How to protect integrity of 2^{75} blocks using the Magma cipher and a single key?

Vitaly Kiryukhin

LLC “SFB Lab”, JSC “InfoTeCS”

CTCrypt 2024
4 June 2024

vitaly.kiryukhin@sfblaboratory.ru

СФБ ЛАБ

Comments about the title

1. We construct pseudorandom functions (PRF),

   not just message authentication codes (MAC)

2. The numerical estimate (2^{75} blocks) is not the best achieved

3. Magma is used only as a well-known example

   of a secure 64-bit block cipher

PRF as MAC

$$\Pr(\text{single forgery in } v \text{ attempts}) \le \mathrm{Adv}^{PRF}(t, q+v, \ell) + \frac{v}{2^{\tau}} \le \pi_{\mathrm{mac}}$$

$\pi_{\mathrm{mac}} = 2^{-10}$	maximum allowable probability of a forgery
q	number of protected messages
ℓ	maximum block-length of message
$\tau = n = 64$	tag length = block length
$v \ll q$	number of forgery attempts
t	attacker's computational resources are reasonably limited (say, 2^{128})

Motivation

Conditions

- Kuznyechik (128-bit) is much more **secure** than Magma (64-bit),

but at the same time, Magma is:

- **no slower** on any platform;
- **much faster** in a **low-resource** environment;

and also multiplication in GF(2^n) is:

- **comparable** to encryption;
- **much faster** in many cases.

Performance and security

СФБ ЛАБ

- By performance:
  - Magma is preferable than Kuznyechik
  - Multiplication in GF is preferable to encryption
- We need the block cipher **modes** that provide

  an **acceptable level of security** even with a **low-resource cipher**

Multilinear hash function

Multilinear hash function

$$\mathrm{MH}(H, B) = H_1 \otimes B_1 \oplus H_2 \otimes B_2 \oplus \ldots \oplus H_\ell \otimes B_\ell = \bigoplus_{i=1}^{\ell} \left(H_i \otimes B_i\right)$$

- ℓ-block key $H = (H_1, \ldots, H_\ell)$
- ℓ-block message $B = (B_1, \ldots, B_\ell)$
- "$\otimes$" – multiplication in $GF(2^n)$

Collision probability

СФБ ЛАБ

Lemma

*Let $H_1,\ldots,H_\ell$ be sampled **with replacements** (random **function**) from $GF(2^n)$. The collision probability of two different ℓ-block messages A and B*

$$\Pr\left[\mathrm{MH}(H, A) = \mathrm{MH}(H, B)\right] = \frac{1}{2^n}.$$

Main observation

Lemma

*Let $H_1,\ldots,H_\ell$ be sampled **without replacements** (random **permutation**) from $GF(2^n)$. The collision probability of two different ℓ-block messages A and B*

$$\Pr\left[\mathrm{MH}(H, A) = \mathrm{MH}(H, B)\right] \leq \frac{1}{2^n - \ell + 1}.$$

Main observation

Lemma

*Let $H_1,\ldots,H_\ell$ be sampled **without replacements** (random **permutation**) from $GF(2^n)$. The collision probability of two different ℓ-block messages A and B*

$$\Pr\left[\mathrm{MH}(H, A) = \mathrm{MH}(H, B)\right] \leq \frac{1}{2^n - \ell + 1}.$$

Only slightly more than in the previous case!

Main observation

СФБ ЛАБ

Lemma

*Let $H_1,\ldots,H_\ell$ be sampled **without replacements** (random **permutation**) from $GF(2^n)$. The collision probability of two different ℓ-block messages A and B*

$$\Pr\left[\mathrm{MH}(H, A) = \mathrm{MH}(H, B)\right] \leq \frac{1}{2^n - \ell + 1}.$$

Only slightly more than in the previous case!

Similarly with preimages, “special” collisions, etc.

