# Message Authentication Codes (MACs)

#### Tung Chou

#### Technische Universiteit Eindhoven, The Netherlands

#### October 8, 2015

Tung Chou (Tony)

Tung Chou (Tony)

• Ph.D. student of Daniel J. Bernstein & Tanja Lange

Tung Chou (Tony)

- Ph.D. student of Daniel J. Bernstein & Tanja Lange
- Research topics: Post-quantum crypto, ECC, MAC design.

Tung Chou (Tony)

- Ph.D. student of Daniel J. Bernstein & Tanja Lange
- Research topics: Post-quantum crypto, ECC, MAC design.
- Email: t.chou@tue.nl

#### • Introduction

- Introduction
- HMAC

- Introduction
- HMAC
- Universal-hash based MACs
  - Poly1305
  - security issues
  - software implementation issues

- Introduction
- HMAC
- Universal-hash based MACs
  - Poly1305
  - security issues
  - software implementation issues

• Diffie-Hellman key exchange

# What are MACs?

#### What are MACs?

• On Wikipedia:

"a message authentication code (often MAC) is a short piece of information used to authenticate a message and to provide integrity and authenticity assurances on the message. Integrity assurances detect accidental and intentional message changes, while authenticity assurances affirm the message's origin"

• Construction:



• Construction:



• Usage:

• Construction:



- S computes h and the SIGN<sub>sk</sub>(h).
- S sends (m, s).

• Construction:



- S computes h and the SIGN<sub>sk</sub>(h).
- S sends (m, s).
- V gets (m', s').

• Construction:



- S computes h and the SIGN<sub>sk</sub>(h).
- S sends (m, s).
- V gets (m', s').
- V computes and check  $hash(m') = VERIFY_{pk}(s')$ .

• Construction:



- S computes h and the SIGN<sub>sk</sub>(h).
- S sends (m, s).
- V gets (m', s').
- V computes and check  $hash(m') = VERIFY_{pk}(s')$ .
- Security

• Construction:



- S computes h and the SIGN<sub>sk</sub>(h).
- S sends (m, s).
- V gets (m', s').
- V computes and check  $hash(m') = VERIFY_{pk}(s')$ .
- Security
  - attacker should not be able to forge a valid  $\left(m,s\right)$  pair

• Construction:



- S computes h and the SIGN<sub>sk</sub>(h).
- S sends (m, s).
- V gets (m', s').
- V computes and check  $hash(m') = VERIFY_{pk}(s')$ .
- Security
  - attacker should not be able to forge a valid  $\left(m,s\right)$  pair
  - attacker might have collected many  $\left(m,s\right)$  pairs

• "Keyed hash function":

message 
$$(m) \longrightarrow MAC$$
 algorithm  $\longrightarrow tag/authenticator  $(t)$   
 $\uparrow$   
shared secret key  $(r)$$ 

• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
  
shared secret key (r)



• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
 $\uparrow$   
shared secret key (r)

Usage:

• S computes  $t = MAC_r(m)$  and sends (m, t).

• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
 $\uparrow$   
shared secret key (r)

- S computes  $t = MAC_r(m)$  and sends (m, t).
- R gets (m', t').

• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
  
shared secret key (r)

- S computes  $t = MAC_r(m)$  and sends (m, t).
- R gets (m', t').
- R computes and checks  $MAC_r(m') = t'$ .

• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
  
shared secret key (r)

- S computes  $t = MAC_r(m)$  and sends (m, t).
- R gets (m', t').
- R computes and checks  $MAC_r(m') = t'$ .
- Security

• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
  
shared secret key (r)

- S computes  $t = MAC_r(m)$  and sends (m, t).
- R gets (m', t').
- R computes and checks  $MAC_r(m') = t'$ .
- Security
  - attacker should not be able to forge a valid (m, t) pair

• "Keyed hash function":

message (m) 
$$\longrightarrow$$
 MAC algorithm  $\longrightarrow$  tag/authenticator (t)  
 $\uparrow$   
shared secret key (r)

