Report on DIMACS[*] Workshop on Cryptography: Theory Meets Practice

Presented under the auspices of the
Special Focus on Communication Security and Information Privacy
and the PORTIA project.

Date of Workshop: October 14 - 15, 2004

Workshop Organizer:

Dan Boneh, Stanford University

Report Author:

Constantin Serban, Dept. of Computer Science, Rutgers University
serban@cs.rutgers.edu

Date of report: March 16, 2005

[*] DIMACS was founded as a National Science Foundation Science and Technology Center. It is a joint project of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, NEC Laboratories America, and Telcordia Technologies, with affiliated partners Avaya Labs, HP Labs, IBM Research, Microsoft Research, and Stevens Institute of Technology.

1 Workshop Focus

Cryptography plays a central role in the design, analysis and implementation of secure systems and communications. Applications of cryptography range from the creation of pure mathematical objects to the detailed engineering specification of complex cryptographic systems.

In contrast to other engineering areas, empirical methods like simulation cannot be used for assessing the security properties of a system. In order to determine the security of cryptographic systems, applied cryptography requires theoretical analysis and mathematical proof based on a careful modeling of the security objectives of the construct and of the attacker capabilities. Cryptography theory faces the challenge of providing both models and constructions that represent in a satisfactory way the needs of actual cryptographic practice. Cryptography thus requires relatively simple and efficient constructions without giving up the essential role of sound mathematical analysis. Due to this close relationship between theory and practice in the cryptography and security areas, in recent years the crypto community has increased its influence in the development of standards and other widely used security systems.

This workshop exposed and encouraged work with a significant theoretical analysis component, that has at the same time meaningful implications and relevance to practical cryptographic and security schemes. It also highlighted new cryptographic requirements of security systems in general. The interaction among the participants of the workshop had a dual effect: it increased the awareness of the need for sound cryptography; and it contributed to a better understanding by the crypto community of the actual needs of practical security systems.

2 Workshop Presentations

2.1 A Cryptographic Model for Access Control

Speaker: Shai Halevi, IBM T.J. Watson Research Center

Dr. Halevi discussed a security model that reconciles cryptography and access control. He began his talk by pointing out the differences between the typical cryptographic model, where the attacker observes and accesses the network, and the reality when the attacker takes over the end hosts of the parties involved in communication. In order to present the differences in security approaches for cryptography and for access control, Dr. Halevi presented several models. He started with two models of cryptography security. He introduced the probabilistic model, then presented the example of the secure channel of communication. He pointed out that the abstraction of a perfect communication channel is never achievable.

Dr. Halevi continued with the access control model: the discretionary model. After presenting the model, he pointed out that the abstraction does not take into account when a participant cooperates with the attacker. Following this model, Dr. Halevi discussed mandatory access control. He made the interesting observation that secrets are not entrusted to persons anymore, but to software running on behalf of persons. He underlined the fact that trusting the person is not equivalent to trusting her software. In order to trust a discretionary scheme, it should be composed of trusted components and the enforcement should be enforced at the entry point in the network. Dr. Halevi gave a number of examples of components that can be trusted (special-purpose network cards) or not trusted (operating systems, applications on top of it). He developed the example of an object storage model, where a number of users access information on a number of shared disks. He gave the example of capabilities with restricted delegation in order to achieve security for such an application.

Finally, Dr. Halevi concluded his talk by emphasizing the need for information flow restriction in conjunction with cryptographic models. In general, the attack models have to align to the realities where the overall software of the user cannot be trusted anymore, instead the user should rely on small pieces of software that are proven to be secure.

2.2 Breakthrough-Resistant Cryptography

Speaker: Adam Stubblefield, Johns Hopkins University

Mr. Stubblefield addressed the issue of building robust cryptographic protocols. These protocols remain secure even in those cases when their underlying primitives fail. He motivated his work by the fact that most common cryptographic protocols rest on the security of a very small set of atomic primitives. As other speakers pointed out during the discussions, these constructions exhibit a number of vulnerabilities and thus represent a single point of failure.

