Secure Affine Domain Extensions

Author: Mridul Nandi

The George Washington University

Computer Science Department.

Presenter: Liting Zhang

Chinese Academy of Sciences

Outline of Talk

- PRF or Pseudorandom Function.
- Study Known Examples.
- Affine Domain Extensions or ADEs
- Collision Relation.
- Secure Affine Domain Extensions or SADEs.
- Improved PRF Analysis.
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

Outline of Talk

- **PRF or Pseudorandom Function.**
- Study Known Examples.
- Affine Domain Extensions or ADEs
- Collision Relation.
- Secure Affine Domain Extensions or SADEs.
- Improved PRF Analysis.
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

Distinguishing/Forgery Attack

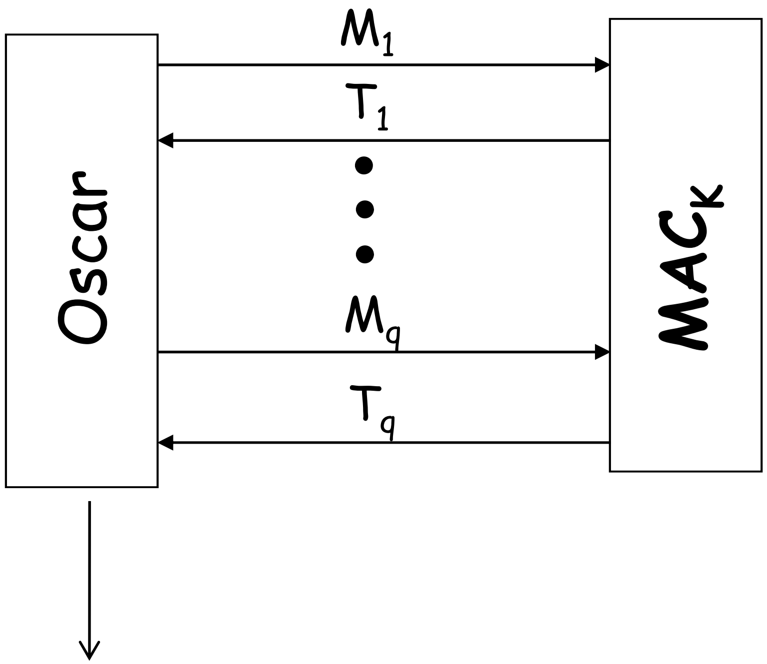
- Pseudorandom function (PRF) is Stronger security notion than unforgeability or unpredictability.

- Oscar makes distinct queries M_1, M_2,...,M_q adaptively and obtains responses T_1,T_2,..., T_q.

  – PRF distinguisher: distinguish (T_1, … ,T_q) from a q-tuple of random strings.

  – Forgery: compute a response T for a different message M.

Distinguishing/Forgery Attack



1. PRF Attack: Is $(T_1,\ldots,T_q)$ completely random?
2. Forgery Attack: Find M different from the messages and its tag.

1. Find some non-random property of $(T_1,\ldots, T_q)$.
2. Find different M and T such that $MAC_K(M) = T$.

PRF Advantage

$Adv^{prf}(Oscar) = |Pr_K[Oscar(\mathbf{T}) = 1 \mid MAC_K] - Pr_{\mathbf{T}}[Oscar(\mathbf{T}) = 1 \mid \text{uniform } \mathbf{T}]|$

- Oscar is interacting with either random function or MAC and finally he has to guess with whom he is interacting. This is also known as distinguishing advantage.

- **$Adv^{prf}(q,t,L,\ldots) = \max \text{prf-}Adv_{MAC}(Oscar)$**, where maximum is over all distinguishers Oscar which makes at most **q queries**, requires **t** and **L blockcipher invocations** to compute q queries and the longest query respectively.