Corollary

- H_1, ..., H_ℓ can be generated by a random **permutation**, rather than a random *function* with the same security
- “PRP-PRF Switching Lemma” and its analogues are NOT needed for for proofs

MGM-PRF

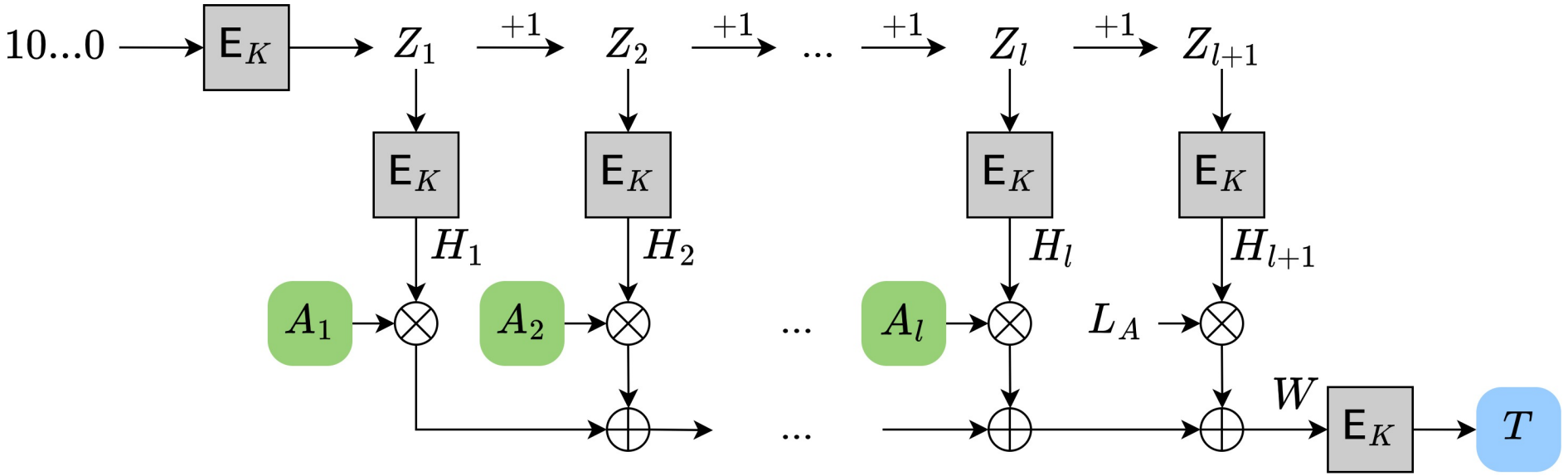
MGM-PRF

Non-standard use of the standardized AEAD-mode MGM:

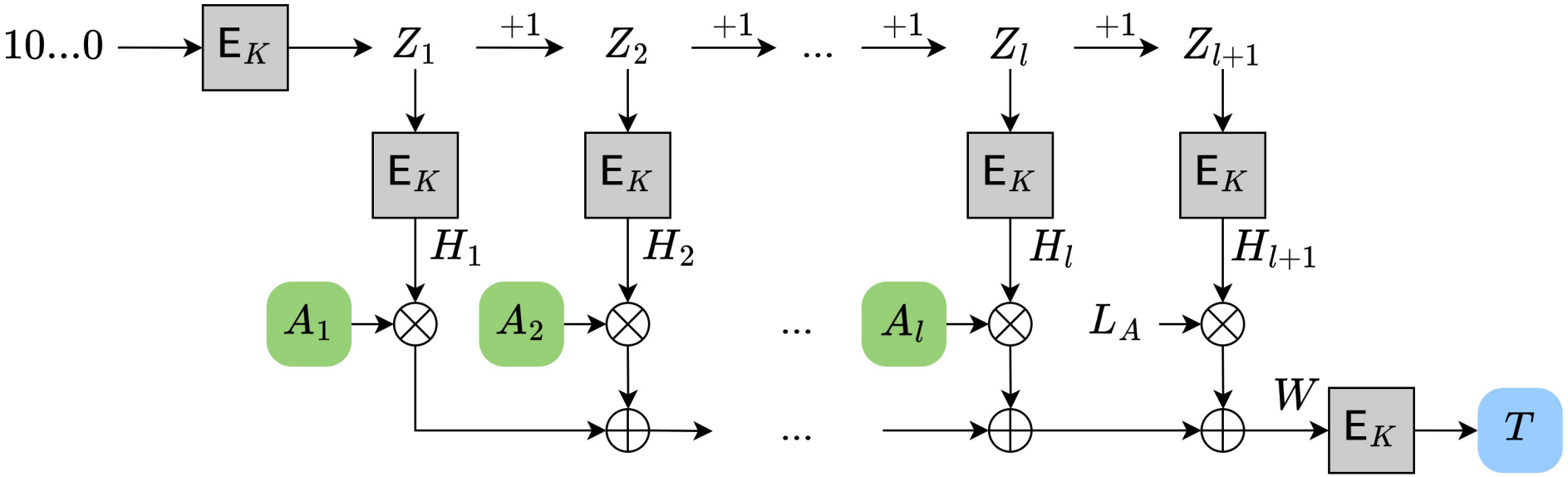
$$\text{MGM-PRF}(K, A) = \text{MGM}(K, 0^{n-1}, A, \varnothing) = (\varnothing, T)$$

