Project Title

A Secure Data Management Framework to Enhance Student Learning

Applicant

<table>
<thead>
<tr>
<th>Name</th>
<th>Rajendra K. Raj</th>
<th>Telephone</th>
<th>475-2595</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dept.</td>
<td>Computer Science</td>
<td>College</td>
<td>GCCIS</td>
</tr>
</tbody>
</table>

Project Summary

Most institutions (governmental, corporate, educational, military, or other) often treat their data as their “crown jewels” and store them in database systems. Keeping this data safe and secure, however, remains one of the weakest areas in information security, database security breaches regularly making the evening news. For example, credit card and other personal financial information are so routinely lost by many enterprises that New York State (as several others) now has an Information Security Breach and Notification Act. Other examples of the need for secure data management come from electronic voting systems (safety, security, and provenance of voting data) and data outsourcing (preservation of confidentiality and integrity of personal and enterprise data). Secure data management thus is a major area of focus for both researchers and practitioners.

I originally developed a course in Secure Database Systems as a required course for the “new” GCCIS-wide MS program in Computing Security and Information Assurance. The course content is divided into two roughly equal parts: (1) practical techniques for security used in current commercial or open-source database systems, and (2) cutting edge research techniques for securing data. Starting in Winter 20032 (before the launch of the new MS program), I offered this course as an elective for both Computer Science graduate and upper-division undergraduate students concentrating in Database Systems or in Security. In Winter 20052, the first batch of the MS Computing and Information Assurance students took this course. Student feedback confirms that the current course content and organization is appropriate for its current targets.

This project is an attempt to enhance the learning experience for the existing student population for the current course, and to make the course content viable for a larger college-wide student population, especially at the undergraduate level. This project will also serve as a proof-of-concept for a grant proposal to be submitted to NSF’s Course, Curriculum, and Laboratory Improvement (CCLI) program, which funds projects that build on prior work and contribute to undergraduate research and practice.

I will develop a suitable database framework that will consist of the following:

- Practical datasets (e.g., credit card information, student information, or electoral results) that can be used to simulate practical real-world situations, along with associated client access interfaces.

- Skeletal database applications that will make use of the developed datasets.
The framework will be designed to permit (1) experimentation with practical techniques for data security available in current database management systems, and (2) adaptation of and experimentation with material from current secure data management research. This framework will be developed with the assistance of a full-time coop student. As the framework is being worked on, I will also develop suitable hands-on laboratory assignments where students can experiment with both practical and “theoretical” techniques for securing data. Topics for these laboratories will be selected from topical secure data management issues, e.g., Hippocratic databases; intrusion detection in databases; database watermarking; database auditing; multi-level database security; multi-database security; enforcing HIPAA-level database privacy; and managing secure data with outsourced databases. Because these laboratories will become the major vehicle for learning in the revised course, I will develop a set of notes to help guide highly interactive and collaborative discussions among students, as well as with the course instructor. The revised course thus will use an active learning format, which will ensure an improved student learning experience.

Given the advanced nature of this project, I plan to use my contacts in relevant companies (specifically healthcare and financial services) and universities to evaluate the developed course material. I will also invite local data security experts to give talks in the class. The project budget includes a provision for honoraria for these experts.

The major deliverables of this project include:

1. The secure database framework containing both datasets and associated client access interfaces.
2. A set of associated laboratory instructions.
3. A summary report containing examples of student work, discussions of student feedback, and comparisons of student evaluations in the current and revised courses, and the overall learning effectiveness of the revised course.
4. A grant proposal to NSF’s Course, Curriculum, and Laboratory Improvement (CCLI) program.
5. A paper discussing the innovations of and experiences with this project will be written and submitted to a premier international pedagogical conference in computing, e.g., ACM SIGCSE (Special Interest Group on Computer Science Education) or IEEE Frontiers in Education (FIE). I also plan to make presentations about the project at universities (including RIT) and conferences.
6. The final PLIG report summarizing all accomplishments.

**Targeted Learners**

The primary audience for this project will be students with a database systems background who want to get hands-on experience with the principles and practices of building secure data management systems. Specifically the typical learner will be an upper-division or graduate student in GCCIS (or another college) who has taken an introductory database system course or worked with databases during a coop.

**Affected Students**

The number of affected students will be around 90 students per year, assuming that the course will be offered twice a year. If the course continues to be offered only once, 45 students per year will be affected.

**Impact on Teaching and Learning**

This project will have a positive impact on teaching and learning in several ways. The students will be
exposed to a holistic university course that covers both commercial practice and current research. The use of active learning techniques will improve students' learning experiences and lead to a deeper understanding of computing security and information assurance. The project will help ensure our students are better trained to meet contemporary workplace needs where secure data management is critical. The revised course will also prepare our students for undergraduate and graduate research in an exciting and highly relevant area of modern computing.

