AE5 Security Notions
Definitions Implicit in the CAESAR Call

Chanathip Namprempre[1] and Phillip Rogaway[2] and Tom Shrimpton[3]

[1] Dept. of Electrical and Computer Engineering, Thammasat University, Thailand
[2] Dept. of Computer Science, University of California, Davis, USA
[3] Dept. of Computer Science, Portland State University, USA

April 29, 2013

Abstract. A draft call for the CAESAR authenticated-encryption competition adopts an interface that is not aligned with existing definitions in the literature. It is the purpose of this brief note to formalize what we believe to be the intended definitions.

1 Introduction

A call for authenticated-encryption (AE) mechanisms, CAESAR, was recently put forward by Dan Bernstein [3]. One surprising feature of the call is the unusual AE interface it describes. The user who wants to encrypt, instead of providing the customary *four* arguments (the key, nonce, associated data, and message), will now provide *five*, the nonce having morphed into a *secret message number* (SMN) and a *public message number* (PMN). The SMN is a *novum* for AE—a secret value, not the plaintext, recoverable from the ciphertext, and for which a single-use requirement may be imposed. The PMN is cousin to the nonce, but the single-use requirement is regarded as optional.

Syntactic adjustments of an AE scheme are never a trivial matter; for starters, when the syntax changes, all security notions must *ipso facto* change. Maybe this doesn't matter if it is obvious how to adjust the security notions to match the changed syntax. But we believe that it is not, that different people would follow different paths, ending up with inequivalent definitions. Thus the purpose of this note is to formalize what *we* take to be the basic definitions for CAESAR-Call-flavor-AE.

The last phrase clearly needs a better name. When we want to be explicit, we'll call it AE5. The number 5 emphasizes that encryption will now take five argument.[4] We claim no particular novelty in formalizing AE5 security; our definitions build on the similar formulations of probabilistic [1, 2, 4] and then nonce-based AE [8, 9], AE with associated data [7], and deterministic and misuse-resistant AE [10].

2 Definitions

NOTATION. If A is a finite set we write $a \xleftarrow{} A$ for the process of sampling uniformly from A and assigning the result to a; if A is distribution, we sample according to it. When a function $F(x_1, \ldots, x_n)$ has multiple arguments, we sometimes write them as subscripts, then superscripts, then parenthesized arguments, in that order. For example, we write $\mathcal{E}_K^{T,A}(S, M)$ in place of $\mathcal{E}(K, T, A, S, M)$,

[4] Additionally, a "5" looks like an "S", and the SMN, whose value we will likewise call S, is the main thing that's new with AE5.

providing a more compact notation.[5] When $\mathcal{A}$ is an adversary that can interact in a game G, we write $\Pr[\mathcal{A}^G \Rightarrow 1]$ for the probability that $\mathcal{A}$ outputs the bit 1 after interacting with the specified game. Strings are finite and over the binary alphabet. The empty string is written as ε.

NEW AE SYNTAX. An AE5 scheme is a function $\mathcal{E}\colon \mathcal{K} \times \mathcal{T} \times \mathcal{A} \times \mathcal{S} \times \mathcal{M} \to \{0,1\}^*$ for sets of strings $\mathcal{K}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{S}$, $\mathcal{M}$, the space of *keys*, *public message numbers*, *associated data*, *secret message numbers*, and *messages*. If $A \in \mathcal{A}$ then all strings of length $|A|$ must likewise be in $\mathcal{A}$. The same holds for $\mathcal{M}$. We require that $\mathcal{K} = \{0,1\}^k$ and $\mathcal{S} = \{0,1\}^s$ and $\mathcal{T} = \{0,1\}^t$, were $k \geq 1$ and $s, t \geq 0$. These conditions can be relaxed as needed.[6] In writing $\{0,1\}^s$ and $\{0,1\}^t$ we follow the customary convention that $\{0,1\}^0 = \{\varepsilon\}$. We insist that $\mathcal{E}$ be injective, by which we mean that $\mathcal{E}_K^{T,A}(S,M) = \mathcal{E}_K^{T,A}(S',M')$ implies that $(S,M) = (S',M')$. For expositional simplicity our basic security definitions assume an *expansionless* scheme: $|\mathcal{E}_K^{T,A}(S,M)| = |M| + |S|$.

For any AE5 scheme, we expect trivial-to-compute algorithms to decide if $A \in \mathcal{A}$ and to decide if $M \in \mathcal{M}$ (equivalently, to decide if their lengths are valid). We leave these algorithms anonymous.

We abbreviate associated data (AD), secret message number (SMN), and public message number (PMN). One may speak of AE5 schemes that have no AD or SMNs or PMNs, but, formally, this means that $\mathcal{A} = \{\varepsilon\}$ or $\mathcal{S} = \{\varepsilon\}$ or $\mathcal{T} = \{\varepsilon\}$ are singleton sets.

