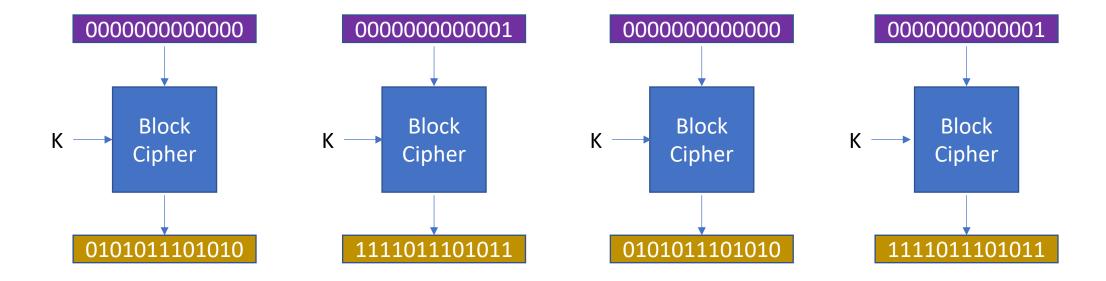
CS 370 Introduction to Security

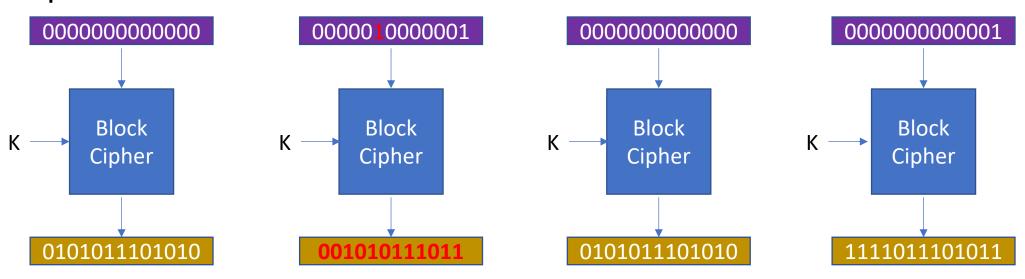
Message Authentication Code and Asymmetric Encryption
Yeongjin Jang



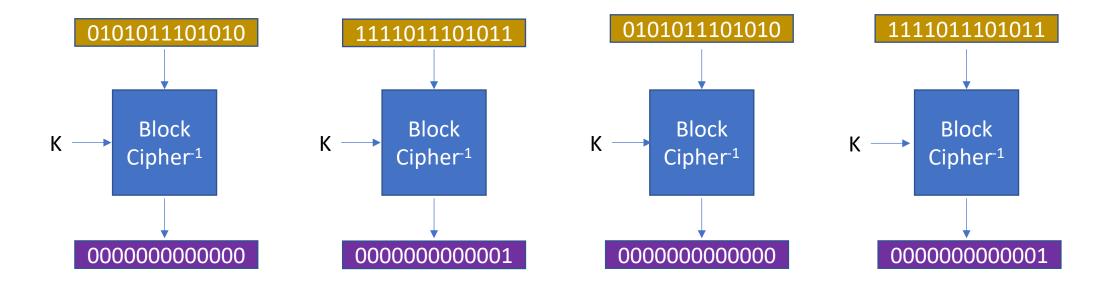
We can run encryption in parallel



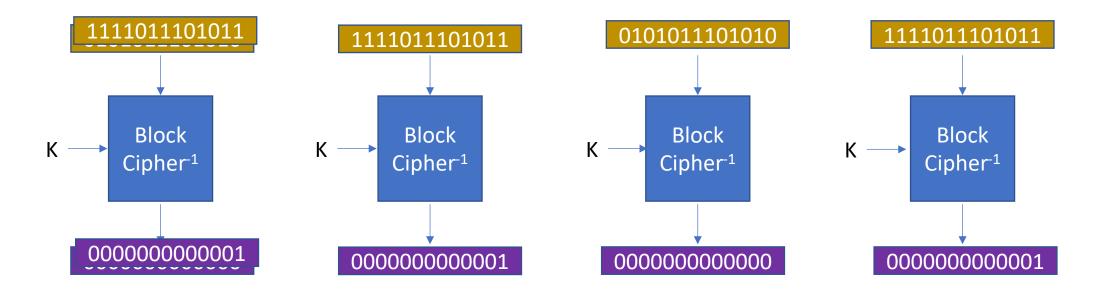
 A specific bit error in ciphertext would result in a random error in plaintext



