EE 488
Introduction to Cryptography Engineering

Yongdae Kim

Homework & Quiz schedules

- Homework due dates
  - 3/12 (note: due is updated)
  - 3/26, 4/9, 5/7, 5/21, 6/4
  - Submit at least 10 minutes before each class
- Quiz
  - 3/17, 3/31, 5/12, 5/26
  - In-class quiz
- You can also check in course website

Math, Math, Math!

Prime

- p $\geq$ 2 is prime if
  - a | p $\Rightarrow$ a = $\pm$ 1 or $\pm$ p
  - Hereafter, p is prime
- [Euclid] p | ab $\Rightarrow$ p | a or p | b
- [Euclid] There are infinite number of primes.
- *Prime number theorem:*
  - let $\pi(x)$ denote the number of prime numbers $\leq$ x, then $\lim_{x \to x \infty} \pi(x)/(x/\ln x) = 1$
- **Euler phi function**: For n $\geq$ 1, let $\phi(n)$ denote the number of integers in [1, n] which are relatively prime to n.
  - if p is a prime then ϕ (p)=p-1
  - if p is a prime, then $\phi(p^r) = p^{r-1}(p-1)$.
  - f is multiplicative. That is if gcd(m,n)=1 then $\phi(m*n) = \phi(n) * \phi(m)$

Fundamental theorem of arithmetic

- Every positive integer greater than 1 can be uniquely written as a prime or as the product of two or more primes where the prime factors are written in order of non-decreasing size

- Examples
  - 100 = 2 * 2 * 5 * 5
  - 182 = 2 * 7 * 13
  - 29820 = 2 * 2 * 3 * 5 * 7 * 71

Pairwise relative prime

- A set of integers $a_1, a_2, \cdots a_n$ are pairwise relatively prime if, for all pairs of numbers, they are relatively prime
  - Formally: The integers $a_1, a_2, \cdots a_n$ are pairwise relatively prime if $\gcd(a_i, a_j) = 1$ whenever $1 \le i < j \le n$.
- Example: are 10, 17, and 21 pairwise relatively prime?
  - gcd(10,17) = 1, gcd (17, 21) = 1, and gcd (21, 10) = 1
  - Thus, they are pairwise relatively prime
- Example: are 10, 19, and 24 pairwise relatively prime?
  - Since gcd(10,24) ≠ 1, they are not

Modular arithmetic

- If a and b are integers and m is a positive integer, then a is *congruent to b modulo m* if m divides a-b
  - Notation: $a \equiv b \pmod{m}$
  - Rephrased: $m \mid a-b$
  - Rephrased: $a \bmod m = b$
  - If they are not congruent: $a \not\equiv b \pmod{m}$

- Example: Is 17 congruent to 5 modulo 6?
  - Rephrased: $17 \equiv 5 \pmod{6}$
  - As 6 divides 17-5, they are congruent
- Example: Is 24 congruent to 14 modulo 6?
  - Rephrased: $24 \equiv 14 \pmod{6}$
  - As 6 does not divide 24-14 = 10, they are not congruent

Example (World of mod n)

…	-2n	-n	0	n	2n	3n	4n	…	0
…	-2n+1	-n+1	1	n+1	2n+1	3n+1	4n+1	…	1
…	-2n+2	-n+2	2	n+2	2n+2	3n+2	4n+2	…	2
…	…	…	…	…	…	…	…	…	…
…	-n-1	-1	n-1	2n-1	3n-1	4n-1	5n-1	…	n-1

More on congruence

- Every integer is either of the form 4k, 4k+1, 4k+2, 4k+3.
- Every integer is either of the form 0 mod 4, 1 mod 4, 2 mod 4, 3 mod 4
- $y^2 - x^2 - 2 \equiv 0$ mod 4 has no solution.

- Let *a* and *b* be integers, and let *m* be a positive integer. Then $a \equiv b \pmod{m}$ if and only if *a* **mod** *m* = *b* **mod** *m*
- Example: Is 17 congruent to 5 modulo 6?
  - Rephrased: does $17 \equiv 5 \pmod{6}$?
  - 17 mod 6 = 5 mod 6
- Example: Is 24 congruent to 14 modulo 6?
  - Rephrased: $24 \equiv 14 \pmod{6}$
  - 24 mod 6 $\neq$ 14 mod 6

Even more on congruence

- Let *m* be a positive integer. The integers *a* and *b* are congruent modulo *m* if and only if there is an integer *k* such that $a = b + km$

- Example: 17 and 5 are congruent modulo 6
  - 17 = 5 + 2*6
  - 5 = 17 -2*6

- Let a, b, c be integers.
  - $a \equiv a \bmod n$
  - $a \equiv b \bmod n \Rightarrow b \equiv a \bmod n$
  - $a \equiv b \bmod n$ and $b \equiv c \bmod n \Rightarrow a \equiv c \bmod n$.

Even even more on congruence

- Let m be a positive integer. If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a+c \equiv (b+d) \pmod{m}$ and $ac \equiv bd \pmod{m}$

- Example
  - We know that 7 ≡ 2 (mod 5) and 11 ≡ 1 (mod 5)
  - Thus, 7+11 ≡ (2+1) (mod 5), or 18 ≡ 3 (mod 5)
  - Thus, 7*11 ≡ 2*1 (mod 5), or 77 ≡ 2 (mod 5)

- An integer x is congruent to one and only one of the integers 0, 1, 2, ⋯, n-1 mod n.

The Caesar cipher

- Julius Caesar used this to encrypt messages
- A function f to encrypt a letter is defined as:
  $f(p) = (p + 3) \bmod 26$
  - Where p is a letter (0 is A, 1 is B, 25 is Z, etc.)
- Decryption: $f^{-1}(p) = (p - 3) \bmod 26$
- This is called a substitution cipher
  - You are substituting one letter with another

Arithmetic Inverse

- Let a be an integer. a^* is an arithmetic inverse of a modulo n, if $a\, a^* \equiv 1 \bmod n$.

- Suppose that gcd(a, n) =1. Then a has an arithmetic inverse modulo n.

- Suppose gcd(a, n) = 1.

  Then $ax \equiv ay \bmod n \Rightarrow x \equiv y \bmod n$.

- $x^2+1 \equiv 0 \bmod 8$ has no solution.

Equations

- $2x \equiv 5 \bmod 3$

  $\Rightarrow 2x \equiv 2 \bmod 3$

  $\Rightarrow 2^{*}\ 2\ x \equiv 2^{*}\ 2 \bmod 3$

  $\Rightarrow x \equiv 1 \bmod 3 \quad (2^{*} \equiv 2 \bmod 3)$

- $3x \equiv 7 \bmod 5$

  $\Rightarrow 3x \equiv 2 \bmod 5$

  $\Rightarrow 3^{*}\ 3x \equiv 3^{*}\ 2 \bmod 5$

  $\Rightarrow x \equiv 4 \bmod 5 \quad (3^{*} \equiv 2 \bmod 5)$

Summary on Congruence

- Notation: $a \equiv b \pmod{m}$
  - Rephrased: $m \mid a\text{-}b$
  - Rephrased: $a \bmod m = b$
  - Rephrased: $a = b + mk_1$ for some integer k_1
- Every integer is either of the form
  - 4k, 4k+1, 4k+2, or 4k+3.
  - 0 mod 4, 1 mod 4, 2 mod 4, or 3 mod 4
- If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then
  - $a+c \equiv (b+d) \pmod{m}$
  - $ac \equiv bd \pmod{m}$
- Suppose that gcd(a, n) =1. Then a has an arithmetic inverse a* modulo n, i.e. $a\, a^* \equiv a^*\, a \equiv 1 \bmod n$.

