BACHELOR THESIS



RADBOUD HONOURS ACADEMY

TBD

Author: Stijn Vandenput Supervisor: Martijn Stam Bart Mennink

20/05/2022

Abstract

Contents

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

1 Introduction

should consist of:

- explaining the challenge
- my contribution

Within the field of cryptography there are a 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 a 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	M	\mathcal{M}	\mathcal{M}	set of all possible message options
message	m	M	m	message the user sends
ciphertext space	С	-	С	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{C}	\mathcal{T}	\mathcal{T}	set off all possible tag options
tag	c	T	t	output of mac function
key space	\mathcal{K}	\mathcal{K}	κ	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{T}	-	\mathcal{L}	set of all possible lock options
lock	t	-	l	nonce that is bound to the user
users	N	-	N	number of users
adversary	A	\mathcal{A}	A	the bad guy
Random sampling	←\$	« —		get a random ellermeth form the set
result of function	←\$	←		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.enc,A.dec), A.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.dec takes a K \mathcal{K} , a tag t in \mathcal{T} and a ciphertext c in \mathcal{C} and outputs a message m in \mathcal{M} or \bot 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 $\mathbf{Adv}_{\mathrm{ADEM},A,N}^{\mathrm{n-moit-ind}}$ = $|\Pr[N\text{-MIOT-IND}_{A,N}^0] - \Pr[N\text{-MIOT-IND}_{A,N}^1]|$, defined by the following game:

```
Game N-MIOT-IND_{A,N}^b
                                                   Oracle Oenc(j, t, m_0, m_1)
                                                                                                           Oracle Odec(j, c)
 0 \rightarrow T
                                                   of if C_j \neq \emptyset: return \perp
                                                                                                           12 if C_j = \emptyset: return \bot
of for all j \in [1..N]:
                                                   or if t \in T: return \bot
                                                                                                          13 if c \in C_j: return \perp
14 m \leftarrow A.dec(K_j, t_j, c)
02 K_j \stackrel{\text{$\ \ }}{\sim} \mathcal{K}
03 C_j \leftarrow \emptyset
04 b' \stackrel{\text{$\ \ \ }}{\sim} \mathsf{A}
                                                   08 T \leftarrow T \cup \{t\}; t_j \leftarrow t
                                                   09 c \leftarrow \mathsf{A.enc}(K_j, t_j, m_b)
                                                                                                          15 return m
                                                   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 of 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)=true. The security of the AMAC is defined with $\mathbf{Adv}_{\mathrm{AMAC,A,N}}^{\mathrm{N-moit-uf}} = \Pr[\mathrm{N-MIOT-UF}_{\mathrm{A,N}}]$, defined by the following game:

