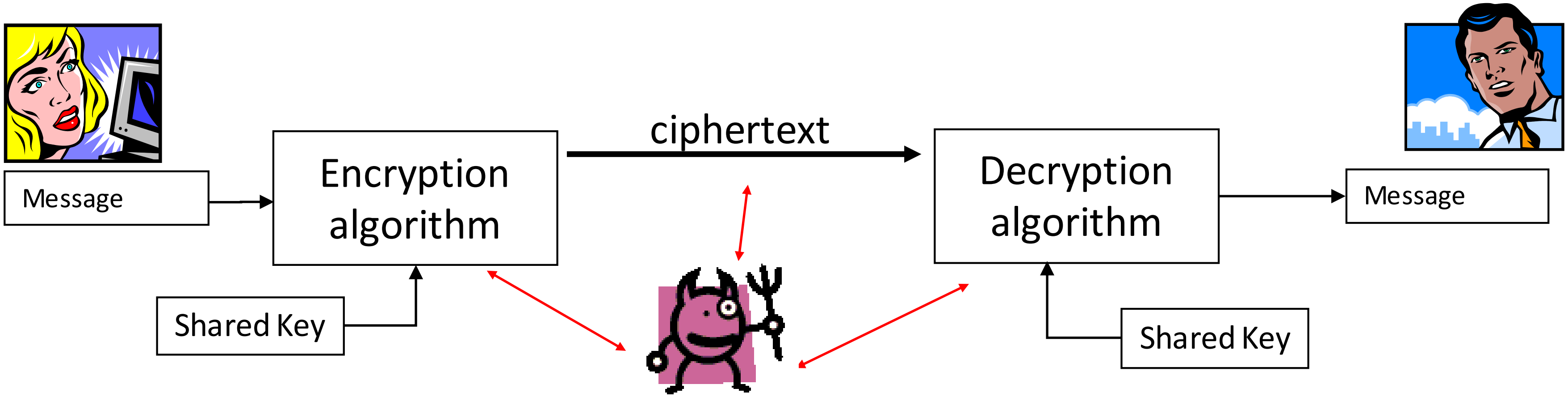
Class CS 6903, Lecture n. 5

- Welcome to Lecture 5!
- In Lectures 1-4 we studied:
  - Classical cryptography, encryption with perfect secrecy
  - Background on algorithms, complexity theory. Modern cryptography: principles, primitives, and a public-key cryptosystem
  - Algorithmic number theory, number theory and cryptographic assumptions, reductions, proofs by reductions, number theory candidates for cryptographic primitives and schemes
  - Pseudo-random generators, functions, permutations and applications

Summary of Lecture 5

- Symmetric Encryption
  - Introduction: classical vs modern cryptography
  - Formal definition: syntax, correctness, security requirements
  - Security notions and schemes satisfying them
    - Perfect secrecy
    - Indistinguishability
    - Indistinguishability with chosen message attack
    - Indistinguishability with chosen ciphertext attack
  - Fundamental concepts in cryptography
    - Experiments
    - Experiment hybrids
    - Simulation

Symmetric (or private-key) encryption



- Goal: Alice and Bob want to communicate privately in the presence of a communication eavesdropper
- Setup: Alice and Bob (somehow) share a secret, random, key
- Applications: widely applied technology in many computer, web services and functions (e.g., online banking, e-mail access, etc.)
- Theory: widely studied research area in cryptography