- nonce is zero **constant**
- NO encryption/decryption, $P = C = \varnothing$
- $H = (H_1, ..., H_\ell)$ can be **precomputed**

MGM-PRF



MGM-PRF



Performance:

		precomputations
E_K	$\ell + 3$	1
$\otimes$	$\ell + 1$	$\ell + 1$

MGM-PRF

Theorem

$$\mathrm{Adv}^{PRF}_{\text{MGM-PRF}}(t, q, \ell) \le \mathrm{Adv}^{PRP}_{\mathsf{E}}(t', q') + \frac{q^2 + q\ell' + 2q + 2\ell'}{2^n},$$

$t' \approx t,\ q' = q + \ell + 1,\ \ell' = \ell + 1.$

Example: Magma, $n = 64$**,** $\pi_{\text{mac}} = 2^{-10}$

$q \le 2^{26}$ messages,

$\ell \le 2^{26}$ maximum blocks per message,

$\sigma \le 2^{52}$ blocks in total

Attacks on MGM-PRF

- As with most PRFs, cipher weaknesses cannot be exploited before a collision. All PT-CT pairs are initially unknown.
- A collision attack with additional adaptive query, $p_1 \leq q^2/2^{n+1}$.
- Collision = linear equation on $H_1,\ldots,H_\ell$. After *two* collisions, the pair (H_i, H_j) can be disclosed, $p_2 \leq (q^2/2^{n+1})^2$.

Attacks on MGM-PRF

- As with most PRFs, cipher weaknesses cannot be exploited before a collision. All PT-CT pairs are initially unknown.
- A collision attack with additional adaptive query, $p_1 \leq q^2/2^{n+1}$.
- Collision = linear equation on $H_1,\ldots,H_\ell$. After *two* collisions, the pair (H_i, H_j) can be disclosed, $p_2 \leq (q^2/2^{n+1})^2$.
- **CMAC**: one collision among known messages leads to a forgery and a partial key recovery (K^*).

MGM-PRF vs CMAC

	MGM-PRF	**CMAC**
$\mathrm{Adv}^{PRF} \lessapprox \frac{\cdots}{2^n}$	$q^2 + q\ell$	$16q^2 + q\ell^2 + 4q\ell$
key capacity	$q \le 2^{26}, \ell \le 2^{26}$	$q \le 2^{24}, \ell \le 2^{14}$
parallel.	Yes	No
precomp.	Yes	No
incremental	Yes	No
"$\otimes$"	$\ell + 1$	0
"E"	$(\ell + 3)$ or 1	ℓ
attacks via one collision	forgery	forgery and K^* recovery

MGM2-PRF

MGM2-PRF

Let's build a PRF from MGM2 in a similar way:

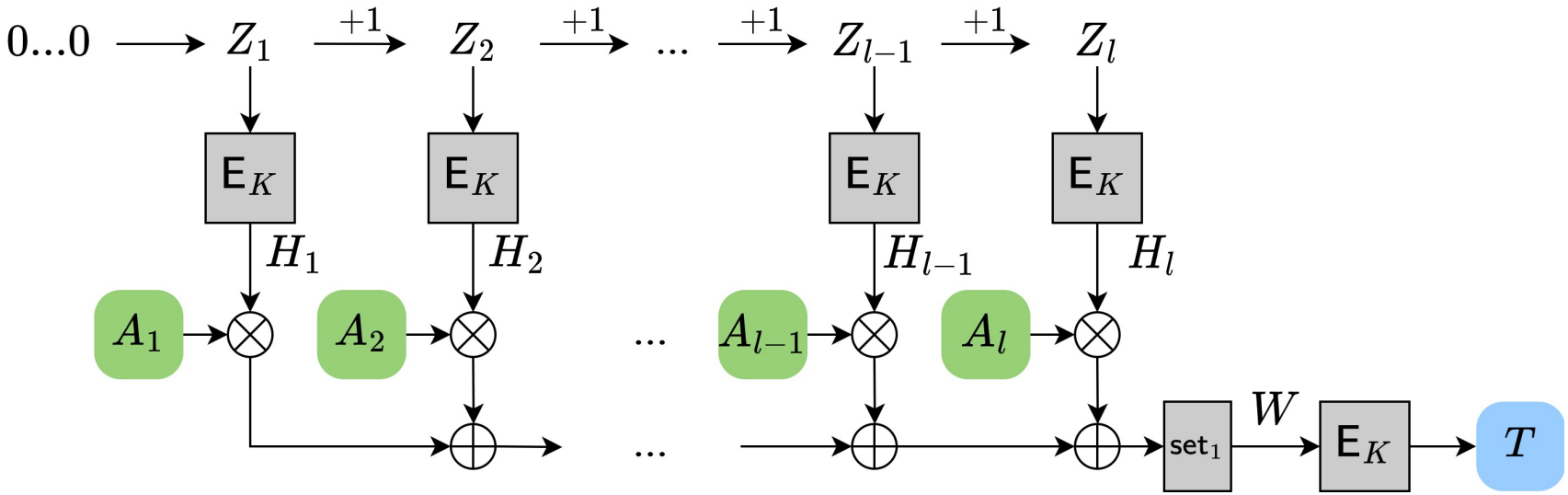
- fix the nonce to a constant
- remove encryption
- slightly change padding and flags
  - at least one bit 10...0 instead of optional 0...0 and $L_A||L_P$
  - nonzero bit is inserted into the most significant position