Mr. Stubblefield analyzed the trustworthiness of common security protocols under the assumption of the failure of some of the underlying primitives.

2.3 The Risks of Electronic Voting

Speaker: Dan Wallach, Rice University

Dr. Wallach talked about the security issues following the pervasive deployment of electronic voting technology. He began his talk by emphasizing that people rely on the fact that the voting machines work correctly: i.e. that every vote is secret, counted and not altered, and that it is important that the model is matched by the reality. He pointed out, however that the reality falls short of the model, because of human factor issues, mechanical flaws and other more severe failures. He then continued by discussing different voting technologies, their comparative advantages and disadvantages.

After emphasizing the simplicity and auditability of the mechanical and optical scan voting systems, Dr. Wallach approached the e-voting schemes. He pointed out that beside the obvious benefits there are serious drawbacks, among them the most important is the lack of an audit trail and no indication that the vote has been counted correctly. In order to address this issue, Dr. Wallach pointed out that the trust relies on certifications from independent authorities, a so-called “faith-based voting.” Dr. Wallach continued his talk by addressing the security issues in several computerized voting machines, and emphasized that hybrid systems that leave a paper trail are the most secure because the result of the election can be verified by hand recounts, thus reducing the trusted computing base (TCB). In revealing the vulnerabilities of the Direct Recording Electronic (DRE) systems, Dr. Wallach presented the findings related to several voting machines currently deployed throughout the country: all users have the same passwords, audit logs can be bypassed, and they can be easily modified. He then showed the example of a smart card protocol that has been designed with no security in mind. He raised the issue of software engineering while devising such systems and remarked that the code quality is well below any “high assurance system”.

Dr. Wallach concluded his talk by emphasizing the importance of these machines for democracy. He reiterated that trust is placed in independent testing authorities, who already have certified poorly designed machines. He encouraged the audience to be active in the electoral process and to question and oversee the procedures.

2.4 Cryptography and the Internet: Where It Is, Where It Isn’t, Where it Should Be - and Why It Isn’t There...

Speaker: Steve Bellovin, AT&T Labs Research

Dr. Bellovin discussed a number of applications that currently use cryptographic techniques as well as applications that require it, and he analyzed a number of causes that mitigate the spread of cryptography to real life applications. Dr. Bellovin began his talk by presenting a number of successful applications of cryptography, SSL and HTTPS, discussing their actual usage and their benefits. Then he continued with the presentation of S/MIME and PGP and pointed out that their use is very low. He drew the same conclusion about IPSEC and IKE, in this case due to poor design and implementations. A particular case he discussed was DNSsec that is difficult to design due to the original DNS scheme. Dr. Bellovin stressed that this example constitutes an argument for designing the protocol and its security mechanisms together. Finally, he presented SSH, and its successful use due to easy deployment.

In the second part of his talk, Dr. Bellovin discussed a number of applications where the use of cryptography is required but no progress has been made so far. He started with routing protocols. In the case of BGP, Dr. Bellovin pointed out that the proposed solutions don’t match the operational reality. In the case of anti-spam applications, he pointed out that the problem is not necessarily the authentication of the source but rather the authorization of the email. At present there is virtually no authorization scheme employed when receiving email, and the source authenticity has not much value since most of the spam comes from hacked computers. With respect to non-repudiation, Dr. Bellovin questioned the problem all together and dubbed it “a cryptographer’s trick.” He pointed out that real signatures are strongly bound to the person and weakly bound to the document, whereas digital signatures are strongly bound to the document and weakly bound to the person.

In the last part of his talk, Dr. Bellovin concentrated on the actual reasons cryptography is not used. He began by pointing out that for most users there is no perceived threat of eavesdropping (the bad guys prefer to hack the servers). The second reason he enumerated referred to the comparative weakness of the endpoints and the communication security. He observed that host security is very weak. With respect to the ease of use of cryptography, he concluded that it is hard to configure and inherently complex, and that users prefer to have security be transparent. As the last reason hindering the deployment of cryptography he cited the operational errors. He pointed out that a successful crypto design has to match closely the operational environment, and has to mirror real life transactions.

