EE817/IS 893

Cryptography Engineering and Cryptocurrency

Yongdae Kim

한국과학기술원

Admin Stuff

- Mar 13 midnight: Homework 1 submission
- Mar 14 morning: Homework 1 solution posting
- Mar 19 class: Quiz 1

- About 2 weeks after: Homework 2, Quiz 2

- About 2 weeks after: Homework 3, midterm, …

Recap

- Math…

- Proof techniques
  - Direct/Indirect proof, Proof by contradiction, Proof by cases, Existential/Universal Proof, Forward/backward reasoning
- Divisibility: a divides b ($a|b$) if $\exists$ c such that $b = ac$
- congruences

Math, Math, Math!

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 $0 \leq r \leq n$ 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.
  - $0 \leq p(s) \leq 1$ 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
  - $P(E^c) = 1 - P(E)$
  - $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
    - $E_1 \cap E_2 = \emptyset \Longrightarrow P(E_1 \cup E_2) = P(E_1) + P(E_2)$
    - For all $i \neq j$, $E_i \cap E_i = \emptyset \Longrightarrow P(\bigcup_i E_i) = \sum_i P(E_i)$

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

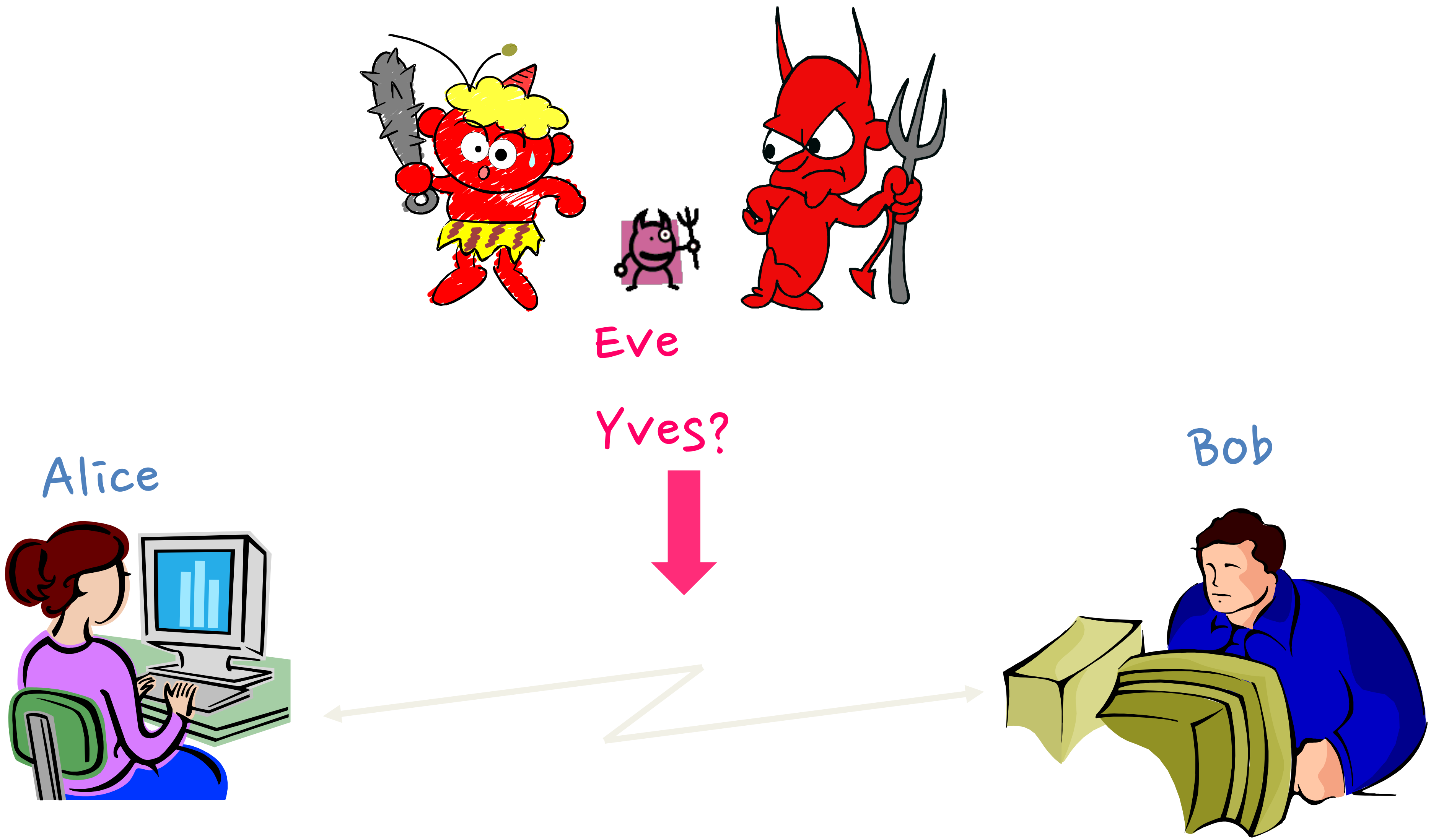
- The expected value of the random variable X(s) on the sample space S is equal to