We can launch a message block substitution attack

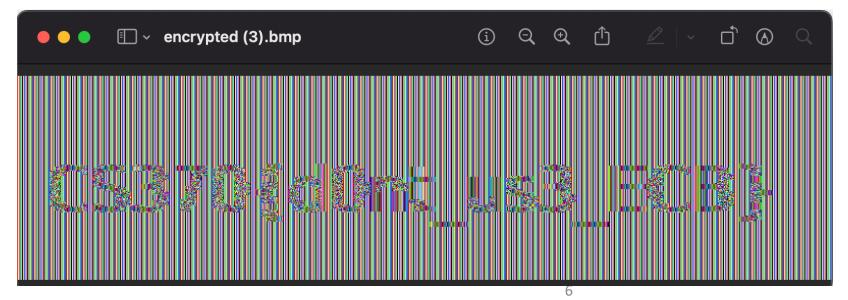


We can launch a message block substitution attack



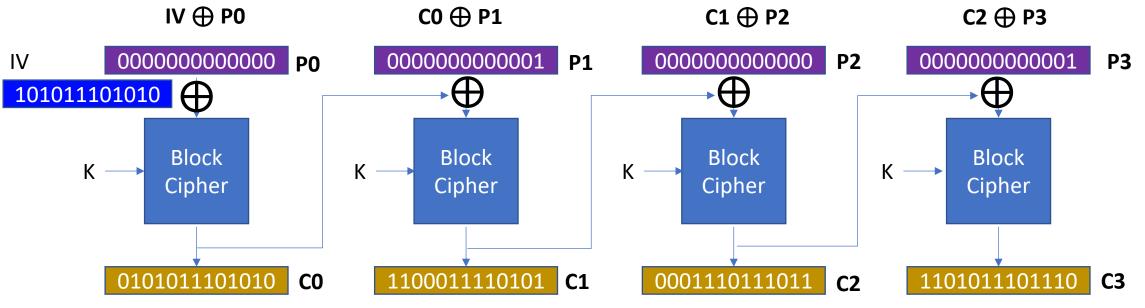
- Can encrypt in parallel
- Can decrypt in parallel

Ciphertext block leaks plaintext block patterns



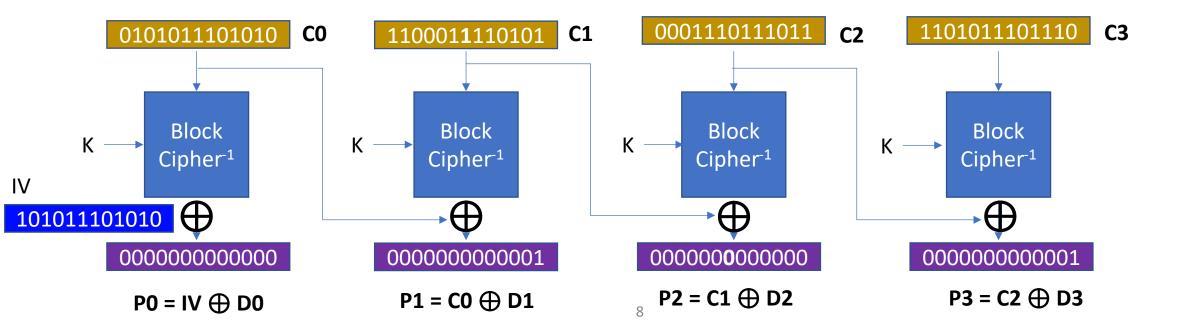
Recap: Cipher Block Chain (CBC)

- Apply XOR between the IV (Initialization Vector) and the plaintext
- Chain the previous ciphertext block to the plaintext with XOR
- Run Encryption on Xor'ed data



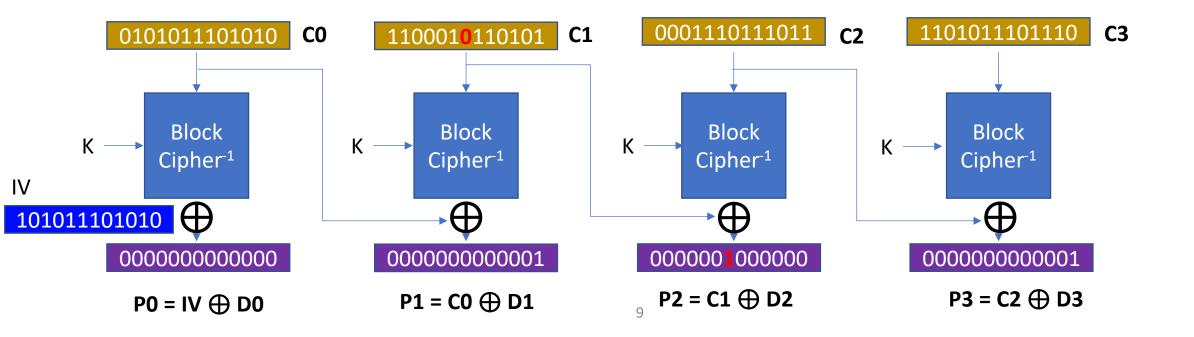
Recap: CBC Attack

 A specific bit error in cipertext-n will be a specific bit error in plaintext-(n+1)



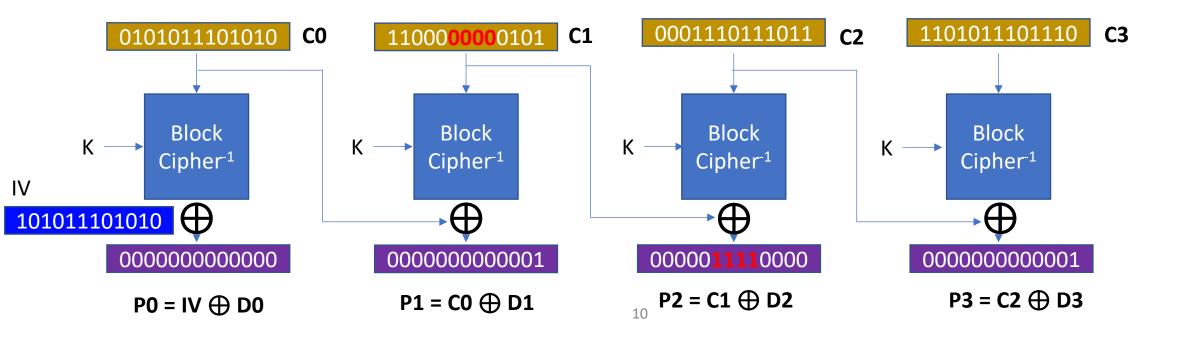
Recap: CBC Attack

 A specific bit error in cipertext-n will be a specific bit error in plaintext-(n+1)



