/thumb prints have been digitized and stored in a database, various methods exist to manipulate such data types and also have it printed on identity documents by the database system. In recent times, identity cards can be produced in the form of Smart cards, incorporating magnetic strips or data chips to store electronic data about the person who is the subject of the card. This data may include bioidentification data. The smart cards can be used with smart card readers and bio-identification readers such as finger print scanners to automatically verify a person's identity. Smart cards can be "read only" cards that simply contain information about the subject or can be "read-write" cards, which have the information contained on the card updated as the cards are used.

The Biometric identification refers to identifying an individual based on his or her distinguishing physiological and/or behavioral characteristics (biometric identifiers). It associates or disassociates an individual with a previously determined identity/identities based on who one is or what one does. Because many physiological or behavioral characteristics are distinctive to each person, biometric identifiers are inherently more reliable and more capable than knowledge-based and token-based techniques in differentiating between an authorized person and a fraudulent impostor. A biometric system is essentially a pattern recognition system that makes a personal identification by establishing the authenticity of a specific physiological or behavioral characteristic possessed by the user. Logically, a biometric system can be divided into the enrollment module and the identification module (Jain et al., 2000). During the enrollment phase, the biometric characteristic of an individual is first scanned by a biometric sensor to acquire a digital representation of the characteristic. In order to facilitate matching and to reduce the storage requirements, the digital representation is further processed by a feature extractor to generate a compact but expressive representation, called a "template." Depending on the application, the template may be stored in the central database of the biometric system or be recorded on a magnetic card or smartcard issued to the individual. During the recognition phase, the biometric reader captures the characteristic of the individual to be identified and converts it to a digital format, which is further processed by the feature extractor to produce the same representation as the template. The resulting representation is fed to the feature matcher that compares it against the template(s) to establish the identity of the individual.

Jain et al. (1997) notes, Biometrics is a rapidly evolving technology that has been widely used in forensics, such as criminal identification and prison security and has the potential to be widely adopted in a very broad range of civilian applications.

There are multitude of biometric techniques either widely used or under investigation module (Jain et al., 2000). These include, facial imaging (both optical and infrared), hand and finger geometry, eye-based methods (iris and retina), signature, voice, vein geometry, keystroke and finger- and palm-print imaging.

Humans have used fingerprints for personal identification for centuries and the validity of fingerprint identification has been well-established module (Jain et al., 2000).

A smart card is a small, tamperproof computer. The smart card itself contains a CPU and some non-volatile storage. In most cards, some of the storage is tamperproof while the rest is accessible to any application that can talk to the card. This capability makes it possible for the card to keep some secrets, such as the private keys associated with any certificates it holds. The card itself actually performs its own cryptographic operations.

Although smart cards are often compared to hard drives, they are store and process information. Smart cards are storage devices with the core mechanics to facilitate communication with a reader or coupler. They have file-system configurations and the ability to be partitioned into public and private spaces that can be made available or locked. They also have segregated areas for protected information, such as certificates, e-purses and entire operating systems. In addition to traditional data storage states, such as read-only and read/write.

In general, registration information must be accurate and complete, must be immune from fraud, be dynamic and up-to-date, be usable and must be easy for eligible individuals to identify, authenticate and perform enrolment.

It has been shown that the three arms of government in Nigeria can be seen as enterprise that is distributed already, at least logically into National, State and local governments and perhaps wards from which it follows that data are distributed already as well because it is expected that each unit of the local government will naturally maintain data that is relevant to its own people and operation.

Client/server architecture is concerned with minimizing processing costs by distributing processing between the server and multiple clients (Watson, 2006). Another factor in the total processing cost equation is communication. The cost of transmitting data usually increases with distance and there can be substantial savings by locating a database close to those most likely to use it. The trade-off for a distributed database is lowered communication costs versus increased complexity.
Distributed database architecture in Fig. 2, describes the situation where a database is in more than one location but still accessible as if it were centrally located. For example each local government could be maintaining data her citizen. If local queries deal with the local situation, communication costs are substantially lower than if the database were centrally located. Furthermore, since the database is still treated as one logical entity, queries that require access to different physical locations can be processed.

RESULTS AND DISCUSSION

The intranet: Intranet standard for exchanging e-mail and publishing web pages are becoming interestingly popular for business use within closed networks called Intranets. A typical intranet is connected to the wider public internet through a firewall with restriction imposed on the types of information that can pass into and out of the intranet, (Connolly and Begg, 2010). Three tier model which solves the problem of enterprise scalability is proposed with the following layers of architecture.

