Purpose-based Access Control Exploits by HDB

Rajneeshkaur K. Bedi
HOD and Assistant Professor
Computer Department,
MIT COE, Paud Road,

Anita M. Thengade
Assistant Professor
Computer Department,
Paud Road,
Pune –411038

ABSTRACT
As organizations increase their reliance on, as a result preserving privacy has become a crucial requirement for operating a business that manages personal data. Private Schools, district schools, and state education agencies maintain a large volume of personnel files. Due to voluminous stores of personal data being held by Education Industry today, preserving privacy has become a crucial requirement for operating a business. Hippocratic databases have been proposed to answer this requirement through a database design that includes responsibility for the privacy of data as a founding tenet. We identify, study, and implement privacy-preserving feature for education industry in Hippocratic databases. This paper includes the support of how current relational database management systems can be transformed into their privacy preserving equivalents.

Categories and Subject Descriptors
INFORMATION SYSTEM - DATABASE MANAGEMENT-LANGUAGES

Subject Descriptors: Query Languages

General Terms

Design, Languages, Security

Keywords
Hippocratic Database, Privacy Preserving, Multiple Policy Version, DDL

1.INTRODUCTION
Database security becomes more crucial as the scale of database for public and private organizations is growing and the various user **access** schemes are required. Recently, most relational database management systems(RDBMS) provide only some limited security techniques, which generally use a policy-based access control[4][6].
Privacy preservation is an important requirement whenever personal data is collected, stored and published. One of the main challenges is to share information while complying with the data-owner privacy preferences. In recent years, several research directions have received substantial attention including Hippocratic databases, anonymization and generalization, privacy-preserving data mining, privacy rules languages, e.g. P3P and EPAL and fine-grained access control techniques in discretionary and mandatory access control.

The notion of Hippocratic databases was introduced to incorporate privacy protection as a founding tenet in relational database systems [1] [2] [4] [6]. Ten guiding principles of Hippocratic databases and initial designs to provide limited disclosure and compliance audition were introduced. One key element of the Hippocratic database architecture is that it makes use of a centralized and standardized definition of privacy rules via a privacy policy. A privacy policy usually is born outside the database system and is expressed using natural language. In order to process this policy more effectively it is expressed using a standard privacy specification language, e.g., P3P or EPAL. The resulting version is translated into its Hippocratic database equivalent, i.e., the policy rules tables inside the database. The great value of this policy-driven approach is that companies that use the Hippocratic database have at their disposal an important tool to comply with privacy laws and guidelines.

2.What is our vision?
2.1 Privacy Legislation
Presently, in India there is no specific legislation which dealing with privacy and data protection. The protection of privacy and data can be derived from various laws like Information Technology Act, 2000(**"IT Act"**), Intellectual Property Laws, Credit Information Companies Regulation Act, 2005("CICRA") etc [14].

Data privacy in the European Union is governed by a very comprehensive set of legislations called the Data Protection Directive [7]. In the United States, privacy protection is achieved through a patchwork of legislation at the federal and state levels. However, privacy has been recognized as a constitutional right and there exists a highly developed system of privacy protection under tort law for the past century [15]. The worldwide phenomenon has ushered in a plethora of privacy-related guidelines and legislations.
To protect student and staff privacy, legislative bodies in many countries have enacted legislation that define personal information and spell out the obligations of the service provider with respect to the privacy.

We now propose the twelve Privacy Principles that is based on ten guiding principles of Hippocratic databases [1] and the Canadian Standards Association's *Model Code for the Protection of Personal Information* [10] recognized as a national standard in 1996; it can be applicable to any Private sector. But here we present especially for education sector. These principles are rooted in the privacy regulations. They articulate what it means for a database system to responsibly manage private information under its control. They also define what a donor of private information can expect if a database system advertises itself to be Hippocratic database.

2.2 Where do you start?

Principle Description

1.Accountability

An organization is responsible for personal information under its control and shall designate an individual or individuals accountable for the organization's compliance with the privacy principles.

