THE JOY OF CRYPTOGRAPHY

Mike Rosulek (rosulekm@eecs.oregonstate.edu)
School of Electrical Engineering & Computer Science
Oregon State University, Corvallis, Oregon, USA

Draft of January 4, 2018
Foreword

These are lecture notes for cs427 at Oregon State University, an introductory course in cryptography at the advanced undergraduate level. By reading and studying these notes, you should expect to learn how to:

▶ State and interpret the standard formal definitions for the most common cryptographic security properties (privacy and authentication).

▶ Formally prove security properties of sound cryptographic constructions, and break the security of unsound ones.

▶ Choose the appropriate cryptographic primitive for a task (block ciphers, hash functions, MACs, public-key encryption, etc.) while avoiding common pitfalls.

Along the way, you will also learn how the most common cryptographic constructions work.

Background

You will get the most out of these notes if you have a solid foundation in standard undergraduate computer science material:

▶ Discrete mathematics [required]: for modular arithmetic, discrete probabilities, simple combinatorics, and especially proof techniques.

▶ Algorithms & data structures [highly recommended]: for reasoning about computations at an abstract level.

▶ Theory of computation (automata, formal languages & computability) [recommended]: for even more experience formally reasoning about abstract processes of computation.

Disclaimer

You are reading an early draft of this book. Of course I make every effort to ensure the accuracy of the content, but the content has not yet benefitted from many critical readers. So, caveat emptor!

More than that, much interesting and hugely important material is not covered. Some more minor missing material is indicated as such and labeled “to-do”, but conspicuously missing major sections are simply absent. Anyone with a pre-existing opinion about cryptography will likely be appalled at what is missing. In my defense, my university schedule

© Copyright 2018 Mike Rosulek. Licensed under Creative Commons BY-NC-SA 4.0. Latest version at eecs.orst.edu/~rosulekm/crypto.
is on the quarter system which necessitates that I generate 10 weeks of course material. Everything beyond that is produced via the much slower labor-of-love mechanism.

I welcome feedback from readers, educators, cryptographers — not only on errors and typos but also on the selection, organization, and presentation of the material.

**Code-based games**

The security definitions and proofs in these notes are presented in a style that is known to the outside world as *code-based games*. I’ve chosen this style because I think it offers significant pedagogical benefits:

- Every security definition can be expressed in the same style, as the indistinguishability of two games. In my terminology, the games are *libraries* with a common interface/API but different internal implementations. An adversary is any calling program on that interface. These libraries use a concrete pseudocode that reduces ambiguity about an adversary’s capabilities. For instance, the adversary controls arguments to subroutines that it calls and sees only the return value. The adversary cannot see any variables that are privately scoped to the library.

- A consistent framework for definitions leads to a consistent process for proving and breaking security — the two fundamental activities in cryptography.

  In these notes, *breaking* a construction always corresponds to writing a program that expects a particular interface and behaves as differently as possible in the presence of two particular implementations of the interface.

  *Proving security* nearly always refers to showing a sequence of libraries (called *hybrids*), each of which is indistinguishable from the previous one. Each of these hybrids is written in concrete pseudocode. By identifying what security property we wish to prove, we identify what the endpoints of this sequence must be. The steps that connect adjacent hybrids are stated in terms of syntactic rewriting rules for pseudocode, including down-to-earth steps like factoring out and inlining subroutines, changing the value of unused variables, and so on.

- Cryptography is full of conditional statements of security: “if A is a secure thingamajig, then B is a secure doohickey.” A conventional proof of such a statement would address the contrapositive: “given an adversary that attacks the doohickey-security of B, I can construct an attack on the thingamajig-security of A.”

  In my experience, students struggle to find the right way to transform an abstract, hypothetical B-attacking adversary into a successful A-attacking adversary. By defining security in terms of games/libraries, we can avoid this abstract challenge, and indeed avoid the context switch into the contrapositive altogether. In these notes, the thingamajig-security of A gives the student a new *rewriting rule* that can be placed in his/her toolbox and used to bridge hybrids when proving the doohickey-security of B.
Code-based games were first proposed by Shoup\textsuperscript{1} and later expanded by Bellare & Rogaway.\textsuperscript{2} These notes adopt a simplified and unified style of games, since the goal is not to encompass every possible security definition but only the fundamental ones. The most significant difference in style is that the games in these notes have no explicit \textsc{Initialize} or \textsc{Finalize} step. As a result, all security definitions are expressed as indistinguishability of two games/libraries, even security definitions that are fundamentally about unforgeability. Yet, we can still reason about unforgeability properties within this framework. For instance, to say that no adversary can forge a MAC, it suffices to say that no adversary can distinguish a MAC-verification subroutine from a subroutine that always returns \textsc{false}.