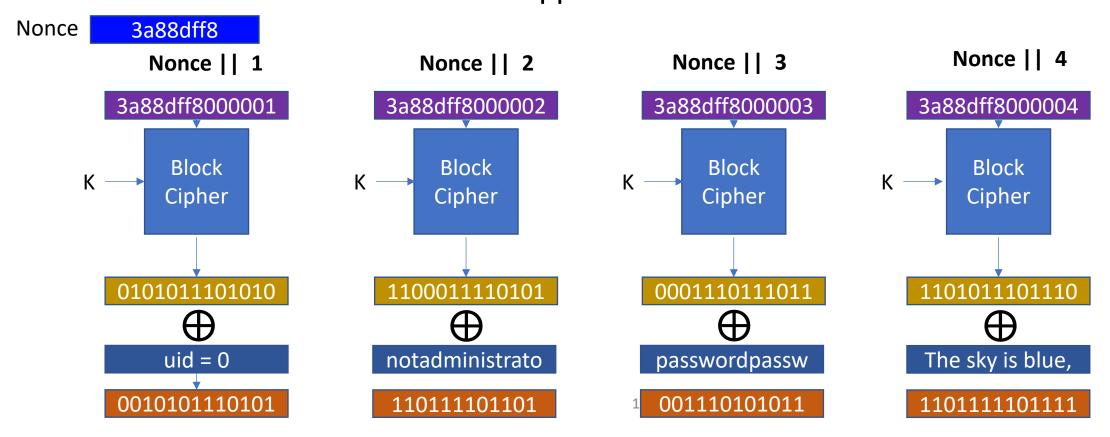
Recap: CBC Attack

 A specific bit error in cipertext-n will be a specific bit error in plaintext-(n+1)



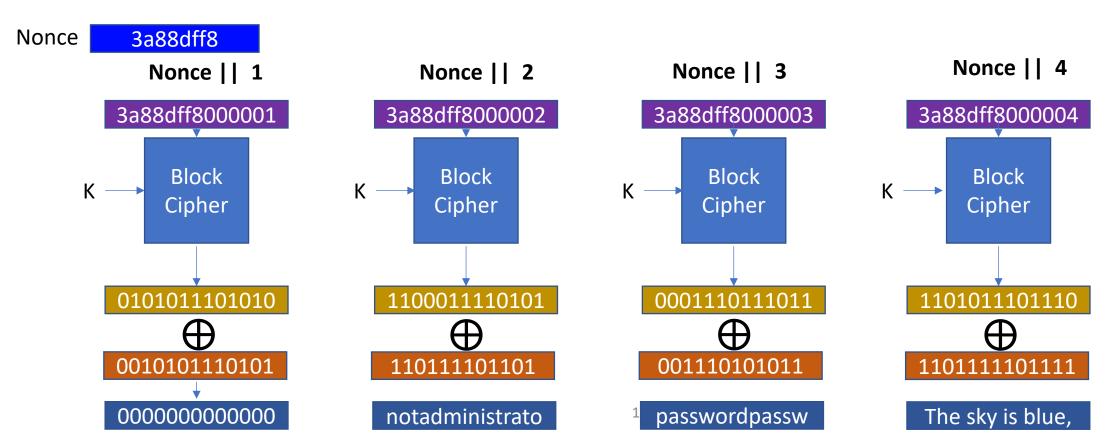
Recap: Counter Mode (CTR)

- CTR (Counter mode)
- Start with a random nonce || counter



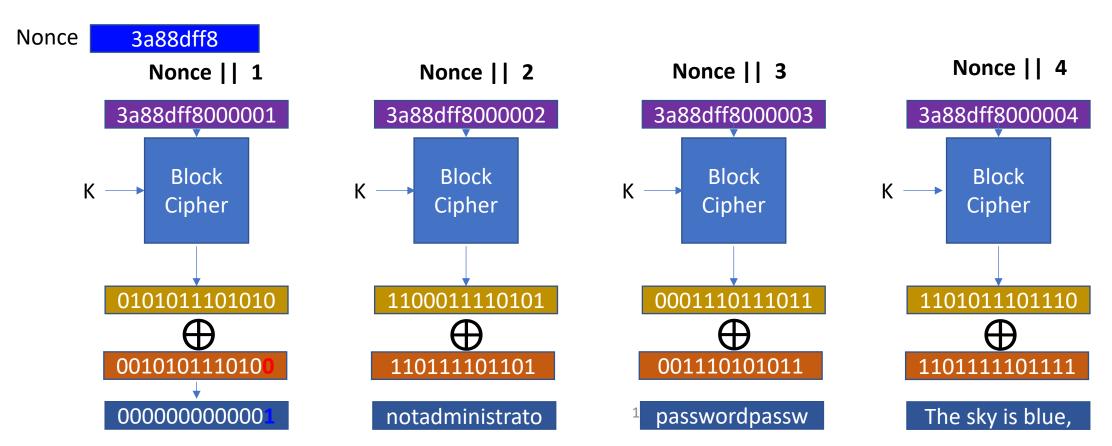
Recap: CTR Attack

 A specific bit error in cipertext-n will be a specific bit error in plaintext-n



Recap: CTR Attack

 A specific bit error in cipertext-n will be a specific bit error in plaintext-n



Block Cipher

- The block cipher itself cannot protect encrypted data modified by attackers
 - ECB, we can substitute blocks to known-plaintext-encrypted-block
 - CBC, we can apply XOR to the ciphertext that is one-block before the plaintext
 - CTR, we can apply XOR to the ciphertext then the result will show on the plaintext
- Why?
 - Block Cipher gives us data confidentiality
 - Not data integrity

ecb-, cbc-, ctr-attack

- Threat model
 - We have a verifier with the secret key
 - We have an encrypted user data, encrypted.user
 - We don't have the key -> cannot arbitrarily generate valid encrypted file
- What can we do?
 - Edit the ciphertext in encrypted.user
- What can we achieve?
 - Change the plaintext at our own will

ecb-, cbc-, ctr-attack

- Threat model
 - We have a verifier with the secret key
 - We have an encrypted user data, encrypted.user
 - We don't have the key
- What can we do?
 - Edit the ciphertext in encrypted.user <- dangerous
- What can we achieve?
 - Change the plaintext at our own will

Is there any cryptographic tool that we can check the integrity of data?

