Conception of Data Base Management System in USB Smart Card Flash Memory: Application for the Cancer Pathology of Medical Information Systems

Walid ATTIAOUI*, Pr. Mohamed BEN AHMED*, Pr. Philippe PUCHERAL** and Dr. Boutheïna CHETALI***

* Laboratoire RIADI-GDL, ENSI, University of La Manouba, Tunisia
Walid.Attiaoui@riadi.rnu.tn
Mohamed.Benahmed@riadi.rnu.tn

** INRIA Rocquencourt, Versailles, France
Philippe.Pucheral@inria.fr

*** AXALTO, Louveciennes, Versailles, France
bchateli@axalto.com

Abstract: Every Smart card equipped with a microcontroller contains hardware-implemented software as well as applications. The introduction of the more powerful technologies and smaller chip structures allows the storage of this software not only in fixed ROM masks (Red Only Memory), but also in re-writable Flash memories. Most smart card controllers today are equipped with ROM. ROM-Memory does not allow for deletion, nor can it be overwritten by the user via program commands. Therefore, all programs and data supposed to be in the ROM have to be installed through the production process. Flash memory is among the top choices for the storage media in ubiquitous computing. With a strong demand of high capacity storage devices, the usages of flash memory quickly grow beyond their original designs. The very distinct characteristics of flash memory introduce serious challenges to researchers the high demand of main-memory space for flash-memory memory and guarantee of security.

The design of Data Base to be resident on portable devices and embedded processors for professional systems requires considering both the device memory and the mobility aspects, which are essential feature of the embedded applications. Moreover, these devices are often part of a larger Information System, comprising fixed and mobile resources. This paper examines the application of USB smart cards in the development of distributed medical information systems. The pocket mobility and security features of smart cards make them an ideal medium for storing the critical medical records of individual. This research proposes a methodology for USB smartcard data base design. The process is strongly driven by the technological issue, which impose a very careful design on the logical and physical data structures in order to meet the constraints the smart card architecture introduces, and to provide satisfactory performance. In this paper we introduce a USB smart card data base methodology and concentrate on the conceptual and logical phases. We present the choice of the analysis dimensions and we take the cancer pathology for the case study. We identified the different actors and their access rights and privileges. We examine a number of rules with the respect to the insertion, search with equality, search with range selection, delete and update operations.

Key words: USB Smart card, Flash memory, Health care, Data Base Management System, Information Systems, Access methods.
1 Introduction

Flash memory is non volatile, shock-resistant, and power economic. With the recent technology breakthroughs in both capacity and reliability, flash-memory storage systems are much more affordable than ever. As result, flash-memory is now among the top choices for storage media in ubiquitous computing. Researchers have been investing how to utilize flash-memory technology in existing storage systems, especially when new challenges are introduced by the characteristics of flash memory.

With the rapid growing of the flash-memory capacity, severe challenges on the flash-memory management issues might be faced, especially guarantee of security. With the advancement of information technology in today’s Internet area, we have seen a steady adoption of health information systems at various levels of health care. However, in comparison to other industries such as banking, manufacturing and airlines for health services has been slow to catch up with the increasing demand and need for distributed sharing of health related information. Traditionally, most health information systems are designed to provide fronted automation of administration for the purpose of patient admissions and building charges. Moreover, these systems are highly autonomous in terms of their ability to integrate and share health information among medical staff in hospitals or clinics. The use of traditional manual charting and recording of vital medical data often results in non-structured data being captured and stored which may vary across hospitals or clinics [1]. As a result, the medical records of patients taken in this manner are often non-reusable and non-portable. Yet, it is recognized that the availability of patient’s medical histories in times of accident and emergency can be vitally important for providing accurate and rapid diagnosis. The advent of smart card technology has paved the way for an individual to carry personal information. With the advantage of being robust and portable, smart cards represent a technology that can be employed in health care to provide distributed storage of a patient’s medical records [1]. Effectively, smart cards provide a distributed storage solution in which patients themselves are responsible for carrying with them their important medical records. This is in contrast to the traditional approach of storing patient records in a central repository, which lacks the means to distribute the information across its local system’s domain [2].

The storage portability and secure features of smart cards represent an integrated solution to support a truly accessible and robust medical information system. Smart cards present an ideal way for individuals to carry their own medical records.

The purpose of our research was to conceive a data base design methodology, which should guide the data base design methodology, which should guide the database designer from the conceptual step, carried out in the traditional way by using one of the well-known conceptual design models, to the semiautomatic choice of the data structures, access methods, and memory allocation, based on the analysis of the storage devices currently offered by smart card technology.