DEFINING DECRYPTION FROM ENCRYPTION. Given an AE5 scheme $\mathcal{E}$, the injectivity requirement makes decryption well-defined: it is inherited from $\mathcal{E}$. In particular, for any AE scheme $\mathcal{E}\colon \mathcal{K} \times \mathcal{T} \times \mathcal{A} \times \mathcal{S} \times \mathcal{M} \to \{0,1\}^*$ there is a corresponding function $\mathcal{D}\colon \mathcal{K} \times \mathcal{T} \times \mathcal{A} \times \{0,1\}^* \to (\mathcal{S} \times \mathcal{M}) \cup \{\perp\}$ defined by saying that $\mathcal{D}_K^{T,A}(C) = (S,M)$ if $\mathcal{E}_K^{T,A}(S,M) = C$ for some $S \in \mathcal{S}$ and $M \in \mathcal{M}$, and $\mathcal{D}_K^{T,A}(C) = \perp$ otherwise.

AE5 SECURITY. To capture the security provided by an AE5 scheme $\mathcal{E}\colon \mathcal{K} \times \mathcal{T} \times \mathcal{A} \times \mathcal{S} \times \mathcal{M} \to \{0,1\}^*$ we define two games. Both begin by choosing a random $K \twoheadleftarrow \mathcal{K}$. Then:

- Game $\text{Real}_{\mathcal{E}}$:
  - A query $(T,A,S,M) \in \mathcal{T} \times \mathcal{A} \times \mathcal{S} \times \mathcal{M}$ is answered by $\mathcal{E}_K^{T,A}(S,M)$.
  - A query $(T,A,C) \in \mathcal{T} \times \mathcal{A} \times \{0,1\}^*$ is answered by $\mathcal{D}_K^{T,A}(C)$.
- Game $\text{Rand}_{\mathcal{E}}$:
  - A query $(T,A,S,M) \in \mathcal{T} \times \mathcal{A} \times \mathcal{S} \times \mathcal{M}$ is answered with $|S| + |M|$ uniformly random bits.[7]
  - A query $(T,A,C) \in \mathcal{T} \times \mathcal{A} \times \{0,1\}^*$ is answered by $\perp$.

Any query not in one of the specified domains is answered by $\perp$.

As a mnemonic, we write Enc before four-argument queries, as though directed to an encryption oracle, and we write Dec before three-argument queries, as though directed to a decryption oracle. In this way, adversarial queries would be written as $\mathsf{Enc}(T,A,S,M)$ or $\mathsf{Dec}(T,A,C)$.

Given an adversary $\mathcal{A}$ and an encryption scheme $\mathcal{E}$, we define the advantage that $\mathcal{A}$ gets in attacking $\mathcal{E}$, the real number $\mathbf{Adv}^{\text{ae}}_{\mathcal{E}}(\mathcal{A})$, as $\Pr[\mathcal{A}^{\text{Real}_{\mathcal{E}}} \Rightarrow 1] - \Pr[\mathcal{A}^{\text{Rand}_{\mathcal{E}}} \Rightarrow 1]$.

Some adversaries can trivially win the game we have defined. A *valid* adversary $\mathcal{A}$ obeys the following rules:

[5] Foreshadowing a bit, the superscripts to $\mathcal{E}$ get integrity protection, while the arguments to $\mathcal{E}$—things inside the parentheses—get privacy and integrity protection.

[6] For reference experiments defining ideal security when the message space is infinite, it is necessary to be more general with the key space $\mathcal{K}$, permitting it be an infinite set endowed with a distribution.

[7] This is where we are using the assumption that $\mathcal{E}$ is expansionless. If it is not, one should compute $C \leftarrow \mathcal{E}_K^{T,A}(S,M)$ and return $c = |C|$ uniformly random bits.

– It makes no query $\mathsf{Dec}(T, A, C)$ if it earlier asked a query $\mathsf{Enc}(T, A, S, M)$ that returned C.
– It never repeats an encryption query, twice asking $\mathsf{Enc}(T, A, S, M)$.

A *nonce-respecting* adversary $\mathcal{A}$ is one that, in addition,

– asks no query $\mathsf{Enc}(T, A, S, M)$ following an earlier query of $\mathsf{Enc}(T, A', S, M')$.

In words, the pair (T, S) is a *nonce*—a value used at most once.

Basic goals. We describe two "basic" security goals for an AE5 scheme.