- S computes  $t = MAC_r(m)$  and sends (m, t).
- R gets (m', t').
- R computes and checks  $MAC_r(m') = t'$ .
- Security
  - attacker should not be able to forge a valid (m,t) pair
  - attacker might have collected many (m,t) pairs

	MACs	Signatures
Integrity	yes	yes
Authenticity	yes	yes
Non-repudiation	no	yes
Key	secret-key	public-key

	MACs	Signatures
Integrity	yes	yes
Authenticity	yes	yes
Non-repudiation	no	yes
Key	secret-key	public-key

"Non-repudiation is about Alice showing to Bob a proof that some data really comes from Alice, such that not only Bob is convinced, but Bob also gets the assurance that he could show the same proof to Charlie, and Charlie would be convinced, too"

	MACs	Signatures
Integrity	yes	yes
Authenticity	yes	yes
Non-repudiation	no	yes
Key	secret-key	public-key

"Non-repudiation is about Alice showing to Bob a proof that some data really comes from Alice, such that not only Bob is convinced, but Bob also gets the assurance that he could show the same proof to Charlie, and Charlie would be convinced, too"

secret-key crypto is "fast"

#### HMAC
• Build MAC from hash functions

- Build MAC from hash functions
- A naive construction:

 $t = H(r \mid\mid m)$ 

- Build MAC from hash functions
- A naive construction:

$$t = H(r \mid\mid m)$$

 Merkle–Damgård construction based hashes (e.g., MD5, SHA1)



- Build MAC from hash functions
- A naive construction:

$$t = H(r \mid\mid m)$$

 Merkle–Damgård construction based hashes (e.g., MD5, SHA1)



• Length extension attack:  $h' = f(h, m_{\ell+1})$ 

• Another construction:

$$t = H(m \mid\mid r)$$

• Another construction:

$$t = H(m \mid\mid r)$$

$$t = H\left((r \oplus p_o) || H((r \oplus p_i) || m)\right)$$

• Another construction:

$$t = H(m \mid\mid r)$$

$$t = H\left((r \oplus p_o)||H((r \oplus p_i)||m)\right)$$

#### • HMAC-SHA1

• Another construction:

$$t = H(m \mid\mid r)$$

• HMAC:

$$t = H\left((r \oplus p_o)||H((r \oplus p_i)||m)\right)$$

- HMAC-SHA1
  - widely used in Internet applications

• Another construction:

$$t = H(m \mid\mid r)$$

• HMAC:

$$t = H\left((r \oplus p_o) || H((r \oplus p_i) || m)\right)$$

- HMAC-SHA1
  - widely used in Internet applications
  - 5.18 Sandy Bridge cycles/byte

Another construction:

$$t = H(m \mid\mid r)$$

• HMAC:

$$t = H\left((r \oplus p_o) || H((r \oplus p_i) || m)\right)$$

- HMAC-SHA1
  - widely used in Internet applications
  - 5.18 Sandy Bridge cycles/byte

Reality: the most commonly used scheme might not be the best

SHA3

The "Sponge" construction:



http://en.wikipedia.org/wiki/SHA-3

• Why?

• Why?

• provides information theoretic security

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic
  - better performance than HMAC

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic
  - better performance than HMAC

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic
  - better performance than HMAC

Construction

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic
  - better performance than HMAC

- Construction
  - "universal" hash function + one-time pad:

 $h_r(m_n) \oplus s_n$ 

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic
  - better performance than HMAC

- Construction
  - "universal" hash function + one-time pad:

 $h_r(m_n) \oplus s_n$ 

• universal hash: low differential probability

- Why?
  - provides information theoretic security
  - usually involves field/ring arithmetic
  - better performance than HMAC

- Construction
  - "universal" hash function + one-time pad:

 $h_r(m_n) \oplus s_n$ 

- universal hash: low differential probability
- one-time pad hides all information about the key

• Construction:

$$t = (((m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) \mod 2^{130} - 5) + s) \mod 2^{128}$$

• Construction:

$$t = (((m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) \mod 2^{130} - 5) + s) \mod 2^{128}$$

- $2^{130} 5$  is a prime
- r, s are shared secret 128-bit values

#### Construction:

$$t = (((m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) \mod 2^{130} - 5) + s) \mod 2^{128}$$

- r, s are shared secret 128-bit values
- $m_{i < \ell}$  is the *i*th 128-bit block of *m* padded by 1.
- $m_{\ell}$  is the "remainder" of m padded by 1.