2 USB Smart Card Technology

Smart cards are essentially devices that allow information storage and processing. The microcontroller used in smart card applications contains a central processing unit (CPU) and blocks of memory, including RAM, ROM, and re-programmable non-volatile memory (usually EEPROM or Flash-EEPROM). RAM is used to store executing programs and data temporarily, while ROM is used to store executing programs and data temporarily, while ROM is used to store the operating system, fixed data, and standard routines. The re-programmable non-volatile memory is used to store information that to be retained when power is removed, but that must also be alterable to accommodate data specific to individual that has to be retained when power is removed, but that must also be alterable to accommodate data specific to individual cards or any changes possible over their lifetimes; more specifically, the smart card Non Volatile Memory (NVM) constitutes the data storage for the database. Based on this consideration, technology issues have a significant impact on the overall system performance. Today’s micro-controllers contain a CPU, memory including about of 196 KB of ROM, 8 KB of RAM and 128 KB of EEPROM [3]. The operating system is typically stored in ROM, the CPU uses RAM as its working and most of the data is stored in EEPROM. We must also notice that the particular features of flash memories make smart card databases differ from main memory databases, where flash memory acts as a true secondary storage with very different read/write proprieties which respect the RAM main memory, making it look more like a traditional Data Base Management System (DBMS)[3],[4].

AXALTO (Schlumberger Sema) is innovative E-gate USB smart cards combine ISO and Universal Serial Bus (USB) versatility with leading edge security functionality. It embeds the USB interface electronics normally found in smart card reader within the card itself. This feature enables the USB smart card to simply plug into the USB port on the computer using either a token form factor connector or a desktop form factor (in standard smart card) [5].

3 Flash memory characteristics

There are two major types of flash memory: NAND flash and NOR flash. The NAND flash memory is specially designed for data storage, and the NOR flash is for EEPROM replacement.

A NAND flash memory is organized in terms of blocks, where each block is of a fixed number of pages. A block is the smallest unit for erase operations, while reads and writes are processed in terms of pages. The typical block size and the page size of a NAND flash memory are 16 KB and 512B, respectively. There is a 16-byte “spare area” appended to every page, where out-of-band data could be written.
to the spare areas.

When a portion of free space on flash memory is written, the space is no longer available unless it is erased. Out-place-updating is usually adopted to avoid erasing operations on every update. The effective copy of data is considered as live, and old versions of the data are invalidated and considered as dead. Pages which store live data and dead data are called “live pages” and “dead pages”, respectively. “Free pages” constitute free space on flash memory. After the processing of a large number of page writes, the number of free pages on flash memory would be low.

Most of the traditional approaches adopt static tables to deal with address translation and space management. Because the granularity of the flash-memory management unit is a fixed system parameter, the granularity size would be an important factor on the trade off between the main-memory overheads and the system performance[6].

Because out-place update is adopted, we need a dynamic address translation mechanism to map a given LBA (Logical Block Address) to the physical address where the valid data reside. Note that a “logical block” usually denotes a disk sector. To accomplish this objective, a RAM-resident translation table is adopted. The translation table is indexed by LBA’s, and each entry of the table contains the physical address of the corresponding LBA. If the system reboots, the translation table could be re-built by scanning the flash memory[7].

4 A Smart Card Database Design Methodology

We introduce our running example and briefly outline the general methodology for smart card database design. The next sections are devoted to the details of those methodology steps that are the subject matter of this paper.

We start by outlining the methodology steps:

- The relevant information is chosen and modelled; this is done by singling out homogeneous information areas, with the corresponding conceptual views, regardless of the target storage media,
- Dictionary data and Views are integrated into a global conceptual schema and possible representation and semantic conflicts are resolved; logical entities (for instance relational tables) are designed,
- Logical entities are fragmented; possibly both horizontally and vertically, in order to single out the first need information to be allocated on the smart card. Indeed, as already noted, smart card data base design criteria must be very similar to those adopted for distributed database design, since the smart card database will often be part of a larger information system, possibly allocated to several fixed machines besides the smart card itself,
- A strict estimate of the fragment cardinalities is made at this step,
- “First need” fragments of the appropriate size are allocated to the smart card, taking into account the card memory size constraint, while the other fragments are allocated to other site of the information system; the interest of this phase of the methodology is better understood reasoning on personal information systems, for instance a personal medical record, where the most recent clinical tests are kept on card,
- Access rights are defined for each fragment and for each user class, and the relevant constraints are included in the view definitions: for instance, in the case of a car accident, the first aid personnel of an ambulance should read from the medical card the patient’s blood pressure record, but not his possible insurance policies,
- The type of access mode (read only, read/write) and the volatility (for example in terms of update/query ratio) are estimated for each fragment.
- Memory protection mechanisms are to be designed in order to prevent unauthorised users to access sensitive information,
- The most convenient data structures are chosen,
- Possibly an encrypting algorithm is chosen for some very sensitive data,
- Access methods are chosen for the data structures defined in last step under the constraints coming before [3].