– The first security goal is to have $\mathbf{Adv}^{\mathrm{ae}}_{\mathcal{E}}(\mathcal{A})$ small for any *nonce-respecting* and computationally reasonable adversary. In essence, one is demanding of the scheme's user that she select (PMN, SMN) pairs that comprise a nonce. We call this *conventional* or *nonce-base* AE.
– A different and stronger security goal is to expect that $\mathbf{Adv}^{\mathrm{ae}}_{\mathcal{E}}(\mathcal{A})$ be small for *any* valid and computationally reasonable adversary. Thus (PMN, SMN) pairs may well repeat. When they do, repetitions of (associated data, plaintext) pairs will be manifest in the ciphertexts, but no other information about plaintexts, beyond their length, will be leaked. Following Rogaway and Shrimpton, we call this *misuse-resistant* AE [10].

For the descriptions above, we do not define "small" or "computationally reasonable." We favor linguistic and technical choices that make conjectures and reductions explicit and constructive in describing relationships between adversarial resources and advantage measures.

3 Discussion

1. The fact that we have defined and named two AE5 security goals, nonce-based security and misuse-resistant security, does not mean that these are the *only* two security goals of interest. To the contrary; we believe that there is an interesting landscape of goals *between* these two extremes, and an interesting set of *orthogonal* aims as well.
2. McGrew's RFC 5116 [5] describes an AE interface that is different from AE5 [3]. Encryption under RFC 5116 takes in a four-tuple, (key, nonce, AD, message), and the process may be randomized or stateful. Length-0 nonces are allowed and, when used, eclipse the single-use requirement for the nonce. While nothing like the SMNs is present in RFC 5116, that work discusses "partially implicit" nonces, which support overlapping cryptographic aims. Also see the discussion on unpredictable and partially implicit IVs in draft-mcgrew-iv-gen-02 [6].
3. Our definitions employ indistinguishability-from-random-bits privacy [9] and an all-in-one approach for defining adversarial advantage [10]. Certainly neither choice is essential.
4. There is support in CAESAR call [3] for the view that a user may *independently* impose single-use (nonce) restrictions on the SMN and PMN. But this view seems unsupported at other points in the document [3], and the intended semantics of having exactly one of the SMN and PMN being a nonce would seem arcane. For this reason, we have not defined these aims.
5. The notation $\mathcal{E}^{T,A}_K(S, M)$, as opposed to $\mathcal{E}^{S,T,A}_K(M)$, emphasizes that the SMN S is encrypted—it is recoverable from the ciphertext—rather than casting (S, T) as a surrogate nonce N.
6. For misuse-resistant AE5, there is, definitionally, no security-relevant distinction between the SMN and the plaintext, nor between the PMN and the IV.
7. This note has taken a non-judgmental stance on whether the AE5 syntax is something *desirable* for the CAESAR call. In fact, while we find the SMN/PMN innovation interesting, we are skeptical that its value merits its cost in familiarity, minimalism, or conceptual complexity.

Acknowledgments

The authors gratefully acknowledge the support of the NSF under grants CNS 0845610, CNS 0904380, and CNS 1228828.

References

1. M. Bellare and C. Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In T. Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 531–545. Springer, Dec. 2000.
2. M. Bellare and P. Rogaway. Encode-then-encipher encryption: How to exploit nonces or redundancy in plaintexts for efficient cryptography. In T. Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 317–330. Springer, Dec. 2000.
3. D. Bernstein. Cryptographic competitions: CAESAR call for submissions, draft 1. April, 2013. Available at http://competitions.cr.yp.to/caesar-call-1.html. Draft 2 at http://competitions.cr.yp.to/caesar-call-2.html.
4. J. Katz and M. Yung. Unforgeable encryption and chosen ciphertext secure modes of operation. In B. Schneier, editor, *FSE 2000*, volume 1978 of *LNCS*, pages 284–299. Springer, Apr. 2000.
5. D. McGrew. Interface and algorithms for authenticated encryption. IETF RFC 5116, Jan 2008.
6. D. McGrew. Generation of deterministic initialization vectors (IVs) and nonces. IETF Internet-Draft, draft-mcgrew-iv-gen-02.txt, Aug 2012.
7. P. Rogaway. Authenticated-encryption with associated-data. In V. Atluri, editor, *ACM CCS 02*, pages 98–107. ACM Press, Nov. 2002.
8. P. Rogaway. Nonce-based symmetric encryption. In B. K. Roy and W. Meier, editors, *FSE 2004*, volume 3017 of *LNCS*, pages 348–359. Springer, Feb. 2004.
9. P. Rogaway, M. Bellare, J. Black, and T. Krovetz. OCB: A block-cipher mode of operation for efficient authenticated encryption. In *ACM CCS 01*, pages 196–205. ACM Press, Nov. 2001.
10. P. Rogaway and T. Shrimpton. A provable-security treatment of the key-wrap problem. In S. Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 373–390. Springer, May / June 2006.