Cute Exercise

- A number is divisible by 3, if sum of the all digits is divisible by 3. Why does this work?

Z_n, Z_n^*

- $Z_n = \{0, 1, 2, 3, \cdots, n-1\}$
- $Z_n^* = \{x \mid x \in Z_n \text{ and } \gcd(x, n) = 1\}$.

- $Z_6 = \{0, 1, 2, 3, 4, 5\}$
- $Z_6^* = \{1, 5\}$

- For a set S, |S| means the number of element in S.

- $|Z_n| = n$
- $|Z_n^*| = \phi(n)$

Cardinality

- For finite (only) sets, cardinality is the number of elements in the set

- For finite and infinite sets, two sets *A* and *B* have the same cardinality if there is a one-to-one correspondence from *A* to *B*

Counting

- Multiplication rule
  - If there are n_1 ways to do task1, and n_2 ways to do task2
    - Then there are $n_1 n_2$ ways to do both tasks in sequence.
  - Example
    - There are 18 math majors and 325 CS majors
    - How many ways are there to pick one math major and one CS major?
- Addition rule
  - If there are n_1 ways to do task1, and n_2 ways to do task2
    - If these tasks can be done at the same time, then…
    - Then there are $n_1 + n_2$ ways to do one of the two tasks
  - How many ways are there to pick one math major or one CS major?
- The inclusion-exclusion principle
  - $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$

Permutation, Combination

- An r-permutation is an ordered arrangement of r elements of the set: P(n, r), ${}_nP_r$
  - How many poker hands (with ordering)?
  - $P(n, r) = n\,(n-1)(n-2)\cdots(n-r+1)$
    $= n!\,/\,(n-r)!$
- Combination: When order does not matter…
  - In poker, the following two hands are equivalent:
    - » A♦, 5♥, 7♣, 10♠, K♠
    - » K♠, 10♠, 7♣, 5♥, A♦
  - The number of r-combinations of a set with n elements, where n is non-negative and $0 \le r \le n$ is:
    $C(n, r) = n!\,/\,(r!\,(n-r)!)$
  - $(x+y)^n$

Probability definition

- The probability of an event occurring is:
  $p(E) = |E| / |S|$
  - Where E is the set of desired events (outcomes)
  - Where S is the set of all possible events (outcomes)
  - Note that $0 \leq |E| \leq |S|$
    - » Thus, the probability will always between 0 and 1
    - » An event that will never happen has probability 0
    - » An event that will always happen has probability 1

What's behind door number three?

- The Monty Hall problem paradox
  - Consider a game show where a prize (a car) is behind one of three doors
  - The other two doors do not have prizes (goats instead)
  - After picking one of the doors, the host (Monty Hall) opens a different door to show you that the door he opened is not the prize
  - Do you change your decision?
- Your initial probability to win (i.e. pick the right door) is 1/3
- What is your chance of winning if you change your choice after Monty opens a wrong door?
- After Monty opens a wrong door, if you change your choice, your chance of winning is 2/3
  - Thus, your chance of winning doubles if you change
  - Huh?

Assigning Probability

- S: Sample space
- p(s): probability that s happens.
  - $0 \le p(s) \le 1$ for each $s \in S$
  - $\sum_{s \in S} p(s) = 1$
- The function p is called probability distribution
- Example
  - Fair coin: p(H) = 1/2, p(T) = 1/2
  - Biased coin where heads comes up twice as often as tail
    - » p(H) = 2 p(T)
    - » $p(H) + p(T) = 1 \Rightarrow 3\,p(T) = 1 \Rightarrow p(T) = 1/3,\ p(H) = 2/3$

More…

- Uniform distribution
  - Each element $s \in S$ ($|S| = n$) is assigned with the probability $1/n$.
- Random
  - The experiment of selecting an element from a sample space with uniform distribution.
- Probability of the event E
  - $p(E) = \sum_{s \in E} p(s)$.
- Example
  - A die is biased so that 3 appears twice as often as others
    - $p(1) = p(2) = p(4) = p(5) = p(6) = 1/7$, $p(3) = 2/7$
  - p(O) where O is the event that an odd number appears
    - $p(O) = p(1) + p(3) + p(5) = 4/7$.

Combination of Events

- Still
  - $p(E^c) = 1 - p(E)$
  - $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$
    - » $E_1 \cap E_2 = \varnothing \Rightarrow p(E_1 \cup E_2) = p(E_1) + p(E_2)$
    - » For all $i \neq j$, $E_i \cap E_i = \varnothing \Rightarrow p(\cup_i E_i) = \sum_i p(E_i)$

Conditional Probability

- Flip coin 3 times
  - all eight possibility are equally likely.
  - Suppose we know that the first coin was tail (Event F). What is the probability that we have odd number of tails (Event E)?
    - Only four cases: TTT, TTH, THT, THH
    - So 2/4 = 1/2.
- Conditional probability of E given F
  - We need to use F as the sample space
  - For the outcome of E to occur, the outcome must belong to $E \cap F$.
  - $p(E \mid F) = p(E \cap F) / p(F)$.

Bernoulli Trials & Binomial Distribution

- Beronoulli trial
  - an experiment with only two possible outcomes
  - i.e. 0 (failure) and 1 (success).
  - If p is the probability of success and q is the probability of failure, $p + q = 1$.
- A biased coin with probability of heads 2/3
  - What is the probability that four heads up out of 7 trials?

Random Variable

- A random variable is a function from the sample space of an experiment to the set of real numbers.
  - Random variable assigns a real number to each possible outcome.
  - Random variable is not variable! not random!
- Example: three times coin flipping
  - Let X(t) be the random variable that equals the number of heads that appear when t is the outcome
  - X(HHH) = 3, X(THH) = X(HTH) = X(HHT) = 2, X(TTH) = X(THT) = X(HTT) = 1, X(TTT) = 0
- The distribution of a random variable X on a sample space S is the set of pairs (r, p(X=r)) for all $r \in X(S)$
  - where p(X=r) is the probability that X takes value r.
  - p(X=3) = 1/8, p(X=2) = 3/8, p(X=1) = 3/8, p(X=0) = 1/8

