

BACHELOR THESIS



RADBOD HONOURS ACADEMY

TBD

Author:
Stijn Vandenput

Supervisor:
Martijn Stam
Bart Mennink

20/05/2022

Abstract

Contents

1	Introduction	1
2	Preliminaries	1
2.1	notation table	2
3	Existing AE/DEM notations in more detail	2
3.1	Existing notation from Hybrid Encryption in a Multi-user Setting, revised	2
3.1.1	notation	2
3.1.2	used primitives	2
3.1.3	goal	3
3.1.4	Security model	3
3.2	Existing notation from Generic Composition Reconsidered	4
3.2.1	notation	4
3.2.2	used primitives	4
3.2.3	goal	4
3.2.4	Security model	4
4	New Definition	5
4.1	notation	5
4.2	goal	5
4.3	Security model	5
4.4	choices made	6
5	Constructions	6
5.1	used primitives	6
5.2	construction	7
6	Use cases	9
7	Related Work	9
8	Conclusion	9
9	Appendix	10

1 Introduction

should consist of:

- explaining the challenge
- my contribution

Within the field of cryptography there are two different fields, symmetric and asymmetric crypto. The former is about situations where the users have a shared secret already, the latter is about situations where this is not the case. Sometimes work in asymmetric crypto uses constructions that are more common in symmetric crypto, and in this fashion, Hybrid Encryption in a Multi-user Setting, revised uses a construction that is very similar to Authenticated Encryption following the generic encrypt-then-MAC construction. This construction has since been revised in Generic Composition Reconsidered to be better applicable to common use cases. The aim of this thesis is to apply the knowledge of Generic Composition Reconsidered to the setting of Hybrid Encryption in a Multi-user Setting, revised and in doing so, create a better primitive for authenticated encryption in an asymmetric crypto setting.

2 Preliminaries

should consist of:

- general notation
- generic AE construction
- generic PKE schemes
- nonces vs locks

2.1 notation table

Name	pkc	GCrec	my notation	rough meaning
message space	\mathcal{M}	\mathcal{M}	\mathcal{M}	set of all possible message options
message	m	M	m	message the user sends
ciphertext space	\mathcal{C}	-	\mathcal{C}	set of all possible ciphertext options
ciphertext	c	C	\mathcal{C}	encrypted message
associated data space	-	\mathcal{A}	\mathcal{A}	set of all possible associated data options
associated data	-	A	a	data you want to authenticate but not encrypt
tag space	\mathcal{T}	\mathcal{T}	\mathcal{T}	set off all possible tag options
tag	t	T	t	output of mac function
key space	\mathcal{K}	\mathcal{K}	\mathcal{K}	set of all possible key options
key	k	K	k	user key
nonce space	-	\mathcal{N}	\mathcal{N}	set of all nonce options
nonce	-	n	n	number only used once
lock space	\mathcal{L}	-	\mathcal{L}	set of all possible lock options
lock	l	-	l	nonce that is bound to the user
users	N	-	N	number of users
adversary	A	\mathcal{A}	A	the bad guy
Random sampling	$\leftarrow^{\$}$	\leftarrow	\leftarrow	get a random element from the set
result of function	$\leftarrow^{\$}$	\leftarrow	\leftarrow	get the result of a function with given inputs
concatination	$a b$	$a b$	$a b$	concatanation of two strings

3 Existing AE/DEM notations in more detail

should consist of:

- the definition of the two paper, if possible already brought more toward one notation standard
- explain how both relate to the preliminaries and overall story

3.1 Existing notation from Hybrid Encryption in a Multi-user Setting, revised

3.1.1 notation

\mathcal{M} is a message space, \mathcal{K} is a finite key space, \mathcal{T} is a tag space and \mathcal{C} is a ciphertext space. N is the number of users.