  $E(X) = \sum_{s \in S} p(s)\, X(s)$

- 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



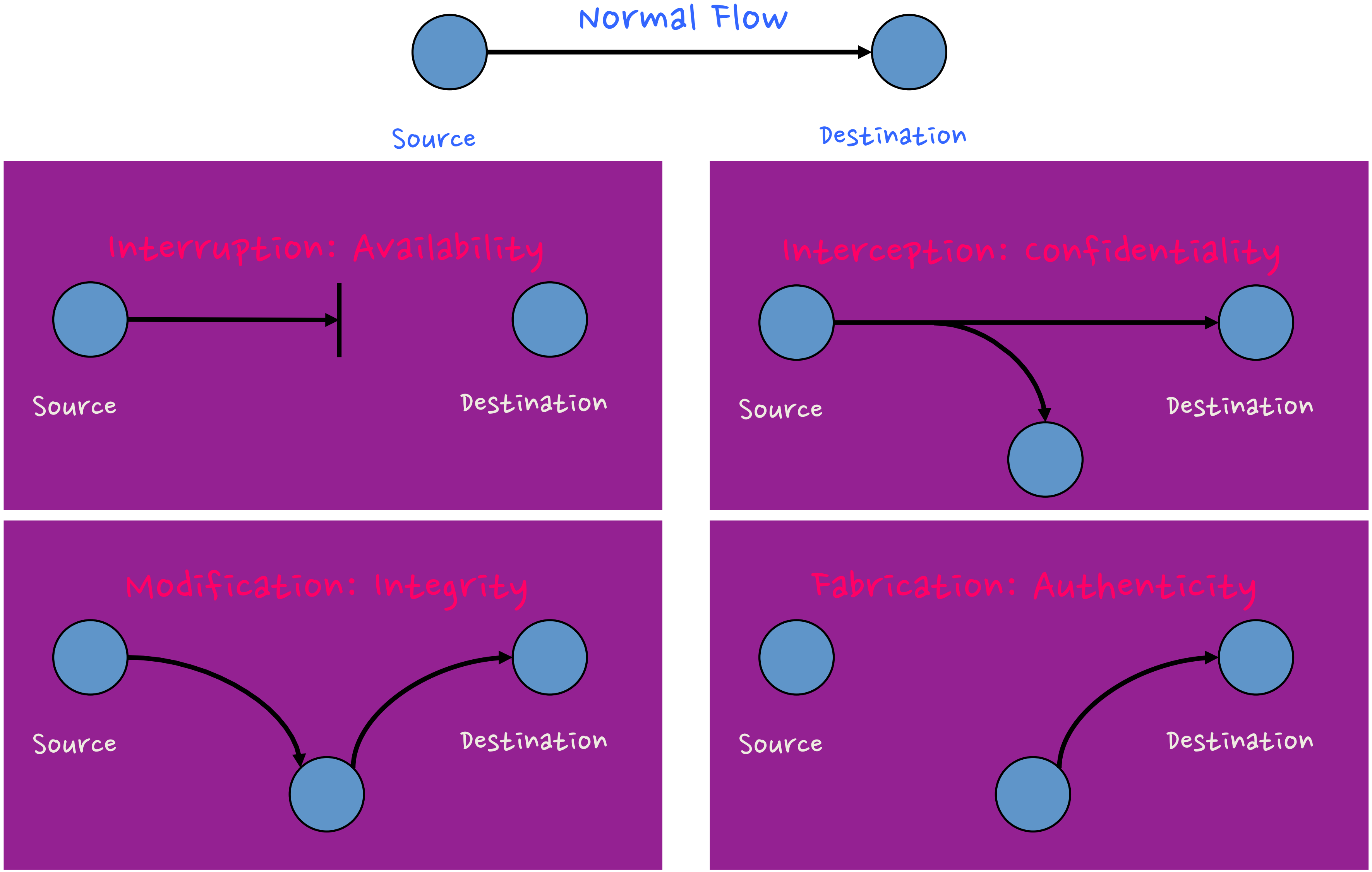
Attacks, Mechanisms, Services

- 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

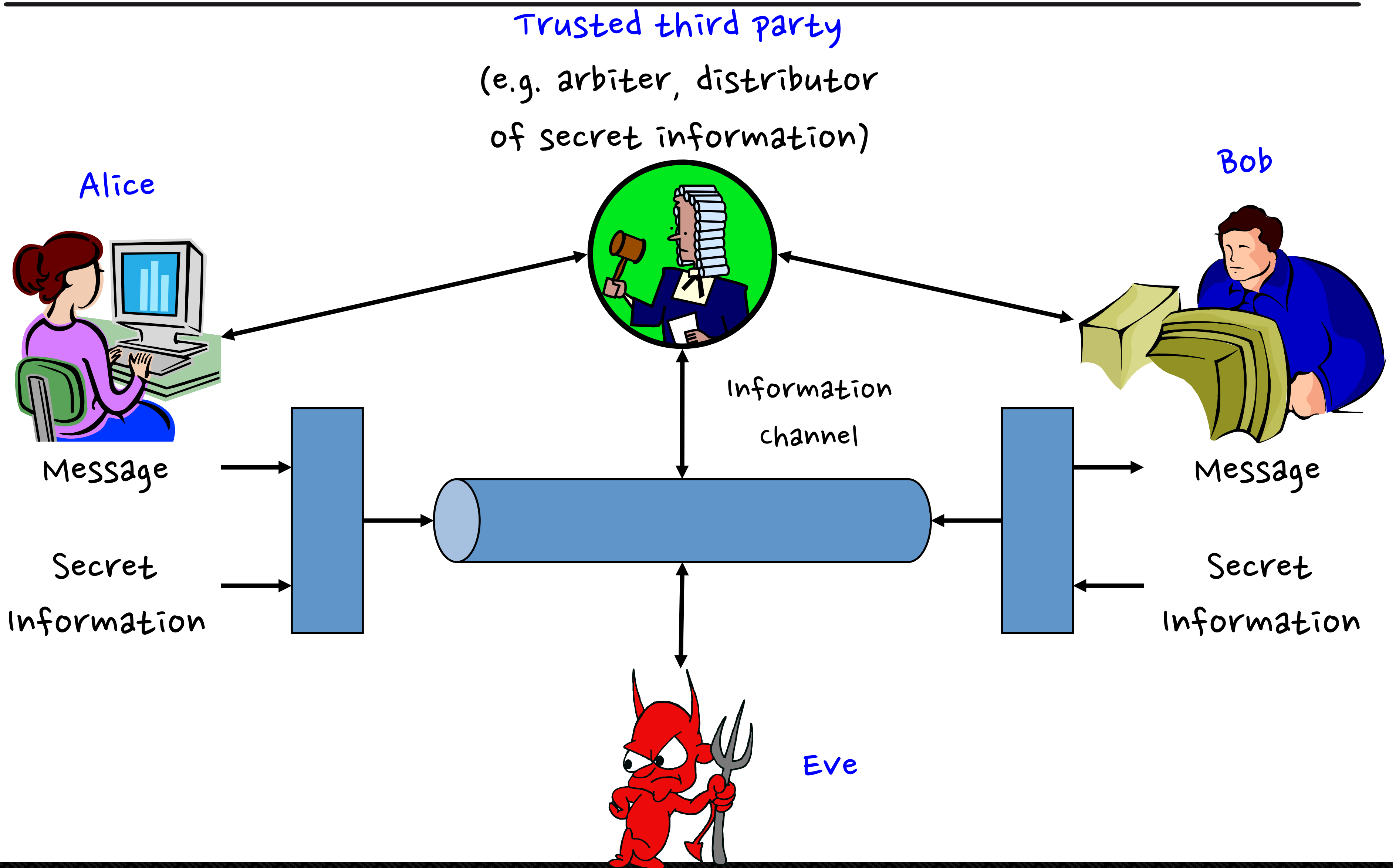