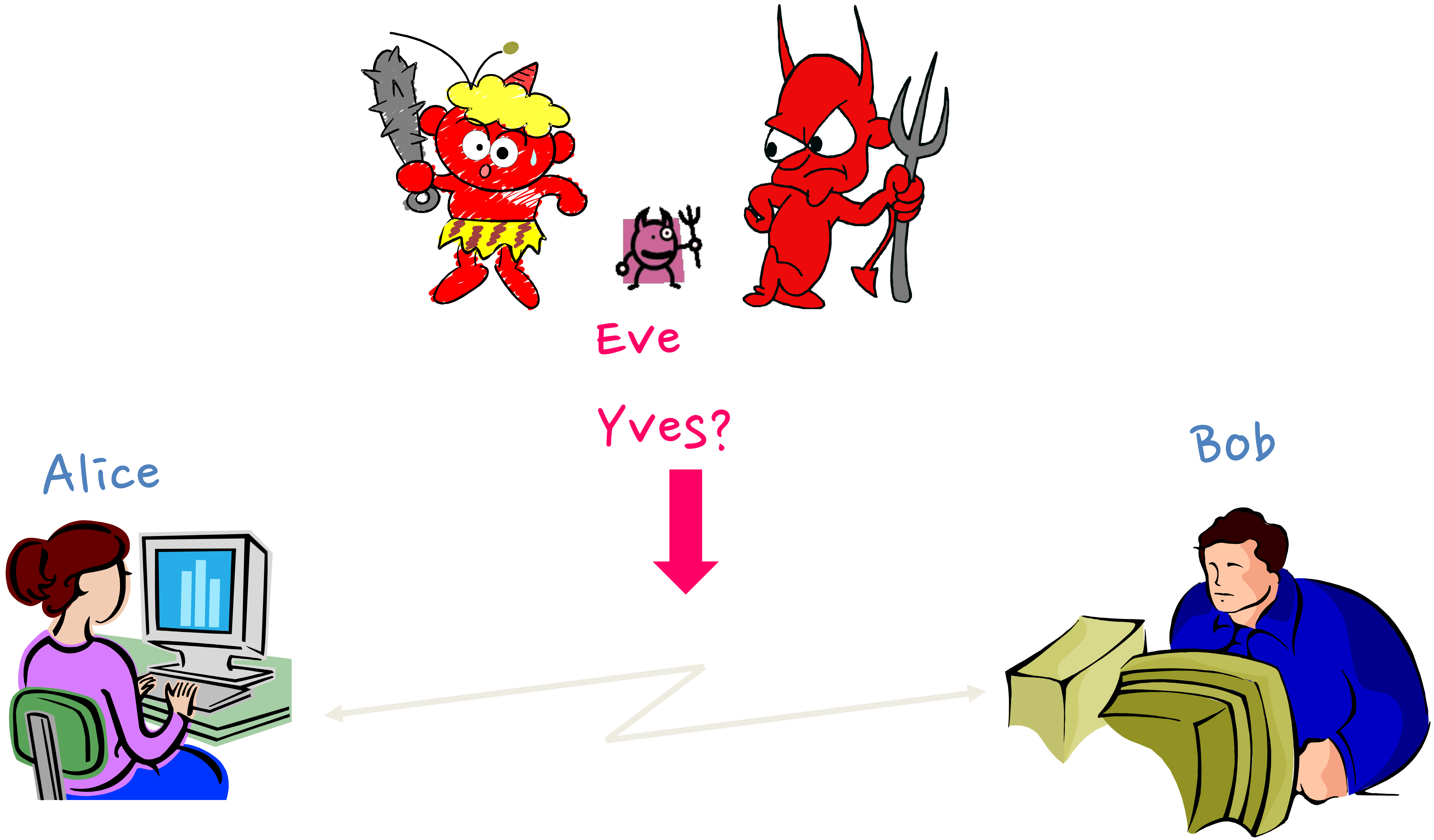
Expected Value

- The expected value of the random variable X(s) on the sample space S is equal to

  $E(X) = \sum_{s \in S} p(s) X(s)$
- Expected value of a Die
  - X is the number that comes up when a die is rolled.
  - What is the expected value of X?
  - $E(X) = 1/6\ 1 + 1/6\ 2 + 1/6\ 3 + \cdots 1/6\ 6 = 21/6 = 7/2$
- Three times coin flipping example
  - X: number of heads
  - $E(X) = 1/8\ 3 + 3/8\ 2 + 3/8\ 1 + 1/8\ 0 = 12/8 = 3/2$

Security: Overview

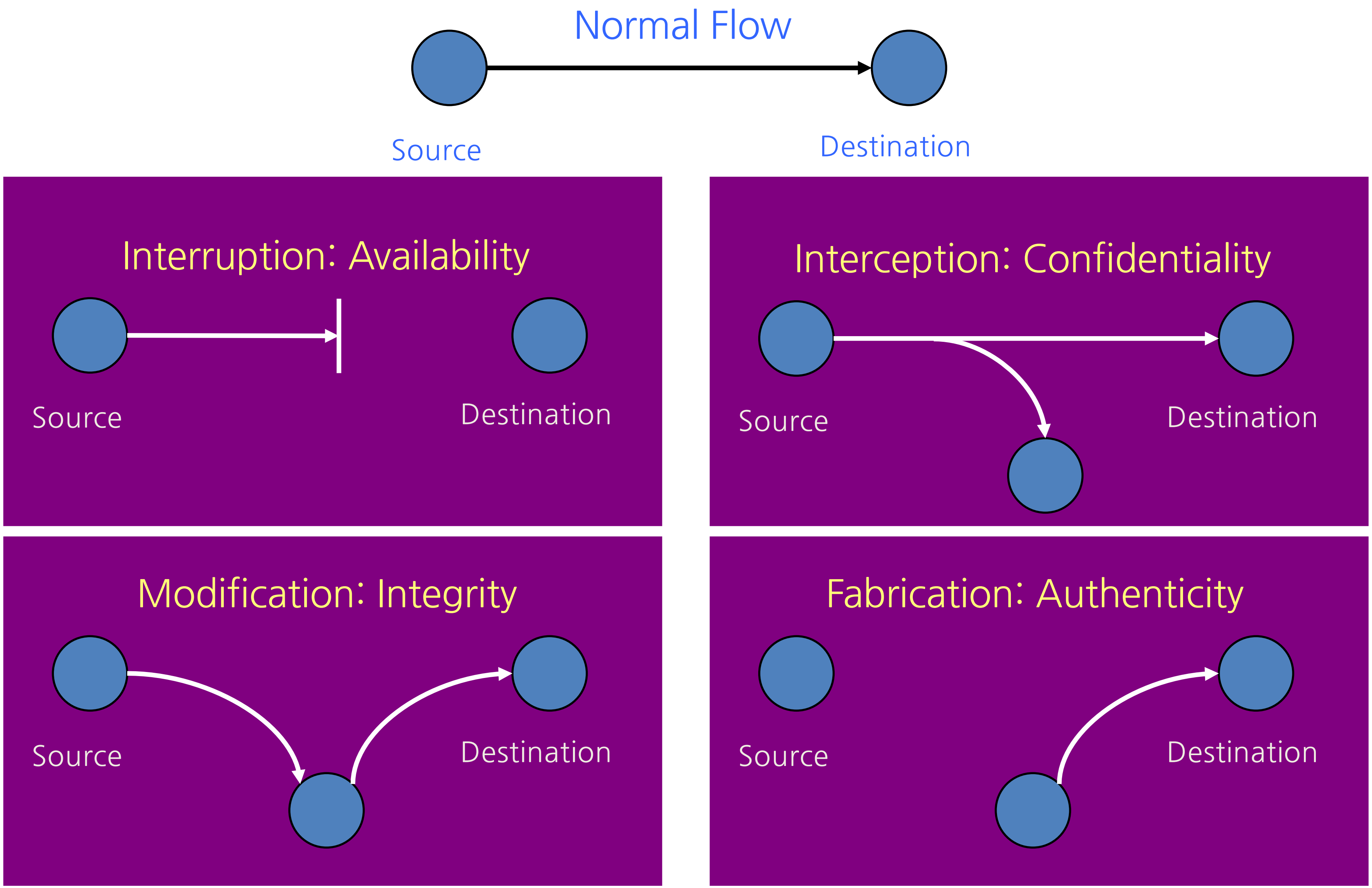
The main players



Attacks, Mechanisms, Services

- Security Attack: Any action that compromises the security of information.
- Security Mechanism: A mechanism that is designed to detect, prevent, or recover from a security attack.
- Security Service: A service that enhances the security of data processing systems and information transfers. A security service makes use of one or more security mechanisms.