2. Identifying Purposes

The purposes for which personal information is collected shall be identified by the organization at or before the time the information is collected.

3. Consent

The knowledge and consent of the individual are required for the collection, use or disclosure of personal information, except when inappropriate.

4. Limiting Collection

The collection of personal information shall be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means.

5.Limited Use

The database shall run only those queries that are consistent with the purposes for which

the information has been collected.

6. Limited Disclosure

The personal information stored in the database shall not be communicated outside the database for purposes other than those for which there is consent from the donor of the information.

7. Limited Retention

Personal information shall be retained only as long as necessary for the fulfillment of the purposes for which it has been collected.

8. Accuracy

Personal information shall be as accurate, complete, and upto-date as is necessary for the purposes for which it is to be used.

9. Safeguards

Security safeguards against theft and other misappropriations shall protect personal information.

10. Openness

An organization shall make readily available to individuals specific information about its policies and practices relating to the management of personal information.

11. Individual Access

Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.

12. Challenging Compliance

A donor shall be able to verify compliance with the above principles. Similarly, the database shall be able to address a challenge concerning compliance.

3.Language Construct

We present the relational database systems with fine-grained access control (FGAC) with retention time and show how they can be used to enforce disclosure control enunciated in the vision for Hippocratic databases [1]. These constructs have been designed to be integrating with the rest of the infrastructure of a relational database system.

We provide constructs that allow restrictions to be specified on access to data in a table at the level of a row, a column, or a cell with support of **retention time**. Privacy policies specified in high-level languages such as P3P can be translated into these constructs, or one could specify the policy directly using these constructs.

3.1 Support of retention time

Retention time means amount of time for provider to keep the information. Limited retention is a principle of Hippocratic databases and a key element of privacy policies. It ensures that data is retained only as long as necessary for the fulfillment of the purposes for which it has been collected. The original architecture of the Hippocratic database [1] suggests the implementation of the *Data Retention Manager, which* basically deletes all data items that have outlived their purpose.

To the best of our knowledge no further mechanism to support retention time with respect to implementation point of view was proposed in the context of Hippocratic databases. The advantage of this approach is that it does not require deleting the information after the allowed retention time. Additionally, using SQL conditions constitutes an exile mechanism to express complex retention restrictions.

Here we propose construct to use of retention time in fine-grained access control (FGAC). We assume there is a table, referred to as primary table ***Restriction_Detail***, which stores basic information of all the restriction and where each row is associated with exactly one restriction on one table. The description of the table is shown in figure 1(a).

Our support of retention time makes use of the ***Signature-Date*** table in which we store the policy signature date i.e. restriction created date for each table and end date of policy signature will get calculated by using the value of retention time mentioned in the restriction. The description of the table is shown in above figure 1(b). During command execution, the translator also builds a condition that ensures that the date in which a command is executed falls in the period between the privacy signature date *sd,* which will probably be different for each restriction, and end date of policy signature.

Figure 1 : Table Description of Restriction_Detail and Signature-Date

a) SQL> describe Restriction_Detail

Rid	NOT NULL	integer
Rname		varchar2(25)
R_date		Date

b) SQL> describe Signature-Date

R_id	integer
R_created_date	Date
Retention_value	integer
Derived_policy_end_date	date

Conceptually, a restriction defines a view of the table in which inaccessible data has been replaced by null values. As discussed in [6], it is possible to use either "table semantics" or "query semantics".

With query semantics, if all the values in a row are hidden by a restriction, then the row is omitted altogether from the view. With table semantics, the row would instead be retained.

Figure 2 gives the syntax of a fine grained restriction command. It states that those in auth-name-1 except those in auth-name-2 are allowed only restricted access to table-x. The keywords **public** (i.e., all users), **group**, **role**, and **user** can be used to qualify the authorized names. Table-x can be any table expression. A restriction can be specified at the level of a column (Section 3.1.1), a row (Section 3.1.2), or a cell (Section 3.1.3). More than one restriction can be specified on a table for the same user.

