PIN-Based scheme for verifying the Physical Entity association

Ass. Prof.. Magdy E. Elhennawy

Abstract --In software engineering, the term data model is used in two related senses. In the one sense, it is a description of the objects represented by a computer system together with their properties and relationships; these are typically "real world" objects such as products, suppliers, customers, and orders. In the second sense, it means a collection of concepts and rules used in defining data models: for example the model uses relations and tuples, while the network model uses records, sets, and fields. The combination between the data model and process model depicts the business model integration which is the primary goal of the software engineering process. The one common requirement for that data is the confidence in its accuracy and completeness. The social security systems and subsidized services are common practice in many countries, today. The need to a physical identity attached to each beneficiary in such systems is crucial. This will guarantee the delivery of the service to the deserved citizens. Family Card System, FCS, is one of such systems, developed in Egypt, to deliver governmental services to deserved people. To guarantee FCS delivering such services to deserved people, each registered person should be supported by his national number, PIN. FSC itself should be sure that each person not only have a right PIN, but his associated PIN. A big difference between both. In this article, we depict the concepts of data accuracy and how to trust its contents. And association to related person. A scheme for safe entry to the PIN to his related person, signing, and verifying its accuracy and the physical association of the identity to appropriate citizen has proposed.

Index Terms --- Smart cards, network communications, Information security.

1. Introduction

The Egyptian government has defined policies to provide subsidies to support its citizens. However, there have been obstacles in defining the citizens who truly deserve the subsidies as well as a need to monitor the allocation of funds more closely, and enhance the level of implementation of the process.

The Family Card System, the electronic system developed in Egypt, would replace the traditional system. The traditional system was no longer efficient enough to accommodate the growing population. The electronic system allowed the registered outlet only replacing goods that have already been claimed by smart card, with all transactions being electronically monitored and documented, thus eliminating waste or illegal transactions.

The electronic system for the family card was designed to provide citizens with the services offered by the government; such as pensions and medical insurance as well as subsidized goods or their monetary equivalent. The smart cards can be used at grocery stores authorized to sell subsidized goods as well as at ATM machines. The family card project has provided a comprehensive database of Egyptian families which can be used by decision makers in defining the families in need, as well as tracing consumption levels and patterns across the country. The implementation has resulted in an over US$175 million monthly in savings through controlling subsidies distribution. The system takes into consideration the need for constant updates to guarantee the accuracy of information. After in-depth study and analysis it was found to be financially feasible to outsource the project to a private sector company, who would be responsible for the design, implementation and maintenance of the technology related to the project as well as applying high level international standards for data and process security.

Ass. Prof.. Magdy E. Elhennawy, High Institute of Computers and Information Technology, Computer Dept., El-Shorouk Academy, Family Card Project Consultant, Ministry of State for Administrative Development, Cairo, Egypt, e-mail: mhennawy@ad.gov.eg

2. Data, Information, and Personal Data

Data, information and knowledge are closely related terms, but each has its own role in relation to the other. Data is collected and analyzed to create information suitable for making decisions**, [1]** while knowledge is derived from extensive amounts of experience dealing with information on a subject. That is to say, data is the least abstract, information the next least, and knowledge the most **[2]**.

On the other hand, personal data, in principle, covers any information that relates to an identifiable, living individual. There are different ways in which an individual can be considered 'identifiable'. A person's full name is an obvious likely identifier. But a person can also be identifiable from other information, including a combination of identification elements such as physical characteristics, pseudonyms occupation,

address etc. It does not matter how the personal data is stored – on paper, on an IT system, on a CCTV system etc **[3]**.

Managing large quantities of structured and unstructured data is a primary function of information systems. Data models describe structured data for storage in data management systems such as relational databases. The information systems, then, can achieve their roles by means of a data model. **Figure 1** depicts this meaning.