3.1.2 used primitives

- ADEM: the ADEM exists of tuple $(A.\text{enc}, A.\text{dec})$, $A.\text{enc}$ take a key K in \mathcal{K} , a tag t in \mathcal{T} and a message m in \mathcal{M} and outputs a ciphertext c in \mathcal{C} . $A.\text{dec}$ takes a K in \mathcal{K} , a tag t

in \mathcal{T} and a ciphertext c in \mathcal{C} and outputs a message m in \mathcal{M} or \perp to indicate rejection. The correctness requirement is that for all K in \mathcal{K} and t in \mathcal{T} and m in \mathcal{M} we have $A.dec(K, t, A.enc(K, t, m)) = m$. The security of the ADEM is defined with $\mathbf{Adv}_{ADEM, A, N}^{n-moit-ind} = |\Pr[N-MIOT-IND_{A, N}^0 = 1] - \Pr[N-MIOT-IND_{A, N}^1 = 1]|$, defined by the following game:

Game N-MIOT-IND _{A, N} ^b	Oracle Oenc(j, t, m_0, m_1)	Oracle Odec(j, c)
00 $T \leftarrow \emptyset$	06 if $C_j \neq \emptyset$: return \perp	12 if $C_j = \emptyset$: return \perp
01 for all $j \in [1 \dots N]$:	07 if $t \in T$: return \perp	13 if $c \in C_j$: return \perp
02 $K_j \xleftarrow{\$} \mathcal{K}$	08 $T \leftarrow T \cup \{t\}; t_j \leftarrow t$	14 $m \leftarrow A.dec(K_j, t_j, c)$
03 $C_j \leftarrow \emptyset$	09 $c \leftarrow A.enc(K_j, t_j, m_b)$	15 return m
04 $b' \xleftarrow{\$} A$	10 $C_j \leftarrow C_j \cup \{c\}$	
05 return b'	11 return c	

Where adversary A has access to oracles Oenc and Odec and the tags in line 9 and 14 are the same.

- AMAC: the AMAC exists of tuple $(M.mac, M.vrf)$. $M.mac$ takes a key K in \mathcal{K} , a tag t in \mathcal{T} , and a message m in \mathcal{M} and outputs a ciphertext c in \mathcal{C} . $M.vrf$ takes a key K in \mathcal{K} , a tag t in \mathcal{T} , a message m in \mathcal{M} and a ciphertext c in \mathcal{C} and returns either true or false. The correctness requirement is that for all K in \mathcal{K} , t in \mathcal{T} and message in \mathcal{M} and c in $[M.mac(K, t, m)]$ we have $M.vrf(K, t, m, c) = \text{true}$. The security of the AMAC is defined with $\mathbf{Adv}_{AMAC, A, N}^{N-moit-uf} = \Pr[N-MIOT-UF_{A, N}]$, defined by the following game:

Game N-MIOT-UF _{A, N}	Oracle Omac(j, t, m)	Oracle Ovrif(j, m, c)
00 forged $\leftarrow 0$	07 if $C_j \neq \emptyset$: return \perp	13 if $C_j = \emptyset$: return \perp
01 $T \leftarrow \emptyset$	08 if $t \in T$: return \perp	14 if $(m, c) \in C_j$: return \perp
02 for all $j \in [1 \dots N]$:	09 $T \leftarrow T \cup \{t\}; t_j \leftarrow t$	15 if $M.vrf(K_j, t_j, m, c)$:
03 $K_j \xleftarrow{\$} \mathcal{K}$	10 $c \leftarrow M.mac(K_j, t_j, m)$	16 forged $\leftarrow 1$
04 $C_j \leftarrow \emptyset$	11 $C_j \leftarrow C_j \cup \{(m, c)\}$	17 return true
05 run A	12 return c	18 return false
06 return forged		

Where Adversary A can access oracles Omac and Ovrif and the tags in line 10 and 15 are the same.

3.1.3 goal

The goal is to make a scheme ADEM' from the ADEM and AMAC which has the same security of the ADEM, but is also secure against active attacks.

3.1.4 Security model

The security is defined by creating new $A.enc'$ and $A.dec'$ calls which are build using the calls the primitives provide us, and placing those in the ADEM game defined earlier:

Proc A.enc'(K, t, m)	Proc A.dec'(K, t, c)
00 $(K_{dem}, K_{mac}) \leftarrow K$	05 $(K_{dem}, K_{mac}) \leftarrow K$
01 $c_{dem} \leftarrow A.enc(K_{dem}, t, m)$	06 $(c_{dem}, c_{mac}) \leftarrow c$
02 $c_{mac} \leftarrow M.mac(K_{mac}, t, c_{dem})$	07 if $M.vrf(K_{mac}, t, c_{dem}, c_{mac})$:
03 $c \leftarrow (c_{dem}, c_{mac})$	08 $m \leftarrow A.dec(K_{dem}, t, c_{dem})$
04 return c	09 return m
	10 return \perp

The new advantage is $\mathbf{Adv}_{\text{ADEM}, A, N}^{\text{n-miot}} \leq 2\mathbf{Adv}_{\text{AMAC}, B, N}^{\text{n-miot-uf}} + \mathbf{Adv}_{\text{ADAM}, C, N}^{\text{n-moit-ind}}$. Where the running time of B is at most that of A plus the time required to run N-many ADEM encapsulations and Qd-many ADEM decapsulations and the running time of C is the same as the running time of A. Additionally, B poses at most Qd-many Ovr queries, and C poses no Odec query.