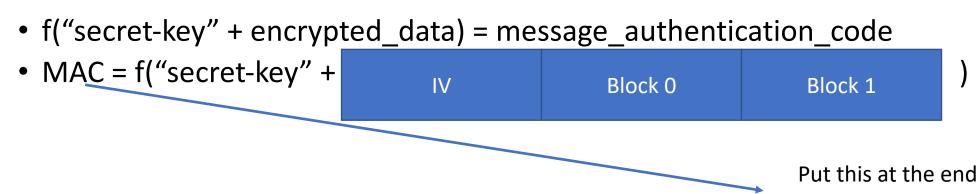
- What we have
 - Encrypted data, e.g.,
 - encrypted.user
- What we want
 - Detect if the attacker has modified the file encrypted.user
- What tool can we use?
 - Cryptographic hash!

An Ideal Hash Function

- Suppose we have a function f(x) that
 - Generates a fixed length message (e.g., 32-byte)
 - You cannot get easily get f⁻¹(y) = x from y
 - A slight value change in x for f(x) will result in drastic change in y
 - So you cannot correlate any f(x') = y' from f(x) = y
- What I will do:
 - f("secret-key" + encrypted_data) = message_authentication_code

Create a MAC (Message Authentication Code)

What I will do:



IV Block 0 Block 1 MAC

Checking a MAC

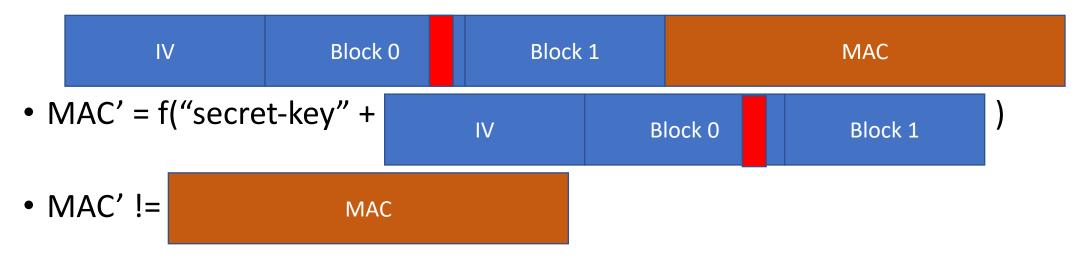
When I read the encrypted data



- MAC = f("secret-key" + IV Block 0 Block 1)
- MAC should be equal to MAC

What Attackers Can Do?

What if they edited data?



- Suppose we have a function f(x) that
 - A slight value change in x for f(x) will result in drastic change in y

They Can Generate Valid MAC if They know the key

What if they edited data?



- But they don't know the secret key; can't generate it
- Suppose we have a function f(x) that
 - You cannot get easily get f⁻¹(y) = x from y
 - A slight value change in x for f(x) will result in drastic change in y

Cryptographic Hash

- A hash function that generates a fingerprint of a data
- SHA256('Hello, world') = 03675ac53ff9cd1535ccc7dfcdfa2c458c5218371f418dc136f2d19ac1fbe8a5

- With characteristics of:
 - One-way function
 - Hard to find x for given y where H(x) = y
 - Hard to find x' for given x,y where x != x', H(x) = y and H(x') = y

SHA256

- Secure Hash Algorithm (SHA)
 - SHA256 is in the SHA2 standard
 - Input can be any length data
 - Output is 256-bit, 32-byte
- SHA256 is a cryptographic hash function that
 - It is one-way function
 - SHA256('Hello, world') = 03675ac53ff9cd1535ccc7dfcdfa2c458c5218371f418dc136f2 d19ac1fbe8a5
 - SHA256⁻¹(03675ac53ff9cd1535ccc7dfcdfa2c458c5218371f418 dc136f2d19ac1fbe8a5) == ???? there could be many..

SHA256 Examples

```
-blue9057@blue9057-vm-ctf1 ~/t <ruby-head>
  $ sha256sum *
9a271f2a916b0b6ee6cecb2426f0b3206ef074578be55d9bc94f6f3fe3ab86aa
4355a46b19d348dc2f57c046f8ef63d4538ebb936000f3c9ee954a27460dd865
53c234e5e8472b6ac51c1ae1cab3fe06fad053beb8ebfd8977b010655bfdd3c3
1121cfccd5913f0a63fec40a6ffd44ea64f9dc135c66634ba001d10bcf4302a2
7de1555df0c2700329e815b93b32c571c3ea54dc967b89e81ab73b9972b72d1d
f0b5c2c2211c8d67ed15e75e656c7862d086e9245420892a7de62cd9ec582a06
87428fc522803d31065e7bce3cf03fe475096631e5e07bbd7a0fde60c4cf25c7
0263829989b6fd954f72baaf2fc64bc2e2f01d692d4de72986ea808f6e99813f
a3a5e715f0cc574a73c3f9bebb6bc24f32ffd5b67b387244c2c909da779a1478
8d74beec1be996322ad76813bafb92d40839895d6dd7ee808b17ca201eac98be
a2bbdb2de53523b8099b37013f251546f3d65dbe7a0774fa41af0a4176992fd4
```