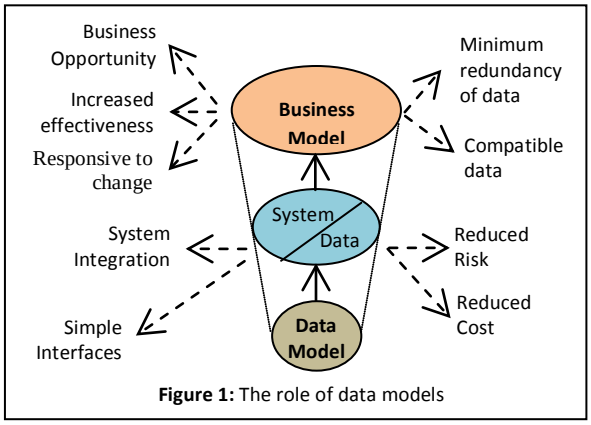

Figure 1: The role of data models

The main aim of data models is to support the development of information by providing the definition and format of data.

If data models were built consistently across systems, then compatibility of data can be achieved. If the same data structures are used to store and access data then different applications can share data. The results of this are indicated above. However, if it was not, then systems and interfaces often cost more than they should, to build, operate, and maintain. They may also constrain the business rather than support it. A major cause is that the quality of the data models implemented in systems and interfaces is poor**[4].** Examples are:

- *When business rules, specific to how things are done in a particular place, are often fixed in the structure of a data model*. This means that small changes in the way business is conducted lead to large changes in computer systems and interfaces **[4]**.
- *When entity types are often not identified, or incorrectly identified*. This can lead to replication of data, data structure, and functionality, together with the attendant costs of that duplication in development and maintenance**[4]**.
- *When data models for different systems are arbitrarily different*. The result of this is that complex interfaces are required between systems that share data. These interfaces can account for between 25-70% of the cost of current systems**[4]**.
- *When data cannot be shared electronically with customers and suppliers, because the structure and meaning of data has not been standardized*. For example, engineering design data and drawings for process plant are still sometimes exchanged on paper**[4]**.

The reason for these problems is a lack of standards that will ensure that data models will both meet business needs and be consistent**[4]**.

Meanwhile, data model is a way finding tool for both business and IT professionals, which uses a set of symbols and text to precisely explain a subset of real information to improve communication within the organization and thereby lead to a more flexible and stable application environment**, [5]** and usually data models are specified in a data modeling language**[6]**.
Communication and precision are the two key benefits that make a data model important to applications that use and exchange data. A data model is the medium which project team members from different backgrounds and with different levels of experience can communicate with one another. Precision means that the terms and rules on a data model can be interpreted only one way and are not ambiguous**[5]**.
A data model can be sometimes referred to as a data structure, especially in the context of programming languages. Data models are often complemented by function models, especially in the context of enterprise models.

A data model may be one of three kinds according to ANSI: **[7]**

- Conceptual data model, which describes the semantics of a domain, being the scope of the model. For example, it may be a model of the interest area of an organization or industry. A conceptual schema specifies the kinds of facts or propositions that can be expressed using the model. The use of conceptual schema has evolved to become a powerful communication tool with business users. Often called a subject area model (SAM) or high-level data model (HDM), this model is used to communicate core data concepts, rules, and definitions to a business user as part of an overall application development or enterprise initiative**[8].**
- Logical data model, which describes the semantics, as represented by a particular data manipulation technology. This consists of descriptions of tables and columns, object oriented classes, and XML tags, among other things.
- Physical data model, which describes the physical means by which data are stored. This is concerned with partitions, CPUs, table spaces, and the like.

The significance of this approach, according to ANSI, is that it allows the three perspectives to be relatively independent of each other. Storage technology can change without affecting either the logical or the conceptual model. The table/column structure can change without (necessarily) affecting the conceptual model. In each case, of course, the structures must remain consistent with the other model. The table/column structure may be different from a direct translation of the entity classes and attributes, but it must ultimately carry out the objectives of the conceptual entity class structure. Types of data models can be: flat model, hierarchical model, network model, relational model, object-relational model, or star schema.

Some important properties of data for which requirements need to be met, as shown in **Figure 2**, are:

- definition-related properties **[4]** which can be:

- *Relevance*: the usefulness of the data in the context of your business.
- *Clarity*: the availability of a clear and shared definition for the data.
- *Consistency*: the compatibility of the same type of data from different sources.