- Passive attacks
  - Eavesdropping
  - Traffic analysis

- 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

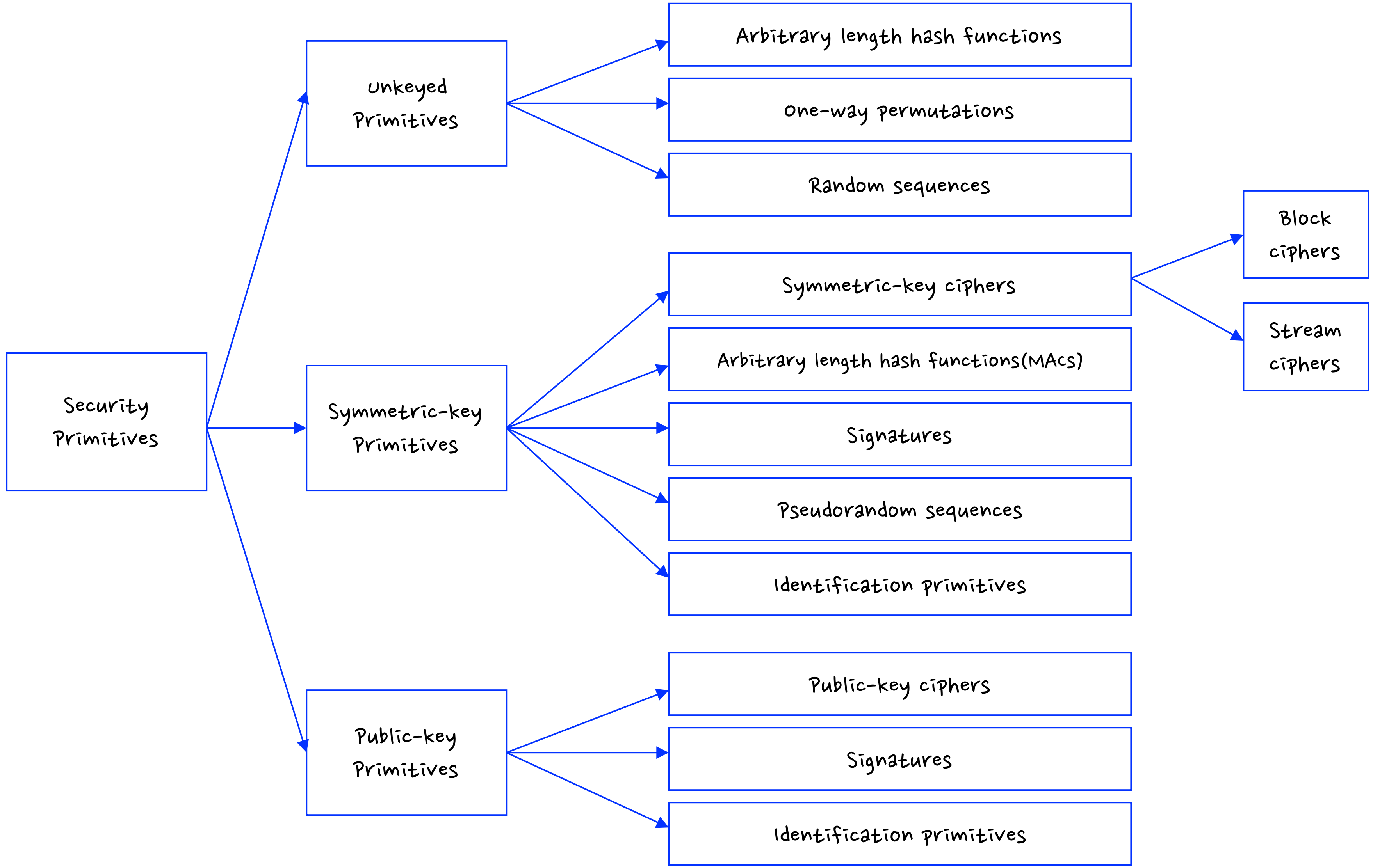
- 1 – 1 if each element in Y is the image of at most one element in X.
- onto if Im(f) =Y
- bijection if f is 1–1 and onto.