SHA256

- SHA256 is a cryptographic hash function that
 - Hard to find x for given y where H(x) = y
 - Find SHA256(x) for
 - This task requires around the 2²⁵⁶ times of search...
- Implication
 - If we know X, it is easy to get SHA256(X) = Y
 - But if we don't know X, even if we know Y, it is hard to calculate X

SHA256

- SHA256 is a cryptographic hash function that
 - Hard to find x' for given x,y where x' != x, H(x) = y, H(x') = H(x)
 - SHA256('Hello, world') = 03675ac53ff9cd1535ccc7dfcdfa2c458c5218371f418dc136f2d19ac 1fbe8a5
 - Can you find another x' that produces SHA256(x') = 03675ac53ff9cd1535ccc7dfcdfa2c458c5218371f418dc136f2d19ac1 fbe8a5
 - Other than 'Hello, world'?
- Implication
 - Even if we know X, Y where SHA256(X) = Y
 - It is hard to find SHA256(X') = Y

Avalanche Effect

 Even with a slight change in input, we want to have a huge change in output to not making attackers correlate output values based on their

inputs...

Cryptographic Hash: Implications

- SHA256 is a cryptographic hash that is included in the SHA2 standard
- SHA256 is a one-way function and
- It is hard to calculate x from y
 - where y = SHA256(x)
- It is hard to calculate x' from x,y
 - where x' != x, SHA256(x) = y, SHA256(x') = y
- It is hard to correlate x and x' from x, y, y'
 - where SHA256(x) = y, SHA256(x') = y'

How can we use this?

- Hash-based Message Authentication Code (HMAC)
 - H = a hash function (e.g., SHA256)
 - HMAC = H(H(K) | | M)
 - K: secret key
 - H(K): hash of the key
 - M: message or data

Encrypt Data with CBC

CBC Data (32-byte blocks)

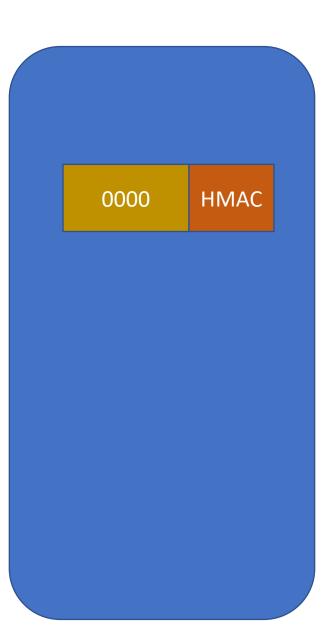
IV	Block 0	Block 1
----	---------	---------

- Suppose you have a hash key = 'asdf'
 - HMAC = SHA256(SHA256('asdf') || encrypted_data)
 - = 7624e1f89ce009f8ec7e6e39781a42c0a27fa38f94db4f05f78b0f301007e06a

IV	Block 0	Block 1	HMAC (key IV+Block0+Block1)
----	---------	---------	--------------------------------

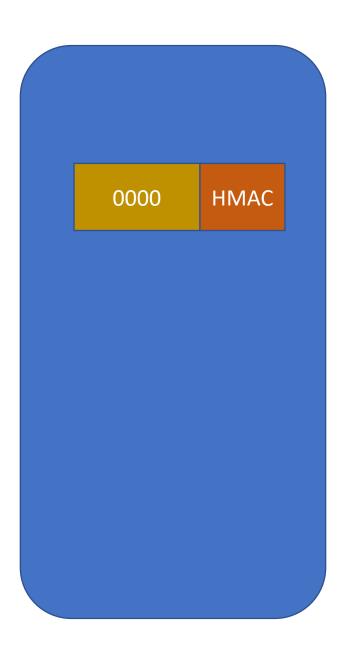


I encrypt data & added HMAC! HMAC(key||0000)





I encrypt data & added HMAC! HMAC(key||0000)

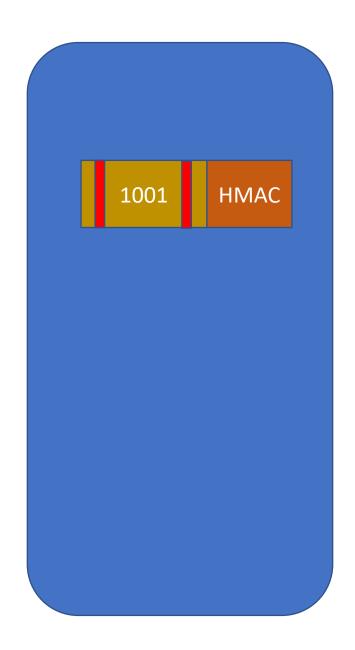


Edit data...



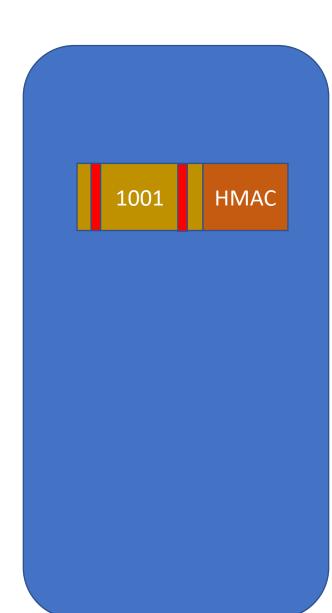


I encrypt data & added HMAC! HMAC(K||0000)



Edit data...