Outline of Talk

- PRF or Pseudorandom Function.
- **Study Known Examples.**
- Affine Domain Extensions or ADEs
- Collision Relation.
- Secure Affine Domain Extensions or SADEs.
- Improved PRF Analysis.
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

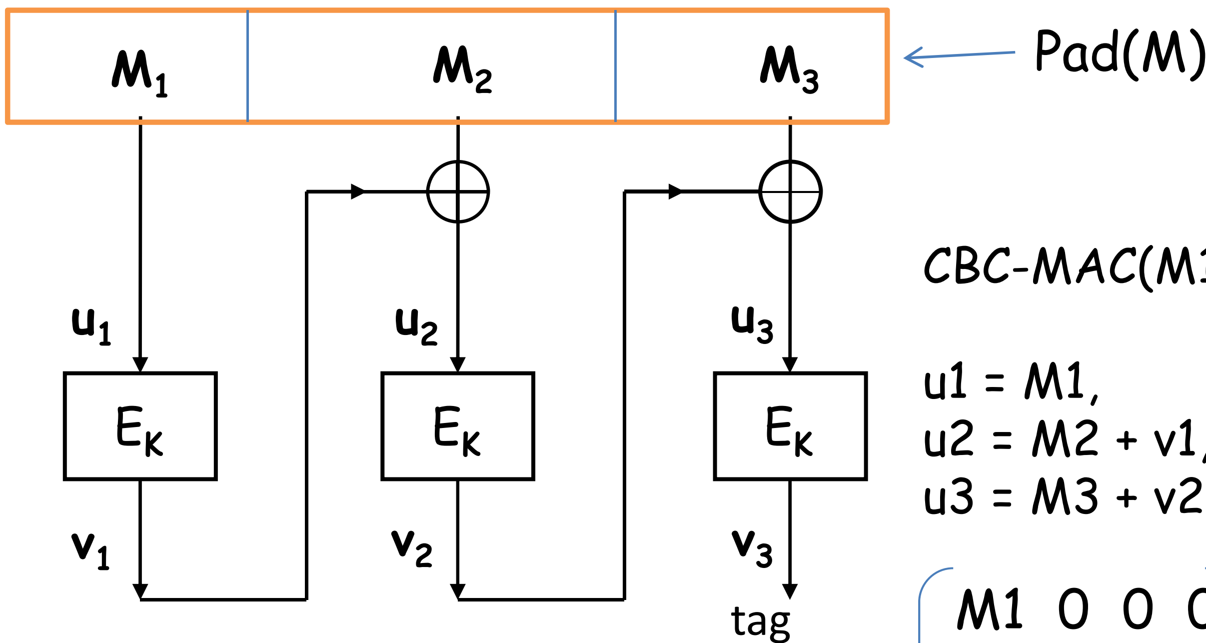
Broad Categories of MAC

- Universal Hash-based: with/without Nonce
  - **Poly1305, UMAC, MMH**, etc.
- Block cipher based
  - Sequential (CBC-type): **CBC-MAC, ECBC, XCBC, TMAC, OMAC**, **GCBC**, etc.
  - Parallel : **PMAC, XOR-MAC, DAG-based-PRF,** etc.
- Hash function (also compression function) based
  - **HMAC, NMAC, EMD, NI, sandwich-MD**,etc.

Broad Categories of MAC

- Universal Hash-based: with/without Nonce
  - **Poly1305, UMAC, MMH**, etc.
- **Block cipher based**
  - Sequential (CBC-type): **CBC-MAC**, **ECBC, XCBC, TMAC, OMAC**, **GCBC**, etc.
  - Parallel : **PMAC**, XOR-MAC, **DAG-based-PRF**, etc.
- Hash function (also compression function) based
  - **HMAC, NMAC, EMD, NI, sandwich-MD**,etc.