- The user interface layer which runs on the end-user’s computer (the client)
- The business logic and data processing layer. This middle tier runs on a server and is often called the application server
- The DDBMS which stores the data required by the middle tier. This tier may run on a separate server called the database server

The implementation language is java. Java is a proprietary language developed by Sun Microsystem and currently marketed by Javasoft. According to (Connolly and Begg 2010). The importance of Java language and its related technologies has been increasing for the last few years. Java (Gosling et al., 1996) is a type-safe, object oriented programming language that is interesting because of its potential for building web application (applets) and server application (serverts).

Java as explicitly defined is a simple, object-oriented, distributed, interpreted, robust secure, architecture neural, portable, high-performance, multi-threaded and dynamic language (Sun, 1997).

Relational mode and query: A formal way of presenting a relation schema. Let $R(f_1, D_1;..., f_n, D_n)$ be a relation schema and for each $f_i, 1 \leq i \leq n$, let Dom$_i$ be the set of values associated with the domain named $D_i$. An instance of $R$ that satisfies the domain constraints in the schema is a set of tuples with $n$ fields:

$$\{<f_1 : d_1, ..., f_n : d_n > | d_1 \in \text{Dom}_1, ..., d_n \in \text{Dom}_n\}$$

Different relation is required for this system, this includes:

- State (statecode, statement)
- Lga (lgacode, lganame, statecode)
- Ward (lgacode, wardname, lgacode)
- Unit (unitcode, unitname, street, wardcode)
- Person info (regno, fname, kname, othenname, gender, place of birth, disability, resaddress, or address
- Indentification info (regno, photograph, signature, fingerprints)

Notations

Number of States = $I$
Number of Local government area = $J$
Number of wards per local government = $K$

Therefore we can represent the state, lga , wards and units as:

States = $S_{i}, i = 1,2,3,...,I$
Local government areas = $L_{ij}, i = 1,2,3,...,I; j = 1,2,..,J$
Wards = $W_{ijk}, i = 1,2,3,...,I; j = 1,2,..,J; k = 1,2,..,K$

Estimates: Total number of registered person in state $i$, lga $j$, in wards $k$ is given as:

$$W_{ijk} = \text{count(total number records in asite database)}$$

Total number of registered person in state $i$, in lga $j$, is given as:

$$L_{ij} = \sum_{k=1}^{K} W_{ijk}$$

Total number of registered person in state $i$ is given as:

$$S_i = \sum_{j=1}^{J} L_{ij}$$
Total number of registered persons in the country is given as: \( N = \sum_{i=1}^{k} s_i \)

**(Sample queries)**

**Case 1- The Database support location transparency**

```sql
SELECT * FROM Okitpupa WHERE lgacode = '16-OKP'
UNION
SELECT * FROM Irele WHERE lgacode = '16-IRE'
UNION
SELECT * FROM Akure_South WHERE lgacode = '16-AKS'
```

**Case 2- The Database support location transparency**

```sql
SELECT * FROM KTP NODE W1 WHERE AGE > = 18
UNION
SELECT * FROM KTP NODE W2 WHERE AGE > = 18
UNION
SELECT * FROM KTP NODE W3 WHERE AGE > = 18
```

**CONCLUSION**

The recent development in the world of information technology has brought great change in the dynamic world. Information can be processed accurately, transmitted from any place to anywhere via the networks, data can be sparsely processed, managed and secured. Here, we have proposed a distributed database model for continuous National Identity registration in Nigeria. The cases of distributed database system is implemented is presented, the state of current registration system in Nigeria is also presented and theoretical background of Distributed database system presented. A model for future continuous registration in Nigeria is proposed and the transaction and algebraic operation on the databases presented. Visual Basic for Web Application has been suggested as the ideal language for the implementation of the system. If the management and control of National Smart Card Identity register can be sparsely managed with embedded forensic application software in a network environment. Then the nation can have a reliable, efficient and effective national identification system which is a necessity of any national government for the proper implementation and execution of its governmental policies and duties.

**REFERENCES**


Ozsu, M.T. and P. Valduriez, 1991. Distributed database systems: Where Are We Now? Computer, 24: 68-78. DOI: 10.1109/2.84879


