# The Development of a Data Security Model for The Collaborative Social and Medical Services System

Ross Dargahi, Dave W. Classen, Risa B. Bobroff, M.S., Cynthia A. Petermann, M.S., Dennis R. Moreau, Ph.D., J. Robert Beck, M.D., and Gregory J. Buffone, Ph.D.

Medical Informatics and Computing Research Program
Baylor College of Medicine
Houston, Texas

## ABSTRACT

*This paper presents the development of the Collaborative Social and Medical Services System's (CSMSS) data security mechanism. This mechanism was synthesized from an analysis of the CSMSS problem domain, and from a study of the methods used by modern operating systems and database management systems.The resulting mechanism is more flexible and expressive than traditional access control methods and is generally applicable to the management of privacy and multi-provider access.*

## INTRODUCTION

According to The World Health Organization, "the road leading to health for all by the year 2000 passes through information" [1]. The truth of this statement is evident in the growing requirement for electronic access to medical record systems, and the explosion of research in the area of computerized patient record systems (CPRS) [2][3][4][5].

For CPRSs to be successfully deployed and used, it is imperative that the needs and expectations of both health care providers and patients be addressed. Generally these expectations mandate that a CPRS be able to manage the rich domain of information required for clinical medicine as well as maintain maximum protection of patient information in both storage and use. That is, the information's confidentiality, reliability, and integrity must not be susceptible to compromise. The importance of privacy with respect to medical information cannot be overstated. According to the American Health Information Management Association, "...Without such assurance (of privacy), the patient may withhold critical information which could affect the quality of care provided, the relationship with the provider, and the reliability of the information maintained" [6].

Protection of patient information, within a given system, is largely provided by system and data security. System security addresses the issue of protection from unauthorized access. This includes provision for hardware, software, communications, and users security [7]. Data security is concerned with the protection of data from accidental or intentional disclosure to unauthorized persons or from unauthorized modification or destruction [8].

Current system security mechanisms provide an acceptable level of system security; for example, Project Athena's Kerberos authentication system [9] provides network based client-server authentication, as well as encryption based on secure keys. Such a system can be incorporated into a CPRS to provide appropriate system security. Unfortunately, there is presently no satisfactory mechanism for providing a complete or even adequate solution to the data security problem faced by CPRSs. In fact, a literature review of confidentiality issues indicates that data security remains a key issue when discussing the comprehensive automation of medical records [3].

In this paper, a methodology for analysing the data security requirements of a CPRS along with its application to the CSMSS problem domain is described. The results of this analysis are the basis for developing a data security model for the CSMSS upon which a spectrum of policy based access control strategies may be built.

0195-4210/94/$5.00 © 1994 AMIA, Inc.

## AN ANALYSIS OF THE DATA SECURITY NEEDS OF AN OUTPATIENT CPRS

CPRSs pose an interesting data security problem. On the one hand is the desire to provide an open system for maximizing information sharing among health care practitioners and limited but sufficient access to researchers; on the other is the very real need to protect patient privacy. Furthermore, since the data is used in the day to day delivery of health care, it must ideally be: rapidly deliverable (there cannot be delays which impair the function of the health care practitioners); available on request (the system must be fault tolerant, that is not subject to the failure of software, hardware or communications); and distributed (information should be available to authorized personnel in different locales). These conflicting goals must be effectively managed by the data security model.

To realize the data security needs of a CPRS, one must understand the methods in which data will be accessed. There are four general use categories: patient care, social services, administration, and research [4]. These use categories exert different requirements upon the CPRS. The four use categories are normally further subdivided on the basis of specific roles. For example, patient care may be partitioned into the organizational roles important for the delivery of health care. Typical roles would be: physician, nurse, and nurse practitioner. Each role performs a defined set of tasks. Each task has a required set of inputs and produces a set of outputs. These relationships play an important part in the analysis model and its application to the CSMSS as described next.

### The Analysis model and its Application to the CSMSS

The problem domain covered by the CSMSS design is the Baylor College of Medicine Teen Health Clinics (THC). The THCs are comprised of five geographically distributed clinics providing teenagers in the Harris County Hospital District with such services as: family planning, screening for and treatment of sexually transmitted diseases, prenatal and postnatal care, and patient education and counseling. A six step process is employed in determining clinic data security needs [10]. Each step will be detailed along with its application to the THCs.