CBC-MAC



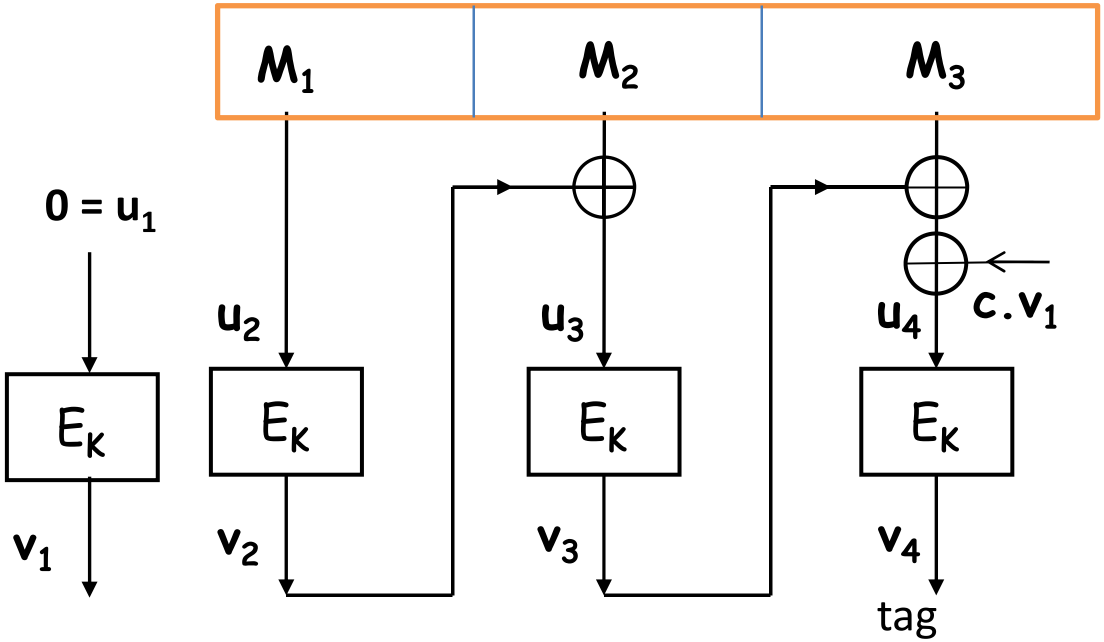
CBC-MAC(M1, M2, M3) = v3 where

u1 = M1, v1= EK(u1),
u2 = M2 + v1, v2 = EK(u2)
u3 = M3 + v2, v3 = EK(u3)

$$\begin{pmatrix} M1 & 0 & 0 & 0 \\ M2 & 1 & 0 & 0 \\ M3 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ v1 \\ v2 \\ v3 \end{pmatrix} = \begin{pmatrix} u1 \\ u2 \\ u3 \end{pmatrix}$$

Coefficient matrix of CBC-MAC for the message (M1,M2,M3).
It is independent of the blockcipher. Associated with each message.

OMAC



$OMAC(M) = v_4$ where

$u_1 = \mathbf{0}$, $v_1 = E_K(u_1)$,
$u_2 = M_1$, $v_2 = E_K(u_2)$
$u_3 = M_2 + v_2$, $v_3 = E_K(u_3)$
$u_4 = M_3 + v_3 + c.v_1$, $v_4 = E_K(u_4)$

c depends on whether message needs padding or not.

Coefficient matrix of OMAC.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ M1 & 0 & 0 & 0 & 0 \\ M2 & 0 & 1 & 0 & 0 \\ M3 & c & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ v1 \\ v2 \\ v3 \\ v4 \end{pmatrix} = \begin{pmatrix} u1 \\ u2 \\ u3 \\ u4 \end{pmatrix}$$

The final E_K output is final output of CBC-MAC and OMAC. Similarly for PMAC

PMAC and GCBC

Coefficient matrix of PMAC.

