Martin Sadler
Hewlett Packard

Trust In a Dotcom World

Is security a done deal? Far from it. The talk will emphasize how little we know about information assurance as we move to a world of virtual corporations, dotcoms, portals, marketplaces and the complex interdependencies between online companies - a context that does not easily respect borders. It will touch upon some of the problems that HP's research laboratories are tackling in this space, but stress the importance of universities, governments and industry working together. HP like many other companies welcomes the leadership that CERIAS provides.

Jens Palsberg
Associate Professor of Computer Sciences

Software Watermarking with Secret Keys

Software watermarking has been studied in the 1990s by Collberg and Thomborsen, and others, and it is the subject of at least four U.S. patents. The previous techniques include approaches to resist attacks by semantics-preserving program transformations. Until now, there has been no good defense against an "open-source'' attacker, that is, an attacker with access to the software for embedding and extracting watermarks. To counter such attackers, we have designed and experimented with an approach to software watermarking that uses a secret key during the embedding and the extraction of a watermark. The secret key makes it harder to locate the watermark for an open-source attacker without access to the key. Our experiments show that watermarking with a secret key can be done efficiently with moderate increases in code size, execution times, and heap-space usage, and give a major disadvantage for open-source attackers.

People involved in this project:

Sowmya Krishnaswamy, Minseok Kwon, Di Ma, Qiuyun Shao, and Yi Zhang

Victor Raskin
Professor of Linguistics and
Coordinator, Natural Language Processing (NLP) Laboratory

Natural Language Processing for Information Security: Establishing a New Paradigm

This paper explores a promising interface between natural language processing (NLP) and information assurance and security (IAS). More specifically, it is devoted to possible applications of the accumulated considerable resources in NLP to IAS. The paper is of a mixed theoretical and empirical nature. Of the four possible venues of applications, (i) memorizing randomly generated passwords with the help of automatically generated funny jingles, (ii) natural language watermarking, (iii) using the available machine translation (MT) systems for (additional) encryption of text messages, and (iv) downgrading, or sanitizing classified information in networks, two venues, (i) and (iv), have been at least partially implemented and the remaining two (ii) and (iii) are being implemented to the proof-of-concept level. We must make it very clear, however, that we have done very limited experimentation or evaluation at this point, though we are moving quickly in that direction. The merits of the paper, if any, are in its venture to make considerable progress in NLP, especially in knowledge representation and meaning analysis, useful for IAS needs. The NLP approach adopted here, ontological semantics, has been developed by one of the coauthors and his associates; watermarking is based on the pioneering research by another coauthor and his associates; most of the implementation of the password memorization software has been done by yet another coauthor.

Other people involved in this project:

Mikhail J. Atallah, Craig J. McDonough, Victor Raskin

John R. Rice
Professor of Computer Sciences

Secure Outsourcing Of Scientific Computations

We investigate the *outsourcing* of numerical and scientific computations using the following framework: A *customer* who needs computations done but lacks the computational resources (computing power, appropriate software, or programming expertise) to do these locally, would like to use an external *agent* to perform these computation. This currently arises in many practical situations, including the financial services and petroleum services industries. The outsourcing is *secure* if it is done without revealing to the external agent either the actual data or the actual answer to the computations. The general idea is for the customer to do some carefully designed local preprocessing (*disguising)* of the problem and/or data before sending it to the agent, and also some local postprocessing of the answer returned to extract the true answer. The disguise process should be as lightweight as possible, e.g., take time proportional to the size of the input and answer. The disguise

preprocessing that the customer performs locally to "hide" the real computation can change the numerical properties of the computation so that numerical stability must be considered as well as security and computational performance. We present a framework for disguising scientific computations and discuss their costs, numerical properties, and levels of security. We show that no single disguise technique is suitable for a broad range of scientific computations but there is an array of disguise techniques available so that almost any scientific computation could be disguised at a reasonable cost and with very high levels of security. These disguise techniques can be embedded in a very high level, easy-to-use system (problem solving environment) that hides their complexity.

People involved in this project:

Mikhail J. Atallah, K.N. Pantazopoulos, John R. Rice, Eugene H. Spafford

Mikhail Atallah
Professor of Computer Sciences

Secure Multiparty Protocols for Approximate Searching

Suppose that A has a database D and that B wants to perform a search query q on D (e.g., "is q in D?"). There are elegant cryptographic techniques for solving this problem under various constraints (such as "A should know neither q nor the answer to the query" and "B should learn nothing about D other than the answer to the query"), while optimizing various performance criteria (e.g., amount of communication). We consider the version of this problem where the query is of the type "is q approximately in D?" for a number of different notions of "approximate", some of which arise in image processing and template matching, while others are of the string-edit type that arise in biological sequence comparisons. New techniques are needed in this framework of approximate searching, because each notion of "approximate equality" introduces its own set of difficulties; using encryption is more problematic in this framework because the items that are approximately equal cease to be so after after encryption or cryptographic hashing. Practical protocols for solving such problems make possible new forms of e-commerce between proprietary database owners and customers who seek to query the database, with privacy and security.