Classical vs Modern Cryptography

Classical Cryptography

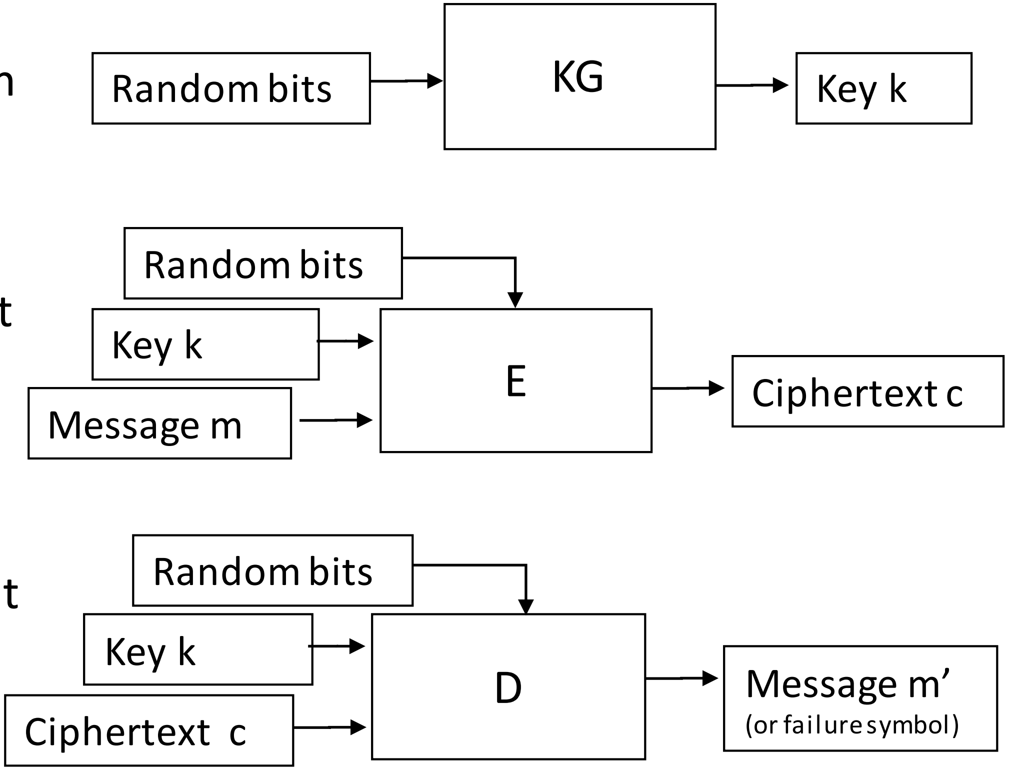
- Mostly an art of designing and breaking “secrecy codes”
- Main strength (or weakness?):
  - secrecy code is known to Alice and Bob but NOT to the eavesdropper
  - Perfect secrecy, one-time pad
- Went on until last century
- Some basic principles:
  - Confusion, diffusion, algorithm obscurity
- Some basic design components:
  - Substitution, transposition, etc.

Modern Cryptography

- New resources: computers, physical tools, etc.
- New paradigms: use of randomness
- Apparent weakness:
  - secrecy code is known to Alice, Bob AND eavesdropper
- Main strengths:
  - security is based on secrecy and randomness of keys
  - secrecy code can be extensively analyzed
  - rigorous definitions of security requirements
  - rigorous proofs that schemes meet security definition

Symmetric Encryption: formal definition

- A symmetric encryption scheme with message space M is formally defined as a triple of efficient (probabilistic) algorithms (KG,E,D) such that:
    - KG, the key-generation algorithm, on input a sufficiently long random string, returns a secret key k
    - E, the encryption algorithm, on input key k and message m in message space M (subset of {0,1}*), returns a ciphertext c
    - D, the decryption algorithm, on input key k and ciphertext c, returns a message m' or a special failure symbol
- Furthermore, it satisfies "correctness" and "security" properties



Random bits → KG → Key k

Random bits, Key k, Message m → E → Ciphertext c

Random bits, Key k, Ciphertext c → D → Message m'
(or failure symbol)

Symmetric Encryption: correctness, security

- (KG,E,D) satisfies "(decryption) correctness" and "security" (or, message privacy, message confidentiality) properties
- Correctness:
  - For any m in message space M, if key k is returned by KG, and ciphertext c is returned by E on input k,m, then D, on input k,c, returns m'=m with probability 1 (or, very high probability; i.e., >=1-d, for some negligible quantity d).
- Security:
  - Intuition: no adversary Adv of a "certain class", defining Adv's "resources", should not be able to obtain any "partial information" about an encrypted message unless with "very small probability"
  - Formalizations of this intuition are quite non-trivial