Dr. Bellovin concluded that most of the problems are not due to the lack of crypto science. The real challenges lay in doing basic engineering, taking into account the human factor, and, most important, binding cryptography to reality. It was also observed that the user needs to be educated regarding the security threats and solutions.

2.5 Efficient Privacy-preserving Information Sharing: Set Intersection and Threshold Set Intersection

Speaker: Dawn Song, Carnegie Mellon University

Dr. Dawn Song discussed the problem of computing the intersection of private sets. She started by offering a brief presentation of the set-intersection problem: given `n` parties, each having a private input set `S`, after engaging in the protocol each party learns the intersection of the sets. She continued by presenting several problems derived from the set-intersection problem: cardinality-set intersection, threshold set-intersection, and over-threshold set-intersection. All these problems present the parties involved in the activity various facets of the intersection set.

In the second part of her talk, Dr. Song presented two adversary models she considered to solve these problems. In the HBC model, an honest but curious adversary cannot acquire any other information but the one officially disclosed by the protocol. In the second model, a malicious adversary can submit any input against the protocol in order to extract more information about the set-intersection and subsequently about the private sets of other parties. She continued by presenting the results of her research in finding more efficient protocols to solve the above problems.

Dr. Song finished her talk by emphasizing the importance of the set-intersection problems to a number of real-life applications. A number of possible applications to private computation have been presented, among them the security check of passenger lists for airlines, and lists of patients filling the same prescription at multiple pharmacies.

2.6 Recent Progress in Anonymous Communication

Speaker: Mike Reiter, Carnegie Mellon University

Dr. Reiter addressed the issue of anonymous communication between parties sharing a communication server. He started his presentation by describing the fragile mixing technique. His model is designed to protect against mixers whose system administrators are not trusted entirely. In order to achieve anonymity of communication, the system administrators are discouraged from revealing mixing information by disclosing an all-or-nothing strategy. The proposed solution works in batch mode, where a set of messages are permuted and transformed prior to sending them such that the secrecy of the mixing depends on each input-output pair. It has been pointed out, however, that such mixing is vulnerable to timing attacks, where an attacker can observe the time patterns of the incoming and outgoing messages in order to find the communicating parties. Dr. Reiter continued to underline the weaknesses in the current solutions to the timing attacks, emphasizing the effects of the dropped messages by the attacker. He proposed a solution called defensive dropping designed to alleviate the timing analysis for low-latency anonymous communication.

In the last part of his talk, Dr. Reiter approached the issue of anonymous push-and-pull communication using a set of database servers. He emphasized that P3 communication achieved private database communication and private retrieval of database records while maintaining asynchrony and oblivious access control.

2.7 Randomness Extraction and Key Derivation Using Common Pseudorandom Modes

Speaker: Hugo Krawczyk, IBM Watson

Dr. Krawczyk addressed the issue of randomness extraction. He based his argument on the fact that most applications today are using well known hash functions in order to extract randomness. This practice, he emphasized, is not appropriate because it represents a strict relaxation of modeling an un-keyed hash function as a random function. The difference between pseudorandom generation and randomness extraction is that the former uses a random secret key while the later uses a random but known key.

Dr. Krawczyk continued his talk by presenting several applications to randomness extraction. One application is the derivation of strong cryptographic keys from non-uniform sources of randomness, like physical noise or event schedulers, while other applications deal with the derivation of pseudorandom keys from a Diffie-Hellman value. He proposed several block-cypher methods for randomness extraction: CBC chaining, Merkle-Damgard cascading, and HMAC, and he presented a result showing that the output of this construction is random and uniform.

At the end of the talk, Dr. Krawczyk reiterated that these types of constructs should be used for randomness extraction since they are more practical than combinatorial extractors and they have proven analytical properties. He pointed out that the well-known hash functions are not designed for such purpose, especially given their vulnerabilities cited by the previous speakers.

2.8 What's the Worst That Could Happen?

Speaker: Eric Rescorla, RTFM, Inc.