Attacks


Normal Flow
Source
Destination
Interruption: Availability
Source
Destination
Interception: Confidentiality
Source
Destination
Modification: Integrity
Source
Destination
Fabrication: Authenticity
Source
Destination

Taxonomy of Attacks

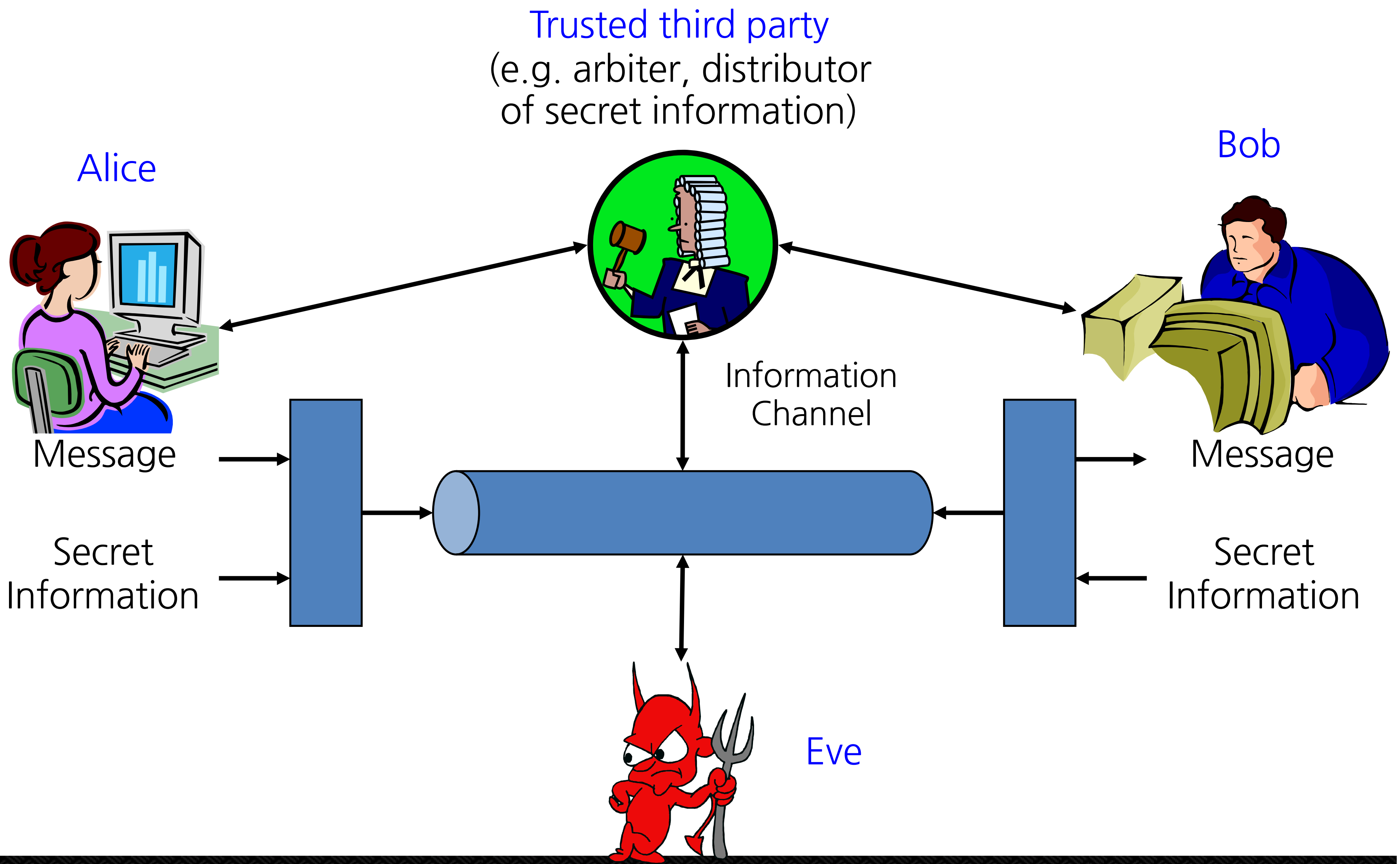
- Passive attacks
  - Eavesdropping
  - Traffic analysis

- Active attacks
  - Masquerade
  - Replay
  - Modification of message content
  - Denial of service

Security Services

- Confidentiality or privacy
  - keeping information secret from all but those who are authorized to see it.
- Data Integrity
  - ensuring information has not been altered by unauthorized or unknown means.
- Entity authentication or identification
  - corroboration of the identity of an entity
- Message authentication
  - corroborating the source of information
- Signature
  - a means to bind information to an entity.
- Authorization, Validation, Access control, Certification, Timestamping, Witnessing, Receipt, Confirmation, Ownership, Anonymity, Non-repudiation, Revocation

Big picture


Trusted third party
(e.g. arbiter, distributor
of secret information)
Alice
Bob
Information
Channel
Message
Secret
Information
Message
Secret
Information
Eve

More details

- Little maths
- Taxonomy
- Definitions

Little Maths :-)

- Function
  - $f : X \rightarrow Y$ is called a function f from set X to set Y.
    - X: domain, Y: codomain.
  - for $y = f(x)$ where $x \in X$ and $y \in Y$
    - y: image of x, x: preimage of y
  - Im(f): the set that all $y \in Y$ have at least one preimage

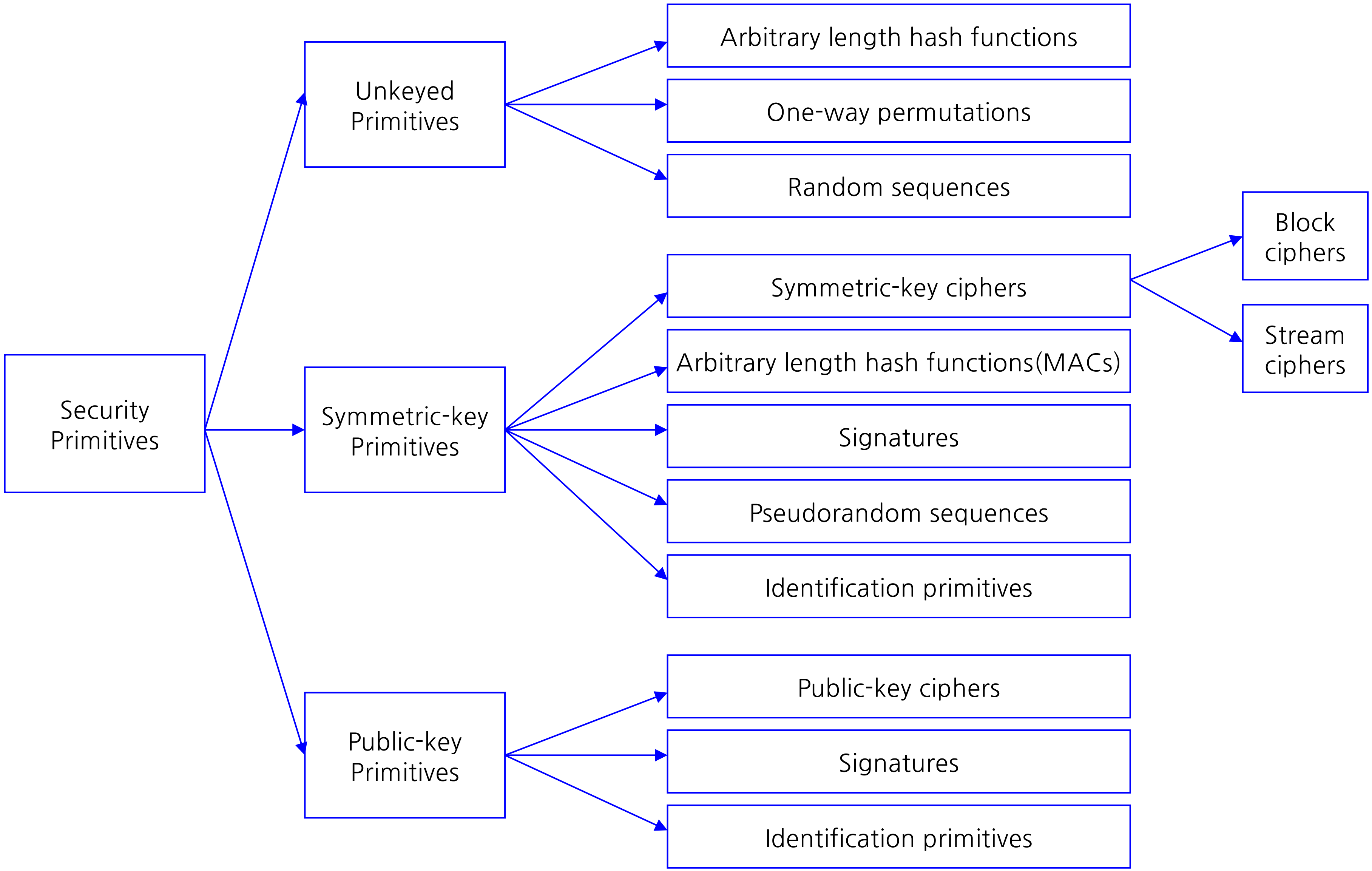
- 1 – 1 if each element in Y is the image of at most one element in X.
- onto if Im(f) =Y
- bijection if f is 1–1 and onto.