The final E_K output is final output of CBC-MAC and OMAC. Similarly for PMAC

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ M_1 & c_1 & 0 & 0 & 0 \\ M_2 & c_2 & 1 & 0 & 0 \\ M_3 & c & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ v1 \\ v2 \\ v3 \\ v4 \end{pmatrix} = \begin{pmatrix} u1 \\ u2 \\ u3 \\ u4 \end{pmatrix}$$

Coefficient matrix of GCBC.

The final E_K output is final output of CBC-MAC and OMAC. Similarly for PMAC

$$\begin{pmatrix} M_1 & 0 & 0 & 0 \\ M_2 & 1 & 0 & 0 \\ M_3 & 0 & c_i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ v1 \\ v2 \\ v3 \end{pmatrix} = \begin{pmatrix} u1 \\ u2 \\ u3 \end{pmatrix}$$

Outline of Talk

- PRF or Pseudorandom Function.
- Study Known Examples.
- **Affine Domain Extensions or ADEs**
- Collision Relation.
- Secure Affine Domain Extensions or SADEs.
- Improved PRF Analysis.
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

Definition of an ADE

- A Blockcipher based PRF is called ADE if there are constants $a_{i,j}$ (depends only on message, not on the blockcipher E_K) and l such that for $1 \leq i \leq l$,
  - $u_i = a_{i0} + a_{i1} v_1 + \ldots + a_{i\,i-1} v_{i-1}$
  - $v_i = E_K(u_i)$

and the final output of PRF is v_l.

Definition of an ADE

$$\begin{pmatrix} a_{10} & a_{11} & \cdots & a_{1l} \\ a_{20} & a_{21} & \cdots & a_{2l} \\ & & & \\ a_{l0} & a_{l1} & \cdots & a_{ll} \end{pmatrix} \begin{pmatrix} 1 \\ v_1 \\ v_2 \\ \\ v_l \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ \\ u_l \end{pmatrix}$$

u_i's and v_i's are intermediate inputs and outputs respectively a_{ij}'s are some constant depend only on the message. The final output is v_l.

Non- ADE

- XOR-MAC: It is the xor of all blockcipher outputs.
- Poly1305, XCBC, TMAC: It requires auxiliary keys other than blockcipher key.
- ECBC: Two independent blockcipher keys.
- However, security analysis of ADE can be used in last two cases. The first case needs a different treatment.

Outline of Talk

- PRF or Pseudorandom Function.
- Study Known Examples.
- Affine Domain Extensions or ADEs
- **Collision Relation.**
- Secure Affine Domain Extensions or SADEs.
- Improved PRF Analysis.
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

Collision Relation

- Collision Relation: Equivalence relation on index set {1,2,…,l} such that i and j are related if and only if $u_i = u_j$.
  - Suppose u1 = u6, u2=u5, u3 = u4 then corresponding collision relation: 1 ~ 6, 2 ~ 5 and 3 ~4.

- $E(u_i) = v_i$ means that $u_i = u_j$ if and only if $v_i = v_j$ It captures the collision pattern without mentioning the actual values of intermediate inputs.

i-isolatied

- Let $\sim$ be a collision relation on {1,…,l} then we say i is **isolated** if no other element is related to i (u_i is fresh different from other inputs).

- If l is isolated then the u_l is fresh, hence the final output (i.e. v_l) is “almost” random.

Collision Relation for Two Messages

- Let $t = l + l'$
- Let M and M' be two messages. Let $u_1, .. u_l$ be intermediate inputs of M and $u_{l+1}, \ldots, u_t$ be intermediate inputs of M'. Similarly for v_i's.
- We similarly define collision relation on [1,t] for all t intermediate inputs/outputs.
- If t is isolated then F(M') is random. Similarly, F(M) is random if l is isolated.

Forced Collision Relation

- There is a unique collision relation $\sim^*$ whose corresponding collisions hold for all permutation. It is called forced collision relation.
- We say F is non-secure ADE if there are messages M and M' such that t is NOT isolated in $\sim^*$ i.e., $F(M') = v_j$ for some $j \neq t$.
- Non-secure ADEs are not "good": They leak some intermediate outputs. Not known how to extend to a generic distinguishing attack.