#### Construction:

$$t = (((m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) \mod 2^{130} - 5) + s) \mod 2^{128}$$

- r, s are shared secret 128-bit values
- $m_{i < \ell}$  is the *i*th 128-bit block of *m* padded by 1.
- $m_\ell$  is the "remainder" of m padded by 1.
- Without proper padding?

#### Construction:

$$t = (((m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) \mod 2^{130} - 5) + s) \mod 2^{128}$$

- r, s are shared secret 128-bit values
- $m_{i < \ell}$  is the *i*th 128-bit block of *m* padded by 1.
- $m_{\ell}$  is the "remainder" of m padded by 1.
- Without proper padding?
  - m = 'FF', m' = 'FF', '00'
  - zero-pad the message obtain a 128-bit block

$$m_1=m_1'=$$
 'FF', '00',  $\ldots$ , '00'

#### Construction:

$$t = (((m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) \mod 2^{130} - 5) + s) \mod 2^{128}$$

- r, s are shared secret 128-bit values
- $m_{i < \ell}$  is the *i*th 128-bit block of *m* padded by 1.
- $m_{\ell}$  is the "remainder" of m padded by 1.
- Without proper padding?
  - m = 'FF', m' = 'FF', '00'
  - zero-pad the message obtain a 128-bit block

$$m_1 = m_1' =$$
 'FF', '00', ..., '00'

• What is wrong with "real" polynomial evaluation?

$$t = m_1 r^{\ell - 1} + m_2 r^{\ell - 2} + \dots + m_\ell + s$$

• What is wrong with "real" polynomial evaluation?

$$t = m_1 r^{\ell - 1} + m_2 r^{\ell - 2} + \dots + m_\ell + s$$

• The attacker forges a valid message-tag pair easily:

$$t + \Delta = m_1 r^{\ell - 1} + m_2 r^{\ell - 2} + \dots + (m_\ell + \Delta) + s$$

• What is wrong with "real" polynomial evaluation?

$$t = m_1 r^{\ell - 1} + m_2 r^{\ell - 2} + \dots + m_\ell + s$$

• The attacker forges a valid message-tag pair easily:

$$t + \Delta = m_1 r^{\ell - 1} + m_2 r^{\ell - 2} + \dots + (m_\ell + \Delta) + s$$

• This does not provide low differential probability

• What is wrong with using the same pad twice?

$$t = m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r + s$$

$$t' = m'_1 r^{\ell} + m'_2 r^{\ell-1} + \dots + m'_{\ell} r + s$$

• What is wrong with using the same pad twice?

$$t = m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r + s$$
  
$$t' = m'_1 r^{\ell} + m'_2 r^{\ell-1} + \dots + m'_{\ell} r + s$$

• The attacker gets information of r by finding roots of

$$t - t' = (m_1 - m_1')r^{\ell} + (m_2 - m_2')r^{\ell-1} + \dots + (m_\ell - m_\ell')r^{\ell-1}$$
### Poly1305: avoiding security issue

• What is wrong with using the same pad twice?

$$t = m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r + s$$
  
$$t' = m'_1 r^{\ell} + m'_2 r^{\ell-1} + \dots + m'_{\ell} r + s$$

• The attacker gets information of r by finding roots of

$$t - t' = (m_1 - m_1')r^{\ell} + (m_2 - m_2')r^{\ell-1} + \dots + (m_\ell - m_\ell')r^{\ell-1}$$