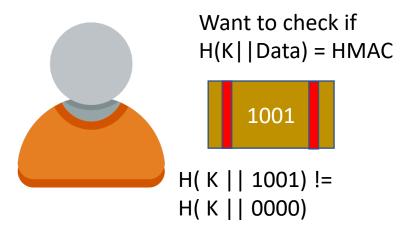


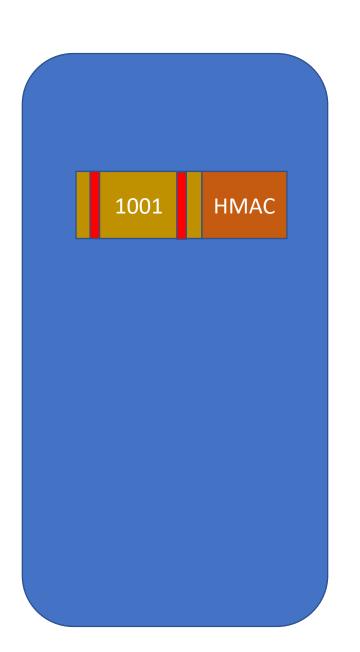




Want to check if H(K||Data) = HMAC

I encrypt data & added HMAC! HMAC(K||0000)

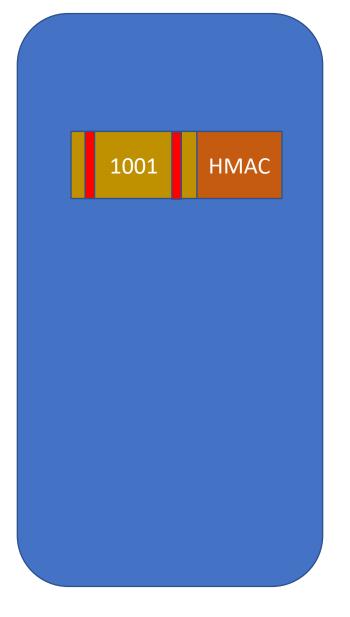




Edit data...



Workflow



Edit data...



Want to check if H(K||Data) = HMAC

1001

H(K||1001)!= H(K||0000)

Reject!

Decrypt Data with CBC

- Suppose you have a hash key = 'asdf'
 - HMAC = SHA256(SHA256('asdf') || encrypted_data)
 - = 7624e1f89ce009f8ec7e6e39781a42c0a27fa38f94db4f05f78b0f301007e06a
- Suppose the attacker changed the encrypted_data



- HMAC = SHA256(SHA256('asdf') || encrypted_data)
- = 389205904d6c7bb83fc676513911226f2be25bf1465616bb9b29587100ab1414
- Mismatch with HMAC!

Cannot edit data because we have HMAC

- Then, can attacker edit HMAC to match that to the edited ciphertext?
- HMAC = SHA256(SHA256('key') || edited_data)
- Attackers don't know the key
 - That's why we need to use key to SHA256.
 - Otherwise, anyone can generate valid MAC!
- Even they know SHA256(SHA256('key')|| encrypted_data)
 - They cannot generate a valid HMAC
 - They cannot correlate that value from this one...

Summary

- Block Cipher modes lets attacker play with ciphertext freely
 - They cannot be secure as we proved in challenges
- That's because Block Cipher protects only the data confidentiality
- Data Integrity left unprotected
- To protect data integrity, we can use cryptographic hash function
 - One way, it is hard to find an inverse of the result...
- HMAC, running cryptographic hash function with key on the data can protect data integrity...

Summary

Encrypt-then-MAC

Encrypted data

HMAC: H(H(key)||Encrypted data)

This is the only secure composition of using MAC with Encrypted data

- You must
 - Encrypt the data, and supply the entire encrypted data to HMAC
- No MAC-then-encrypt
 - Cryptanalysis exists (proven to be insecure)

Public (Asymmetric) Key Cryptography

- There is a scheme that we use different key to encryption and decryption
 - Why is it important? We will discuss this later about the 'key exchange'

- RSA (Rivest, Shamir, Adleman)
 - A public-key cryptography algorithm
 - Based on the difficulty of prime factorization
 - i.e., if the prime factorization of a large prime number is difficult, then the cryptography scheme is secure
 - Can be used for digital signature

How RSA Works?

- Choose two large prime number, p and q
- N = pq
- $\phi = (p-1)(q-1)$

- Choose public key, say, e = 65537 (a prime), that is coprime to ϕ
- Find $de == 1 \pmod{\phi}$
 - d can be efficiently be computed if you know φ
 - Blue: public key, Red: private key
 - Attackers don't know φ, to know φ, you need to factor N

RSA Encryption

• Public key: e, N

• Message: M

Me mod N

RSA Decryption

- Private key: d
- Public key: e N
- Ciphertext = C = M^e
- ed = 1

C^d mod N (M^e)^dmod N M^{ed} mod N M mod N

In reality: RSA-4096

Ν

>>> n
94300162313076688501487627887743414045960390714020238212657873381998561681681396483075239596099727423614481972904672187684996447088678082779896359977776138389400094618
93811971402397228611383901773325321259010777553654865622984724944010604379009841301441764487806140346612303663579718455548165265742251289534980309219758481925957858787
99446168652865946938887547013421958335603541458898859985233102756405213301150045330555227176327316853362195678419436714942441571041724570176803145478446833917315101830
88856681940225944856271324669491185023754645727393394123350591112266076929457503053224635114890484540850553592509484360925706475886219500022922117482666285104831675845
83315533957568315104232320670250600708758347303059147821341336205419089515531617207836027717015263175059127264155564477809166344370523152038595667063337410819626147392
61504146573604212524025625329042730131363602682044377326795554520903135271401816609380989125787711356372203148662219805667048555875256480930486742228216374620641072794
39035295803547382839528902460696618996010706014197280097861310233823234888621192701394780193796900301510396063855789518617879808500828987759386908963639259712107524427
2314777803224755716476436654020264220108971589745725860770832311

