Reading Guide 3:
Authenticated Encryption

Paul D'Avilar, Jeremy D'Errico, Ken Berends, Michael Peck

September 21, 2004

1 Summary

The paper, *Authenticated Encryption: Relations among notions and analysis of the generic composition paradigm*, by Mihir Bellare and Chanathip Namprempre analyzes the security of authenticated encryption schemes designed by generic composition. The term generic composition refers to making black-box use of a given symmetric encryption scheme and a given MAC. This paper combines the materials presented in week 1, Reductionist Cryptography and Pseudorandom Functions, and week 2, Message Authentication. Please refer to those lectures for further details on individual topics.

2 Goal

We want to consider two notions of authenticity of symmetric encryption, integrity of plaintexts and integrity of ciphertexts and relate them to the standard notions of privacy for symmetric encryption. Basically we want to answer what composition scheme provides the best performance and security: Encrypt-and-MAC, MAC-then-Encryption, or Encrypt-then-MAC.

3 Details

CBC Encryption (Cipher Block Chaining) is a mode of encryption where each plaintext block is XORed with the ciphertext of the preceding block. This chaining or dependency method creates a resultant ciphertext that eliminates the repetitive patterns that plague ECB encryption. (This method is very similar to the Iterated Hash Function demonstrated in last weeks *Keying Hash Functions for Message Authentication* presentation).

Section 1 defines several terms related to authenticated encryption. In an authenticated encryption scheme, encryption takes a key and a plaintext and returns a ciphertext. Decryption takes the same key and a ciphertext and returns either the plaintext or some indication that the ciphertext is invalid or unauthentic.

Section 1.1 mentions two privacy goals for symmetric encryption schemes: indistinguishability and non-malleability. These two goals need to be considered under chosen-plaintext and chosen-ciphertext attacks, so four notions of security are considered: IND-CPA, IND-CCA, NM-CPA, and NM-CCA. Two notions of integrity (this term is used interchangeably with authenticity) are defined: INT-PTXT and IND-CTXT. The section makes the difference between authenticity and privacy quite clear: for example, sending an unencrypted message along with a strong MAC achieves INT-CTXT but no privacy. The section considers each of the two notions of authenticity (INT-PTXT and INT-CTXT) coupled with IND-CPA (the weakest notion of privacy).

Section 1.2 $| M_0 |=| M_1 |$ means: M_0 and M_1 have to be the same length. The encryption function should not leak information about the plaintext. When you encrypt a message, the length of the ciphertext has to be at least as long as the length of the plaintext. Therefore, if an adversary has to two messages, M_0 and M_1, of different lengths, encrypting the messages could give the adversary an advantage in determining what message, M_0 or M_1, was encrypted. Section 1.2 presents one method to design an authenticated encryption scheme: generic composition. Generic composition means to combine a standard symmetric encryption scheme with a MAC (it's generic because the encryption scheme and the MAC are each seen as black boxes). The paper looks at 3 possible ways to do this and compares their security. The three possible ways are Encrypt-and-MAC, MAC-then-encrypt, and Encrypt-then-MAC. In Encrypt-and-MAC, the plaintext is encrypted and then concatenated with a MAC of the plaintext. In MAC-then-encrypt, a MAC is concatenated to the plaintext and then this entire string is encrypted.

In Encrypt-then-MAC the plaintext is encrypted and a MAC of the produced ciphertext is concatenated. Section 1.2 also summarizes the results of the paper (which are proved in Section 4). In summary, if the given MAC is strongly unforgeable, Encrypt-then-MAC is secure for the privacy goals IND-CPA, IND-CCA and NM-CPA as well as the integrity goals of INT-PTXT and INT-CTXT.

Section 2.1 The security parameter k is the length of the key used in the key generation algorithm $\mathcal{K}$. Section 2.1 defines the syntax of symmetric encryption schemes and should look familiar from previous lectures. A symmetric encryption scheme consists of three algorithms. First, a key generation algorithm returns a random key. Second, an encryption algorithm takes a key and a plaintext and produces a ciphertext. Third, a decryption algorithm takes a key and a ciphertext and returns either a plaintext or some indication that the plaintext is invalid. It is important to note that the encryption algorithm may return a different ciphertext each time, while the decryption algorithm must be deterministic (always giving the same output for a given input). The $\perp$ symbol is pronounced "bottom".