3.2 Existing notation from Generic Composition Reconsidered

3.2.1 notation

\mathcal{K} is a nonempty key space, \mathcal{N} is a non-empty nonce space, \mathcal{M} is a message space and \mathcal{A} is the associated-data space. \mathcal{M} (and \mathcal{A} ? it's a bit ambiguous in the paper but I assume this part only applies to \mathcal{M}) contain at least two strings, and if \mathcal{M} and \mathcal{A} contain a string of length x, they must contain all strings of length x.

3.2.2 used primitives

- nE: A nonce-based E scheme is defined by triple $(\mathcal{K}, \mathcal{E}, \mathcal{D})$. \mathcal{E} is a deterministic encryption algorithm that takes three inputs (K, N, M) to a value C , the length of C only depends the length of K , N and M . When (K, N, M) is not in $\mathcal{K} \times \mathcal{N} \times \mathcal{M}$, C will be \perp . \mathcal{D} is the decryption algorithm that takes three inputs (K, N, C) to a value M . \mathcal{E} and \mathcal{D} are inverse of each other implying correctness (if $\mathcal{E}(K, N, M) = C \neq \perp$, then $\mathcal{D}(K, N, C) = M$) and tidiness (if $\mathcal{D}(K, N, C) = M \neq \perp$, then $\mathcal{E}(K, N, M) = C$). The security is defined as follows:

$$\mathbf{Adv}_H^{\text{nE}}(\mathcal{A}) = \Pr[\mathcal{A}^{\mathcal{E}(\cdot, \cdot)} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{S}(\cdot, \cdot)} \Rightarrow 1]$$

where $H = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is a nE scheme; $K \leftarrow \mathcal{K}$ at the beginning of each game; $\mathcal{E}(N, M)$ returns $\mathcal{E}_K(N, M)$; $\mathcal{S}(N, M)$ computes $C \leftarrow \mathcal{E}_K(N, M)$, returns \perp if $C = \perp$, and otherwise returns $|C|$ random bits; and \mathcal{A} may not repeat the first component of an oracle query.

- MAC: The MAC is a deterministic algorithm F that takes in a K in \mathcal{K} and a string x and outputs either a n-bit length T or \perp . The domain of F is the set X such that $F(K, x) \neq \perp$. The security of F is defined by $\mathbf{Adv}_F^{\text{pr}} = |\Pr[A^F = 1] - \Pr[A^p = 1]|$. the game on the left selects a random K from \mathcal{K} and provides oracle access to $F(K, \cdot)$ the game on the right selects a uniformly random function p from X to $\{1, 0\}^n$ and provide oracle access to it. With each oracle, queries outside X return \perp .

3.2.3 goal

The end goal is a nonce-based authenticated encryption scheme $(\mathcal{K}, \mathcal{E}, \mathcal{D})$. \mathcal{E} is a deterministic encryption algorithm that takes four inputs (K, N, A, M) to a value C , the length of C value only depends the length of K , N , A and M . When (K, N, A, M) is not in $\mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M}$, C will be \perp . \mathcal{D} is the decryption algorithm that takes four inputs (K, N, A, C) to a value M . \mathcal{E} and \mathcal{D} are inverse of each other implying correctness (if $\mathcal{E}(K, N, A, M) = C \neq \perp$, then $\mathcal{D}(K, N, A, C) = M$) and tidiness (if $\mathcal{D}(K, N, A, C) = M \neq \perp$, then $\mathcal{E}(K, N, A, M) = C$).