Mr. Eric Rescorla discussed the effects of presumptive failures of cryptographic primitives on security protocols. The protocols that are generally evaluated are SSH, SSL/TLS, IPSEC, SMIME and certificates. He began his talk by analyzing the status of four major classes of cryptographic algorithms. He first discussed the key establishing algorithms (RSA and DH) and concluded they are quite sound and less vulnerable in practice. He drew the same conclusion for signature algorithms (RSA and DSS). He continued then with encryption algorithms. He pointed out that 3DES and AES are basically sound, but DES and RC4 are vulnerable due to inherent weaknesses and serious flaws. With respect to hashing functions, he pointed out that MD5 is the most vulnerable.

In the second part of his talk, Mr. Rescorla analyzed a number of possible attacks targeting the most vulnerable algorithms above, and discussed various implications to security protocols. He pointed out that S/MIME is vulnerable to collision attack, but the attack is not plausible: real life signatures and contracts are provable by intention and not just by signatures. With respect to certificates, Mr. Rescorla pointed out that they are protected because of the randomness introduced by the serial number and the validity fields. When assessing the effects of the second preimage attacks, Mr. Rescorla indicated that the certificates are seriously affected, but not the other protocols, due to real-time constraints. He then continued with the attack on the RC4 initial bytes, that can produce serious effects to HTTPS transactions. However, a solution to this problem can be devised shortly.

He finally considered a number of less plausible attacks like remote key recovery and total cypher break. The conclusions were that even though these attacks may became realistic in the near future, their effects will still be minimal on most security protocols.

2.9 Fuzzy Commitment

Speaker: Ari Juels, RSA Laboratories

Dr. Ari Juels started the second day of the workshop by presenting a new approach of using biometric data in cryptography, work called fuzzy-commitment. He began by presenting an example of how biometric data is used as a form of authentication. In traditional cases, Dr. Juels argued that the biometric data need not be kept secret since the authentication happens under the supervision of human officers. He continued by pointing out that there will be more opportunities for spoofing of the biometric data once the authentication process becomes more automatic. Since the revocation of biometric data is hard at best, the biometric has to be kept secret.

In the second part of his talk, Dr. Juels discussed cryptographic tools for password secrecy. He pointed out that traditional cryptographic methods do not apply directly since the biometrics are "approximatively" the same. He continued to present his approach, called “fuzzy-commitment”, that provides an error-tolerant set of cryptographic primitives. Dr. Juels finally discussed some work in progress at RSA Laboratories that successfully extracts 60-bit keys from eye irises. The conclusions of the talk were that biometrics will be used more and more for personal identification, and that fuzzy crypto techniques are a typical place where theory meets practice.

2.10 Secure Fuzzy Extractors

Speaker: Xavier Boyen, Voltage Inc.

Dr. Boyen continued in the same direction as Dr. Juels by exploring biometric authentication topics. His talk addressed the issue of using biometrics in such applications as authenticating to a server. In the first part of his talk he presented the notions of fuzzy sketches and fuzzy extractors that allow the generation of reproducible keys from noisy non-uniform biometrics. He pointed out that these methods alone are not able to provide secure communication with multiple servers or could reveal the secret to malicious servers.

In the second part of his talk, Dr. Boyen presented two extractors that addressed the above problems. He first presented the “reusable” fuzzy extractor, and described an example of zero-storage remote biometric authentication. The vulnerability of this scheme to outsider chosen perturbation (CP) attack is known. Dr. Boyen continued the talk with ’sealed’ fuzzy extractors. These extractors provide a form of tamper resistance which allows a handshake without fear that the biometric secret might be leaked to a cheating server. He brought up the example of an authenticated key exchange.

2.11 Using Biometrics for Secure Network-Based Authentication

Speaker: Jonathan Katz, University of Maryland

Dr. Katz continued the discussion of biometrics for the case of secure remote authentication. He motivated his work by the fact that humans are incapable of the storing and timely handling of secure long cryptographic secrets. He pointed out that biometrics represent a free storage.