Outline of Talk

- PRF or Pseudorandom Function.
- Study Known Examples.
- Affine Domain Extensions or ADEs
- Collision Relation.
- **Secure Affine Domain Extensions or SADEs.**
- Improved PRF Analysis.
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

Secure Affine Domain Extensions

- Definition: **SADE is not non-secure ADE.**
- That is, for all $M \neq M'$ and any fixed i
  - Pr[F(M') = v_i] < 1, v_i is i^{th} intermediate output of F(M').
- No need to be the above probability very small in the definition. However, due to affine relation the probability is either one or close to $1/2^n$.

CBC is NOT SADE

- Let M= (m1,m2) and M’ = m1 then clearly, F(M’) = v1 with probability one. ➔ NOT SADE
- Use the above property to have length extension attack, so it is not PRF.

A variant of OMAC is NOT SADE

- Consider a variant of OMAC in which one of the constant c is 1.
- We have PRF attack and it is not SADE.
- M'=m1, M= (m1,0) then F(M') = v2.

Prefix-free CBC-MAC, GCBC, OMAC, PMAC, DAG-based PRF are SADE

- One can show that there are no trivial collisions between final output and some intermediate output. Hence these are SADE.

Outline of Talk

- PRF or Pseudorandom Function.
- Study Known Examples.
- Affine Domain Extensions or ADEs
- Collision Relation.
- Secure Affine Domain Extensions or SADEs.
- **Improved PRF Analysis.**
- Comparison With Existing Bounds.
- Conclusion and Open Problems.

Main Theorem

- Let N(M,M’) denote number of all accident one collision relations for M and M’ such that one of (l+l’) and l is not isolated.
- $N(t,q) = \textbf{max}\ (N(M_1,M_2) + \ldots N(M_{q-1}, M_q))$ maximum over all q messages which requires t invocations.
- For any SADE D, and any (t,q)-distinguisher A the PRF advantage:
  - $ADV^{prf}(A) = O(N(t,q)/2n + tq/2^n)$ and hence
  - $ADV_D(t,q) = O(N(t,q)/2n + tq/2^n)$.

Accidents of Collision Relation

- Not all collisions are ``unexpected''.
- There are some collision which are
  - known before hand (e.g. forced collisions occurs due to choice of messages) or
  - implied from previous collisions.
- Accident = largest set of unexpected collisions. All Collisions are implied from Accidents.
- Pr[a randomly chosen permutation has accident a] ≈ $1/2^{na}$.