```
Game N-MIOT-UF<sub>A,N</sub>
                                                 Oracle Omac(j, t, m)
                                                                                                      Oracle Ovrf(j, m, c)
 of forged \leftarrow 0
                                                 07 if C_j \neq \emptyset: return \bot
08 if t \in T: return \bot
                                                                                                     13 if C_j = \emptyset: return \bot
14 if (m, c) \in C_j: return \bot
of T \leftarrow \emptyset
02 for all j \in [1..N]:
03 K_j \stackrel{\text{de}}{\leftarrow} \mathcal{K}
04 C_j \leftarrow \emptyset
                                                 og T \leftarrow T \cup \{t\}; t_j \leftarrow t
                                                                                                      15 if M.vrf(K_j, t_j, m, c):
                                                 10 c \leftarrow \mathsf{M.mac}(K_j, t_j, m)
                                                                                                     16 forged \leftarrow 1
                                                 11 C_j \leftarrow C_j \cup \{(m,c)\}
                                                                                                              return true
05 run A
                                                                                                      18 return false
 6 return forged
```

Where Adversary A can access oracles Omac and Ovrf 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:

The new advantage is $\mathbf{Adv}_{\mathrm{ADEM',A,N}}^{\mathrm{n-miot}} \leq 2\mathbf{Adv}_{\mathrm{AMAC,B,N}}^{\mathrm{n-miot-uf}} + \mathbf{Adv}_{\mathrm{ADAM,C,N}}^{\mathrm{n-moit-ind}}$. 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 Ovrf 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}, E, D) . 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 \bot . D is the decryption algorithm that takes three inputs (K, N, C) to a value M. E and D are inverse of each other implying correctness (if $E(K, N, M) = C \neq \bot$, then D(K, N, C) = M) and tidiness (if $D(K, N, C) = M \neq \bot$, then E(K, N, M) = C). The security is defined as follows:

$$\mathbf{Adv}_{II}^{\mathrm{nE}}(\mathcal{A}) = \Pr \left[\mathcal{A}^{\mathcal{E}(\cdot,\cdot)} \Rightarrow 1 \right] - \Pr \left[\mathcal{A}^{\$(\cdot,\cdot)} \Rightarrow 1 \right]$$

where $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is a nE scheme; $K \leftarrow \mathcal{K}$ at the beginning of each game; $\mathcal{E}(N, M)$ returns $\mathcal{E}_K(N, M)$; \$(N, M) computes $C \leftarrow \mathcal{E}_K(N, M)$, returns \bot if $C = \bot$, and otherwise returns |C| random bits; and 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 \bot . The domain of F is the set X such that $F(K,x) \neq \bot$. The security of F is defined by $\mathbf{Adv}_F^{\mathrm{pfr}} = |\Pr[A^F] - \Pr[A^P]|$. the game on the left selects a random K from \mathcal{K} and provides oracle access to F(K,.) 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 \bot

3.2.3 goal

The end goal is a nonce-based authenticated encryption scheme (\mathcal{K}, E, D) . 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 \bot . D is the decryption algorithm that takes four inputs (K, N, A, C) to a value M. E and D are inverse of each other implying correctness (if $E(K, N, A, M) = C \neq \bot$, then D(K, N, A, C) = M) and tidiness (if $D(K, N, A, C) = M \neq \bot$, then E(K, N, A, M) = C)

3.2.4 Security model

Security is defined as follows:

$$\mathbf{Adv}_{\pi}^{\mathrm{nAE}}(\mathcal{A}) = \Pr\left[\mathcal{A}^{\mathcal{E}(\cdot,\cdot,\cdot)}, \mathcal{D}(\cdot,\cdot,\cdot) \Rightarrow 1\right] - \Pr\left[\mathcal{A}^{\$(\cdot,\cdot,\cdot)}, \bot(\cdot,\cdot,\cdot) \Rightarrow 1\right]$$

where $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ 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 \$(N, A, M) computes $C \leftarrow \mathcal{E}_K(N, A, M)$, returns \bot if $C = \bot$, and |C| random bits otherwise, and $\bot(N, A, M)$ returns \bot ; and 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. Nis the number of users.

4.2 used primitives

• DEM: a DEM scheme is defined by tuple (E.enc,E.dec). E.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 \bot . E.dec is the decryption algorithm that takes three inputs (k,l,c) to a value m. E.enc and E.dec are inverse of each other implying correctness (if E.enc $(k,l,m) = c \neq \bot$, then E.enc (k,l,c) = m) and tidiness (if E.dec $(k,l,c) = m \neq \bot$, then E.enc(k,l,m) = c). The advantage is calculated with $\mathbf{Adv}_{A,N}^{\mathrm{DEM}} = |\Pr[\mathrm{DEM}_{A,N}^{\mathrm{DEM}}| - \Pr[\mathrm{DEM}_{A,N}^{\$}]|$. $\mathrm{DEM}_{A,N}^{\mathrm{DEM}}$ is defined below, $\mathrm{DEM}_{A,N}^{\$}$ is the game $\mathrm{DEM}_{A,N}^{\mathrm{DEM}}$, where E.enc in line 10 is replaced with $\$.\mathrm{enc}(k,l,m)$. $\$.\mathrm{enc}$ calls $c = \mathrm{E.enc}(k,l,m)$ then outputs \bot if c is \bot or |c| random bits otherwise.

Figure 1: 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 \bot . The domain of M.mac is the set X such that $\mathrm{M.mac}(k,l,m) \neq \bot$. The advantage is calculated with $\mathbf{Adv}_{A,N}^{\mathrm{MAC}} = |\Pr[\mathrm{MAC}_{A,N}^{\mathrm{MAC}}] - \Pr[\mathrm{MAC}_{A,N}^{\$}]|$. $\mathrm{MAC}_{A,N}^{\mathrm{MAC}}$ is defined below, $\mathrm{MAC}_{A,N}^{\$}$ is the game $\mathrm{MAC}_{A,N}^{\mathrm{MAC}}$, where M.mac in line 10 is replaced with \$.mac. \$.mac(k,l,m) computes $t=\mathrm{m.mac}(k,l,m)$ then outputs \bot if t is \bot or |t| random bits otherwise.

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

Figure 2: MAC game

4.3 goal

The end goal is to build a Authenticated Encryption scheme (EA). The AE scheme is defined by tuple (AE.enc,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 \bot . AE.dec is the decryption algorithm that takes three inputs (k,n,c) to a value m. AE.enc and E.dec are inverse of each other implying correctness (if AE.enc(k,l,m) = $c \ne \bot$, then AE.dec(k,l,c) = m) and tidiness (if AE.dec $(k,l,c) = m \ne \bot$, then AE.enc(k,l,m) = c).

4.4 Security model

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

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

Figure 3: AE game

4.5 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:

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

The AE schemes should be 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 initial schemes can be found, followed by the AE.enc and AE.dec calls that can we construct following these schemes.

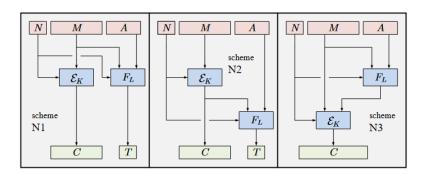


Figure 4: Original N schemes from Generic Composition Reconsidered

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

Figure 5: Calls based on N1

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

Figure 6: Calls based on N2

AE.	enc(k,l,m)	AE.	$\operatorname{dec}(k,l,c)$
0:	$(k1, k2) \leftarrow k$	5:	$(k1, k2) \leftarrow k$
1:	t = M.mac(k2, l, m)	6:	m' = E.dec(k1, l, c)
2:	m' = mt	7:	$(m,t) \leftarrow m'$
3:	c = E.enc(k1, l, m')	8:	t' = M.mac(k2, l, m)
4:	$\mathbf{return}\ c$	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