(Trap-door) One-way function

- one-way function if
  - f(x) is easy to compute for all $x \in X$, but
  - it is computationally infeasible to find any $x \in X$ such that $f(x) = y$.

- trapdoor one-way function if
  - given trapdoor information, it becomes feasible to find an $x \in X$ such that $f(x) = y$.

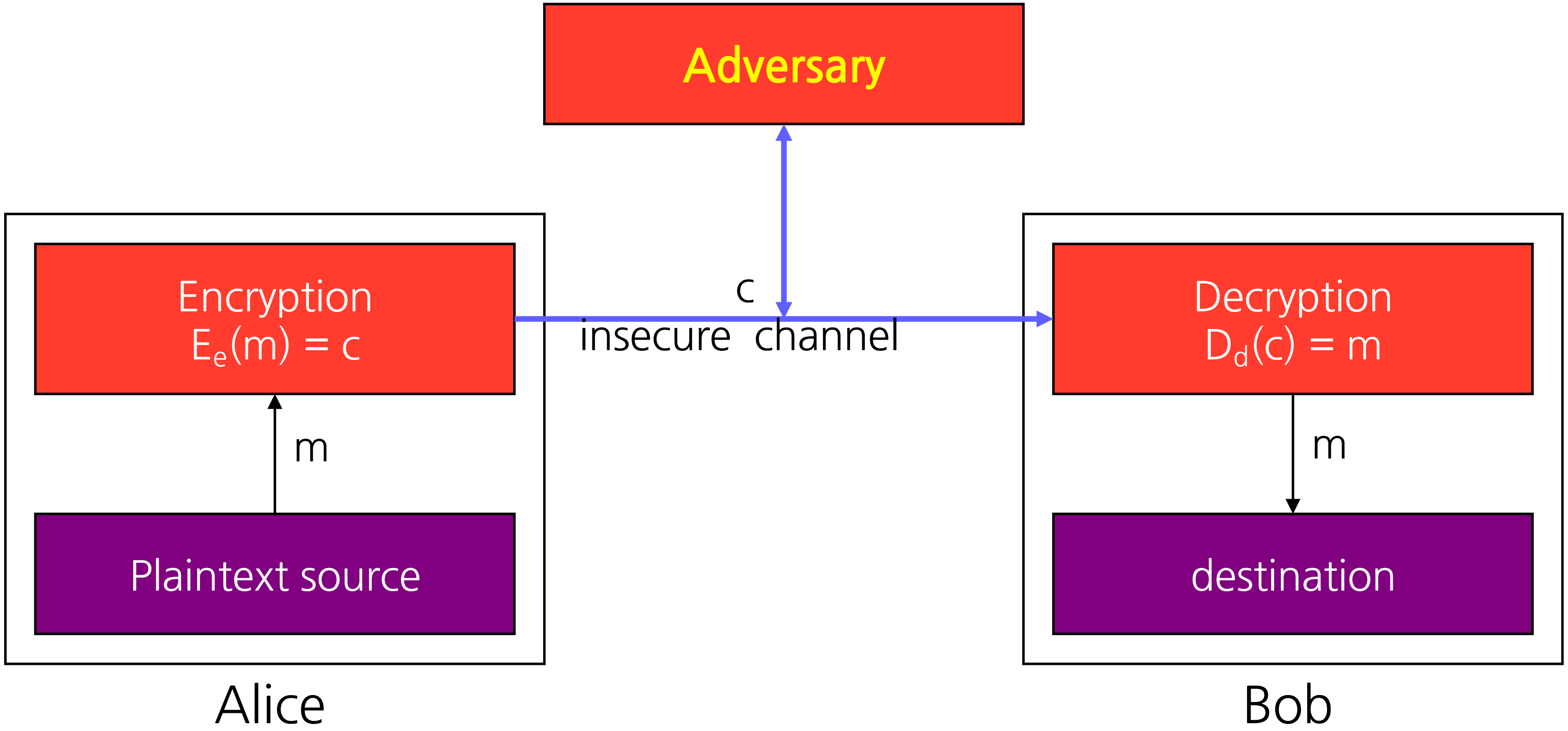
Taxonomy of crypto primitives



Terminology for Encryption

- M denotes a set called the *message space*
  - M consists of strings of symbols from an alphabet
  - An element of M is called a *plaintext*
- C denotes a set called the *ciphertext space*
  - C consists of strings of symbols from an alphabet
  - An element of C is called a *ciphertext*
- K denotes a set called the *key space*
  - An element of K is called a *key*
- E_e is an *encryption function* where $e \in K$
- D_d called a *decryption function* where $d \in K$