The first task involves discovering and classifying the roles to be considered for the CPRS. This process examines the clinic's current organizational structure and extracts any roles that will require use of the CPRS. These roles are then broadly classified under the four use categories previously described. Interviews are conducted with staff members for each role in order to further understand its function and appropriateness as a part of the CPRS. The interviews may also result in the definition of new roles suitable only in the context of a CPRS; for example, a need for a data entry or transcriber role might be revealed. These emergent roles are defined by classification and further interviews. This iterative process continues until all the potential roles have been discovered. At the THCs, the roles identified are typical of many outpatient facilities and are illustrated in Table 1.

TABLE 1. Baylor College of Medicine Teen Health Clinic Roles

| **Clinic Roles** | **Patient Care** | **Social Services** | **Administrative** | **Research** |
|---|---|---|---|---|
| Administrative Secretary | | | ■ | |
| Billing Clerk | | | ■ | |
| Clerk | | | ■ | |
| Community Services Aid | | ■ | ■ | |
| Director | | | ■ | ■ |
| Health Educator | | ■ | ■ | |
| Medical Assistant | ■ | ■ | ■ | |
| Nurse | ■ | | ■ | |
| Nurse Manager | ■ | | ■ | |
| Physician/Nurse Practitioner | ■ | | ■ | |
| Project Coordinator | | | ■ | |
| Research Assistant | | | | ■ |
| Social Services Coordinator | | ■ | ■ | |
| Social Worker | | ■ | ■ | |

As previously discussed, each role in a clinic has a set of tasks which the person engaged in that role is responsible for performing. The goal of the second step in the analysis process is to define the tasks associated with each role. Observations of how the clinic staff work together to perform a single patient centered function, as well as interviews with staff members are used to define and refine the role specific tasks. The goal of the interviews is to derive, for each role, a set of tasks with crisp boundaries and minimal functional overlap (except in cases where it is appropriate for the performance of the task). In addition, the interviews should identify any role overlap by exposing those tasks which are performed by more than one role. It should be noted that it is often permissible for there to be role overlap and task redundancy; in many cases this is required for maintaining patient safety, meeting legal requirements, or ensuring the efficiency of clinic operations. In the case of the THCs, it was frequently found that one role performed the tasks of another due to the limited availability of staff to perform the required task.

With both the clinic roles and tasks defined, the input and output for each task must be analyzed. Questionnaires serve an important capacity in this step. Questions are centered around the data requirements of a task (task input), and the data produced by the execution of a task (task output), and they are framed in terms of a metaphor familiar to the clinic staff. Questionnaires for the THCs used forms, logs, and reports as the entities for uncovering task data and access requirements. Task input was specified in terms of the forms accessed to provide the input for a task, and task output was specified in terms of the forms, logs and reports generated or updated by the execution of a task.

With a knowledge of the roles, the tasks associated with those roles, and the input, output, and data access requirements for those tasks, it is possible to define the aggregate data requirements of a given role within the clinic. For the THCs the data requirements for a given role were determined by gathering all forms, logs, and reports required by the set of tasks associated with that role.

The next step is to discover any exceptions which may be applicable in the given health care environment. This process involves identifying the scenarios that would require the normal data security mechanism to be overridden, and defining any consequent processing, such as audit trails, which would be necessary in such instances. No exceptions were discovered at the THCs

Finally, the findings and conclusions are compiled into a matrix and are then verified with the clinic staff. Rows in the matrix represent the roles defined for the clinic, and columns represent the data requirements of each role. In the case of the THCs, this would be the forms, logs, and reports used by a task. The intersection of each row and column indicates the access permissions for the respective role/data element combination. The matrix is iteratively refined with the clinic staff until an acceptable policy is agreed upon.

The authors believe that the process presented above represents a generic six step process for analyzing the data security requirements in any health care setting. The next section examines the advantages and disadvantages of existing data security mechanisms for use in the CSMSS.