Akhmetzyanova L., Alekseev E., Babueva A., Bozhko A., Smyshlyaev S.

Misuse-resistant MGM2 mode

CTCrypt 2021

MGM2-PRF



- MSB of counters $Z_1,\ldots,Z_\ell$ is always **zero**
- MSB of hash value W is always **one** due to set$_1$
- Guaranteed: NO collisions between $Z_1,\ldots,Z_\ell$ and W
  - Adv^{PRF} is almost independent of the message length (ℓ)

MGM2-PRF

Theorem

$$\mathrm{Adv}^{PRF}_{\text{MGM2-PRF}}(t, q, \ell) \le \mathrm{Adv}^{PRP}_{\mathsf{E}}(t', q + \ell) + \frac{5q^2 - 5q}{2^{n+1}},$$

$t' \approx t,\ q + \ell \le 2^{n-1}$.

Example: Magma, $n = 64$**,** $\pi_{\text{mac}} = 2^{-10}$

$q \le 2^{26}$ messages,

$\ell < 2^{63}$ maximum blocks per message,

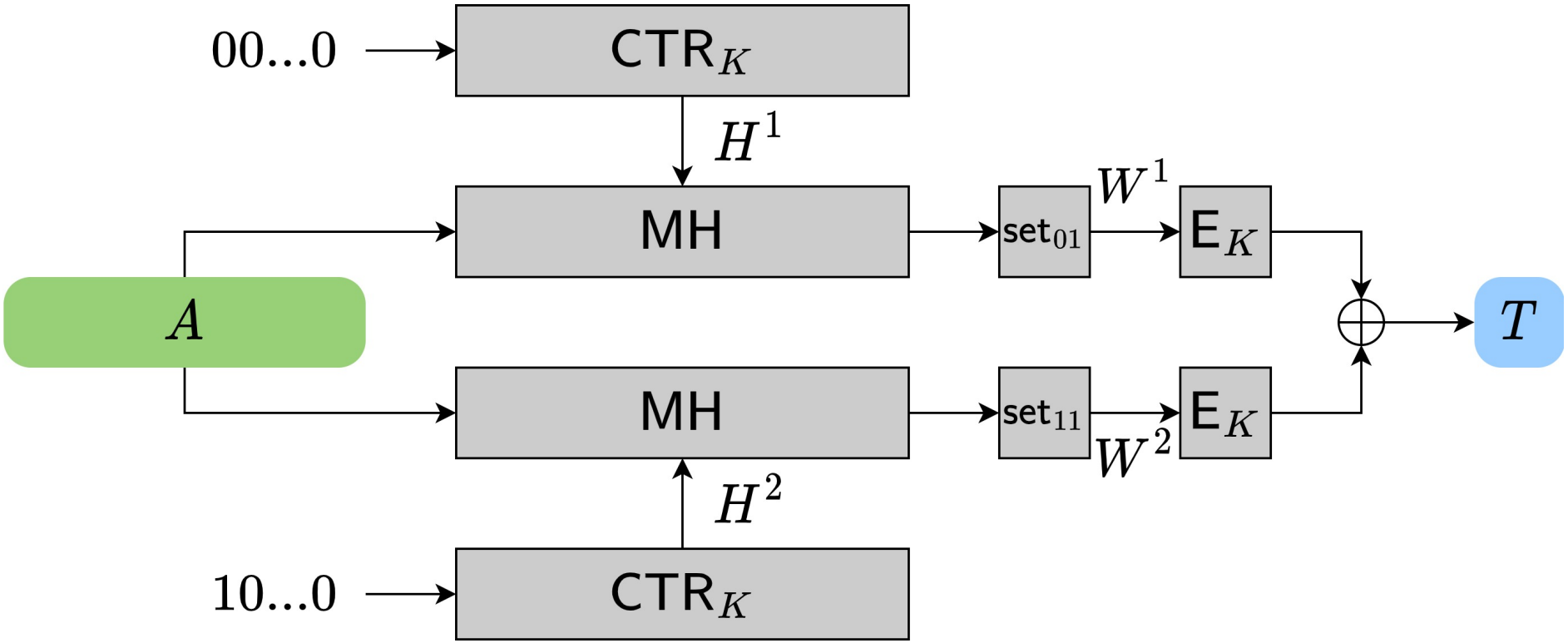
$\sigma < 2^{89}$ blocks in total

SUM-MGM

Let's go beyond the birthday bound (BBB)!

SUM-MGM

Double **b**lock – **H**ash - **t**hen - **S**um (DbHtS)



SUM-MGM: Properties

1. Single encryption key, parallel, online
2. Domain separation
   - 00 and 10 for counters,
   - 01 and 11 for hash values
3. Message is always padded with 10...0
4. Hash keys H^0 and H^1 can be **precomputed**
5. Performance:

		precomputations
E_K	$2\ell + 2$	2
$\otimes$	2ℓ	2ℓ

SUM-MGM

СФБ ЛАБ

Theorem

$$\mathrm{Adv}^{PRF}_{\text{SUM-MGM}}(t, q, \ell) \le \mathrm{Adv}^{PRP}_{\mathrm{E}}(t', q') + \frac{37q^3 + 5q\ell^2 + 22q^2\ell}{2^{2n}} + \frac{q}{2^n},$$

$t' \approx t$, $q' = (2q + 2\ell) \le 2^{n-4}$.

Example: Magma, $n = 64$, $\pi_{\text{mac}} = 2^{-10}$

$q \le$	2^{37}	messages,
$\ell \le$	2^{38}	maximum blocks per message,
$\sigma \le$	2^{75}	blocks in total