Encryption

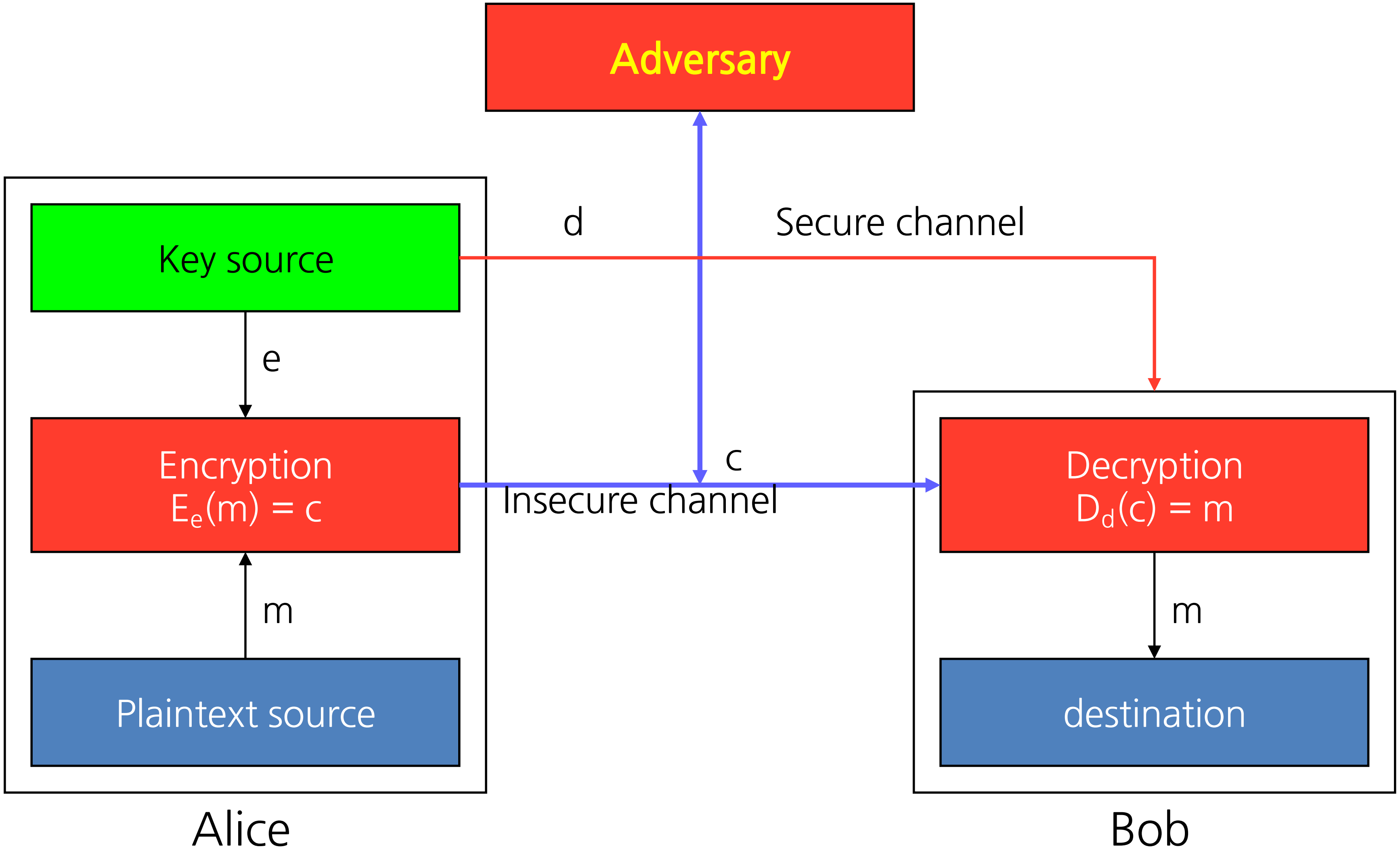


- Why do we use key?
  - Or why not use just a shared encryption function?

Symmetric-key encryption

- Encryption scheme is symmetric-key
  - if for each (e,d) it is easy computationally easy to compute e knowing d and d knowing e
  - Usually e = d
- Block Cipher
  - Breaks plaintext into block of fixed length
  - Encrypts one block at a time
- Stream Cipher
  - Takes a plaintext string and produces a ciphertext string using keystream
  - Block cipher with block length 1

SKE with Secure channel



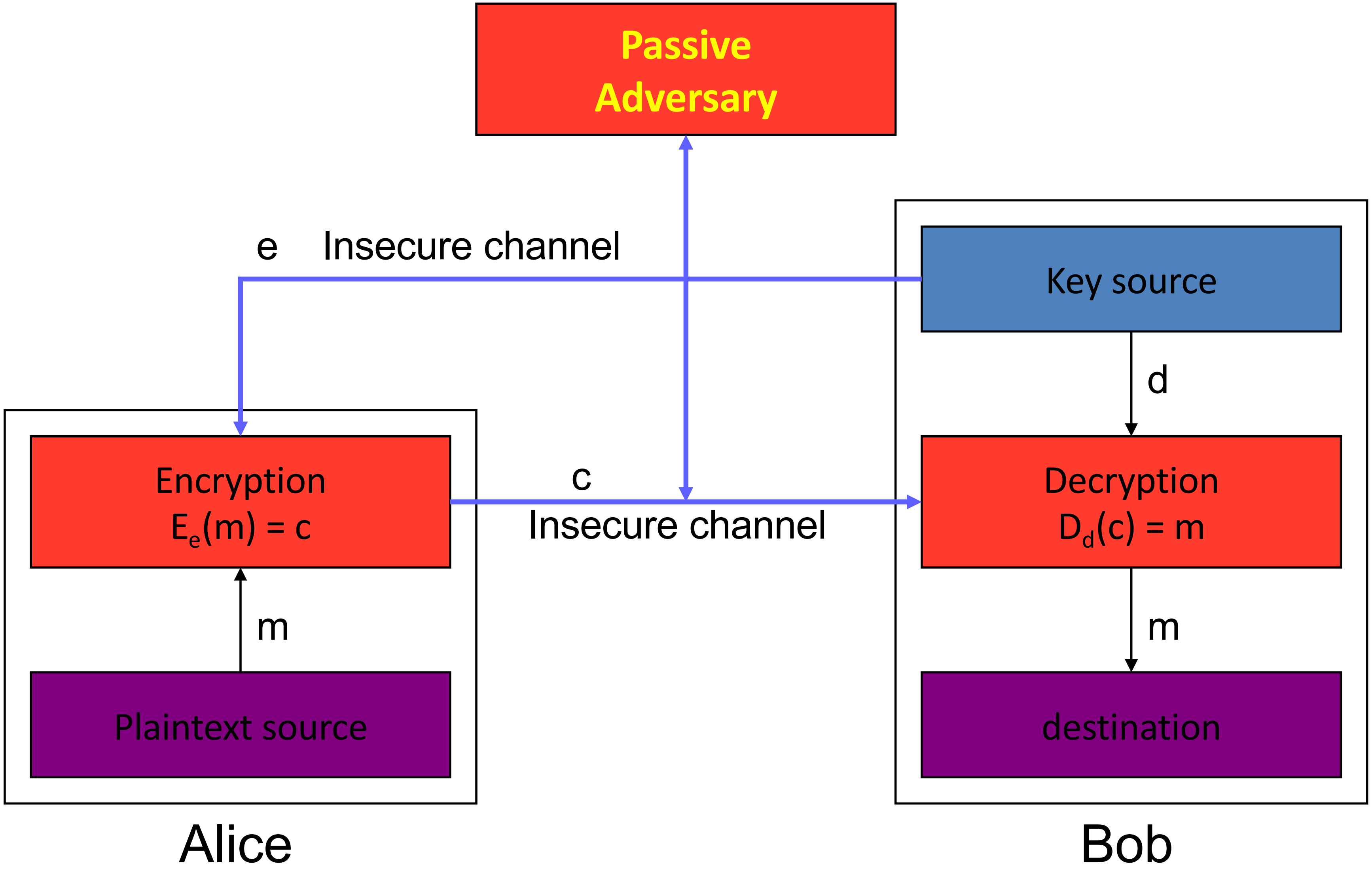
SysSec System Security Lab

Public-key Encryption (Crypto)

- Every entity has a private key SK and a public key PK
  - Public key is known to all
  - It is computationally infeasible to find SK from PK
  - Only SK can decrypt a message encrypted by PK