Dr. Katz started his talk by presenting two models for remote authentication. He first presented a plug-in solution for the cases when data from the server may be tampered with. This solution is proven secure in the random oracle model. Secondly, he presented a specific solution for key exchange proven secure in the standard model.

He pointed out that these solutions tolerate more general errors, achieve mutual authentication, and offer improved bounds on entropy loss. The main observation is that there are solutions using biometrics in the standard model.

2.12 Cryptographic Mechanisms to Secure Routing Protocols

Speaker: Adrian Perrig, Carnegie Mellon University

Dr. Perrig started his talk by addressing the need for secure routing protocols. He pointed out that the routing protocols assumed a trusted environment, while the reality shows that even misconfigurations can severely disrupt communication. Dr. Perrig's talk focused on techniques to prevent malicious routing. He also pointed out that securing the protocols requires detection and recovery mechanisms as well as the use of techniques to reduce the impact of attacks. He addressed two protocol domains: ad-hoc networks and Internet infrastructure.

In the second part of the talk, he continued by addressing secure routing in ad-hoc networks. After a brief explanation of ad-hoc routing, he drew attention to the fact that ad-hoc networks are vulnerable to a multitude of attacks with severe results. He illustrated it by showing wormhole and rushing attacks. After briefing the audience on Distance Vector Routing, he presented the Secure Efficient Distance Vector Routing (SEAD) protocol. He explained that the sequence numbers and metrics are protected against forgery by chain hashing, which renders the claim of a lower sequence number by an attacker impossible. A number of other attacks that can be prevented by a slightly modified scheme were discussed. Dr. Perrig also offered a hash tree chain solution to this problem that proves to be cheaper than SEAD.

In the third part of his talk, Dr. Perrig addressed routing in the Internet. He started by presenting the essentials of BGP, and outlined three possible attacks. He pointed out some weaknesses of the S-BGP protocol. He continued by presenting the ASPATH Protector, a scheme that protects against an attacker modifying the encoded ASPATH. He also showed an example of its usages. At the end of the talk there was a discussion on the issue of securing Internet protocols. It was followed by a discussion on the effects of the routing disruption on Internet communication.

2.13 Cryptographic Hashing: Blockcipher-based Constructions Revisited

Speaker: Tom Shrimpton, Portland State University

Dr. Shrimpton proposed a discussion about cryptographic hashing. The grounds for his talk were the new findings affecting the hash functions: reported near collisions in SHA-0, collisions in SHA-0, MD5, RIPEMD, etc. Dr. Shrimpton started his talk with a discussion of the desirable properties of hashes: second, preimage, and collision resistance. He also proposed a further discussion about the near-collision property of hashes. The focus of his talk then shifted towards hash building methods. He argued for the use of blockciphers for building hashes, based on the shortcomings of DES. Dr. Shrimpton continued by presenting a number of provably secure compression functions. He then proposed to model the blockciphers as random permutations, pointing out that the PRP model is not adequate. Dr.

Shrimpton then discussed the practical aspects of hash constructions. He presented aspects of the cascaded constructions, while studying collision properties.

From the conclusions of the talk it followed that hash functions represent a big opportunity for research in several directions. One direction is the formalization and the study of the properties of hashes. It has been pointed out that the ideal cipher model requires proofs while PRP does not. Another area for future research is the analysis of MDC2.

2.14 Privacy-Preserving Bayesian Network Structure Computation on Distributed Heterogeneous Data

Speaker: Rebecca Wright, Stevens Institute of Technology

Dr. Wright continued the session by discussing the issue of privacy-preserving computation on distributed heterogeneous data. In her talk, Dr. Wright proposed an algorithm to compute a Bayesian Network structure out of data held in two databases maintained by two different organizations. She motivated her work by the fact that data-mining applications are continuously growing, while the data set is increasingly distributed, posing serious challenges to privacy issues. She cited several examples of such applications from the technical and academic domain, among them a genetic database linked to a patient health record database.