(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 $f(x) = y$.

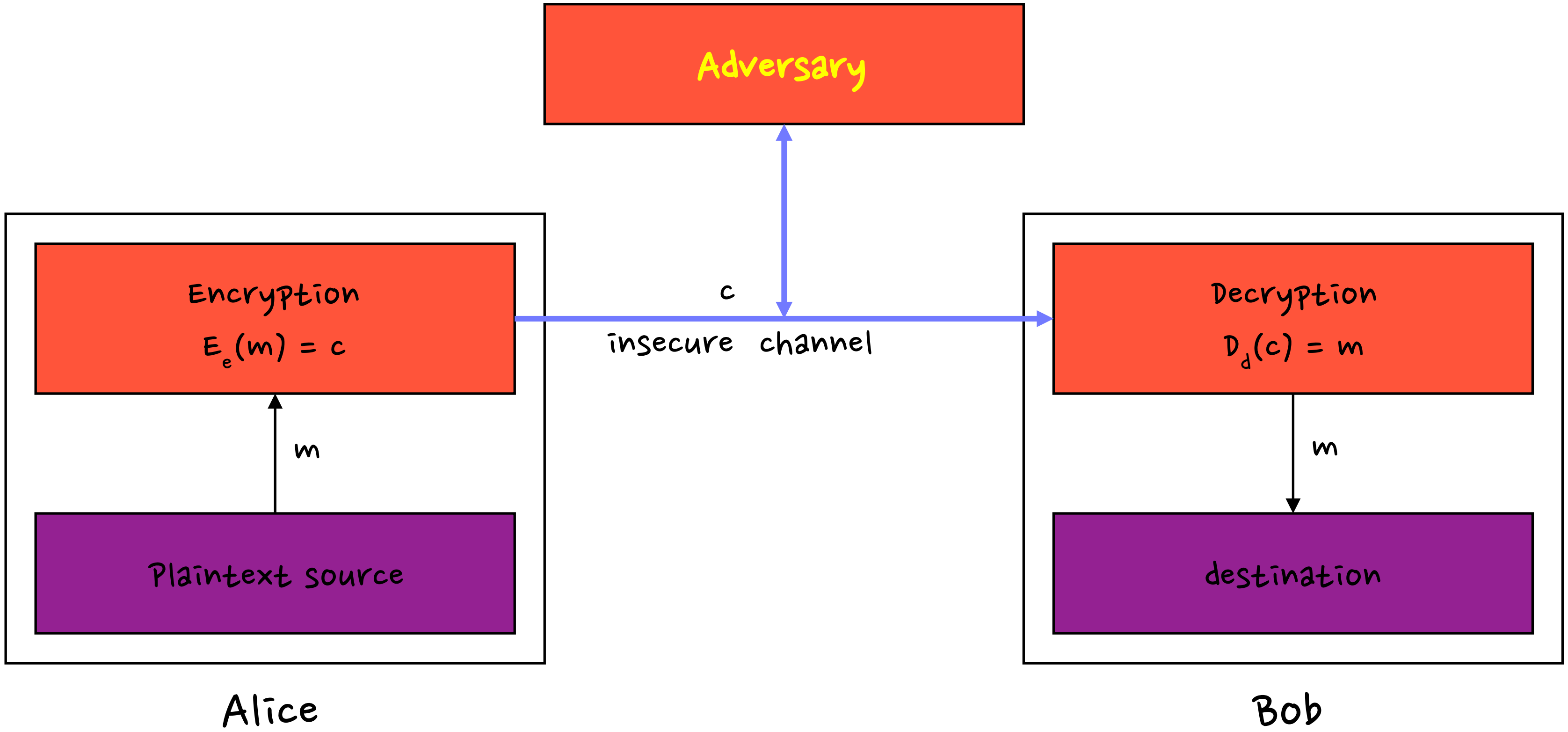
Taxonomy of crypto primitives