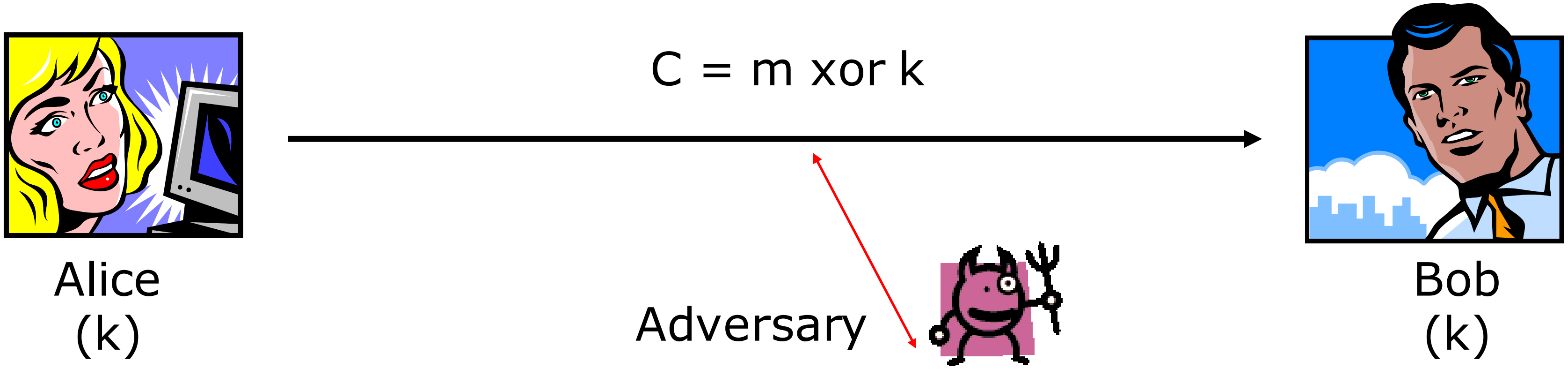
Summary of Lecture 5

- Symmetric Encryption
  - Introduction: classical vs modern cryptography
  - Formal definition: syntax, correctness, security requirements
  - Security notions and schemes satisfying them
    - Perfect secrecy
    - Indistinguishability
    - Indistinguishability with chosen message attack
    - Indistinguishability with chosen ciphertext attack
  - Fundamental concepts in cryptography
    - Experiments
    - Experiment hybrids
    - Simulation

Symmetric Encryption: security notions

- Only security notion before modern cryptography:
  - Information-theoretic security or perfect secrecy
- Most studied security notions in modern cryptography:
  - (1) message indistinguishability (IND) against eavesdroppers and its equivalent notion of semantic security
  - (2) IND + adversary also performs a chosen plaintext attack (IND-CPA)
  - (3) IND-CPA + adversary also performs a chosen cipertext attack (IND-CCA)
- Theorem 1:
  - When x>y, if a symmetric encryption scheme satisfies security notion x, then it is also satisfies security notion y, for x,y in {1,2,3}
- Theorem 2:
  - When x<y, if there exists a symmetric encryption scheme satisfying security notion x, there exists one that satisfies security notion x but not security notion y, for x,y in {1,2,3}

Perfect Secrecy, one-time pad: review



C = m xor k

Alice
(k)

Adversary

Bob
(k)

- Notion: probability(message was encrypted) remains same after seeing ciphertext
  - Specifically, for any message m in message space M, and any ciphertext c in ciphertext space C such that Prob[c] > 0, it holds that Prob[m|c] = Prob[m]
- One-time pad scheme (perfectly secrecy proved in Lecture 1):
  - Message, key and ciphertext space are all equal to $\{0,1\}^L$ for some integer L>0
  - KG generates a random string k from the key space
  - To encrypt L-bit message m, E returns c = k XOR m
  - To decrypt ciphertext c, D returns m’ = c XOR k
- Pros: natural notion, equivalent to other natural notions, provides secrecy against computationally unlimited adversaries
- Cons: secrecy against computationally unlimited adversary implies limitations, such as, the amount of randomness being >= message length

The (Message) Indistinguishability notion (in the presence of an eavesdropper)

- A denotes an efficient algorithm, called adversary
- Experiment Exp(A,ind):
  - Let k be returned by algorithm KG
  - A chooses 2 equal-length messages m(0), m(1)
  - b←{0,1}
  - Ciphertext is computed as c=E(k, m(b))
  - A, on input m(0),m(1),c, returns a guess bit d