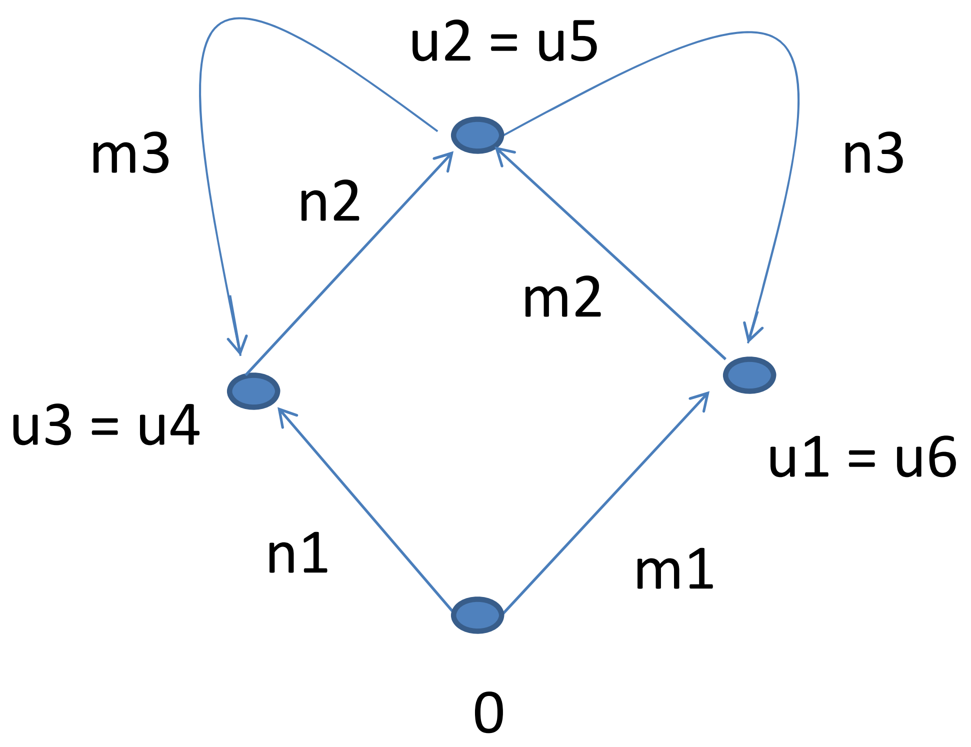
An Example

M = (m1, m2, m3), M' = (n1, n2, n3) such that m1 ⊕ m3 = n1 ⊕ n3

collision relation: 1 ~ 6, 2 ~ 5, 3 ~ 4

$$
\begin{pmatrix}
m1 & 0 & 0 & 0 & 0 & 0 & 0 \\
m2 & 1 & 0 & 0 & 0 & 0 & 0 \\
m3 & 0 & 1 & 0 & 0 & 0 & 0 \\
n1 & 0 & 0 & 0 & 0 & 0 & 0 \\
n2 & 0 & 0 & 0 & 1 & 0 & 0 \\
n3 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
1 \\ v1 \\ v2 \\ v3 \\ v3 \\ v2 \\ v1
\end{pmatrix}
=
\begin{pmatrix}
u1 \\ u2 \\ u3 \\ u3 \\ u2 \\ u1
\end{pmatrix}
$$

The graphical representation of the Example



PRF Bounds for Some Popular Examples

- For CBC, OMAC, GCBC and PMAC
  - $N(M,M') \leq c(l + l')$ for some constant c.
  - Hence $N(t,q) \leq tq$ and we prove our bound.
- For any SADE $N(M,M') \leq c(l + l')^2$. Hence
  - $N(t,q) = O(t^2)$.

PRF Bound Comparison

Mode	#BC	Known PRF-bound	**PRF-bound [this paper]**
CBC	m	$Lq^2/2^n$	**$tq/2^n$**
GCBC	m	$t^2/2^n$	**$tq/2^n$**
OMAC	m+1	$tq/2^n$	**$tq/2^n$**
PMAC	m+1	$tq/2^n$	**$tq/2^n$**
DAG-based	m	$t^2/2^n$	**-**
SADE [this paper]	-	-	**$N(t,q)/2^n + tq/2^n$**

Some Notes on Our Bounds

- $tq \le Lq^2$ since $t \le Lq$.
- Sometimes Lq^2 can be worse. E.g., when $t/2 = q = L$ (all message have one block except one which has q blocks) then
  - $tq = 2q^2$, $t^2 = 4q^2$, $Lq^2 = q^3$.
- $N(t,q) < t^2$. But, sometimes $N(t,q) < tq$. We will talk later.

Conclusion

- We characterize a PRF secure class of blockcipher based construction: SADE.
- We provide a security analysis which can potentially give improved bounds $O(tq/2^n)$.
- In particular we have the improved bounds for CBC, GCBC.

Open Questions

- Is $N(t,q) = O(tq)$ for all SADE?
- Are all non-SADE insecure?
- Are there some interesting SADE which are not proposed yet?

Thank you very much for your attention.

Please send your questions and comments to
mridul.nandi@gmail.com