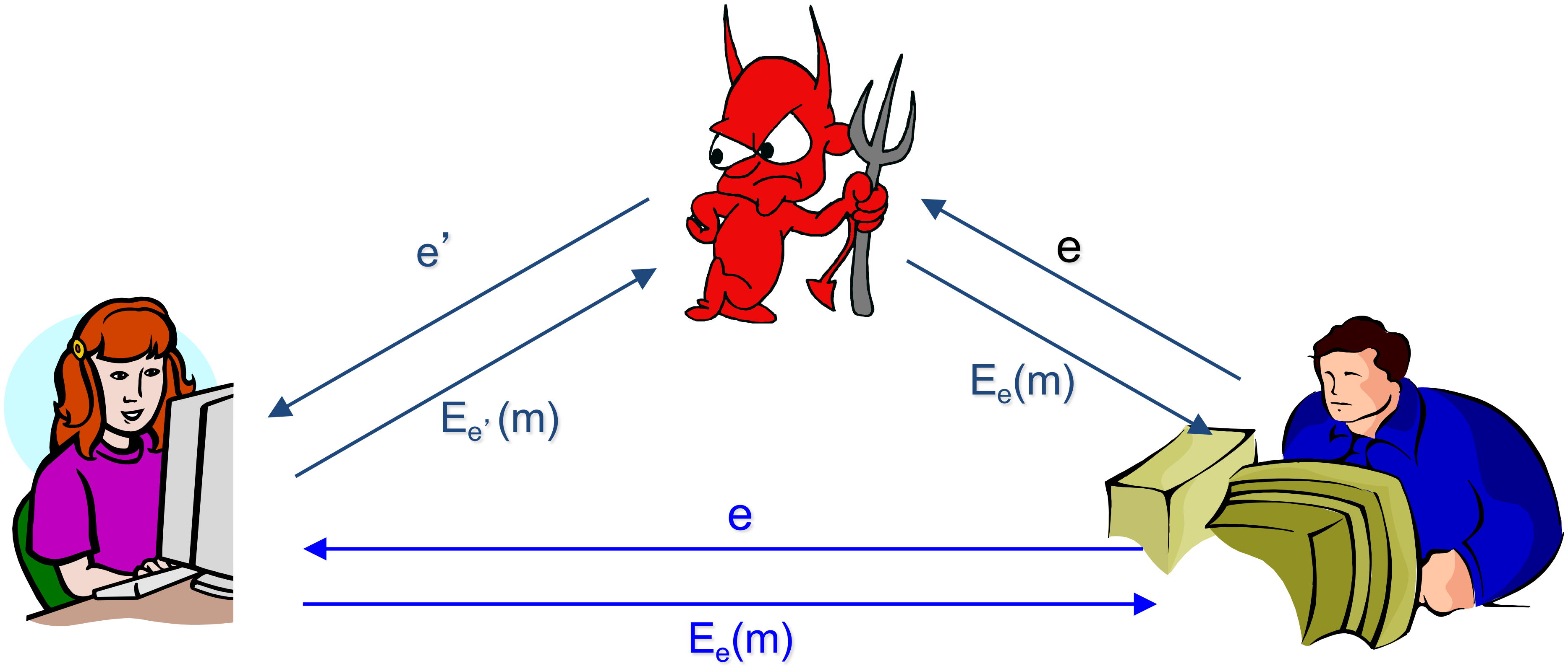
- If A wishes to send a private message M to B
  - A encrypts M by B's public key, $C = EB_{PK}(M)$
  - B decrypts C by his private key, $M = DB_{SK}(C)$

PKE with Insecure Channel



SysSec
System Security Lab

Public Key should be authentic!



Digital Signatures

- Primitive in authentication and non-repudiation
- Signature
  - Process of transforming the message and some secret information into a tag
- Nomenclature
  - M is set of messages
  - S is set of signatures
  - S_A is signature transformation from M to S for A, kept private
  - V_A is verification transformation from M to S for A, publicly known

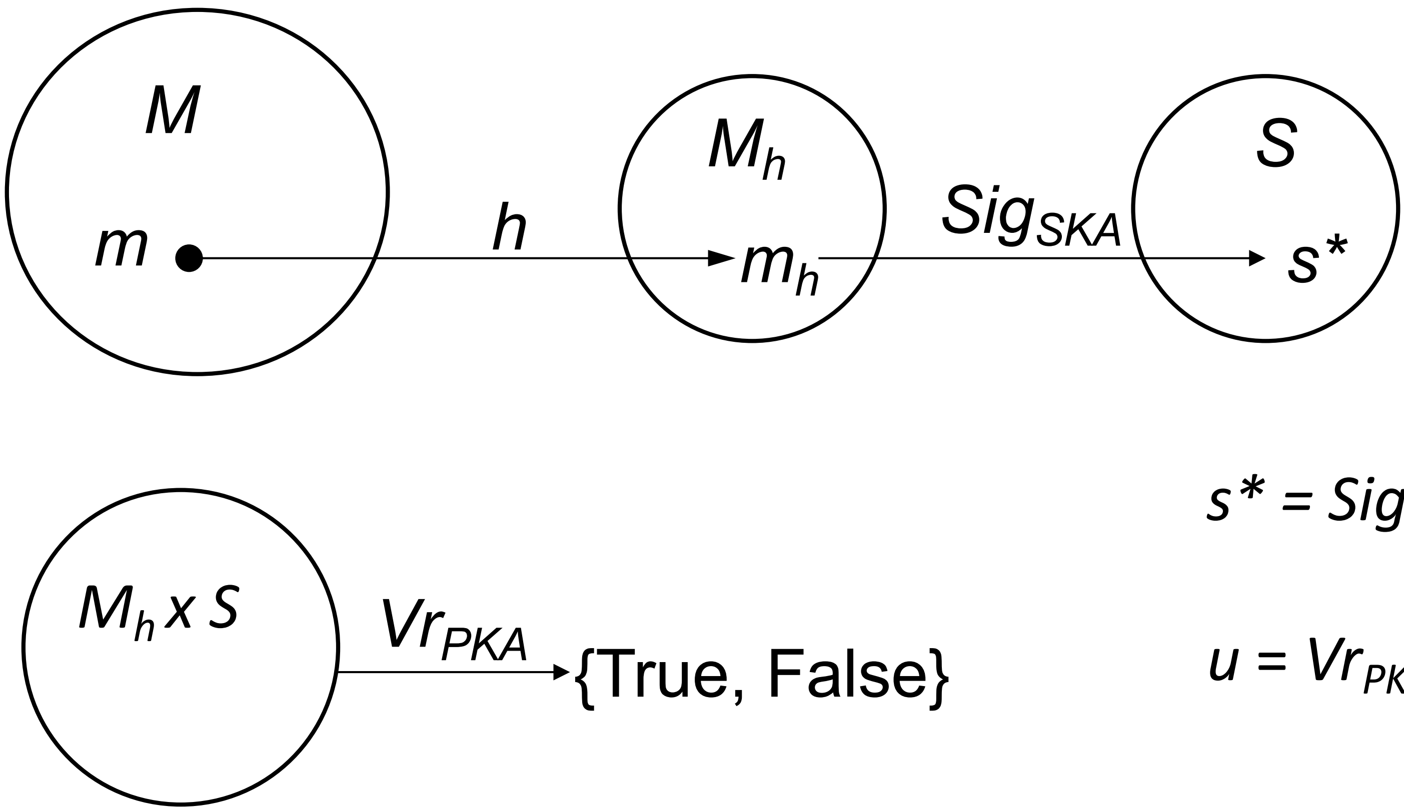
Definitions

- Digital Signature - a data string which associates a message with some originating entity
- Digital Signature Generation Algorithm - a method for producing a digital signature
- Digital signature verification algorithm - a method for verifying that a digital signature is authentic (i.e., was indeed created by the specified entity).
- Digital Signature Scheme - consists of a signature generation algorithm and an associated verification algorithm

Digital Signature with Appendix

- Schemes with appendix
  - Requires the message as input to verification algorithm
  - Rely on cryptographic hash functions rather than customized redundancy functions
  - DSA, ElGamal, Schnorr etc.