3.2.4 Security model

Security is defined as follows:

$$\mathbf{Adv}_H^{\text{nAE}}(\mathcal{A}) = \Pr[\mathcal{A}^{\mathcal{E}(\cdot, \cdot, \cdot), \mathcal{D}(\cdot, \cdot, \cdot)} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{S}(\cdot, \cdot, \cdot), \perp(\cdot, \cdot, \cdot)} \Rightarrow 1]$$

where $H = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is an nAE scheme; $K \leftarrow \mathcal{K}$ at the beginning of each game; $\mathcal{E}(N, A, M)$ returns $\mathcal{E}_K(N, A, M)$ and $\mathcal{D}(N, A, C)$ returns $\mathcal{D}_K(N, A, C)$; and $\mathcal{S}(N, A, M)$ computes $C \leftarrow \mathcal{E}_K(N, A, M)$, returns \perp if $C = \perp$, and $|C|$ random bits otherwise, and $\perp(N, A, M)$ returns \perp ; and \mathcal{A} may not repeat the first component of an encryption (=left) query, nor make a decryption (=right) query (N, A, C) after C was obtained from a prior encryption (=left) query (N, A, M) .

We define the scheme secure if there is a tight reduction from breaking the nAE-security of the scheme to breaking the nE-security and the PRF security of the underlying primitives.

4 New Definition

should consist of:

- syntax of the primitive (input,output,correctness,tidiness, expected bounds)
- game based code
- explanation of the choices made
- formal comparison with other choices

4.1 notation

\mathcal{K} is a nonempty key space, \mathcal{N} is a non-empty nonce space and \mathcal{M} is a message space. \mathcal{M} contain at least two strings, and if it contain a string of length x , it must contain all strings of length x . N is the number of users.

4.2 goal

The end goal is to build a Authenticated Encryption scheme (EA). The AE scheme is defined by tuple $(\text{AE.enc}, \text{AE.dec})$. AE.enc is a deterministic encryption algorithm that takes three inputs (k, l, m) to a value c , the length of c only depends on the length of k , l and m . When (k, l, m) is not in $\mathcal{K} \times \mathcal{L} \times \mathcal{M}$, c will be \perp . AE.dec is the decryption algorithm that takes three inputs (k, n, c) to a value m . AE.enc and AE.dec are inverse of each other implying correctness (if $\text{AE.enc}(k, l, m) = c \neq \perp$, then $\text{AE.dec}(k, l, c) = m$) and tidiness (if $\text{AE.dec}(k, l, c) = m \neq \perp$, then $\text{AE.enc}(k, l, m) = c$).

4.3 Security model

The advantage is calculated with $\text{Adv}_{A,N}^{\text{AE}} = |\Pr[\text{AE}_{A,N}^{\text{AE}} = 1] - \Pr[\text{AE}_{A,N}^{\$} = 1]|$. $\text{AE}_{A,N}^{\text{AE}}$ is defined below, $\text{AE}_{A,N}^{\$}$ is the game $\text{AE}_{A,N}^{\text{AE}}$, where AE.enc in line 10 is replaced with $\$.enc$. $\$.enc(k, l, m)$ calls $c = \text{AE.enc}(k, l, m)$ then outputs \perp if c is \perp or $|c|$ random bits otherwise. EA.dec in line 15 is replaced with $\$.dec(k, l, c)$ that always returns \perp .

Game $\text{AE}_{A,N}^{\text{AE}}$	Oracle $\text{Oenc}(j, l, m)$	Oracle $\text{Odec}(j, c)$
0 : $L = \emptyset$	6 : if $C_j \neq \emptyset$: return \perp	13 : if $c_j \neq \emptyset$: return \perp
1 : for $j \in [1..N]$:	7 : if $l \in L$: return \perp	14 : if $c \in C_j$: return \perp
2 : $k_j \leftarrow \mathcal{K}$	8 : $L = L \cup \{l\}$	15 : $m \leftarrow \text{AE.dec}(k_j, l_j, c)$
3 : $C_j = \emptyset$	9 : $l_j = l$	16 : return m
4 : $b' \leftarrow A$	10 : $c \leftarrow \text{AE.enc}(k_j, l_j, m)$	
5 : return b'	11 : $C_j = c$	
	12 : return c	

Figure 1: AE game

4.4 choices made

In this security model, there are a lot of choices made, in this section I will elaborate on these. Firstly, the used MAC primitive is required to be indistinguishable from RO. This choice was made as it is required for AE constructions from Generic Composition Reconsidered. I also choose to have "locks" instead of nonces. In this setting, this results in the adversary not being able to make decryption queries for any incorrect lock values. I choose this as a setting with locks will suffice the setting of hybrid encryption while probably being easier to prove. For this reason I also choose to not allow multiple encryption calls for one user, as well as not allowing decryption calls to a user which encryption call is not yet made. In my security game for the DEM, there is no decryption oracle considered. When distinguishing from a random function, this is equivalent to a situation in which a decryption oracle is considered (I am not 100% sure about his part). Lastly, we choose to indistinguishably from a random function as a security model, instead of indistinguishably of two encrypted messages. This is a stronger security notion in the current setting as the length of the ciphertext only depends on the length of the inputs. Indistinguishably of two encrypted messages probably suffices in the current setting so I might still change to this if proving indistinguishably from a random function seems infeasible.