Student involved in the research: Kevin Du

Edward J. Delp
Professor of Electrical Engineering

Scene Adaptive Video Watermarking

The security of multimedia data has become an important research area with the recent growth of new digital imaging technologies. This growth has created a need for techniques that can be used for copyright protection of digital images and video. Our research has focused on algorithms for image and video authentication and forgery prevention known as digital watermarking.

In this project we are developing watermarking techniques for digital video sequences that exploit human visual system models and are compliant with the MPEG compression standards. Our previous research has indicated that watermarking techniques for still images that use visual models are more robust against attack.

Deborah Bennett
Assistant Professor of Educational Studies

Data Security Issues in Electronic Educational Assessment Systems

The Indiana Assessment System of Educational Proficiencies (IASEP) represents the first statewide attempt to collect, store, and transmit electronic performance ratings and multimedia documentation. This presentation will focus on the collaborative efforts of CERIAS and the School of Education at Purdue in applying data security models and principles to protect sensitive IASEP educational records. An analysis of IASEP and the creation of a customized security architecture will be discussed. The process of matching security controls with security risks and the selection of appropriate security software, hardware, and user guidelines will also be highlighted. Finally, the importance of user error in managing educational data will be discussed, and a teacher training protocol will be presented.

People involved in this project:

Stephanie Miller, former CERIAS Master's Student
Steve Hare, CERIAS Associate Director
Deb Bennett, SOE Assistant Professor of Educational Studies
Jennifer Radecki, SOE Master's Student and Research Assistant

Clay Shields
Assistant Professor of Computer Sciences

Denial of Service, Traceback and Anonymity

With the widespread adoption of the Internet as a communication medium, network security has become a research topic for academia and industry alike. For companies and government organizations, the often issues include the secure dissemination of information, detection and trace back of network intruders, and understanding and prevention of denial-of-service attacks. Individuals are often concerned with maintaining their privacy throughout network transaction. This talk will cover on-going and future research efforts at CERIAS that address these important topics.

Ahmed Elmagarmid
Professor of Computer Sciences

Association Rule Hiding

Presented by: Yucel Saygin and Elena Dasseni
Department of Computer Sciences, Purdue University

The security of data, against unauthorized access, has been a long-term goal for the database security research community and the government statistical agencies. Recent advances, in data mining and machine learning algorithms, have increased the security risks encountered when releasing data to outside parties. However, issues related to data mining and security have not been investigated until recently.

The association rule discovery is one of the major application areas of data mining and related technologies. The association rules form a special category of inference rules -- very similar to deductive rules and functional dependencies -- that indicates the degree of coupling between a pair of disjoint patterns. If the information or the knowledge content of a certain association rule is above a certain privacy threshold, then this rule is characterized as sensitive. Sensitive rules should not be disclosed to the public for a number of different reasons.

In this project, we investigate the disclosure limitation of sensitive association rules by tuning their support and their confidence. A transformation of the original database is performed in order to bring the support and the confidence of these sensitive rules below a certain threshold. We refer to such a transformation as "sanitization" of the database.

Finally we prove that the optimal solution of the sanitization problem is NP-hard and we also devise heuristic algorithms for obtaining a computationally satisfactory solution that maximizes the information and knowledge content of the disclosed database. Analytical and experimental results indicate that the proposed heuristics outperform the time and space complexity of other naive approaches.

People involved in this project:

M. Atallah
CERIAS and Department of Computer Sciences, Purdue University

E. Bertino
Dipartimento di Scienze dell'Informazione, Universita degli Studi di Milano, Italy

A. Elmagarmid
Department of Computer Sciences, Purdue University

V. Verykios
College of Information Science and Technology, Drexel University

Josh Boyd
Assistant Professor of Communication

Community Is Security: Online Security Communication At Ebay

This presentation focuses on how eBay has rhetorically created its auction site as a safe place for people to do business. In spite of people's uncertainties about online auctions--unknown sellers, merchandise seen only in pictures--eBay has persuaded over 10 million people to become registered users. EBay has argued that community is security, and the presentation isolates key elements of community eBay employs to help define the new phenomenon of the online community of commerce. It also examines recent changes to eBay's system, suggesting that so-called improvements might actually weaken the community security already in place--a warning to other sites that might imitate eBay's community security approach. Graduate students Josh Clark and Alex Miranda assisted me on this project.

Mr. Michael Fleming
Chief, Information Assurance Solutions Office
National Security Agency

Dr. Vic Maconachy
Program Manager, National Information Assurance Education and Training Program
National Security Agency

Information Assurance Challenges for the 21st Century

In today's increasingly dependent and interdependent global information society, information assurance for systems is gaining tremendous importance. Individuals, governments, and societies are insisting on secure and safe communications environments. The solution to providing those assurances lies in the formation of partnerships between and among business, academia and government. Mr. Fleming will present an overview of a model for such partnerships, to include critical elements for the success of those joint ventures. Dr. Maconachy will briefly discuss the cognitive and social implications related to moving towards a global security-based enterprise.