pq = N

>>> p
32068656324685270630293083389512032043113021564958908104074082522961856175155506846871162451521328939255338658509282518527589856290667522356720088237485538660191295524
66763080165651305435511312859214781036375072132558927069941823437366007834867296636112996194202733617438183660982840127747653230903017526198432501279953198838940252489
073973798439710809380710667964584394881725673204473174843215662005677203445933236655588112762119550999547645761253947551466759960870349272433501174519655039023335818685
463158445435736255026949066782447079073069611763861822177996113062890795417277837248384413086920855329944403808273
>>> q
29405710472654850958671060473587380272033499777349031897972774267355140300355329654820633532129222264029138740441499983175231104660614751766804322996494698305495888552
75830646534910098483628642910172110351604197235439018569363033804440884203852197669902443119162881776394993390408980913691022609261096791810408703245340903925873967271
31915873792822333318309536841669683057109843537055633086239079548671678418825449122417399716632535551668617851522599535745800792690727833156308278625231616345078778788
62534221887593666478665101149415792855605030513596595792804362412144174514663486222431955289564169364164960140372807

True_

True_

e,d

>>> e 55537 >>> d 18849384718575836845896027058446964676474069889583977304207060764212294704213288583979822675876320692881116222058550218329698543675506704371830299154811288476755744686 33122309922202781555596519725201973060912784929930859509301614305911758508400276501043488636804559844677375802873099344914795687645505468558344783111507028292873460032 1845850144731287524450961847316519911189719384644008890527082980271738228698980203454999061232737047023478870918523805193489875712666735033175924888974589217149348121427 5561533993647801818235736077590268892093299752855415205175538617224904417010075667176161820778219898456478814438777554613519848613000196946332897123652219249131858620 56419600479189059635429627251759483101314923287035484839408409780479225523102668224739334748234253874639242903082590778667288728135476436491806847150495168056124486222 56426969696979130604479458312909282114054641455801867990797517633

In reality: RSA-4096

Encryption

```
>>> c = pow(m, e, n) # m ** e % n

>>> c

44378470637778735239608611810942282378207997450147968063174353524631847658659086953707365508877463886993450049298512874492100304888894329853348670989505877423737248417

80400795758744992904133748470628082245089418796820314345284578141769246629011746745886476950463573078507843823434081211222204304995686000554834475890467697637847246844

8632211687399982137712399982644387191965538893966531625400636903013509472170170783645161654401139964366004445993093041228905914489501117462826954207948823222431789196358093

93698243007398395836444485292409163123893335382268476709587920108064020611921281974647321167361411263176436550983106315712512550381458866081739361427686243767474405671

696921253146645179015461600858262605157304413159663632253076388977366485979730059180207738237383458006267025213769309185729231546159544178725517773424229933552200952804

22942476481060100821238941445817250493239517641421074872678593347890043728988899931048881456864826850140750986137042474998371731154209998675414707669442684372571942321

22531302749839804996479430073226620101139728472052638053446458176003508443044486697874544730775786131613929373770832871564550265529076972479414717864262176092017206844

5561283568170818845183097678931422241453739518372479767621293111
```

Decryption

```
>>> _c = pow(c, d, n) # c ** d % n == m ** (de) % n == m % n == m >>> _c  
12345  
>>> ■
```

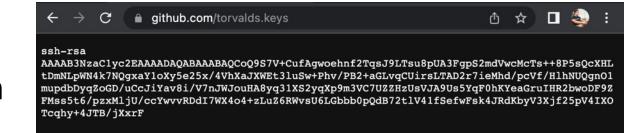
Public Key Cryptography

- We can use separate key for encryption and decryption
 - Encryption key: public key (e, N)
 - Decryption key: private key (d)
- Attackers cannot guess the private key from the public key
 - In RSA, attacker must factor the prime number N = pq
 - In creating the key, we choose p and q as a big prime number
 - Factoring a multiplication of two big prime numbers is a difficult task

Characteristics

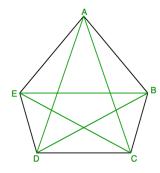
- We use a public key for encryption
 - We can publicize this key
 - If you publish your key, anyone who can access that can encrypt message
 - (e, N) is public, me mod N!
 - Only the holder of the private key can decrypt the message
 - d is private, m^{ed} == m¹ == m (mod N)

- Why is this important?
 - Let's talk about the key exchange problem



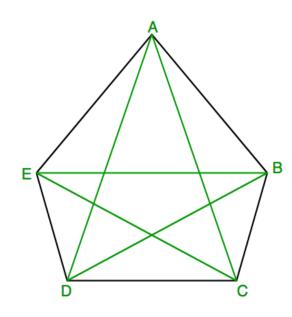
Key Exchange

- Suppose we have 5 people, A, B, C, D, E
 - How many keys do we need to have to make them communicate securely?
 - E.g., if A talks to B, C or others must not see the message
 - But anyone should be able to talk to anyone...
- A block cipher
 - We need 1 key for A and B can talk securely
- How many keys do we need to let them communicate securely?
 - A-B, A-C, A-D, A-E
 - B-C, B-D, B-E
 - C-D, C-E
 - D-E
 - 10 keys (5*4/2 = 10)



Symmetric Key Cryptography

- Encryption and the decryption operations are using the same key
 - Block Cipher encryption key == decryption key
 - You cannot share that other than 2 people
- Key exchange complexity
 - We need 1 key per each pair of people
 - N (N − 1) / 2
 - O(N²)

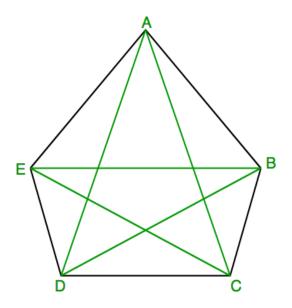


Asymmetric Key Cryptography

- Can we use a different key for the encryption and decryption?
 - K = k1, k2
 - Enc(k1, M) = C, Dec(k2, C) = M?
- Preferably, can we publish the encryption key to public?
 - While keeping the decryption key secret
- Then we need O(N) keys
 - Each member's public key, that's it.