A tries to guess which message was encrypted

- Exp(A,ind) is successful if d=b

- Advantage(A) = |Prob[Exp(A,ind) is successful] -1/2|

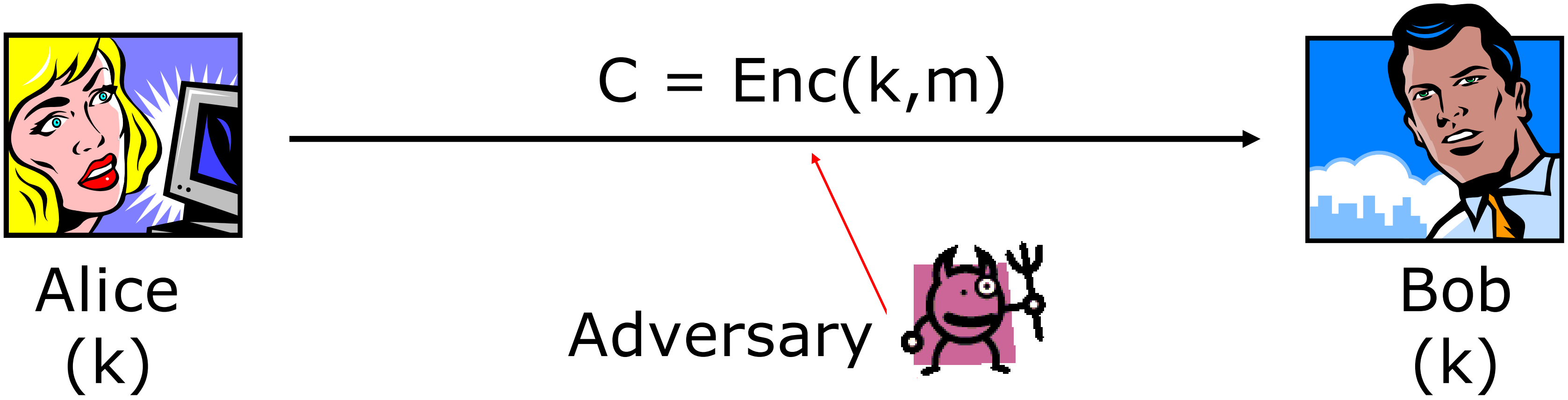
Better than a random guess?

- A symmetric encryption scheme satisfies security in the sense of (message) indistinguishability if for any efficient adversary algorithm A, Advantage(A) is negligible.

Only difference with a secrecy (adversarial indistinguishability) notion in Lecture 1

- Note: eavesdropping (sniffing, snooping) is a practical attack

A symmetric encryption scheme satisfying (message) indistinguishability



C = Enc(k,m)

Alice
(k)

Adversary

Bob
(k)

- Recall one-time pad definition: Enc(k,m) = k xor m, Dec(k,c) = k xor c
  - Properties: perfect secrecy (+), key length = message length (-)
- A new scheme using a pseudo-random generator G:
  - KG returns a random key k
  - Enc(k,m) = G(k) xor m
  - Dec(k,c) = G(k) xor c
- Properties: key length << message length (+), computational secrecy (-), security in the sense of indistinguishability (to be proved soon)

A symmetric encryption scheme satisfying (message) indistinguishability: result and proof sketch

- Theorem: G pseudo-random generator –> symmetric encryption scheme from previous slide satisfies (message) indistinguishability notion
- Proof sketch (using experiment hybrids): Recall that

  KG returns a random key k, Enc(k,m)=G(k) xor m, Dec(k,c)=G(k) xor c
- Need to show for this scheme that for any efficient algorithm A, Experiment(A,ind), also denoted as E0, is successful with probability at most negligible
- Consider 2 modified experiments:
  - E1: like E0, but c is computed as r xor m_b, for some random string r

    (intuition: E1 replaces the output of the pseudo-random generator G in experiment E0 with a random string of the same length)
  - E2: like E1, but c is computed as s, for some random string s

    (intuition: E2 replaces the (one-time-pad-like) ciphertext c in experiment E1 with a random string of the same length)
- Logic in rest of proof: we prove that
  - A's advantage in E2 is 0 as A's view in E2 does not depend on b
  - experiment E1 is indistinguishable from E2 from A's point of view;
  - experiment E0 is indistinguishable from E1 from A's point of view (if G is a pseudo-random generator);
  - Thus, A cannot succeed in E0 (unless with negligible probability)

A symmetric encryption scheme satisfying (message) indistinguishability: result and proof sketch