5. The Conceptual and Logical Phases

The conceptual phase can be decomposed into the following steps:

5.1 Application information modelling

This is done by the usual techniques for conceptual database design, taking into account all the information relevant to the application at hand, regardless of the target of the target storage media. The Pico Data Base design must be merged within the design of the distributed database it belongs to.

Contrary to the banking sector where a clerk is very often allowed to get access to client-related personal data upon data upon request simply by being a clerk of this bank, the health care and welfare sector is managed in a rather different way. Even doctors within the same department of clinical or hospital are by law not allowed to get access to the same amount of personal data even for the same patient. It depends on the roles they are engaged to play within a very specific care and treatment process, and the related rules and policies.

For the access point of view, a lot of people sharing only a few rather general data items about a patient (even only administrative ones, for the process of exchanging bills between the clinic or hospital and the insurance company the patient belongs to). A
diagnostic or therapeutic team has to share much more information about the specific patient in order to treat him well. But even this amount of patient related information is strictly limited to the real process of care they are obliged to deliver. And at the bottom of this triangle, we find the one and the only responsible doctor with access to all medical and administrative patient-related information. So this access aspect could be seen as the major problem providing in order to achieve the highest possible quality in the care and treatment process:

- The right amount of patient-related medical and administrative information (correctness from a medical point of view and integrity from the security point of view),
- To the right person (the person with explicit access rights to read these information items),
- At the right time[8],

This may seem to be rather simple but isn’t. Even when it comes to the example explained before (doctors in the same department of a clinic) it very much depends on their roles within the care process. If one of them became the caring doctor of a specific patient, he has of course access to all data. This is both a legal and ethical requirement. If the other doctor is just another doctor of this department he may have access to data, which are relevant for the services he has to provide for the department. But as soon as he might become the deputy chair of the department (if the chair is absent for a surgery or a meeting) he has to have access to the medical data. In this case, he (as the deputy chair) is responsible for all activities provided by his team.

This example might have illuminated that the rights of professionals in health care consist of both static and dynamic permission and admissions. Basically, static permissions are issued by health authorities, dynamic permissions may be issued even by employers (hospitals, clinics).

In this system we propose the use of USB Smart cards for individuals to store their personal and medical information. The pocket mobility of smart card and allows users to carry along with them their own medical history. A multi-level security password protection system is used to protect the smart card data and allow access by different people and professionals. For example, a doctor may have the maximum access right to gain access to his patient’s medical history, while a pharmacist may view only information pertaining to the prescription of drugs.

5.2 Choice of the analysis dimensions

Analysis dimensions provide the different perspectives the mobile device is viewed from, and are used:

- The holder dimension refers to the type of users carrying the Pico device, whose views over the whole information system can be quite different. For example, in a medical application, doctors will hold information about all their patients, while patients will only hold information related to them, may be at a finer level of detail.
- The interest topic dimension refers to the particular aspect/subject the user might be interested in at a certain moment.
- The situation dimension refers the user may access different views of data, being able to perform different operations. For instance, in the personal medical information system, examples of situations are the regular (ordinary patient’s state) as opposed to a temporary hospitalized situation.

The dynamic progress of communication and cooperation technologies and structures with increasing immediate needs to spread information systems, particularly medical networks, is a great challenge for the health care and welfare sector. A part from economic and organizational aspects, security and privacy requirements have to be taken more and more in account. As health professionals at all mostly work with highly sensitive medical data appropriate administrative, technical, ethical, and legal solutions have to be developed and implemented.

In the following we intend to explain the basic statements of this paper by introducing the example of cancer pathology[9], [10], [11], [12].

The actors of the system are:

- The patient: He is the person who presents symptoms. He addresses to the general doctor with his medical USB Smart card.
- The general doctor: He carry out examinations and asks complementary examinations (biological and radiological). He can determine the diagnosis. He has the following access right:
  - Consultation of all data (select),
  - Insertion of data about the case of a late stage (insert).
- The specialist doctor (lung specialist, gastroenterologist...): He will push the investigations (for example the gastroenterologist will carry out other examinations like a fiberscope, coloscopy). If he discovers in this examination a tumour of the stomach for example, then he carries out the biopsy. It will be addressed to the histological examination. If the diagnosis is not obvious then the specialist doctor can supplement the investigations (complementary examinations) by requiring an ictography or the IRM "imagery by magnetic resonance" or a scanner.
- The surgeon who will carry out a surgical operation. For instance, in the personal medical information system, examples of situations are the regular (ordinary patient’s state) as opposed to a temporary hospitalized situation.