Dr. Wright continued her talk by presenting a privacy preserving protocol for learning the bayesian networks, based on the `K2` protocol. She focused on the description of the scoring function and gave special attention to the scalar product protocol. She finished the analysis of the protocol by noting that its complexity is linear in the number of records in the database and the number of attributes and exponential on the number of possible parents for every node.

At the end of her talk, Dr. Wright raised a number of open issues related to this protocol. In particular, she questioned the effects of the leak of intermediary data to the adversary parties (the relative order of the scoring function), as well as the privacy of data in the case of malicious attackers that can deviate arbitrarily from the protocol.

2.15 Error Correction in the Bounded Storage Model

Speaker: Yan Zong Ding, Georgia Institute of Technology

In his talk, Dr. Ding presented his work on the Maurers bounded storage model for those cases when the involved parties have inconsistent views of the public random source due to transmission or other types of errors. At the beginning he argued that all the previous protocols do not function properly in the presence of errors and the private-key encryption scheme of Aumann, Ding and Rabin only tolerates a limited number of errors. Dr. Ding pointed out that the new scheme tolerates a constant fraction of errors, and attains the near optimal parameters achieved by Vadhans construction in the errorless case.

Dr. Ding continued his talk by presenting the construction. First he showed that any local fuzzy extractor yields a secure and error-resilient cryptosystem in the model. He then showed the construction of efficient local fuzzy extractors by extending Vadhans sample-then-extract paradigm. He mentioned that the main ingredients to this are averaging samplers, randomness extractors, error correcting codes, and fuzzy extractors.

2.16 Smart Theory Meets Smartcard Practice

Speaker: Jean-Jacques Quisquater, Universite Catholique de Louvain, Belgium and CNRS, France

Dr. Quisquater started his talk by offering a short history of smart cards. He presented a hardware implementation model of the smart card and proposed a discussion about how tamperproof crypto theory matches the real world of working with constrained objects. Dr. Quisquater gave a number of examples of passive attacks and active fault attacks, then he continued with a presentation of a tamperproof model. While discussing the usefulness of the tamperproof model, he pointed out that it is useful in simulating public key crypto in closed systems, but that we don't know how to translate tamperproof into trapdoor in a crypto function.

In the second part of his talk, Dr. Quisquater discussed the issue of security with two chips: either an unsecure fast processor, or an unsecure memory. He concluded his talk by pointing out that this issue poses a number of open questions and that in general cryptography with strongly constrained objects sets a number of problems that can have practical results.

3 Conclusions and Future Research Challenges

This two day workshop brought together researchers from the computer science community both from academia and from industry. The discussions following each talk and the final discussion that concluded the workshop addressed the existing gap between the cryptographic primitives and their actual implementation and usage by the software community. Several ideas were prevalent regarding a number of open problems and new research directions in applied cryptography. Some of these open problems are summarized below:

- How do we use cryptographically imperfect biometric data for generating reproducible keys used in authentication systems?

- Given its non-revocation properties, how do we use biometric data in authentication without fear of leakage?

- How do we efficiently and correctly extract cryptographically strong random keys based on nonuniform sources of randomness?

- What are the effects and what are the countermeasures to be taken if a number of basic cryptographic primitives become vulnerable or fail to deliver?

- How do we build and incorporate new hashing functions into protocols? We need an in-depth study of hash function properties.
- How do we share data without exposing private aspects of it? Aspects of this problem include privacy-preserving data mining on distributed sets of data, set intersection and mixers.
- Why are the existing security primitives not implemented in existing protocols and what are the effects of not implementing them?
- What changes are required for cryptographic primitives in order to make them more suitable for larger classes of applications?

4 Acknowledgments

The author of this report wants to thank Dr. Brenda Latka, Associate Director of DIMACS, for her valuable comments. The author is grateful to all the speakers for sharing their slides used for presentation at the workshop. The author also wants to apologize to those authors whose work was not fully or accurately described in this report, due to limitations of the author's knowledge and availability of materials.

The author and the DIMACS Center acknowledge the support of the National Science Foundation under grant number CCR 03-14161 to Rutgers University.