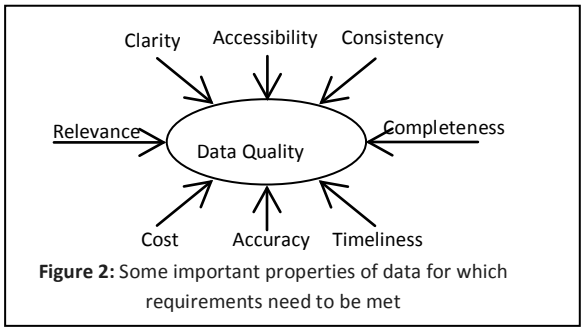

Figure 2: Some important properties of data for which requirements need to be met

- content-related properties, which can be:
  - *Timeliness*: the availability of data at the time required and how up to date that data is.
  - *Accuracy*: how close to the truth the data is.
- properties related to both definition and content, which can be:
  - *Completeness*: how much of the required data is available.
  - *Accessibility*: where, how, and to whom the data is available or not available (e.g. security).
  - *Cost*: the cost incurred in obtaining the data, and making it available for use.

The family card system is intended to build a complete system for achieving a set of objectives. Such objectives are: guaranteeing the delivery of the support services to the deserved people, provide a families database to support the services delivery and help decision maker, allow civilized environment to provide the service in a civilized mode. The system was composed of a set of basic components as shown in **Figure 3**, namely: centralized database hosted in a specialized data center, set of applications to achieve the system functionalities, call center to allow interaction with the citizens, technical support center to guarantee the system security and service provision continuity, supply offices automation to supervise system activities, service centers to conduct families data manipulation. To allow delivering more than one service to the citizen using the same smart card, the system employs a multi-application smart card technology.

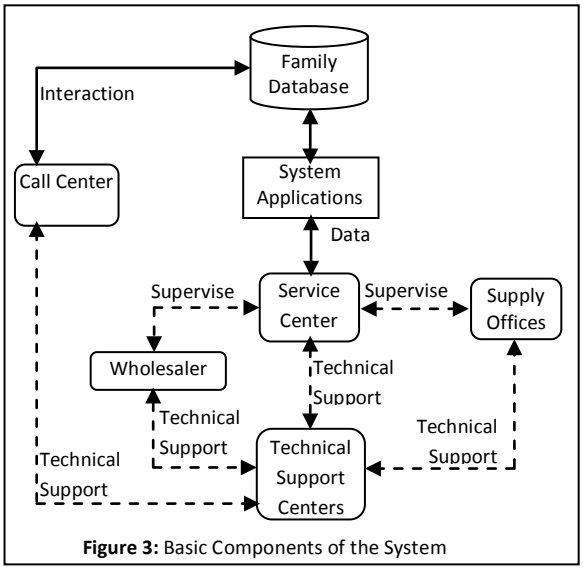

Figure 3: Basic Components of the System

Meanwhile, The term data model can have two meanings: a data model *theory* or data model *instance*. A data model *theory*, which is a formal description of how data may be structured and accessed. A data model theory has three main components: **[9]**

- The structural part: a collection of data structures which are used to create databases representing the entities or objects modeled by the database.
- The integrity part: a collection of rules governing the constraints placed on these data structures to ensure structural integrity.
- The manipulation part: a collection of operators which can be applied to the data structures, to update and query the data contained in the database.

The family card system has been proposed, studied, analyzed, designed, contracted and is currently operationally monitored by the Ministry of State for Administrative Development, MSAD. The stakeholders of the system include the Ministry of Supply and Interior Trade (MOSIT), Ministry of Social Solidarity (MoSS), Ministry of Petroleum (MOP), MSAD, and the Egyptian society.

MSAD has outsourced the implementation of the system to a consortium which was responsible for the following: building the system components, technical support to system programs and applications, hosting of family card database, availing service provision centers, availing call center, applying networks and communication lines, training of civil servants responsible for managing and operating the system, system management, system operation and maintenance, and availing of necessary applications and tools.

3. Identification and Physical Association

Physical association of the identity is crucial when we deal with social funds transfer to under poverty people and the delivery of support services. To deliver such fund transfer and other support services to deserved people, you should insure the eligibility of these people who receive such services. It is a more critical issue if the decision to go the cash instead of commodities. One example of such systems, as stated above is the launching of the Family Card System (FCS), in Egypt, in 2006 **[10] [11]**.