In this case, the surgeon carries out a palliative
operation. The specialist doctor has the following access rights:

- Consultation of all the data.
- Modification of the data concerning the diagnosis, the symptoms, the examinations and the treatments of the tumour.
- The insertion of data about the diagnosis, the symptoms, the examinations and the treatments of the tumour.

- The specialist doctor in histology: He determines nature, size and the stage of the tumour. The specialist doctor in histology has the following rights of access:
  - Consultation of all the data,
  - The insertion and modification of the data concerning the diagnosis and the stage of the tumour.

- The doctor specialist in chemotherapy or radiotherapy intervenes to treat the tumour. He has the following access rights:
  - Consultation of all the data,
  - The insertion and modification of the data concerning the diagnosis and the treatment of the tumour.

- A medical personnel for example nurses, technicians of laboratories, anaesthetists: At the time of the hospitalization, the male nurses have an access limited to the medical file for which they are authorized. They will be able to reach the consultations of the administrative data of the patient and those concerning the diagnosis, the nature of the tumour, previous pathologies, and nature of the treatment and with the regulations.

- The pharmacist has the right to consult prescriptions.

- The administrator: He is the person who manages the system. He carries out the modifications of the passwords and gives the access rights to different users.

- The local data base: It is the data base of the secondary server which is in the hospitals and polyclinics, it contains all information concerning the patients, the doctors, the equipment and the drugs.

- The general data base: It is the data base of the primary server, it centralizes all information about the hospitals and polyclinics.

As an output of this step, the identified dimensions are collected which drives the actual choice of the information to be kept on the USB Smart card. The dimensions used in this paper form a three positions array model: <holder, interest-topic, situation>.

### 5.3 Dimensions conceptual design

In this phase, on conceptual schema is built for each dimension value, we build build one view for the patient, one for each doctor and one for each of the possible other values of this dimension (hospital administrator).

### 5.4 Conceptual view merge

Here we present examples of array schema in this Medical Data Base:

As a conclusion of this phase we assemble different array schemas (different views), in order to define which is the information that must be stored on one single device. For example, normally a patient’s USB smart card will contain all the array schemas related to the patient’s (regular) situation plus those related to his chronic diseases, previous pathologies, allergies and prescriptions. When the patient is at the hospital, the “regular” array schema will be removed to leave on all the array schemas related to different situations at all times [4].

### 5.5 Logical view merge

Here different activities are carried out:

- Logical design of the global database:
  - Administrative_folder_Patient (code-patient, patient_name, patient_first_name, birth_date, birth_place, address, city, profession, telephone, sex, nationality, civil_State, code_group).
  - Previous pathology (previous-name-pathology, previous_type-antecedent, classification).
  - Prescription (code_prescription, name_prescription).
  - Patient_antecedent (code-patient, code-doctor, doctor_name, doctor_first_name, doctor_phone, speciality).
  - Consultation (patient_code, code-doctor, date_consult, time_consult).
  - Blood group (code-group factor_rhesus).
  - Vaccine (code_vaccine, vaccine, date_vaccine).
  - Allergy (patient_code, name_allergy).
  - Tumour (code_tumour, name_tumour).
  - Hospitalization (code_hospitalization, date_start_hospitalization, date_end_hospitalization).
  - Result_hospitalization.
  - Code_patient, code_responsible_doctor.
  - Code_service.
  - Diagnosis (code_patient, code_tumour, code_doctor, date_diagnosis).
  - Pharmacology (code_pharmacology, type_pharmacology).
  - Pharmacology_Hospitalization (code_pharmacology, code-hospitalization).
  - Examination (name_examination, characteristic_examination).
  - Patient_Examination (patient_code, name_examination, date_examination).
Logical array schema production: the array schemas are defined as logical views over the global logical database produced above; for example, <Patient, Pharmacology_Hospitalization, hospital>.