Why O(N)?

• We need O(N²) keys for symmetric encryption



Why O(N)?

Suppose each will generate public and private key

- Public A, Private A
- Public_B, Private_B
- Public_C, Private_C
- Public_D, Private_D
- Public_E, Private_E

Why O(N)?

- Each will have their own private key, and then,
 - publish all their public keys
- A: Private_A
- B: Private B
- C: Private C
- D: Private_D
- E: Private_E
- Public keys: Public_A, Public_B, Public_C, Public_D, Public_E

Can A Send an Encrypted message to B?

- Can A send an encrypted message to B?
 - Yes, encrypt data using Public_B; only B (holder of Private_B) can decrypt it
- Can C send an encrypted message to E?
 - Yes, encrypt data using Public_E; only E (holder of Private_E) can decrypt it
- Can X send an encrypted message to Y?
 - Yes, if X knows the public key of Y
- We only need to know the receiver's public key
 - Sender does not matter, that's why we have O(N)
 - Suppose we have N = 200, we need 19900 keys in symmetric, and we need 400 keys for asymmetric

RSA: Digital Signature

- RSA can be used as a digital signature scheme
- What is that?

- In RSA, encryption is applying the public exponent to the message
 - M^e mod N
- In RSA, decryption is applying the private exponent to the message
 - C^d mod N

RSA: What will be the meaning of private encrypt?

- Suppose A encrypts the following message with her private key
 - "I would like to donate \$100 to OSU if I get A from CS 370"
- M = 5315140633361125709395629341158475998805322893872442710 1859883089254119711739486837784167497839141764612450119 856395995171455585519613744
- C = m^d mod N

RSA: What will be the meaning of private encrypt?

- M =
 5315140633361125709395629341158475998805322893872442710
 1859883089254119711739486837784167497839141764612450119
 856395995171455585519613744
- C = m^d mod N
- Anyone can have e. That means, anyone can decrypt C
 - $C^e == m^{de} == m^1 == m \pmod{N}$

RSA: What will be the meaning of private encrypt?

- M = 5315140633361125709395629341158475998805322893872442710 1859883089254119711739486837784167497839141764612450119 856395995171455585519613744
- C = m^d mod N
- Anyone can have e. That means, anyone can decrypt C
 - $C^e == m^{de} == m^1 == m \pmod{N}$
 - m =
 53151406333611257093956293411584759988053228938724427101859883
 08925411971173948683778416749783914176461245011985639599517145
 5585519613744
 - "I would like to donate \$100 to OSU if I get A from CS 370"

RSA: What will be the meaning of privat

• M = 5315140 1859883 8563959

We can verify that the encrypted content C contains

The ciphertext that only the holder of private key can generate.

We all have public key, and if that is decrypted to

"I would like to donate \$100 to OSU if I get A from CS 370",

then, we know that the holder of private key 'endorsed it'

- C = m^d mod N
- Anyone can have e. That means, anyone can decrypt C
 - $C^e == m^{de} == m^1 == m \pmod{N}$
 - m =
 53151406333611257093956293411584759988053228938724427101859883
 08925411971173948683778416749783914176461245011985639599517145
 5585519613744
 - "I would like to donate \$100 to OSU if I get A from CS 370"

RSA: private_encrypt

- RSA Encryption using the private key is so-called as 'Signing'
- Why?
 - The ciphertext will be decrypted as a plaintext using the public key
 - Anyone can decrypt!
 - But the ciphertext can only be generated with the private key
 - Only the private key owner can generate it!

Implication

- Holder of the private key generated a ciphertext message of message M
- M is signed, endorsed by the holder's private key
- (Because it can only be generated with the private key)

RSA Summary

- Public/Private key Scheme
 - We can publish the public key encryption key
 - We must hide the private key decryption key
- Based on the difficulty of prime factorization
 - You cannot correlate the private key from the public key unless
 - You can factor a big number (a multiple of 2 big prime numbers)
- Anyone can encrypt message to the private key owner
 - Enc(public_key, message)
- Only the private key owner can decrypt message
 - Dec(private_key, encrypted_message)

RSA Summary

- Encryption with private key could be a 'digital-signature'
 - Signed_message = Enc(private_key, message)
 - Message = Dec(public_key, signed_message)
- The correctly decrypted message using public key means that the private key holder have endorsed ('encrypted') the data
 - Anyone can verify this using the public key

Symmetric vs. Asymmetric

- Can we use symmetric key as digital signature?
 - A-B, A-C, A-D, A-E
 - B-C, B-D, B-E
 - C-D, C-E
 - D-E
- If a message was encrypted with key D-E
 - Then either D or E generated the message -> ambiguity
 - Only D or E can verify that -> result is not public
 - They must leak the key D-E for verification -> key need to be leaked to verify

Symmetric vs. Asymmetric

- In asymmetric key scheme
 - Public_A, Private_A
 - Public_B, Private_B
 - Public_C, Private_C
 - Public D, Private D
 - Public_E, Private_E

- If a message was encrypted with Private_D
 - We can decrypt the data using Public_D -> generated by D -> not ambiguous
 - Anyone can verify this -> result is public
 - Does not need to leak the private key -> don't leak the key for verification