### Existing Data Security Mechanisms

Data security is of importance to the designers of both operating systems (OS) and database management systems (DBMS). Clearly, such systems must provide an acceptable level of security for the resources they manage. CPRSs must address these same concerns as well as additional issues, as they provide similar services in the context of health care information management. The fundamental requirements of a data security model for a CPRS should also include:

- Flexibility. The model must be able to support a broad range of security policies. In particular it must provide both a generalized means of control that can be based on user role and required access, as well as provide for exception handling at the data element level.
- Speed. The model must not hinder the timely delivery of health care information. In addition, both emergency and multi-location access should be feasible.
- Ease of administration. The model must be easy to use and administer. Given the requirements for clinical consultation, practice coverage, and resident and fellow education, the system must be able to maintain an appropriate level of access and privacy.

The feasibility of using existing security models to address these issues is described below.

### Operating Systems

The prevailing strategies for operating system data security are: access control lists, capability lists, multi-level security, and access matrices. The strengths and weaknesses of each will briefly be discussed.

**Access Control Lists (ACL)** have the advantage of providing fine grained access. However, with ACLs searching for a user's access privileges can be computationally expensive; furthermore ACLs are not easy to administer, as it is difficult to determine system-wide access rights for a given user since the rights data is localized by data element[12].

**Capability Lists (C-Lists)** also have the advantage providing fine grained access. C-Lists also localize permissions by user [12]; this makes data security administering on a per user basis a lot simpler. However this localization has the disadvantage of making it particularly difficult to revoke access to a given data element. [11].

**Multilevel Security** has the advantage that it ensures only upward information flow through the model.

This makes administering this model very easy. However, this method is also quite inflexible, since it is not possible to make exceptions when using this model. Due to its rigidity multilevel security can only express limited types of policy [13].

**Access Matrices** have the advantage of providing a very flexible approach to security by permitting a localized view of access either by data element, or user. The disadvantage of access matrices is that they tend to be sparse, and can become quite large and wasteful of system resources [11][12].

### Database Management Systems (DBMS)

Although operating systems generally provide data security for the information stored in their file systems, DBMSs usually need much finer grained access control than provided at the granularity of a file. For this reason many commercial DBMSs provide their own data security mechanism.

**Relational DBMS (RDBMS)** provide security through the Grant/Revoke and view mechanisms. Grant/Revoke has the advantage that it can restrict access to a table or view (discussed next) by restricting the query language[14]. It's disadvantages are that it is not able to limit access to the individual rows in a table, and it does not scale well to a language with more than a small set of terms. Views can confine the user's view of the rows of any given set of tables [15]. Unfortunately, the restriction that views place on table rows tend to be too static and too coarse grained for the level of flexibility required by a CPRS. Views are also sensitive to changes in their constraining clauses and to changes in the underlying data model, thus they can be quite difficult to administer.

**Object Oriented DBMS (OODBMS)** are a relatively new entry into the commercial DBMS market. Consequently, OODBMS vendors are not offering particularly innovative solutions to security issues. However, there is ongoing research in the area, including extending the relational security model to OODBMSs [16] and using security constraints to enforce mandatory and discretionary security [17].

From the foregoing discussion it is clear that none of the traditional data security mechanisms, in and of itself, is sufficient for meeting the requirements of the CSMSS, as established in the analysis of the THC project. Furthermore, they do not support a level of flexibility consistent with the intention of constructing a domain independent CPRS infrastructure. Consequently, in order to meet the needs of the CSMSS, a hybrid data security strategy has been synthesized. This strategy is presented in the next section.

## CSMSS DATA SECURITY STRATEGY

The CSMSS strategy is based on an extended access matrix model hybridized with parameterized role assertion and segmentation to facilitate access to dynamically allocated data elements. Specifically, the rows of the access matrix represent roles and the columns represent data segments.

Segmentation of data resources permits the logical collection of data element types into related groups, thus allowing them to be managed as a single entity. Since there are likely to be far fewer segments than data element types, this adaptation pre-empts one of the chief disadvantages of access matrices, large matrix size.

In order to represent finer grained data security without generating additional segments and roles, parameterized roles have been introduced. Parameterized roles allow the specification of a constrained relationship between roles and data resources. For example, consider restricting access to a certain patients medical records to only her physician; this is possible by parameterizing the physician role by the patient identifier.