Figure2. Fine grained restriction syntax

create restriction restriction-name
on table-x
for auth-name-1 [**except** auth-name-2]
(((**to columns** column-name-list)
| (**to rows** [**where** search-condition])
| (**to cells** (column-name-list [**where** search-condition])+)
)
[**for purpose** purpose-list]
[**for recipient** recipient-list]
[**for retention time** time-interval**]**
)+
command-restriction

A restriction may additionally specify purposes and/or recipients for which the access is allowed. If no purpose or recipient is specified, then the restriction applies to all purposes and recipients respectively. If a purpose or recipient is specified, the user's access is limited to only the specified purpose-recipient combinations.

If value of retention time element is not specified, then access is allowed without time restriction. If time interval is specified, there is use of *Signature-Date* table in which the policy signature date and calculated end of policy signature date is available.

Akin to the database system variable **user** that can be referenced in queries and returns the id of the user issuing the query, the new system variables **purpose** and **recipient** return the list of purposes and recipients from the current query context [6]. These values in turn determine the restrictions for the current query.

The proposed facility is complimentary to the current table level authorization mechanisms provided by commercial database systems using the **grant** command [2]. While grant controls whether a user can access a table at all, the proposed constructs define the subset of the data within a table that the user is allowed to access.

The command-restriction that appears as the last element of the syntax has the following form and states that access can be restricted to any combination of select, delete, insert,

or update commands:

restricting access to (**all** | (**select** | **delete** | **insert** | **update**)+)

The discussion below will use, for illustration, the student table with the following schema:

Student(id **integer**, name **char**(32), phone **char**(32), father-income **integer ,** mentor-id **integer**).

Student _marks (sid integer , year char(32) , marks integer)

Staff(id **integer**, name **char**(32), designation **char**(32), student-feedback **char(32),** salary **integer)**

3.1.1 Column Restriction

A column restriction specifies a subset of the columns in table-x that auth-name-1 is allowed to access only within specified time interval. The following restriction, named r1, ensures that only the father-income column of student is accessed by user smith only for 6 month from when restriction get created. Whenever restriction get created at same time restriction creation date and end date of policy will get added into *Signature-Date table* :

create restriction r1
on table student
for user Sonia(HOD)
to columns father-income
for retention time 6month
restricting access to all

The restriction r2 below ensures that members of the administrator group have only select access to columns name and phone without time restriction or it can be implemented at more finer level.

create restriction r2
on staff
for group admin
to columns name, phone
restricting access to select

3.1.2 Row Restriction

A row restriction gives the subset of rows in table-x that auth-name-1 is allowed to access. This subset is specified using a search-condition over table-x. The restriction r3 below ensures that every access to Staff is qualified by the predicate, name = **user**.

create restriction r3
on Staff
for public
to rows where name = **user**
restricting access to all

If user Priyanka issues **select * from** Staff, she would see id, name and phone for those rows where name equaled Priyanka.

3.1.3 Cell Restriction

A cell restriction defines the row-column intersections that auth-name-1 is allowed to access. It is possible to specify multiple column-name lists, each possibly annotated with a search-condition. A search-condition is a correlated subquery with an implicit correlation variable *t* defined over the tuples of table-x.

Access to the columns in column-name-list for each individual row identified by *t* is conditionally granted depending upon the result of the search condition. If no search-condition is given, then access is granted to all column values in column-name-list in table-x. If the search condition ignores correlation variable *t*, then access is granted or denied to all columns values in column-name-list in table-x, depending upon the result of the search-condition.