5 Constructions

should consist of:

- the old primitives
- how to construct the new primitive from old primitives
- security bounds + proof
- comparison with existing alternatives

5.1 used primitives

\mathcal{K} is a nonempty key space, \mathcal{N} is a non-empty nonce space and \mathcal{M} is a message space. \mathcal{M} contain at least two strings, and if it contain a string of length x , it must contain all strings of length x . N is the number of users.

- DEM: a DEM scheme is defined by tuple $(E.\text{enc}, E.\text{dec})$. $E.\text{enc}$ is a deterministic encryption algorithm that takes three inputs (k, l, m) to a value c , the length of c only depends on the length of k , l and m . When (k, l, m) is not in $\mathcal{K} \times \mathcal{L} \times \mathcal{M}$, c will be \perp . $E.\text{dec}$ is the decryption algorithm that takes three inputs (k, l, c) to a value m . $E.\text{enc}$ and $E.\text{dec}$ are inverse of each other implying correctness (if $E.\text{enc}(k, l, m) = c \neq \perp$, then $E.\text{dec}(k, l, c) = m$) and tidiness (if $E.\text{dec}(k, l, c) = m \neq \perp$, then $E.\text{enc}(k, l, m) = c$). The advantage is calculated with $\text{Adv}_{A,N}^{\text{DEM}} = |\Pr[\text{DEM}_{A,N}^{\text{DEM}} = 1] - \Pr[\text{DEM}_{A,N}^{\$} = 1]|$. $\text{DEM}_{A,N}^{\text{DEM}}$ is defined below, $\text{DEM}_{A,N}^{\$}$ is the game $\text{DEM}_{A,N}^{\text{DEM}}$, where $E.\text{enc}$ in line 10 is replaced with $\$.enc(k, l, m)$. $\$.enc$ calls $c = E.\text{enc}(k, l, m)$ then outputs \perp if c is \perp or $|c|$ random bits otherwise.

Game $\text{DEM}_{A,N}^{\text{DEM}}$	Oracle $\text{Oenc}(j,l,m)$
0 : $L = \emptyset$	6 : if $C_j \neq \emptyset$: return \perp
1 : for $j \in [1..N]$:	7 : if $l \in L$: return \perp
2 : $k_j \leftarrow \mathcal{K}$	8 : $L = L \cup \{l\}$
3 : $C_j = \emptyset$	9 : $l_j = l$
4 : $b' \leftarrow A$	10 : $c \leftarrow \text{E.enc}(k_j, l_j, m)$
5 : return b'	11 : $C_j = c$
	12 : return c

Figure 2: DEM game

- MAC: The MAC is a deterministic algorithm M.mac that takes in a fixed length k in \mathcal{K} , a fixed length l in \mathcal{L} and a variable length message m in \mathcal{M} and outputs either a n -bit length tag or \perp . The domain of M.mac is the set X such that $\text{M.mac}(k,l,m) \neq \perp$. The advantage is calculated with $\text{Adv}_{A,N}^{\text{MAC}} = |\Pr[\text{MAC}_{A,N}^{\text{MAC}} = 1] - \Pr[\text{MAC}_{A,N}^{\$} = 1]|$. $\text{MAC}_{A,N}^{\text{MAC}}$ is defined below, $\text{MAC}_{A,N}^{\$}$ is the game $\text{MAC}_{A,N}^{\text{MAC}}$, where M.mac in line 10 is replaced with $\$.mac$. $\$.mac(k,l,m)$ computes $t = \text{M.mac}(k,l,m)$ then outputs \perp if t is \perp or $|t|$ random bits otherwise.