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

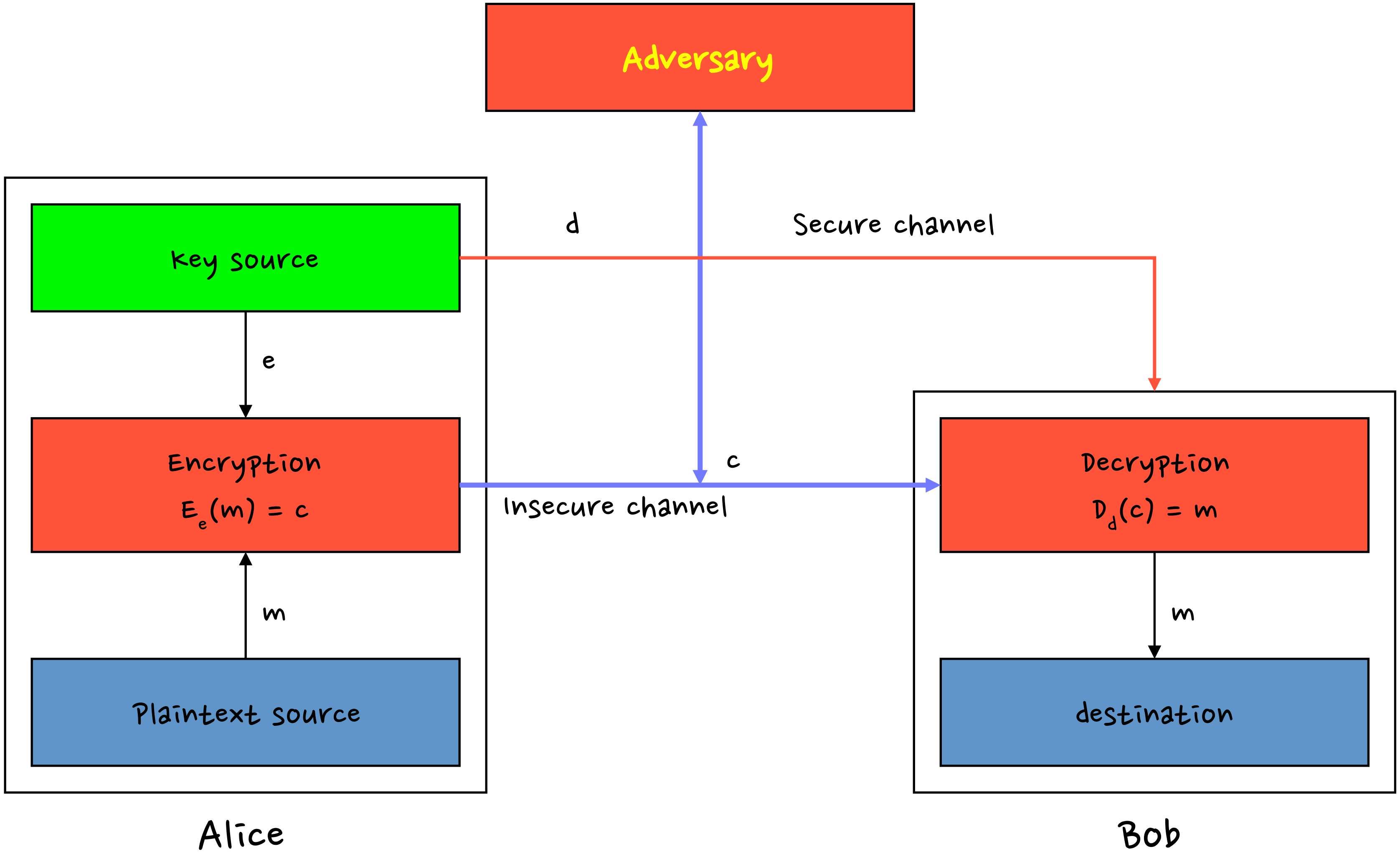


- Why do we use key?
  - Or why not use just a shared encryption function?