• "nonce-misuse" issue

### Poly1305: avoiding security issue

• What is wrong with using the same pad twice?

$$t = m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r + s$$
  
$$t' = m'_1 r^{\ell} + m'_2 r^{\ell-1} + \dots + m'_{\ell} r + s$$

• The attacker gets information of r by finding roots of

$$t - t' = (m_1 - m_1')r^{\ell} + (m_2 - m_2')r^{\ell-1} + \dots + (m_{\ell} - m_{\ell}')r^{\ell}$$

- "nonce-misuse" issue
  - In practice s is usually replaced by stream cipher output, e.g.,  ${\sf AES}_k(n)$  for  $m_n$

## Poly1305: avoiding security issue

• What is wrong with using the same pad twice?

$$t = m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r + s$$
$$t' = m'_1 r^{\ell} + m'_2 r^{\ell-1} + \dots + m'_{\ell} r + s$$

• The attacker gets information of r by finding roots of

$$t - t' = (m_1 - m_1')r^{\ell} + (m_2 - m_2')r^{\ell-1} + \dots + (m_\ell - m_\ell')r^{\ell-1}$$

- "nonce-misuse" issue
  - In practice s is usually replaced by stream cipher output, e.g.,  ${\sf AES}_k(n)$  for  $m_n$
  - HMAC does not use nonce

## Poly1305: polynomial evaluation Consider $m_1r^8 + m_2r^7 + \cdots + m_8r$

Consider  $m_1r^8 + m_2r^7 + \cdots + m_8r$ 

• Horner's rule:



Consider  $m_1r^8 + m_2r^7 + \cdots + m_8r$ 

• Horner's rule:



*n* multiplications (and *n* − 1 additions)

Consider  $m_1r^8 + m_2r^7 + \cdots + m_8r$ 

• Horner's rule:



- n multiplications (and n-1 additions)
- The issue of being "on-line"

• The NIST-standard authenticated encryption scheme GCM

- The NIST-standard authenticated encryption scheme GCM
  - Galois Counter Mode

- The NIST-standard authenticated encryption scheme GCM
  - Galois Counter Mode
  - Special hardware support for AES-GCM in high-end CPUs

- The NIST-standard authenticated encryption scheme GCM
  - Galois Counter Mode
  - Special hardware support for AES-GCM in high-end CPUs
- Polynomial evaluation MAC:

$$t = (m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) + s$$

- The NIST-standard authenticated encryption scheme GCM
  - Galois Counter Mode
  - Special hardware support for AES-GCM in high-end CPUs
- Polynomial evaluation MAC:

$$t = (m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) + s$$

• Based on arithmetic in

$$\mathbb{F}_{2^{128}} = \mathbb{F}_2[x] / (x^{128} + x^7 + x^2 + x + 1)$$

- The NIST-standard authenticated encryption scheme GCM
  - Galois Counter Mode
  - Special hardware support for AES-GCM in high-end CPUs
- Polynomial evaluation MAC:

$$t = (m_1 r^{\ell} + m_2 r^{\ell-1} + \dots + m_{\ell} r) + s$$

• Based on arithmetic in

$$\mathbb{F}_{2^{128}} = \mathbb{F}_2[x]/(x^{128} + x^7 + x^2 + x + 1)$$

#### Binary fields: better in hardware

## GCM



http://en.wikipedia.org/wiki/Galois/Counter\_Mode

# GMAC: speeds

reference	platform	PCLMUQDQ	cycles per byte
Käsper–Schwabe 2009	Core 2	no	14.40
	Sandy Bridge	no	13.10
Krovetz–Rogaway 2011	Westmere	yes	2.00
Gueron 2013	Sandy Bridge	yes	1.79
	Haswell	yes	0.40

### $Auth256^{*}$

#### • Construction

### • Construction

• a *pseudo-dot-product* MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

### Construction

• a pseudo-dot-product MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

- Construction
  - a pseudo-dot-product MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

