Research in Cryptography, Information Security and Algorithm Development
9807-12 & 26

Progress Report: July 1, 2000—December 31, 2000

Shafi Goldwasser, Ronald L. Rivest and Michael Sipser

Project Overview

Recent NTT-sponsored research has focussed on both designing secure cryptographic protocols and showing their optimality by exhibiting impossibility of stronger cryptographic protocols in both multi-party(distributed) and two-party scenarios.

The emphasis of the research which is expressed in each one of our projects is is on the development of novel frameworks and theoretical provably secure solutions to problems arising from applications.

Progress Through December 20000

We highlight progress made during this period on several ongoing projects.

A. AUTHOR IDENTIFICATION BY COMPRESSION METRICS

An intriguing question brought up by the presence of the Internet, is how to identify whether different texts were generated by the same person. This question of author attribution, of course is not new, and linguists have long debated of whether famous bodies of literature were written by Shakespeare or many authors (to name one such debate). However, with the onset of the Internet we both have access to much more data by many more authors, and great incentive to embark on the task of identifying authors not merely for intellectual pursuit, but as a means to ensure tracking of individuals who violate the law over the net. Moreover, the question of author attribution is a special case of a wider class of "filtering" one type of data (say technical papers) vs. another (pornography).

This year Nitin Thaper and Shafi Goldwasser have embarked on a new project of identifying whether texts have been written by the same author, via the use of text compression algorithms.

The idea is, that if you try to compress texts written by one author using a compression model built by another text of the same author, a much better compression rate would be achieved then by using a compression model built by another text of another author. The idea to use compression algorithm comes from the essential "equivalence" between prediction and compression -- the ability to build a good succinct predictor of text, yields a good compression algorithm, and every compression algorithm can be viewed as a predictor.

Excellent implementations of compression algorithms are available, and Nitin Thaper has used several to test our hypothesis in his Masters thesis [N]. Indeed, the idea seems to work with a great degree of accuracy.

The system is divided into 2 phases: a training phase and a testing phase. During the training phase texts by known different authors are used to build models which are then used in the testing phase to aid in compression of texts which wish to be classified as having been written by one author or another. Finally, Thaper used clustering algorithms to cluster works of the "same" author in a setting for which there is no known a-priori authors which you are trying to matcah against, but rather just vast amount of texts for which it is not known whether they were written by one, two, or many authors.

[N] Using Compression for Source Based Classification of Text, by Nitin Thaper Master Thesis, filed February 5th 2001.

B. AMORTIZED ELECTRONIC CASH

We present an e-cash scheme which provides a trade-off between anonymity and efficiency. In many e-cash schemes, the provable anonymity of an e-coin is achieved by means of a zero-knowledge protocol, while the authenticity of an e-coin is achieved by a digital signature. Therefore, the computation involved in generated each e-coin is quite expensive. We show how to amortize the cost of a zero-knowledge protocol and a signature by using them to generate an arbitrary number of e-coins.

Our work actually solves an open problem of Okamoto (CRYPTO '95) in "divisible e-cash." Namely. we achieve results similar to those of Okamoto, but (1) in a simple fashion, (2) based on simple and traditional complexity assumptions (rather than ad hoc ones), and (3) within a much crisper definitional framework that highlights the anonymity properties (rather than intuitively leaving them to the reader to deduce).

[LM] Amortized E-Cash by Moses Liskov and Silvio Micali (Submitted.)

C. FORWARD-SECURE SIGNATURE SCHEMES

A signature scheme is said to be forward-secure if compromise of the secret signing key does not compromise the security of previously signed messages. To make a signature scheme work, one clearly needs to periodically update the secret signing key without having to change the public signature verification key. Ross Anderson, and then Mihir Bellare and Sara Miner, have developed interesting proposals for forward-secure digital signature schemes.

Recently, we have developed a new forward-secure digital signature scheme that minimizes the size of the secret key information that the signer must keep, while maintaining the security of this scheme comparable to that of the previous proposals.

This is work by Leo Reyzin and Michel Abdalla.

D. EXPOSURE-RESILIENT CRYPTOGRAPHY

In this paper, we consider the question of adaptive security for two related cryptographic primitives: all-or-nothing transforms (AONT) and exposure-resilient functions (ERF). Both are concerned with retaining security when an intruder learns some bits of a string which is supposed to be secret, and have a variety of applications.

In the strongest possible setting (so-called "perfect security"), adaptive and non-adaptive security are equivalent. We give a new lower bound for AONT's in this setting. This proves the near-optimality of some known constructions; it also provides a new lower bound on ramp secret-sharing schemes.

With the weaker but more realistic requirement of "statistical security", we show that adaptivity adds strictly more power. We relate and reduce the construction of adaptive ERF's to that of "almost-perfect resilient functions" (introduced by Kurosawa et al.), where the adversary can actually set some of the input positions. We give a simple probabilistic construction of these functions based on universal hash functions which is essentially optimal and improves on previous constructions. As a result, we get nearly optimal adaptively secure ERF's and AONT's.

[DSS] On perfect and adaptive security in exposure-resilient cryptography, by Y. Dodis, A. Sahai and A. Smith. To appear in Proceedings of EUROCRYPT '01, May 2001.