FCS has been expanded than before and is now responsible for delivering four different services to more than 19 million families, covering more than 76 million citizens. FCS depends mainly on identifying the citizen with the national number, PIN. The system is supported by a family database contains enough data about the families and their members, all are identified with the PIN. Basic issues that should be fulfilled in the family database are the identification, correct association of such

identification to appropriate persons, completeness, and the continual updating. The correct association of such identification to appropriate persons guarantees the delivery of the service to correct and alleged families and individuals **[10] [11]**. Obviously, in Egypt, each citizen, has assigned a PIN as it born, and he can issue a PIN identification card at 16 years old.

A research on the database of the FCS indicates that about 2.6% of the populations registered in the database are associated with wrong persons or not correct at all, and about 6% of them have no PIN registered yet. It is required to revise such association and completeness to be correctly completed. Other issues, continual updating is not the interest of this research. That is why we are interest in this research to establish an appropriate mechanism that guarantees the correct association and prevent any forgery process relating to it during registration process.

4. The Need to Physical Identity in FSC System

FCS was delivering about four services, the ration commodities, the LPG, Social pensions, and the supported bread. Rations were delivered to about 19 million families, selected according to MSIT rules. They represent about 85% of the Egypt's population. There should be some rules to exit families that do not deserve such service. Recently, the Economic Justice Unit, EJU, has been formed in the ministry of finance to take over the role of defining the exiting rules. The more important than the foundation of such rules is to be sure you can apply such rules correctly. To guarantee the correct application of these rules, the correct association of the PIN to its appropriate persons is crucial.

On the other hand, since the supported bread is critical issue to all Egypt's population, it was delivered to all people. The government of Egypt assigns a big budget to such item, so we should be very sure when we allow some people to receive such service. To guarantee the delivery of such service to the correct and deserved person, the correct association of the PIN to its appropriate persons is crucial. Other services can behave so similar.

The social justice dictates the above need, which means:

- Each individual in FCS database should have his PIN number registered in his the database.
- The registered PIN should be correct in both its value and in its association to the appropriate physical individual.

To overcome and achieve this requirement, a review to the above cases could be executed to achieve the following objectives:

- Insuring the PIN fill-in of the correct PIN with its appropriate person. This can be achieved by scanning the PIN number from his PIN identification card and store it in his database record. This can be done in any of the related service centers distributed all over Egypt's governorates. The legal employee doing this job should be identified and registered together with his PIN number in the same database record. The same process, scanning the employee PIN number, can be done the same way.
- Signing this record with a **dual** signature, containing both the signature of the person himself and the related employee. This signature will be stored together with the person database record for any later on violation that may need to be verified.

5. Authentication and Dual-Digital Signature

Authentication or message authentication is a procedure to verify that received messages come from the alleged source and have not been altered. Message authentication may also verify sequencing and timeliness. Identification and digital signature are the basic authentication mechanisms. **[12]**

User Identification, whenever a would – be user seeks the services of a computer (essentially, when the identity of the user is a key case), how can the latter make sure the former is not forging a false identity? The classic solution to this problem is through the use of passwords. The key idea to overcome the threat's raised from this technique can be implemented using one–way function

Due to the vulnerability to interception either as it goes through the communication links to the computer or inside the computer memory before the one–way function, has irreversibly transformed it, the **challenge/response technique** can be used

Digital Signatures, In situations where there is no complete trust between sender and receiver, the digital signature is needed. Digital signature is a message–dependent quantity that can be computed only by the sender of the message on the basis of some private information. It allows authentication of messages by guaranteeing that no one can forge the sender's signature and the sender cannot deny a message he sent. Digital signature scheme are usually classified into one of two categories: *Direct* digital signature and *Arbitrated* digital signature **[12]. The direct Digital Signature** involves only the communicating parties (source, and destination). It may be formed by encrypting the entire message with the sender's private key or by encrypting a hash code of the message with the sender's private key. In the a**rbitrated Digital Signature,** every signed message from a sender A to a receiver B goes first to an arbitrator X, who subjects the message and its signature to a number of tests to check its origin and contents **[13].**