Digital Signature with Appendix



$s^* = Sig_{SKA}(m_h)$

$u = Vr_{PKA}(m_h, s^*)$

Hash function and MAC

- A hash function is a function h
  - compression — h maps an input x of arbitrary finite bitlength, to an output h(x) of fixed bitlength n.
  - ease of computation — h(x) is easy to compute for given x and h
  - Properties
    - one-way: for a given y, find x' such that h(x') = y
    - collision resistance: find x and x' such that h(x) = h(x')

- MAC (message authentication codes)
  - both authentication and integrity
  - MAC is a family of functions h_k
    - ease of computation (if k is known !!)
    - compression, x is of arbitrary length, $h_k(x)$ has fixed length
    - computation resistance: given $(x', h_k(x'))$ it is infeasible to compute a new pair $(x, h_k(x))$ for any new $x \neq x'$

Authentication

- How to prove your identity?
  - Prove that you know a secret information
- When key K is shared between A and Server
  - A ➔ S: $HMAC_K(M)$ where M can provide freshness
  - Why freshness?
- Digital signature?
  - A ➔ S: $Sig_{SK}(M)$ where M can provide freshness
- Comparison?

Key Management Through SKE

- Each entity A_i shares symmetric key K_i with a TTP
- TTP generates a session key K_s and sends $E_{K_i}(K_s)$
- Pros
  - Easy to add and remove entities
  - Each entity needs to store only one long-term secret key
- Cons
  - Initial interaction with the TTP
  - TTP needs to maintain *n* long-term secret keys
  - TTP can read all messages
  - Single point of failure

Authentication

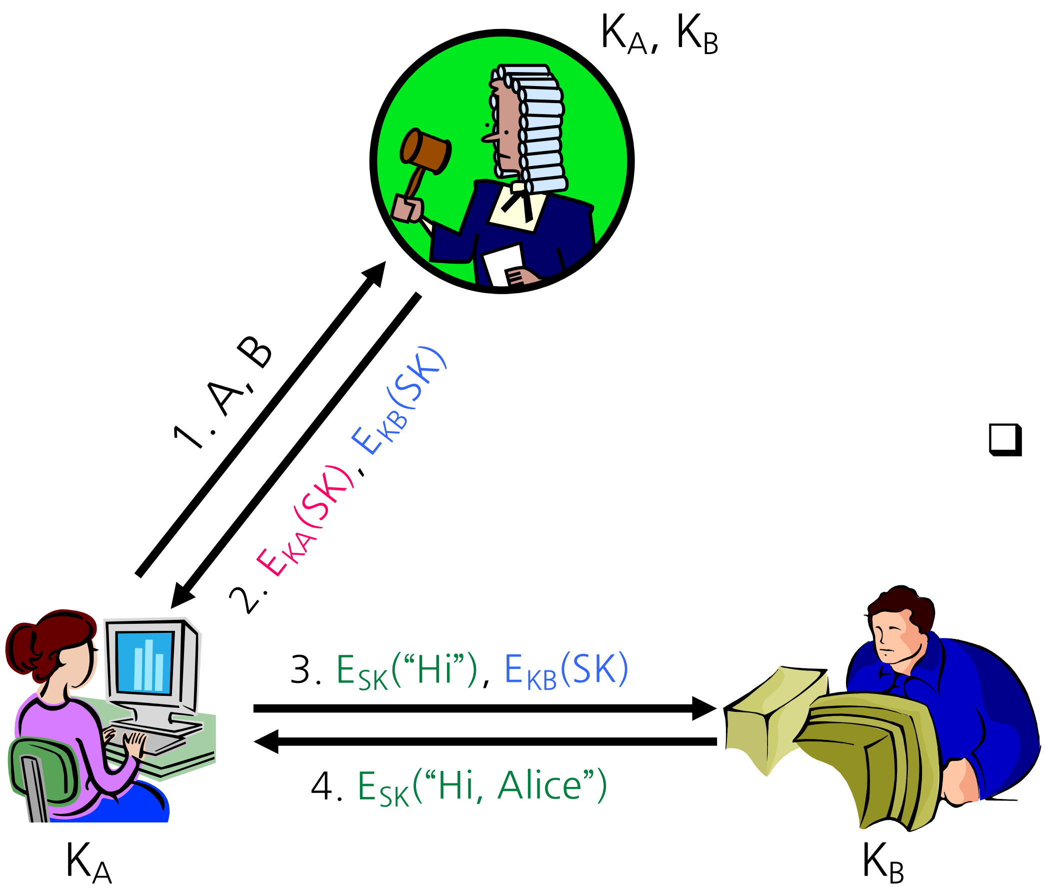
- Authentication
  - Message (Data origin) authentication
    - provide to one party which receives a message assurance of the identity of the party which originated the message.
  - Entity authentication (identification)
    - one party of both the identity of a second party involved, and that the second was active at the time the evidence was created or acquired.

Key Management

- Key establishment
  - Process to whereby a shared secret key becomes available to two or more parties
  - Subdivided into key agreement and key transport.

- Key management
  - The set of processes and mechanisms which support key establishment
  - The maintenance of ongoing keying relationships between parties

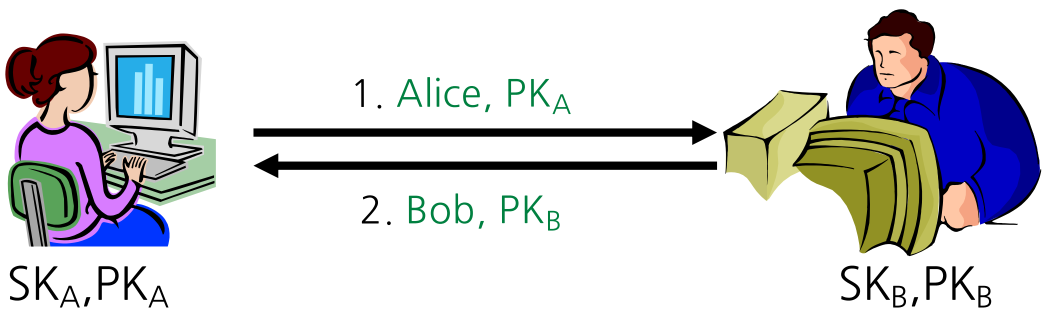
Key Management Through SKE



- Pros
  - Easy to add and remove entities
  - Each entity needs to store only one long-term secret key
- Cons
  - Initial interaction with the TTP
  - TTP needs to maintain *n* long-term secret keys
  - TTP can read all messages
  - Single point of failure

Key Management Through PKE

0xDAD12345	Alice
0xBADD00D1	Bob



- Advantages
  - TTP not required
  - Only *n* public keys need to be stored
  - The central repository could be a local file
- Problem
  - Public key authentication problem
- Solution
  - Need of TTP to certify the public key of each entity

Public Key Certificates

- Entities trust a third party, who issues a certificate
- Certificate = (data part, signature part)
  - Data part = (name, public-key, other information)
  - Signature = (signature of TTP on data part)
- If B wants to verify authenticity of A's public key
  - Acquire public key certificate of A over a secured channel
  - Verify TTP's signature
  - If signature verified A's public key in the certificate is authentic

Questions?

- Yongdae Kim
  - email: yongdaek@kaist.ac.kr
  - Home: http://syssec.kaist.ac.kr/~yongdaek
  - Facebook: https://www.facebook.com/y0ngdaek
  - Twitter: https://twitter.com/yongdaek
  - Google "Yongdae Kim"