The following is an example of a cell restriction used to enforce individual user's privacy preferences expressed as opt-in/out choices. Assume that for the purpose of official,

Sita is allowed to see name, but his access to phone is allowed only if the user has opted-in to revealing her phone number when name is Rajiv.

create restriction r4
on Staff **for user** Sita (Hod),
to cells name,
(phone **where name** = 'Rajiv'))
for purpose official
for recipient others
restricting access to select

The above restriction specifies cell restrictions for two column-name-lists: The first list contains the name column, and the second contains the phone column. The restriction allows Smith access to name, only if the variable **purpose** includes official, and **recipient** includes others. Otherwise, all values of the name column will be null for Smith.
The second list of columns has a search-condition associated with it since access to phone is dependent upon individual user choices.

4.Related Work

4.1 Data Protection law in India: The Trips Analysis

The concerns and apprehensions of the MNCs regarding lack of data protection in India are far-fetched and unwarranted. The TRIPS Agreement, the Copyright Act,1957 and and the IT Act, 2000 provide sufficient safeguards for preventing violations of electronic and paper based databases of MNCs. The brightest and the positive aspect of this situation is that even non-data items are also protected, both under the TRIPS Agreement and the Copyright Act, 1957. Similarly, the IT Act, 2000 sufficiently protects the electronic data property and there is no need of further amendments. Further, the explanation to section 43 defines and protects computer database. The enforcement aspect of data protection is also adequately covered under the IT Act, 2000. For instance, the IT Act, 2000 provides for both civil and criminal liabilities in the form of "contraventions" and "offenses". Thus, the present framework of the data protection regime is sufficient to accommodate the mandates of both the Constitution of India and the TRIPS Agreement. The ultimate solution to any problem is not to enact a plethora of statutes but their rigorous and dedicated enforcement.

5.CONCLUSION:

Hippocratic Database technologies are well-suited to enable the transition to the 21st century electronic records in education sector to preserve privacy. These technologies offer efficient methods of managing, auditing, sharing, and analyzing electronic records that preserve

the privacy of student, staff and management committee member also. We have introduced a support of retention time in fine grained access control to allow restrictions to be specified on access to data in a table at the level of a row, a column, or a cell.

We hope these construct will serve as a model for future research and development of useful education sector information management technologies that respect individual privacy.

6.REFERENCE

[l] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu. "Hippocratic databases". VLDB 2002.

[2] K. LeFevre, R. Agrawal, V. Ercegovac, R. Ramakrishnan and Y. Xu. "Limiting disclosure in Hippocratic databases". VLDB 2004.

[3] R. Agrawal, R. Bayardo, C. Faloutsos, J. Kiernan, R. Rantzau, and R. Srikant. "Auditing compliance with a Hippocratic database". VLDB 2004.

[4] R. Agrawal, P. Bird, T. Grandison, J. Kiernan, S. Logan, W. Rjaib. "Extending Relational Database Systems to Automatically Enforce Privacy Policies". Industrial paper. ICDE 2005

[5] T. Kyte. Fine-grained access control. Technical report, Oracle Corporation, 1999.

[6] Yasin Laura-Silva, Walid Aref "Realizing Privacy-Preserving Features in Hippocratic Databases"

[7] George Yee and Larry Korba "Privacy Policy Compliance for Web Services"

[8] R. Agrawal , Christopher Johnson,T. Grandison, J. Kiernan "Enabling the 21st Century Health Care Information Technology Revolution" Communications Of The Acm February 2007/Vol. 50, No. 2

[9] European Union, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the

protection of individuals with regard to the processing of personal data and on the free movement of such data".

[10] Canadian Standards Association, "Model Code for the Protection of Personal Information",

[11] Min-A Jeong', Jung-Ja Kim**',** and Yonggwan Won2 **"**A Flexible Database Security System using Multiple Access Control Policies"

[12] Julia B. Earp, Fay C. Payton "Data Protection in the University Setting: Employee Perceptions of Student Privacy"

[13] Industry Canada, "Privacy and the Information Highway, Regulatory Options for Canada", chapter 6, from: http://strategis.ic.gc.ca/SSG/ca00257e.html#6

[14] PRAVEEN DALAL "Data Protection law in India : The Trips Analysis"