One of the basic mechanisms in today's protection and signing mechanisms is the dual signature for example, the Secure Electronic Transaction, **SET** was a communications protocol standard for securing credit card transactions over insecure networks, specifically, the Internet, uses the dual signature . SET was not itself a payment system, but rather a set of security protocols and formats that enabled users to employ the existing credit card payment infrastructure on an open network in a secure fashion **[14].** To meet the business requirements, SET incorporates the following features: 1) Confidentiality of

information, 2) Integrity of data, 3) Cardholder account authentication, and 4) Merchant authentication.

As described in (Stallings 2000):

An important innovation introduced in SET is the *dual signature*. The purpose of the dual signature is to link two messages that are intended for two different recipients. In a case like e-commerce, the customer wants to send the order information (OI) to the merchant and the payment information (PI) to the bank. The merchant does not need to know the customer's credit-card number, and the bank does not need to know the details of the customer's order. The customer is afforded extra protection in terms of privacy by keeping these two items separate. However, the two items must be linked in a way that can be used to resolve disputes if necessary. The link is needed so that the customer can prove that this payment is intended for this order and not for some other goods or service.

The message digest (MD) of the OI and the PI are independently calculated by the customer. The dual signature is the encrypted MD (with the customer's secret key) of the concatenated MD's of PI and OI. The dual signature is sent to both the merchant and the bank. The protocol arranges for the merchant to see the MD of the PI without seeing the PI itself, and the bank sees the MD of the OI but not the OI itself. The dual signature can be verified using the MD of the OI or PI. It doesn't require the OI or PI itself. Its MD does not reveal the content of the OI or PI, and thus privacy is preserved.

Meanwhile, duo to the wide spread of using handheld devices, it offers an opportunity for mobile devices to be used as a universal payment method. However, some issues impede the widespread acceptance of mobile payment; for example: privacy protection, limited capability of mobile devices, and limited bandwidth of wireless networks. In ecommerce payment, Secure Socket Layer (SSL) protocol has been used to establish a secure channel between customers and merchants to secure the payment and the order information. Both SSL and SET assume the existence of Public Key Infrastructure (PKI) where extensive computations are carried out. In mobile payment, the same protocols of ecommerce payment are used but their application is limited due to heavy computations over wireless and GSM networks. A Modified Secure Electronic Transaction (MSET) protocol is proposed to minimize the extensive computations of SET protocol through replacing time consuming public key encryption and decryption algorithms by symmetric key cryptography. **[15]**

6. Proprietary Dual-Digital Signature Scheme

Used terms:

CP ….. The concatenation of the customer PIN number and name.

EP ….. The concatenation of the employee PIN number and name.

H ….. One way hash function.

H(CP)… the hash value of the CP.

H(EP)… the hash value of the EP.

$[H(CP)]_{ENC}$ … the encryption of the hash value of the CP

$[H(EP)]_{ENC}$ … the encryption of the hash value of the EP

SIG_{CE} … The final citizen record signature.

The proposed scheme apply the dual signature for both the citizen that we keep his record data and try to keep these data correct and complete, and the employee that is responsible for making the scanning of the PIN of the appropriate citizen. This scheme try to achieve and detect any forgery that may discovered between that citizen and the legal employee. **Figure 4** depicts the signing and verifying cycles constituted in the scheme.