**Impact on Student Success**

The secure data framework will make this course more stimulating, interesting, and relevant to the targeted students. The innovations envisaged here conform to RIT’s goal of providing students with learning experiences that are immediately useful in industry or research. The planned improvements will help students learn a set of advanced concepts in a hands-on setting, which should help improve their coop experiences, find challenging jobs in industry, or pursue further advanced studies. All of these improvements should in turn help with retention.

**Measuring the Impact**

Evaluations of the teaching and learning innovations of this project will follow a four-part procedure:

1. Measurements will be made to compare the effectiveness of the revised course over the existing course. The focus will be on comparing student feedback, satisfaction, and performance between students taking current and revised versions of the course.

2. Likert scale measures of student perceptions of each component (datasets, laboratories, discussion guides) of the revised course will be obtained via weekly surveys when the revised course is offered. The focus will be on evaluating the new components.

3. Feedback from invited speakers and reviews of paper submissions to peer-reviewed conferences will provide additional evaluations of the curricular innovations of this project.

4. Project success should reflect in increases in students who complete subsequent Honors or MS capstone projects in database systems or security, and decreases in the average time taken to complete these capstone projects. Because these numbers are lagging indicators of project success, I plan to make such measurements approximately one year after project completion, along with Likert scale measures of student perception of the role played by the revised course.

I will disseminate the project deliverables to GCCIS and other faculty interested in database systems, or computing security and information assurance. The deliverables will also be published on a website accessible to the external computing community, and the website will be publicized using venues such as ACM SIGMOD’s (Special Interest Group on Management of Data) DBWorld mailing list. In addition, I will make presentations about the findings to GCCIS and other faculty, and submit presentation proposals for international conferences such as SIGCSE or FIE, and at regional conferences such as CCSCNE.

**Project Rationale**

a. Few universities worldwide offer courses in secure data management; the ones that do either offer graduate level courses that explore research in this area, or undergraduate courses that cursorily introduce basic security features available in commercial database systems. No attempt has been made, at RIT or elsewhere, to develop a secure data management framework that covers both commercial practice and current research in a holistic manner as proposed here. The project innovations require faculty time and effort that is well beyond the scope of regular college business.
b. This project lies in the core competencies of the Computer Science department and GCCIS to offer courses in computing security and databases. As stated previously, this course is required for the college-wide MS program in Computing Security and Information Assurance. Additionally, the Security and Data Management clusters in the MS Computer Science program currently include this course as an elective, and many students in both clusters typically take this course.

c. Once the secure data management framework and the associated laboratories are in place, I will work with faculty in the IT and NSSA departments in GCCIS to explore possibilities for its use in teaching security issues, both in a purely database systems setting and in a networked client-server setting.

d. I believe I have the required credentials and experience for this project. As mentioned, I developed the original Secure Database Systems course proposal, and have taught the current course three times. I regularly teach the Computer Science courses in introductory and advanced database systems, and have chaired over a dozen MS capstone projects involving database systems. I am also involved in Computing Security having been a member of the group that helped develop the GCCIS-wide MS program in Computing Security and Information Assurance. I have taught a seminar course on Language-Based Security twice, and have jointly developed a new seminar course in Secure Software Systems that will be offered next year. Finally, I have worked in the financial industry for 9 years where I designed, developed, and supported state-of-the-art secure database systems.

e. As stated in part (a) above, this project represents innovation in this discipline. Therefore, the project deliverables include an NSF CCLI grant proposal and a paper discussing its innovations to be written and submitted to an international pedagogical conference in computing.

**Project Timeline**

| Development of framework and laboratories | Summer 20054 | Fall 20061 | Winter 20062 | Spring 20063 | Summer 20064 |
| Revisions to framework and course content | √                 |             |              |              |              |
| First offering of revised course |                                    | √             |              |              |              |
| Evaluation of student learning and of the framework, labs, and course content |                                | √             | √             | √             |              |
| Dissemination—papers, presentations, follow-on grant proposals, and PLIG report |                                | √             | √             | √             |              |

The above table summarizes the project development plan. I plan to work over Summer 20054 with a coop student to develop the proposed course material, and then make needed revisions in Fall 20061. I will use the developed material in Winter 20062 when the course will be offered next. Data security experts external to RIT will be invited to give talks in the class or evaluate the developed course material. I will start collecting assessment data in Winter 20062; analyze the data, and write reports, papers, and follow-on grant proposals in Spring 20063; and complete all deliverables in Summer 20064.