An index of security definitions has been provided at the end of the book.

One instance where the approach falls short, however, is in defining collision resistance. I have not been able to define it in this framework in a way that is both easy to use and easy to interpret (and perhaps I achieved neither in the end). See Chapter 12 for my best attempt.

**Supplementary material**

Security proofs in this book follow a standard pattern: We start from one “library” and perform a sequence of small, cumulative modifications. Each modification results in a separate hybrid library that is indistinguishable from the previous one.

I have prepared PDF slide decks to supplement the security proofs contained in the book. They are available from the course website. The slides allow the reader to step forward and backward through the proof’s sequence of logical steps, seeing only the current hybrid library at any time (with changes highlighted and annotated).

**Acknowledgements**

Some financial support for writing this book has been kindly provided by the National Science Foundation (award #1149647) and the Oregon State University Open Textbook Initiative.

**About the cover**

The cover design consists of assorted shell illustrations from \textit{Bibliothèque conchyliologique}, published in 1846. The images are no longer under copyright, and were obtained from the Biodiversity Heritage Library (http://biodiversitylibrary.org/bibliography/11590). Like a properly deployed cryptographic primitive, a properly deployed shell is the most robust line of defense for a mollusk. To an uniformed observer, a shell is just a shell, and crypto is just crypto. However, there are a wide variety of cryptographic primitives, each of which provides protection against a different kind of attack. Just as for a seasoned \textit{conchologist}, the joy is in appreciating the unique beauty of each form and understanding the subtle differences among them.

\begin{itemize}
\end{itemize}
Copyright

This work is copyright by Mike Rosulek and made available under the Creative Commons BY-NC-SA 4.0 license. Under this license, you are free to:

**Share:** copy and redistribute the material in any medium or format.

**Adapt:** remix, transform, and build upon the material.

The licensor cannot revoke these freedoms as long as you follow the following license terms:

**Attribution:** You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**NonCommercial:** You may not use the material for commercial purposes.

**ShareAlike:** If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.
Contents

0  Review of Concepts & Notation ......................................... 1
   0.1  Modular Arithmetic ................................................. 1
   0.2  Strings ............................................................. 2
   0.3  Functions .......................................................... 2
   0.4  Probability .......................................................... 2
   0.5  Asymptotics (Big-O) ................................................. 4

1  One-Time Pad .................................................................. 5
   1.1  One-Time Pad Definition ............................................ 5
   1.2  Properties of One-Time Pad .......................................... 6

2  The Basics of Provable Security ..................................... 12
   2.1  Library-Based Security Reasoning ................................. 12
   2.2  A General-Purpose Security Definition for Encryption ... 17
   2.3  How to Prove Security with the Hybrid Technique ......... 21
   2.4  How to Demonstrate Insecurity with Attacks ................ 24

3  Secret Sharing ................................................................ 30
   3.1  Definitions ............................................................. 30
   3.2  A Simple 2-out-of-2 Scheme ....................................... 32
   3.3  Polynomial Interpolation ............................................ 35
   3.4  Shamir Secret Sharing ............................................... 36
   3.5  Visual Secret Sharing ............................................... 39

4  Basing Cryptography on Limits of Computation ............... 44
   4.1  Polynomial-Time Computation ..................................... 45
   4.2  Negligible Probabilities .............................................. 46
   4.3  Indistinguishability ................................................... 48
   4.4  Sampling with Replacement & the Birthday Bound ......... 49

5  Pseudorandom Generators ............................................ 54
   5.1  Definition ............................................................. 54
   5.2  Application: Shorter Keys in One-Time-Secret Encryption 57
   5.3  Taking the Contrapositive Point-of-View ......................... 60
   5.4  Extending the Stretch of a PRG ................................... 61

6  Pseudorandom Functions ............................................. 67
   6.1  Definition ............................................................. 67
   6.2  Attacking Insecure PRFs ............................................. 69
   6.3  A Theoretical Construction of a PRF from a PRG ........... 70

7  Pseudorandom Permutations ......................................... 77
   7.1  Definitions ............................................................ 77
   7.2  Switching Lemma .................................................... 78
   7.3  Feistel Ciphers ....................................................... 79
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.3 ElGamal Encryption</td>
<td>185</td>
</tr>
<tr>
<td>15.4 Hybrid Encryption</td>
<td>188</td>
</tr>
<tr>
<td>Index of Security Definitions</td>
<td>192</td>
</tr>
</tbody>
</table>