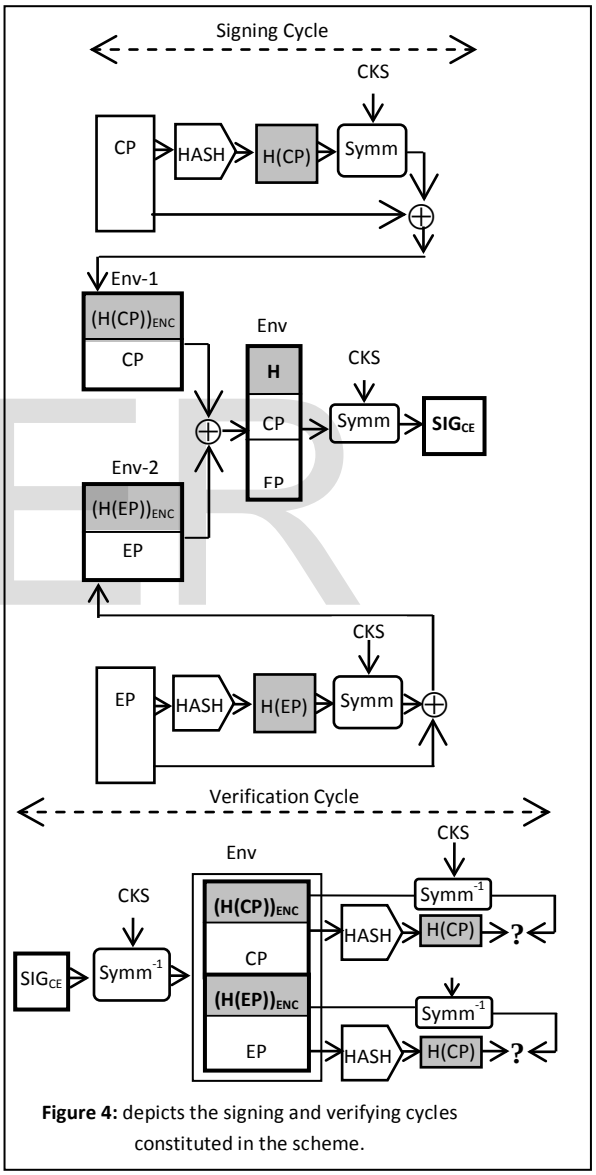

Figure 4: depicts the signing and verifying cycles constituted in the scheme.

Signing Cycle:

- The hash of CP, H(CP) is calculated, then encrypted with the symmetric-key cryptosystem by system session key, CKS, to generate the citizen signature, $[H(CP)]_{ENC}$.
- The encrypted H(CP), $[H(CP)]_{ENC}$ is concatenated with the full text of CP, the name and PIN of the citizen, generating Env-1.

- The hash of EP, H(EP) is calculated, then encrypted with the symmetric-key cryptosystem by system session key, CKS, to generate the citizen signature, $[H(EP)]_{ENC}$.
- The encrypted H(EP), $[H(EP)]_{ENC}$ is concatenated with the full text of EP, the name and PIN of the citizen, generating Env-2.
- The Env1 is concatenated with Env2 to generate Env.
- The Env is signed with the symmetric-key cryptosystem by system session key to generate the record signature, $\mathbf{SIG_{CE}}$ **.**

Verification Cycle

- The record signature, $\mathbf{SIG_{CE}}$ is decrypted with the symmetric-key cryptosystem by the system session key, CKS, to generate the envelop Env.
- Noting that the envelop Env contains the concatenation of Env1 and Env2.
- The encrypted H(CP), $[H(CP)]_{ENC}$ is selected from the envelop Env1 and decrypted with the symmetric-key cryptosystem by the system session key, CKS, to generate the hash value, H(CP).
- The concatenation of the citizen name and PIN number, CP, is selected from the envelop Env1 and hashed with the hash function, to generate the hash value, H(CP).
- The two hash values, then, are compared if they are equal, it means that they have no any forgery and the name and PIN of the citizen are the correct, else it means that there is modification have been happened to the record signature.
- If there is mismatch has happened between the name and his PIN in the citizen record, that means that there is some forgery happened between the citizen and legal employee, we can have the employee and citizen name to deal with to resolve such mismatch.

Identification is another crucial issue. It is normal to identify any record by a unique number, the key. In some contexts, we are in need to be sure not only just identifying the records, but very sure it is the correct property of that record. Sometimes, one can forge the key of another to have the right to get the benefit of another. The problem of physical impersonation of another by

Accordingly, insuring the correct identification is a vital issue in such systems. The proposed solution needs to install HW and SW on sites then applying the above stated scheme on registration of the PIN numbers together with its associated name to allow the verification process to be done whenever needed. It goes through two steps, the first is the SW and HW preparation, the second is to follow the following algorithm procedures.