- Theorem: G pseudo-random generator –> symmetric encryption scheme from previous slide satisfies (message) indistinguishability notion
- Proof sketch (using experiment hybrids): Recall that KG returns a random key k, Enc(k,m)=G(k) xor m, Dec(k,c)=G(k) xor c

- (continues from previous slide):

- Fact 1: E2 is successful with probability ½
  - Proof sketch: Use the fact that A's view in E2 is independent on b and b is a random bit
- Fact 2: E1 is successful with probability ½
  - Proof sketch: One-time pad secrecy → no algorithm A can distinguish E1 from E2; thus, E1 is successful with the same probability as E2, which is ½
- Fact 3: If G is pseudo-random, E0 is successful with probability < ½ + negligible
  - Proof sketch: Only difference between E0 and E1 is that c uses G(k) in E0 and r in E1. Thus, if G is pseudo-random, A can distinguish c in E0 from c in E1 with probability at most negligible. Thus, E0 is successful with probability at most ½ + negligible
- This implies that the scheme satisfies message indistinguishability in the presence of an eavesdropper

Question set 12

- Consider the notion of perfect secrecy (the adversarial indistinguishability variant), from Lecture 1, and the notion of security (or message privacy, or computational secrecy) in the sense of indistinguishability in the presence of an eavesdropper, from this Lecture.
    - What are the differences between the two notions with respect to the attack experiment?
    - Because of these differences, does every scheme satisfying one notion also satisfy the other?
    - Mention schemes satisfying the two notions. What is the difference between these schemes with respect to length of the shared keys?

- Prove, using a hybrid argument, that the input-output behavior of a pseudo-random function is computationally indistinguishable from the input-output behavior of a random permutation

Summary of Lecture 5

- Symmetric Encryption
  - Introduction: classical vs modern cryptography
  - Formal definition: syntax, correctness, security requirements
  - Security notions and schemes satisfying them
    - Perfect secrecy
    - Indistinguishability
    - Indistinguishability with chosen message attack
    - Indistinguishability with chosen ciphertext attack
  - Fundamental concepts in cryptography
    - Experiments
    - Experiment hybrids
    - Simulation

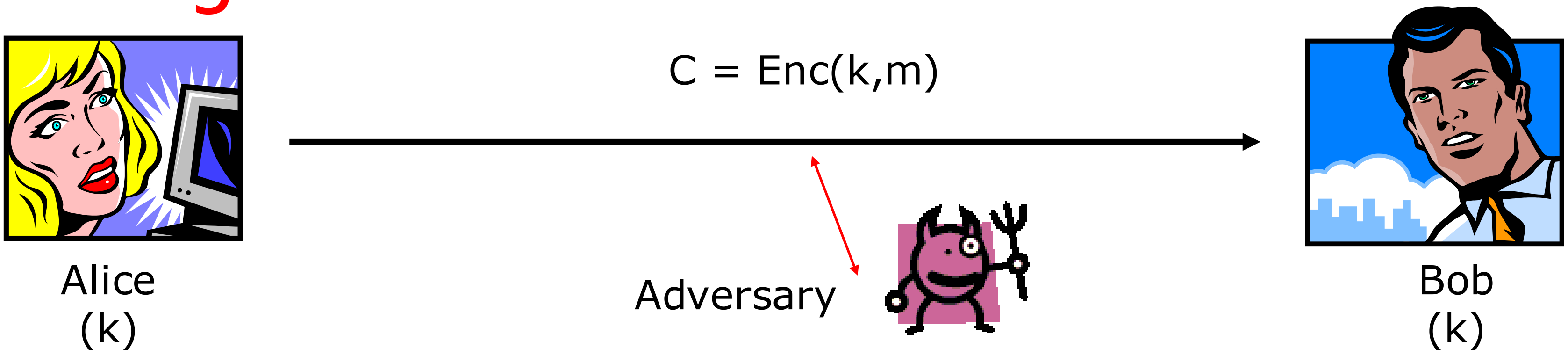
The (Message) Indistinguishability in the presence of a chosen message attack (IND-CMA) notion