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

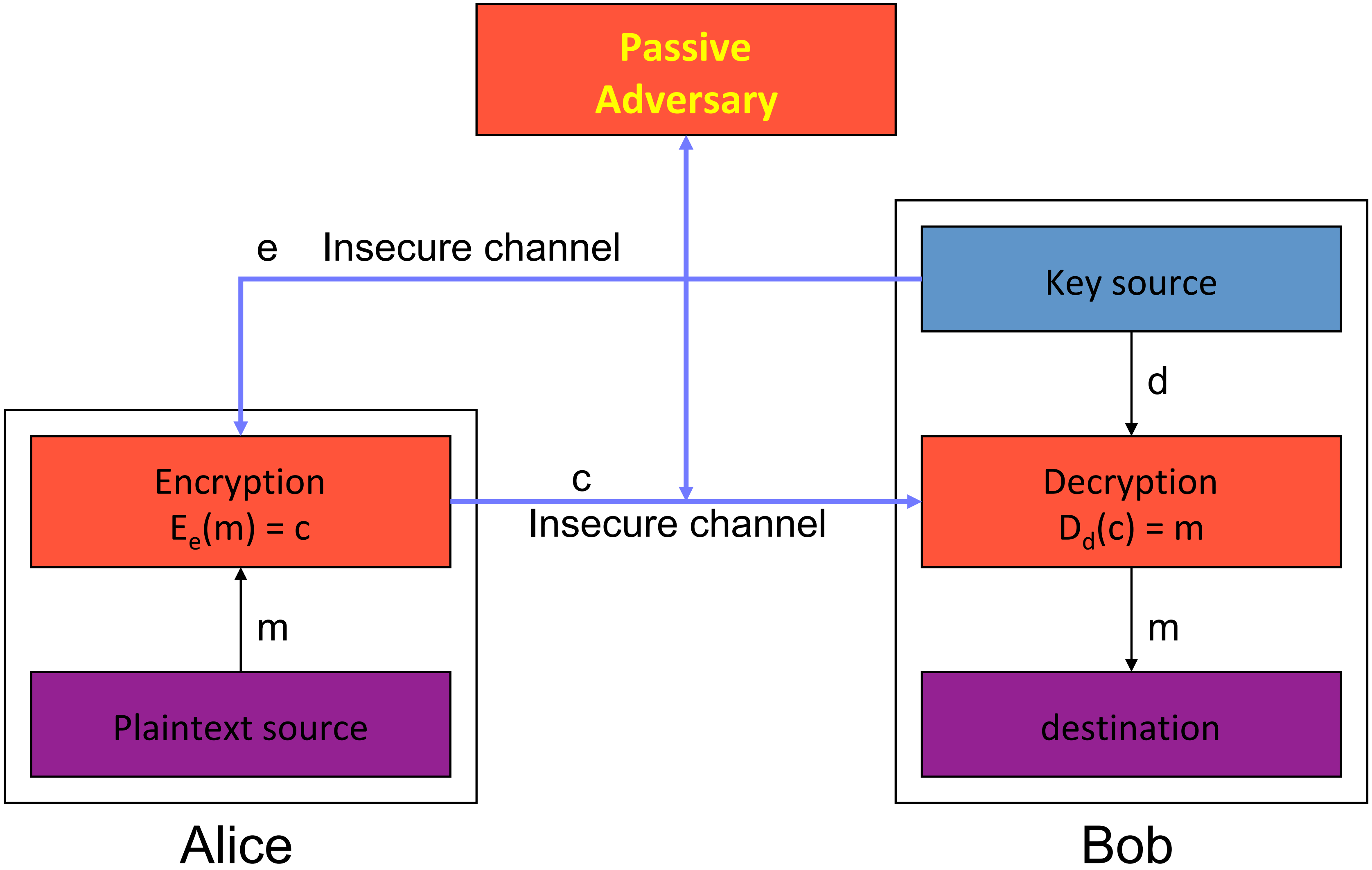


Public-key Encryption (crypto)

- 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

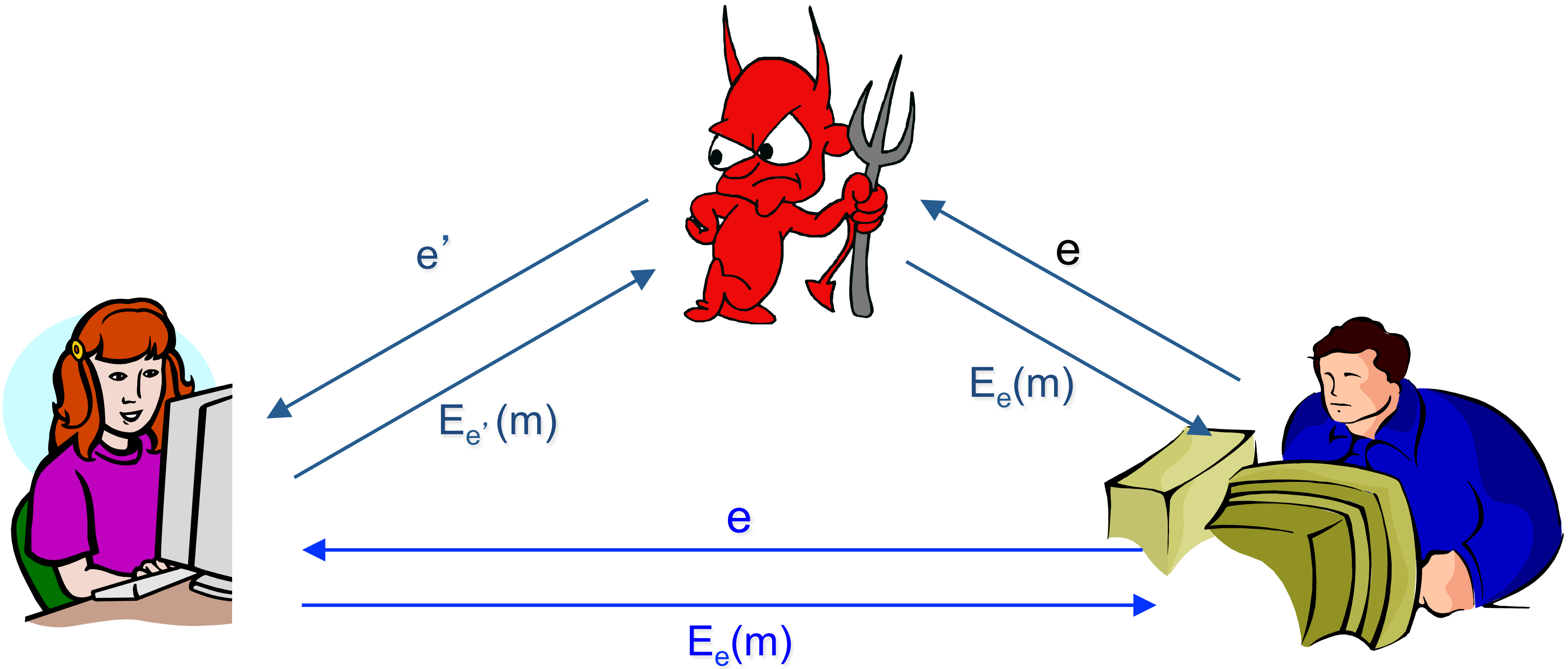
- If A wishes to send a private message M to B
  - A encrypts M by B's public key, $c = E_{B_{PK}}(M)$
  - B decrypts c by his private key, $M = D_{B_{SK}}(c)$