The folder_pathology_cancer is a view. It is a virtual table calculated from other tables by query. A view has many advantages:

- The simplification of the access to the data by masking the operations of joints, presentation of same data in various forms adapted to the various particular users and a reinforcement of the data security by masking of the lines and the columns sensitive to the not competent users. Example:
  
  Create View folder_pathology_cancer As
  Select *
  From Treatment_tumour, Administrative_folder_Patient, Hospitalization, Pharmacology_Hospitalization, Diagnosis
  Where Administrative_folder_Patient.code_patient = Treatment_tumour.code_patient
  And Administrative_folder_Patient.code_patient = Hospitalization.code_patient
  And Hospitalization.code_hospitalization = Pharmacology_Hospitalization.code_hospitalization
  And Diagnosis.Code_patient = Administrative_folder_Patient.code_patient;

5.6 Access Types, operations et evaluations

In this section, we perform a comparison of the four adopted data types with respect to the typical operations to be performed on the data of each on-card table.

- Select :
  - Scan: fetch all records in the table.
  - Search with equality selection: fetch all records that satisfy an equality selection.
  - Search with range selection: fetch all records that satisfy a range selection, such as, for example: “find the stage of tumour with date of treatment is between 01/01/2004 and 30/09/2004”.

- Insert: insert a given record into a table. Depending how the record are stored, ordered or not, the operation may require a shift of the records following the position where the new record is inserted, that is, fetch all records, include the new one, and write back all records.

- Delete: search the record and remove it by freeing the space. More precisely, it is necessary to identify the record, fetch the block containing it, modify it, and write the block back. Depending on the record organisation, it may be necessary to fetch, modify, and write back all the records following the one under consideration.

- Update: search a given record, read it and rewrite it.

Examples of rules:

- Regulate 1: The Ali nurse has the permission to manage Administrative_folder_Patient in the service of cancerology.
  Grant privileges one Administrative_folder_Patient To Ali;

- Regulate 2: The doctor practitioner has the permission to consult the entirety of the folder of his patients.
  Definition of an array schema:
  Create View dossier_patient_doctor As
  Select *
  From folder_pathology_cancer
  Where folder_pathology_cancer.code_doctor = CURRENT_USER;
  Grant Select on folder_pathology_cancer To Moncef;

- Regulate 3: The doctor has the permission to update parts from the folder of cancer pathology:
  Grant insert, update on folder_pathology_cancer To Moncef;

- Regulate 4: The doctor has the permission to authorize the nurse to update the prescription in the case of the hospitalisation about cancer tumour.
  Definition of array schema:
  Create View Prescription_hospitalier_cancer As
  Select *
  From folder_pathology_cancer
  Where folder_pathology_cancer.code_doctor = CURRENT_USER;
  Grant Select, update(prescription) On Prescription_hospitalier_cancer To Moncef With Grant Option to Ali; [13]

- Regulate 5: A doctor has the permission to authorize the nurse to update the prescription in the case of the hospitalisation about cancer tumour:
  Definition of an array schema:
  Create View Prescription_hospitalier_cancer As
  Select *
  From folder_pathology_cancer
  Where folder_pathology_cancer.code_doctor = CURRENT_USER;
  Grant Select, update(prescription) On Prescription_hospitalier_cancer To Moncef With Grant Option to Ali; [13]
Then we discuss the storage issues and propose a very compact model based on a combination of flat storage, domain storage, and ring storage. Based on performance evaluation, we derive guidelines to decide the best way to store an attribute. The Pico DBMS manages users which handle different objects according to their privileges [14]. Generic relations, with multiple indices were presented in last paper [14], [15], [16] where we proposed an indexing mechanism for smart cards, to make such indices as compact as possible: the idea applies to indices where attribute domain have very small cardinality. These implements an index structure having a ring form, from the domain values, through all the tuples having such a value.

Conclusion

This research proposes a methodology for USB smartcard data base design. The process is strongly driven by the technological issue, which impose a very careful design on the logical and physical data structures in order to meet the constraints the smart card architecture introduces, and to provide satisfactory performance. In this paper we introduced a USB smart card data base methodology and concentrated on the conceptual and logical phases. We presented the choice of the analysis dimensions and we take the cancer pathology for the case study. We identified the different actors and their access rights and privileges. We examined a number of rules with the respect to the insertion, search with equality, search with range selection, delete and update operations. Future work on the methodology will address the following aspects:

- Find fragmentations and allocation criteria for the different logical units (tables) of the USB smart card database,
- Implement the USB smart card for the cancer pathology using Flash-EEPROM storage and evaluate their performance in order to optimize them,
- Manage transaction to guarantee data consistency, especially useful for situations where a card is disconnected before committing the transaction,

References

[8] WOHLMACHER Petra, PHAROW Peter. Applications in Health Care using Public-Key Certificates and Attribute Certificates.