Thus far the CSMSS data security model permits security restrictions to be placed upon data element types by role and even permits finer grained restrictions via parameterized roles. However the system must provide, for special cases, still finer grained access by permitting restrictions to be placed upon individual data elements. This is imperative for protecting such information as HIV test results or social worker notes. However, defining access constraints to specific data elements within the access matrix would imply creating a new column for each restricted data element. This results in an unconstrained, and therefore unacceptable, growth in the matrix. A solution is to introduce an auxiliary security mechanism for managing exceptional access control to dynamically created data elements. The mechanism selected is based on the access control list paradigm. Since it is being used for exceptional cases, there should be little overhead placed on the system, as a rapid table lookup will determine whether a ACL entry exists for a given data element.

Finally, it should be noted that the CSMSS data security paradigm can support domains which require either a open/closed (information is generally unrestricted) or a closed/open (information is generally restricted) security model. In the former, the access matrix and ACL entries register restrictions; in the latter, they register permissions.

## CONCLUSION

Data security policy is generally quite dynamic. Roles and tasks may evolve over time; new roles and tasks may be created, and old ones may be retired. For example, new medical procedures can redefine the process involved in completing a task as well as the inputs and outputs associated with a task. Organizational constraints can also change the tasks associated with a particular role. Furthermore, legal constraints can place mandatory requirements on specific types of data[5]. For these reasons, a coherent process for evolving the data security policy as well as a flexible architecture that can support change must be in place.

This paper has presented a six step data security analysis methodology that was applied to the THCs to derive the specific needs for this component of the CSMSS. This analysis and a study of the current security models used by OSs and DBMSs resulted in the development of a generalized hybrid data security paradigm synthesized in part from existing security models. It is felt that the flexibility of this model will realize the broad spectrum of policy based access strategies demanded by CPRS including the CSMSS.

### Acknowledgments

This work has been supported, in part, by grant LM04905 from the National Library of Medicine.

The authors wish to thank the staff of the Baylor College of Medicine Teen Health Clinics for their help on this project.

### References

[1] Weiss, W.V., *Health Care: conflicting opinions tough decisions*, NC Press Limited, Toronto, 1992.

[2] Chueh, H.C., Barnett, G.O., "Client-server, Distributed Database Strategies in a Health-care Record System for a Homeless Population", *Journal of the American Medical Informatics Association*, Vol. 1, No. 2, March/April 1994.

[3] Benjamin, C.D., and Baum B., "The automated medical record: A practical realization?", *Topics in Health Record Management*, Vol. 9, No. 1, 1988.

[4] Henkind, S.J., Orlowski, J.M., Skarulis, P.C., "Application of a Multilevel Access Model in the Development of a Security Infrastructure for a Clinical Information System", *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care*, Washington, MD, October/November 1993.

[5] Dick, R.S, Steen, E.B., Eds., *The Computer Based Patient Record: An Essential Technology for Health Care*, National Academy Press, Washington, DC, 1991.

[6] American Medical Record Association. "Confidentiality of patient health information", position statement of the American Medical Record Association, Chicago, 1985.

[7] Martin, J., *Managing the database environment*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[8] Martin, J., *Computer Data-Base Organization*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[9] Miller, S.P., Neuman, B.C., Schiller, J.I., and Saltzer, J.H., *Section E.2.1: Kerberos Authentication and Authorization System*, M.I.T. Project Athena, Cambridge, MA, December 21, 1987.

[10]Orr, G.A., Brantley, B.A., "Development of a Model of Information Security Requirements for Enterprise-Wide Medical Information Systems", *Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care*, Baltimore, MD, November 1992.

[11]Peterson, J.L., Silberschatz, A., *Operating System Concepts*, Addison Wesley, Reading, MA, 1986.

[12]Tanenbaum, A.S., *Modern Operating Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

[13]Bell, D., Grimson, J., *Distributed Database Systems*, Addison-Wesley, Reading, MA, 1992.

[14]Date, C.J., An Introduction to Database Systems, Volume 1, Addison-Wesley, Reading, MA, 1986.

[15]Pratt, P.J., *Database Systems Management and Design*, Boyd and Fraser, Boston, MA, 1987.

[16]Kim, W., "Architectural Issues in Object-Oriented Databases", *Journal of Object-Oriented Programming*, Vol.2, No. 6, March/April 1990.

[17]Thuraisingham, M.B., "Security in Object-Oriented Database Systems", *Journal of Object-Oriented Programming*, Vol. 2, No. 6, March/April 1990.