E. SIGNATURE ALGEBRAS

We have been investigating the question as to whether it is possible to design signature schemes with interesting algebraic properties.

Most specifically, the properties are of the form: Can the verifier compute sign(x op y) from sign(x) and sign(y), for interesting algebraic operators "op". That is, a verifier Bob who doesn't know Alice secret signing key is given Alice's signature on x, sign(x), and also given Alice's signature on y, sign(y), and we would like Bob to be able to produce, on his own, Alice signature on the quantity (x op y).

Of course, this requires some modification of the usual notion of security for a digital signature scheme, since computing a signature for (x op y) should be considered a feature and not a forgery.

We have produced signature schemes that have such interesting properties:

(1) A transitive signature scheme, wherein if Bob has Alice's signature on an edge (u,v) of a graph, and also Bob has Alice's signature on an edge (v,w) of the graph, then Bob can compute on his own Alice's signature on the edge (u,w) of the graph. This problem is motivated by various applications in network management and public-key infrastructure. Our solution works for undirected graphs.

(2) A prefix aggregation scheme, wherein if Bob has Alice's signature on the bit-string x0 and Bob also has Alice's signature on the bit-string x1, then Bob can compute on his own Alice's signature on the bit-string x. The problem is motivated by problems in authenticating network routing tables.

This work is by Ron Rivest, Silvio Micali, Tal Rabin, and Suresh Chari.

F. MIN-ROUND RESETTABLE ZERO KNOWLEDGE

In STOC 2000, Canetti, Goldreich, Goldwasser, and Micali put forward thestrongest notion of zero-knowledge to date, *resettablezero-knowledge* (RZK) and implemented it in constant rounds in a new model, where the verifier simply has a public key registered before any interaction with the prover.

To achieve ultimate round efficiency, we advocate a slightly stronger model. Informally, we show that, as long as the honest verifier does not use a given public key more than a fixed-polynomial number of times, there exist 3-round (which we prove optimal) RZK protocols for all of NP.

[RM] Min-round resettable zero knowledge in the public-key model. by Leonid Reyzin and Silvio Micali. Eurocrypt 2001, to appear.

G. SOUNDNESS IN THE PUBLIC-KEY MODEL

The public-key model for interactive proofs has proved to be quite effective in improving protocol efficiency [CGGM00]. We argue, however, that its soundness notion is more subtle and complex than in the classical model, and that it should be better understood to avoid designing erroneous protocols. Specifically, for the public-key model, we

* identify four meaningful notions of soundness;
* prove that, under minimal complexity assumptions, these four notions are distinct;
* identify the exact soundness notions satisfied by prior interactive protocols; and
* identify the round complexity of some of the new notions.

[CGGM00] Resettable Zero Knowledge. by Ran Canetti, Oded Goldreich, Shafi Goldwasser, and Silvio Micali.

[RM] Soundness in the Public-Key Model. by Leonid Reyzin and Silvio Micali. In submission.

Research Plan for the Next Six Months

Over the next six months, we plan to continue our work on the both the fundamentals and the applications of cryptography to information security. We tend to work on a wide variety of problems simultaneously; it is hard to predict exactly which problems and approaches are likely to be fruitful. However the following directions will certainly be emphasized over the next six months.

-- Signature algebras: We hope to extend our work on transitive signature schemes to directed graphs. Similarly, we would like to extend our work on prefix aggregation schemes to handle arbitrary "and" and "or" constructions on values in a monotone circuit (e.g. for access control).

-- Electronic voting: Given the problems seen in the U.S. Presidential election of November 2000, we will be spending some time re-considering the potential of electronic voting technology. In particular, the previous NTT-supported work on voting based on the Fujioka-Okamoto-Ohta protocol will be reviewed in light of the evolving requirements for voting, and new protocols may be designed. Receipt-freeness is turning out to be an interesting and challenging design goal.

-- Credential systems: We plan to continue our work on pseudonyms and credential systems. Efficiency is still something of a problem with the known techniques; can we do better?

-- Proof techniques for cryptographic protocols: We continue to explore expanded models of adversarial capability, to design new cryptographic protocols to defeat such adversaries, and to invent new proof techniques for proving such protocols to be secure. Some of this work may go in the direction of the recent Abadi-Rogaway proposal to unify classic cryptographic proof techniques with more standard program-proving techniques from logic; issues such as using keys to encrypt themselves are particularly hard to handle.

-- User authentication in pervasive computing environments: This problem is particularly difficult when the basic components (e.g. cell phones or laptops) may be anonymous commodity-like devices that the user picks up or borrows as needed. We are studying schemes for user authentication involving user-worn beacons that broadcast cryptographic information to these insecure devices that can then act on the user's behalf.

-- Fundamental cryptographic primitives: There remains a rich store of open problems in the area of fundamental cryptographic primitives. Recent advances, e.g. in all-or-nothing transformations and block ciphers, suggest many open problems. For example, there are generalizations of the Luby-Rackoff paradigm that remain intriguing and unsolved: when does a generalized network of Feistel-like transformations yield a pseudo-random permuation or a super pseudo-random permutation?