- A denotes an efficient algorithm, called adversary
- Experiment Exp(A,ind-cma):
  - Let k be returned by algorithm KG
  - A can use E(k,.) as an oracle (i.e., A issues polynomially many queries {u(i)} and obtains {E(k,u(i))})
  - A chooses 2 equal-length messages m(0), m(1)
  - b←{0,1}
  - Ciphertext is computed as c=E(k, m(b))
  - A, on input queries {u(i)}, answers {E(k,u(i)))}, and m(0),m(1),c, returns his guess bit d
- Experiment Exp(A,ind-cma) is successful if d=b
- Advantage(A) = |Prob[Exp(A,ind-cma) is successful] -1/2|
- A symmetric encryption scheme satisfies security in the sense of (message) indistinguishability in the presence of a chosen message attack if for any efficient adversary algorithm A, Advantage(A) is negligible
- Practical CMA scenarios: see WW2, client-server examples in [KL, pp. 83,84]

This bullet was not in previous IND definition

Compared to previous IND definition, A also uses queries and answers to guess which message was encrypted

A symmetric encryption scheme satisfying (message) indistinguishability against chosen message attacks



C = Enc(k,m)

Alice
(k)

Adversary

Bob
(k)

- One-time pad: Enc(k,m) = k xor m, Dec(k,c) = k xor c
  - Satisfies perfect secrecy but is randomness-inefficient
- Using a pseudo-random generator G:
  - Enc(k,m) = G(k) xor m, Dec(k,c) = G(k) xor c, |k| much smaller than |m|
  - Does not satisfy security against chosen message attacks (exercise)
- Using pseudo-random functions F:
  - KG returns a random key k
  - Enc(k,m) = (r, F(k,r) xor m), where r is a random string
  - Dec(k,(c1,c2)) = F(k,c1) xor c2
- Properties: key length << message length, security against chosen message attack (to be proved soon)

Symmetric encryption with (message) indistinguishability against chosen message attack: result and proof sketch (no CMA first)

- Theorem: F pseudo-random function –> symmetric encryption scheme from previous slide satisfies (message) indistinguishability notion
- Proof sketch (using experiment hybrids): Recall that

  KG returns a random key k, Enc(k,m)=(r, F(k,r) xor m), Dec(k,(c1,c2))=F(k,c1) xor c2
- Need to show for this scheme that for any efficient algorithm A, Exp(A,ind-cma), denoted as E0, is successful with probability at most ½ + negligible
- We first consider case E0 = Exp(A,ind); consider 2 modified experiments:
  - E1: like E0, but c=(c1,c2) is computed as (r, s xor m_b), for some random strings r,s
  - E2: like E1, but c =(c1,c2) is computed as u,v, for some random strings u,v
- Fact 1: E2 is successful with probability ½ (as A's view in E2 is independent on b)
- Fact 2: No A can distinguish E1 from E2 (follows from one-time pad secrecy). Thus, E1 is successful with the same probability as E2, which is ½
- Fact 3: Only difference between E0 and E1 is that c uses F(k,r) in E0 and s in E1. Thus, if F is a pseudo-random function, A can distinguish c in E0 from c in E1 with probability at most negligible. Thus, E0 is successful with probability at most ½ + negligible.
- This implies that the scheme satisfies message indistinguishability (but not yet message indistinguishability with chosen message attack)

Symmetric encryption with (message) indistinguishability against chosen message attack: intuitions for the final proof sketch

- In previous slide we proved that this scheme satisfies indistinguishability without chosen message attack; is that enough to prove that it satisfies indistinguishability with chosen message attack?
- No, because A could learn k or even partial information about k while asking queries to the oracle E(k,.) in a chosen message attack
- If A learns k or partial information about k, the previous facts that c is indistinguishable from random strings might not be true
- On the other hand, if A learns nothing new about k (or, better, nothing new at all), then the previous facts on the indistinguishability of c are still true and A's chosen message attack is useless
- Thus, one approach would be to prove that A learns nothing new at all while asking queries to the oracle E(k,.) in a chosen message attack
- How do we prove that A learns nothing new?
- By simulation, a very important concept in cryptography
- If we could simulate oracle E(k,.) in the chosen message attack by a simulator that does not use k (or any secret information at all), then A's chosen message attack is ineffective in learning anything new about the key (or anything new at all)
- Next slide shows how to simulate oracle E(k,.) without k and in polynomial time

Symmetric encryption with (message) indistinguishability with chosen message attack: completing proof sketch