Game $\text{MAC}_{A,N}^{\text{MAC}}$	Oracle $\text{Omac}(j,l,m)$
0 : $L = \emptyset$	6 : if $T_j \neq \emptyset$: return \perp
1 : for $j \in [1..N]$:	7 : if $l \in L$: return \perp
2 : $k_j \leftarrow \mathcal{K}$	8 : $L = L \cup \{l\}$
3 : $T_j = \emptyset$	9 : $l_j = l$
4 : $b' \leftarrow A$	10 : $t \leftarrow \text{M.mac}(k_j, l_j, m)$
5 : return b'	11 : $T_j = t$
	12 : return t

Figure 3: MAC game

5.2 construction

The AE schemes we will look at are constructed from the DEM and the MAC. Following Generic Composition Reconsidered, three ways to construct this AE are of interest, namely the ones following from the N1, N2 and N3 scheme. One thing to keep in mind with this that these schemes would originally use associated data. For now we can discard this but it is not proven that the same security results would also follow from this case without associated data. Down here the schemes, adjusted to our setting, can be found, followed by the AE.enc and AE.dec calls that can we construct following these schemes. These calls can be plugged into $\text{AE}_{A,N}^{\text{AE}}$ to find their respective security games.

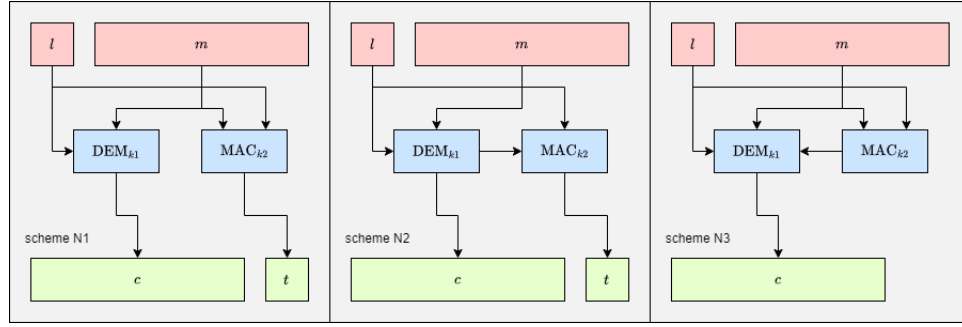


Figure 4: Adjusted N schemes from Generic Composition Reconsidered

$\text{AE.enc}(k, l, m)$	$\text{AE.dec}(k, l, c)$
0 : $(k1, k2) \leftarrow k$	5 : $(k1, k2) \leftarrow k$
1 : $c' \leftarrow \text{E.enc}(k1, l, m)$	6 : $(c', t) \leftarrow c$
2 : $t \leftarrow \text{M.mac}(k2, l, m)$	7 : $m \leftarrow \text{E.dec}(k1, l, c')$
3 : $c \leftarrow (c', t)$	8 : $t' \leftarrow \text{M.mac}(k2, l, m)$
4 : return c	9 : if $t = t'$: return m
	10 : else : return \perp

Figure 5: Calls based on N1

$\text{AE.enc}(k, l, m)$	$\text{AE.dec}(k, l, c)$
0 : $(k1, k2) \leftarrow k$	5 : $(k1, k2) \leftarrow k$
1 : $c' \leftarrow \text{E.enc}(k1, l, m)$	6 : $(c', t) \leftarrow c$
2 : $t \leftarrow \text{M.mac}(k2, l, c')$	7 : $m \leftarrow \text{E.dec}(k1, l, c')$
3 : $c \leftarrow (c', t)$	8 : $t' \leftarrow \text{M.mac}(k2, l, c')$
4 : return c	9 : if $t = t'$: return m
	10 : else : return \perp

Figure 6: Calls based on N2

AE.enc(k, l, m)	AE.dec(k, l, c)
0 : $(k1, k2) \leftarrow k$	5 : $(k1, k2) \leftarrow k$
1 : $t \leftarrow \text{M.mac}(k2, l, m)$	6 : $m' \leftarrow \text{E.dec}(k1, l, c)$
2 : $m' \leftarrow m \parallel t$	7 : $(m, t) \leftarrow m'$
3 : $c \leftarrow \text{E.enc}(k1, l, m')$	8 : $t' \leftarrow \text{M.mac}(k2, l, m)$
4 : return c	9 : if $t = t'$: return m
	10 : else : return \perp

Figure 7: Calls based on N3

The scheme is considered secure when there is a tight reduction from breaking the AE-security of the scheme to breaking the defined security of the underlying primitives.

6 Use cases

should consist of:

- possible use cases

7 Related Work

Location not final yet

8 Conclusion

9 Appendix