Attacks on SUM-MGM

- A collision between tags does NOT lead to forgery!
- The attack can be mounted using the “zero sum” property

“Zero sum“ attack

- Find four messages X, Y, Z, W with $(T_X \oplus T_Y \oplus T_Z \oplus T_W) = 0$ and two more similar conditions ($3n$-bit filter) by using generalized BP
- The success probability $\approx q^4/2^{3n}$

Leurent G., Nandi M., Sibleyras F. – CRYPTO 2018

Generic Attacks against Beyond-Birthday-Bound MACs

“Zero sum“ attack

- Find four messages X, Y, Z, W with $(T_X \oplus T_Y \oplus T_Z \oplus T_W) = 0$ and two more similar conditions ($3n$-bit filter) by using generalized BP
- The success probability $\approx q^4/2^{3n}$
- The key capacity “by the best known attack” is $q \leq 2^{45}$ messages
- The provable security bound is slightly untight

Leurent G., Nandi M., Sibleyras F. – CRYPTO 2018

Generic Attacks against Beyond-Birthday-Bound MACs

Comparison with other BBB-PRFs

Many BBB-PRFs exist:

- SUM-ECBC, 3kf9, PMAC+, LightMAC+, mPMAC+, mLightMAC+, etc

Comparison with other BBB-PRFs

Many BBB-PRFs exist:

- SUM-ECBC, 3kf9, PMAC+, LightMAC+, mPMAC+, mLightMAC+, etc

Advantages of **SUM-MGM**:

- Single key
- Effective precomputation
- Good key capacity “in messages”
- One of the best key capacity “in blocks”

Comparison with other BBB-PRFs

Many BBB-PRFs exist:

- SUM-ECBC, 3kf9, PMAC+, LightMAC+, mPMAC+, mLightMAC+, etc

Advantages of **SUM-MGM**:

- Single key
- Effective precomputation
- Good key capacity “in messages”
- One of the best key capacity “in blocks”

Disadvantages of **SUM-MGM**:

- multiplication by arbitrary values instead of “$\otimes$ 2”

Conclusion

- Multilinear hash is awesome, even if the coefficients are generated by a random *permutation* rather than a function
- We propose effective and secure PRF
  that are based on MH and n-bit block cipher:
  - MGM-PRF
  - MGM2-PRF
  - SUM-MGM

Conclusion

- MGM-PRF is non-standard use of the standardized MGM:
  - Nonce is fixed, no encryption
  - Better than CMAC in many cryptographic and operational properties

Conclusion

- MGM-PRF is non-standard use of the standardized MGM:
  - Nonce is fixed, no encryption
  - Better than CMAC in many cryptographic and operational properties
- MGM2-PRF:
  - technically slightly better than MGM-PRF
  - secure even with very long 2^n-block messages

Conclusion

- MGM-PRF is non-standard use of the standardized MGM:
  - Nonce is fixed, no encryption
  - Better than CMAC in many cryptographic and operational properties
- MGM2-PRF:
  - technically slightly better than MGM-PRF
  - secure even with very long 2^n-block messages
- SUM-MGM is a new “beyond the birthday bound” PRF:
  - “double“ MGM2-PRF
  - key capacity is $\approx 2^{\frac{2}{3}n}$ messages, totally $\approx 2^{\frac{4}{3}n}$ blocks
  - effective precomputation

Conclusion

- MGM-PRF is non-standard use of the standardized MGM:
  - Nonce is fixed, no encryption
  - Better than CMAC in many cryptographic and operational properties
- MGM2-PRF:
  - technically slightly better than MGM-PRF
  - secure even with very long 2^n-block messages
- SUM-MGM is a new “beyond the birthday bound” PRF:
  - “double“ MGM2-PRF
  - key capacity is $\approx 2^{\frac{2}{3}n}$ messages, totally $\approx 2^{\frac{4}{3}n}$ blocks
  - effective precomputation
- We will gratefully accept verification of the proposed proofs or the disproofs (attacks)!

See in the next episode...

- MGM-like single-key AEAD-mode
- Compared to MGM:
  - Similar security properties (including MRAE-int)
  - The same performance in the general case
  - Twice as fast with precomputations
  - Much faster with precomputations in “MAC only” mode

See in the next episode...

- MGM-like single-key AEAD-mode
- Compared to MGM:
  - Similar security properties (including MRAE-int)
  - The same performance in the general case
  - Twice as fast with precomputations
  - Much faster with precomputations in “MAC only” mode
- Key capacity with Magma:

  $\approx 2^{45}$ **encrypted blocks** or

  $\approx 2^{45}$ **packets** in “MAC only” mode

Thank you for attention!
Questions?

Vitaly Kiryukhin

LLC “SFB Lab”, JSC “InfoTeCS”

CTCrypt 2024

4 June 2024

vitaly.kiryukhin@sfblaboratory.ru

Why do we need BBB-modes if there is an external re-keying?

Usually, the probability of a threat is computed for **one specific** key

$\Rightarrow$ an external re-keying solves all problems

Why do we need BBB-modes if there is an external re-keying?

СФБ ЛАБ

Usually, the probability of a threat is computed for **one specific** key

$\Rightarrow$ an external re-keying solves all problems

Multi-user (multi-key) setting:

- μ independent secret keys (users)
- Forgery/threat for **any** one of μ users
- The success probability **increases linearly** with μ

Why do we need BBB-modes if there is an external re-keying?

СФБ ЛАБ

Example:

- μ independent secret keys
- each key protects only $q = 16 = 2^4$ short packets
- total $\mu \cdot q$ packets
- the probability of forgery must be less than 2^{-10}

Why do we need BBB-modes if there is an external re-keying?

СФБ ЛАБ

Example:

- μ independent secret keys
- each key protects only $q = 16 = 2^4$ short packets
- total $\mu \cdot q$ packets
- the probability of forgery must be less than 2^{-10}

How many packets can be protected in total?

	$\mathrm{Adv}(q)$	μ	$\mu \cdot \mathrm{Adv}(q)$	$\mu \cdot q$
CMAC	2^{-51}	2^{41}	2^{-10}	2^{45}
SUM-MGM	2^{-110}	2^{100}		2^{104}

Why do we need BBB-modes if there is an external re-keying?

СФБ ЛАБ

⇒ If a multi-user (multi-key) threat model is used,

then we must apply at least one of the following:

- strict limitation of the total key capacity
- Kuznyechik
- BBB-secure cipher modes