- Lemma: The oracle E(k,.) responses to A’s queries can be simulated by an efficient algorithm S (with no access to k)
- Proof sketch: Recall that KG returns a random key k
  Enc(k,m)=(r, F(k,r) xor m), Dec(k,(c1,c2))=F(k,c1) xor c2
- Thus, when asking queries u(i) to E(k,.), A obtains as output
  E(k,u(i))=(r(i), F(k,r(i)) xor u(i)), for random and independent r(i), for i=1,…,n
- Algorithm S simulates these answers as
  (t(i),v(i)), for random and independent t(i) and v(i), for i=1,…,n
- The proof that this is a valid simulation is by a hybrid argument; consider distributions:
  - D0 = {(u(i), r(i), F(k,r(i)) xor u(i)), i=1,…,n}
  - D1 = {(u(i), r(i), s(i) xor u(i)), for random s(i), i=1,…,n}
  - D2 = {(u(i), t(i), v(i)), for random t(i), i=1,…,n}
- Fact 1: D2 is the same as S’s simulation output
- Fact 2: D1 is equal to D2 (follows from one-time pad secrecy)
- Fact 3: If F is a pseudo-random function, A can distinguish D0 from D1 with probability at most negligible.
- Thus, S’s output is a valid simulation of oracle E(k,.)

The (message) indistinguishability in the presence of (adaptive) chosen message attack (adaptive CMA) notion

- Experiment Exp(A,ind-ad-cma):
  - KG returns key k
  - A can use E(k,.) as an oracle (i.e., A issues polynomially many queries u(i) and obtains E(k,u(i)))
  - A chooses 2 equal-length messages m(0), m(1)
  - b←{0,1}
  - Ciphertext is computed as c=E(k, m(b))
  - A can use E(k,.) as an oracle (that is, A issues a sequence of polynomially many queries v(i) and obtains E(k,v(i)))
  - A, on input all queries and answers, m(0),m(1), and c, returns his guess bit d

Only updated parts with respect to previous definition

- Experiment Exp(A,ind-ad-cma) is successful if d=b
- A symmetric encryption scheme satisfies security in the sense of (message) indistinguishability in the presence of an adaptive chosen-message attack if for any efficient adversary algorithm A, Advantage(A) = |Prob[Exp(A,ind-ad-cma) is successful] -1/2| is negligible.
- Note: Previous scheme based on a PRF does not satisfy this definition

Various Remarks

- We will define in a later Lecture the notion of chosen ciphertext attack, and study a scheme that satisfies indistinguishability in the presence of this attack, assuming the existence of one-way functions
- The notion of (message) indistinguishability can be showed to be equivalent to other natural notions, including semantic security, requiring (informally speaking) that the ciphertext does not help computation any more than a random plaintext
- We have seen symmetric encryption schemes satisfying different security notions under the existence of one-way functions; how do we use symmetric encryption in most applications?
- Next lecture: Block ciphers, block cipher theory

Question set 13

- A deterministic symmetric encryption scheme is a symmetric encryption scheme where algorithm E(k,m) does not use any random bits other than the key k shared by Alice and Bob. A fixed-key scheme is a scheme where the same fixed key k is used to encrypt all messages. Consider an arbitrary deterministic and fixed-key symmetric encryption scheme.
  - Can it satisfy the IND security notion?
  - Can it satisfy the IND-CMA security notion?
  - (Hint: to answer "yes", give an example of a deterministic scheme that does satisfy the notion; to answer "no", show that the attack in the notion's security experiment is successful against any deterministic encryption scheme)
- Let G be a pseudo-random generator. Determine what security notion is satisfied by the following deterministic symmetric encryption scheme:
  - Alice and Bob share a long enough random key K
  - Enc(K,m) = pick the next unused k from K and return c = G(k) xor m
  - Dec(K,c) = pick the next unused k from K and return m'= G(k) xor c

  Why is this scheme not contradicting what found in the first question from this slide?
- Fill a 3x3 table whose entries indicate which of the statements

  "if an encryption scheme satisfies notion A then it satisfies notion B"

  is true or unknown or false, where A and B are taken from set {the adversarial indistinguishability variant of perfect secrecy, IND-security, IND-CMA-security}

Class CS 6903, End of Lecture n. 5

Reference → Topic ↓	[KL]	[MOV]	[FSK]
Symmetric encryption	3.2, 3.4, 3.6.2, 3.7	1.5	2.1, 2.6, 2.7