The SW and HW preparation consists of the following:

- Site will be equipped with a PC with appropriate specs.
- Scanner that can scan at least the identity card with appropriate speed.
- Scanner will be connected to the PC.
- The OCR SW will be installed on the PC.
- An appropriate interface will be built to allow to interact between the scanner and the OCR.

The algorithm procedure consists of the following:

- Entering the personal data in an electronic form.
- Building a buffer database that containing such entered data.
- Against the legal employee, the citizen can verify his data by providing his identity card the employee, which scans the identity number imaged on the identity card.

7. Proposed Technique

Physical Identification and Orientation, Our Technique

acquiring its identity is frequently happened, especially in systems that deliver benefits to citizens The stated above figures for a study on the FCS indicates that about 2.6% of the populations registered in the database are associated with wrong persons or not correct at all, and about 6% of them have no PIN registered yet.

- The system scans the identity number imaged on the identity card, digitize it and convert it into integer number that can be processed. It, then compares it to the number entered in the electronic form, and verifies its correctness. If correct, sign the record in the buffer database as accepted record and store it in the system database, and register both the identity of the citizen and legal employee as the both responsible legally about the correctness of that record.
- In any time if it is proved that the identity is not correct, both will be legally face the appropriate penalties.

8. Conclusions

In the above scheme, we have present the way to associate the PIN to individual person in the FCS database. This will allow the correct completeness to the PIN to all missed PIN persons in FCS database. Moreover, we introduce a new proposal for a proprietary dual-signature based authentication scheme to enable the registration of the person registration signature which depends on both the PIN of the individual and the legal employee responsible for such registration. This signature will allow, later on, to authenticate the person who do the registration and eventually responsible about any forgery happened in that record.

REFERENCES

[1] "Joint Publication 2-0, Joint Intelligence". *Defense Technical Information Center (DTIC)*. Department of Defense. 22 June 2007. pp. GL–11. Retrieved February 22, 2013.
[2] Akash Mitra (2011). "Classifying data for successful modeling".
[3] http://www.dataprotection.ie/docs/What-is-Personal-Data-/210.htm
[4] Matthew West and Julian Fowler (1999). Developing High Quality Data Models. The European Process Industries STEP Technical Liaison Executive (EPISTLE).
[5] "Data Modeling Made Simple 2nd Edition", Steve Hoberman, Technics Publications, LLC 2009

[6] Michael R. McCaleb (1999). "A Conceptual Data Model of Datum Systems". National Institute of Standards and Technology. August 1999.
[7] American National Standards Institute. 1975.*ANSI/X3/SPARC Study Group on Data Base Management Systems; Interim Report*. FDT (Bulletin of ACM SIGMOD) 7:2.
[8] "Data Modeling for the Business", Steve Hoberman, Donna Burbank, Chris Bradley, Technics Publications, LLC 2009
[9] Beynon-Davies P. (2004). Database Systems 3rd Edition. Palgrave, Basingstoke, UK. ISBN 1-4039-1601-2
[10] Magdy E. Elhennawy, M. Amer, Abdelhafeez, "Health Care Implementation by Means of Smart Cards", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011, ISSN (Online): 1694-0814, www.IJCSI.org
[11] M. Elhennawy, T. Saad, S. Bedair, and A. AbdelWahab, "Adapting Family Card System by Means of Smart Cards", 11th European Conference on eGovernment – ECEG 2011, Faculty of Administration, University of Ljubljana, Ljubljana, Slovenia 16-17 June 2011.
[12] Cryptographic Engineering Technology for the Internet Security", A dissertation submitted in partial fulfillement of the requirements pf ph.D Degree in Computer & System Engineering, **April** 2002.
[13] Stalling W., Ph.D "*Network and Ineternetwork Security*", Printic Hall, (1995).
14] Mark S. Merkow, "Secure Electronic Transactions (SET)". In Hossein Bidgoli. *The Internet Encyclopedia*. John Wiley & Sons. pp. 247–260. ISBN 978-0-471-22203-3, (2004).
From Wikipedia, the free encyclopedia.
[15] Sabrina M. Shedid, Magdy El-Hennawy and Mohamed Kouta, "Modified SET Protocol for Mobile Payment: An Empirical Analysis", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.11, November 2008