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

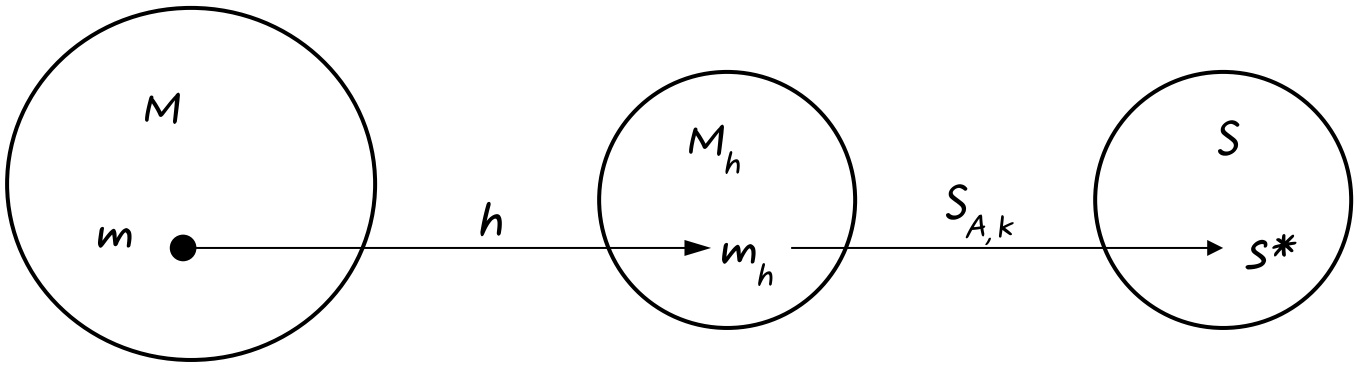
Definitions

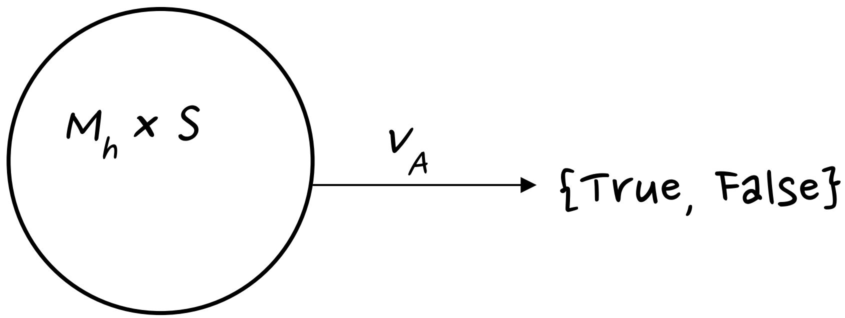
- Digital Signature – a data string which associates a message with some originating entity
- Digital Signature Generation Algorithm – a method for producing a digital signature
- Digital signature verification algorithm – a method for verifying that a digital signature is authentic (i.e., was indeed 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





$$s^* = S_{A,k}(m_h)$$

$$u = V_A(m_h, s^*)$$

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 $x \neq x'$

Message Authentication code MAc

- 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 $x \neq x'$

- Typical use
  - A ➔ B: $(x, H = h_k(x))$
  - B: verifies if $H = h_k(x)$

- Properties
  - Without k, no one can generate valid MAc.
  - Without k, no one can verify MAc.
  - both authentication and integrity

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?