Section 2.2 The "Left-or-Right model" is a definition for security of symmetric encryption under which the notion of indistinguishability is test/analyze against CPA and CCA. Basically under the left-or-right model, an adversary is allow queries of the form (M_0, M_1), where M_0, M_1 are equal length messages. Two games are played/ considered: In the first game, each query is answered by encrypting the left message (M_0), and in the second game, each query is answered by encrypting the right message (M_1). The scheme is similar to Adam's lecture where the adversary has to determine what world he is in or whether M_0 or M_1 was encrypted. As was mentioned from Adam's lecture, we consider the encryption scheme to be "good" if a "reasonable" adversary cannot obtain significant advantage in distinguishing the cases b = 0 and b = 1 given access to the left-or-right oracle.

Section 2.4 defines message authentication schemes. Message authentication schemes consist of three algorithms. A key generation algorithm returns a random key K. A tagging algorithm takes a key and a message and returns a tag (the tag may be different each time the algorithm is invoked). The verification algorithm takes a key, a message, and a candidate tag and determines whether or not the tag is valid. This section also defines weak forgery of MACs as well as strong forgery of MACs, as defined below.

Section 2.5 provides notation for a reduction, where adversary A' executes another adversary A within itself and responds to A's oracle queries.

Section 3 formalizes the results summarized in Figure 1 of Section 1.2. Theorem 3.1 states that if an encryption scheme is INT-CTXT secure, it must be INT-PTXT secure as well. Theorem 3.2 states that if an encryption scheme is INT-CTXT and IND-CPA secure, it must be IND-CCA secure. Proposition 3.3 states that IND-CCA does not necessarily imply INT-PTXT: if given a symmetric encryption scheme which is IND-CCA secure, one can construct another symmetric encryption scheme which is also IND-CCA secure but is not INT-PTXT secure. The constructed symmetric encryption scheme's encryption function puts a 0 in front of the ciphertext produced by the original encryption function. The constructed decryption scheme returns the result of the original decryption scheme if the ciphertext begins with a 0. Otherwise, it decrypts the ciphertext to 0. Proposition 3.4 states that INT-PTXT and IND-CPA does not necessarily imply NM-CPA.

Theorem 3.1 is very similar to the PRP-CPA and PRP-CCA theorems (Section 3.4.3 from week 1's *Pseudorandom Functions* paper). This relation is intuitive since the it can't be any harder to break the scheme when given an additional oracle (the one for the ciphertexts).

Section 4 presents formal security results through the use of proofs for Encrypt-and-MAC, MAC-then-Encrypt, and Encrypt-then-MAC. Encrypt-then-MAC is shown to be theoretically secure against the considered forms of attack when a strongly unforgeable MAC is used.

4 Definitions

Authentication Encryption Refers to a shared-key based transform whose goal is to provide both privacy and authenticity.

CPA - Chosen Plaintext Attack An attacker can obtain the ciphertext for any provided plaintext (but does not have the key).

CCA - Chosen Ciphertext Attack An attacker can also, in addition to being able to perform a CPA, obtain the plaintext for any provided ciphertext (but does not have the key).

IND - Indistinguishability The advantage of a reasonable adversary determining what message was sent, M_0 or M_1, - Stubblefield, Lecture 2-3.

IND-CPA Indistinguishability under a Chosen Plaintext Attack

IND-CCA Indistinguishability under a Chosen Ciphertext Attack

NM - Non-malleability Is the advantage of a reasonable adversary being able to change the message to be meaningful - Group 1, Lecture 4.

NM-CPA Non-malleability under a Chosen Plaintext Attack

NM-CCA Non-malleability under a Chosen Ciphertext Attack

INT-PTXT - Integrity of Plaintext To be computationally infeasible to produce a ciphertext decrypting to a message that the sender had never encrypted.

INT-CTXT - Integrity of Ciphertext To be computationally infeasible to produce a ciphertext not previously produced by the sender.

WUF - Weakly unforgeable MAC An attacker (with the capability of making chosen message attacks) is incapable of creating a new accepted message and tag pair (the message cannot have previously been queried).

SUF - Strongly unforgeable MAC An attacker (with the capability of making chosen message attacks) is incapable of creating a new accepted message and tag pair (the tag cannot have previously been returned in response to a query - but the message may have been queried).