- base field  $\mathbb{F}_{2^{256}} = \mathbb{F}_{2^8}[x]/(\phi)$ . Tower field construction for  $\mathbb{F}_{2^8}$ .
- Compared to GMAC

- Construction
  - a *pseudo-dot-product* MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

- base field  $\mathbb{F}_{2^{256}} = \mathbb{F}_{2^8}[x]/(\phi)$ . Tower field construction for  $\mathbb{F}_{2^8}$ .
- Compared to GMAC
  - higher security level

- Construction
  - a *pseudo-dot-product* MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

- Compared to GMAC
  - · higher security level
  - 0.5/1 multiplications per block

- Construction
  - a pseudo-dot-product MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

- Compared to GMAC
  - higher security level
  - 0.5/1 multiplications per block
  - larger key size

- Construction
  - a pseudo-dot-product MAC:

$$t = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + s$$

- Compared to GMAC
  - higher security level
  - 0.5/1 multiplications per block
  - larger key size
  - very different field construction for low bit operation count

Wegman-Carter construction: security

- " $\delta\text{-xor-universal hash"}$  : For all distinct (m,m') and  $\Delta,$  we have

$$Pr\left(\mathsf{Hash}_r(m) = \mathsf{Hash}_r(m') \oplus \Delta\right) \le \delta$$

## Wegman-Carter construction: security

- " $\delta\text{-xor-universal hash"}$  : For all distinct (m,m') and  $\Delta,$  we have

$$Pr\left(\mathsf{Hash}_r(m) = \mathsf{Hash}_r(m') \oplus \Delta\right) \leq \delta$$

- The one-time pad hides all information about the key r.
- The best strategy for the attacker is to guess.

## Auth256: Security Proof

Hash values:

$$h = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + (m_{2\ell-1} + r_{2\ell-1})(m_{2\ell} + r_{2\ell}),$$
  
$$h' = (m'_1 + r_1)(m'_2 + r_2) + (m'_3 + r_3)(m'_4 + r_4) + \dots + (m'_{2\ell-1} + r_{2\ell-1})(m'_{2\ell} + r_{2\ell}).$$

## Auth256: Security Proof

#### Hash values:

$$h = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + (m_{2\ell-1} + r_{2\ell-1})(m_{2\ell} + r_{2\ell}),$$
  
$$h' = (m'_1 + r_1)(m'_2 + r_2) + (m'_3 + r_3)(m'_4 + r_4) + \dots + (m'_{2\ell-1} + r_{2\ell-1})(m'_{2\ell} + r_{2\ell}).$$

Then  $h = h' + \Delta$  if and only if  $r_1(m_2 - m'_2) + r_2(m_1 - m'_1) + r_3(m_4 - m'_4) + r_4(m_3 - m'_3) + \cdots$  $= \Delta + m'_1m'_2 - m_1m_2 + m'_3m'_4 - m_3m_4 + \cdots$ 

### Auth256: Security Proof

#### Hash values:

$$h = (m_1 + r_1)(m_2 + r_2) + (m_3 + r_3)(m_4 + r_4) + \dots + (m_{2\ell-1} + r_{2\ell-1})(m_{2\ell} + r_{2\ell}),$$
  
$$h' = (m'_1 + r_1)(m'_2 + r_2) + (m'_3 + r_3)(m'_4 + r_4) + \dots + (m'_{2\ell-1} + r_{2\ell-1})(m'_{2\ell} + r_{2\ell}).$$

Then 
$$h = h' + \Delta$$
 if and only if  
 $r_1(m_2 - m'_2) + r_2(m_1 - m'_1) + r_3(m_4 - m'_4) + r_4(m_3 - m'_3) + \cdots$   
 $= \Delta + m'_1m'_2 - m_1m_2 + m'_3m'_4 - m_3m_4 + \cdots$ 

 $m\neq m'$  implies that there are at most  $|K|^{2\ell-1}$  solutions for r.

### CBC-MAC



http://en.wikipedia.org/wiki/CBC-MAC