Encryption and Authentication

- $E_k(M)$

- Redundancy-then-Encrypt: $E_k(M, R(M))$
- Hash-then-Encrypt: $E_k(M, h(M))$
- Hash and Encrypt: $E_k(M), h(M)$
- MAC and Encrypt: $E_{h1(K)}(M), HMAC_{h2(K)}(M)$
- MAC-then-Encrypt: $E_{h1(K)}(M, HMAC_{h2(K)}(M))$

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

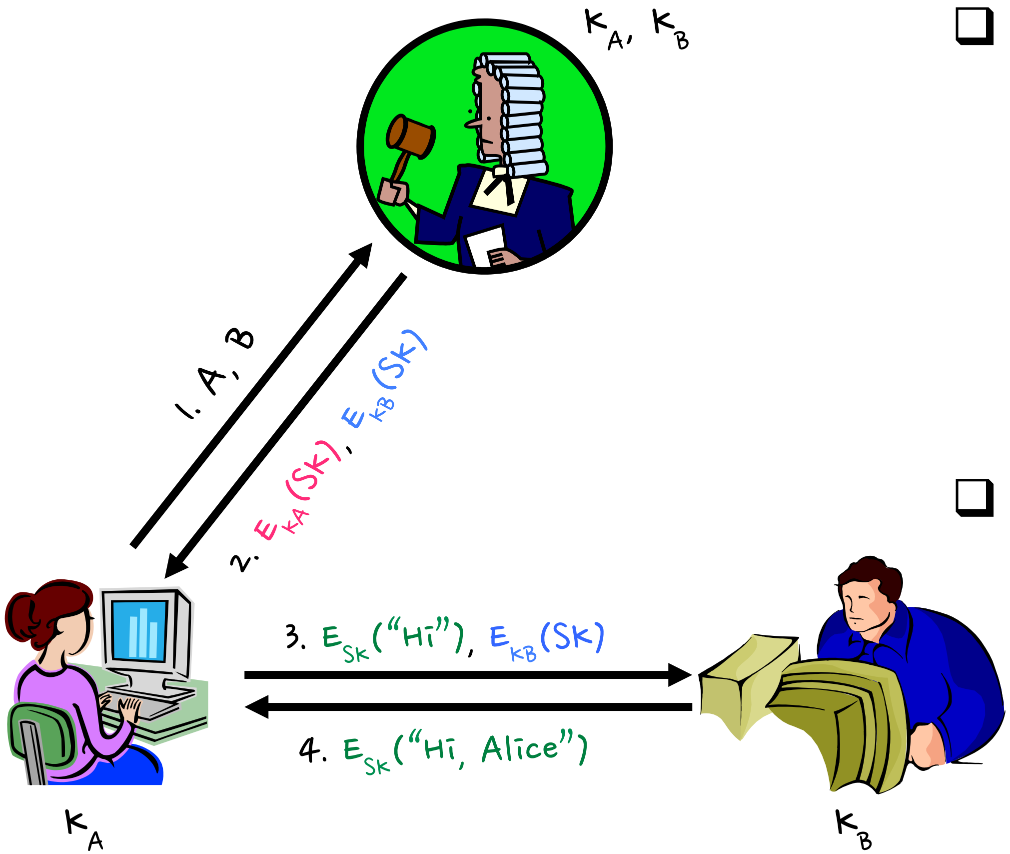
- Authentication
  - Message (Data origin) authentication
    - provide to one party which receives a message assurance of the identity of the party which originated the message.
  - Entity authentication (identification)
    - one party of both the identity of a second party involved, and that the second 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

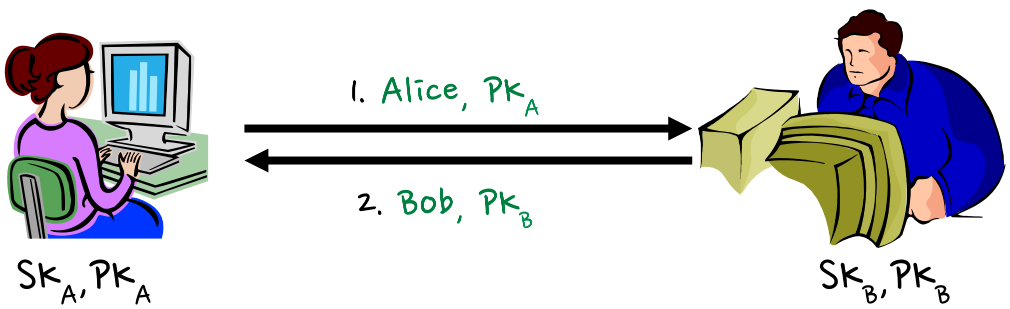
Key Management Through SKE



- 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

Symmetric vs. Public key

	Pros	cons
SKE	■ High data throughput ■ Relatively short key size	■ The key must remain secret at both ends ■ $O(n^2)$ keys to be managed ■ Relatively short lifetime of the key
PKE	■ $O(n)$ keys ■ Only the private key must be kept secret ■ longer key life time ■ digital signature	■ Low data throughput ■ Much larger key sizes

Kerckhoff's Principle

- Security should depend only on the key
  - Don't assume enemy won't know algorithm
    - can capture machines, disassemble programs, etc.
    - Too expensive to invent new algorithm if it might have been compromised
  - Security through obscurity isn't
    - Look at history of examples
    - Better to have scrutiny by open experts

- "The enemy knows the system being used." (Claude Shannon)

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"