Introduction to Cyber Security

Fall 2017 | Sherman Chow | CUHK IERG 4130

Sampling this course

Today's Schedule

- What is meant by "secure communication"?
- What does probability have to do with security?
- RSA: a "simple" encryption scheme based on "number theory"
- DNS: spoofing, poisoning, packet, rebinding attack
- Privilege of a running program
- What happen when a function call is made?

Sampling

- What is meant by “secure communication”?
  - Basic Cryptography
- What does probability have to do with security?
  - em… Not just about Cryptography
- RSA: a “simple” encryption scheme based on “number theory”
  - Some “Mathematics”

Sampling continued

- DNS: spoofing, poisoning, packet, rebinding attack
  - Network security and Web security
- Privilege of a running program
  - System security
- What happen when a function call is made?
  - Application security

Symmetric Key Encryption

- Alice wants to send a message to Bob, s.t. Eve cannot read.
- Eve can wiretap / eavesdrop the network connection.
- Alice encrypts a plaintext message such that Bob can decrypt but Eve cannot
- What differentiate Eve from Bob?
- Alice and Bob have a pre-established share secret.
  - E.g., a 56-bit key.

Cipher: Terminology

- Encryption: converting plaintext to ciphertext
- Decryption: converting ciphertext to plaintext
- Cryptanalysis: to break the code by analyzing the algorithm

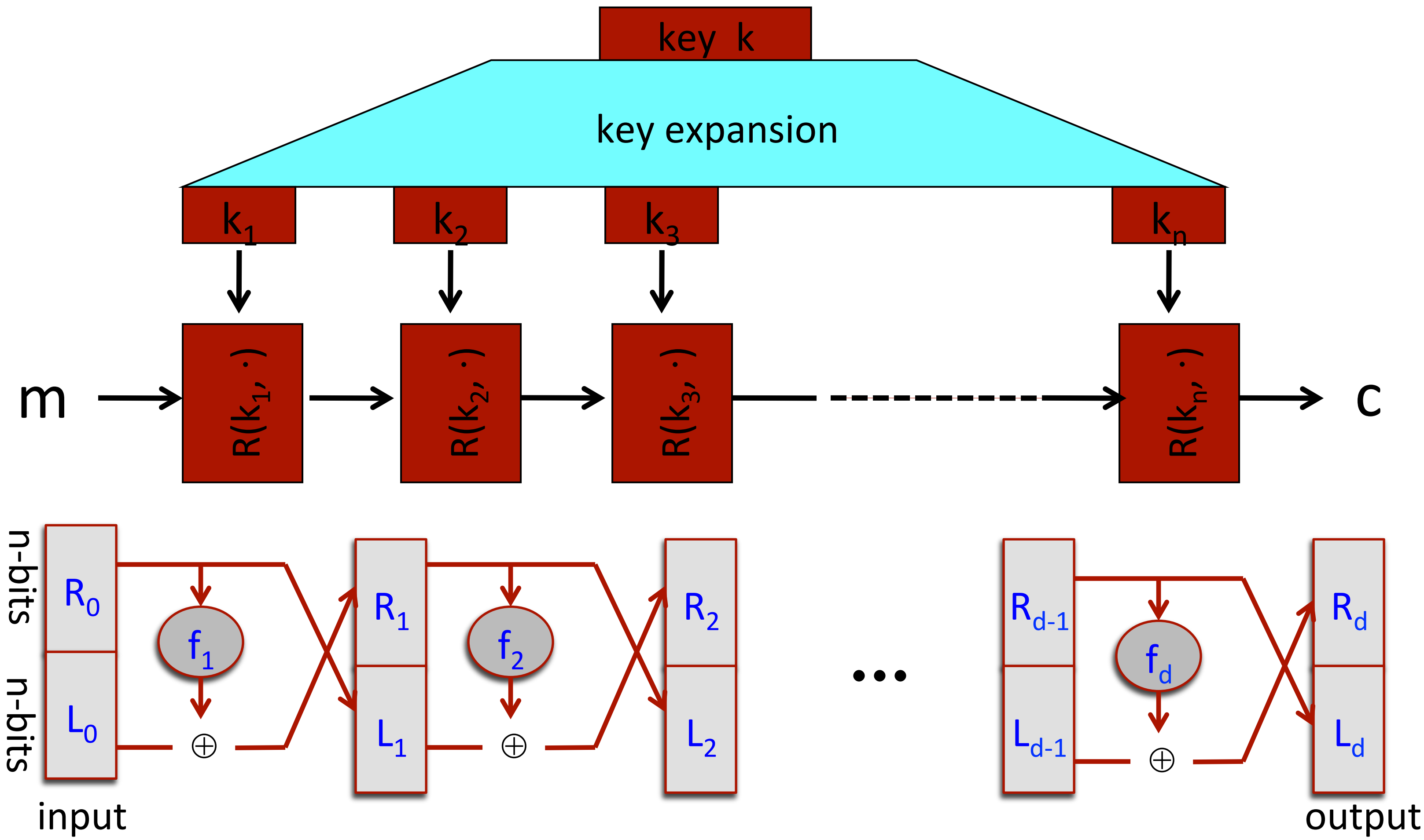
Confidentiality vs. Authenticity

- Alice encrypted “I love you” to Bob.
- Eve knew that Alice loves Bob.
- Eve “changes” Alice’s ciphertext, sends it to Bob as if Alice sent it, s.t. Bob will eventually receive “I hate you” after decryption

DES: The Data Encryption Standard

- Designed in the 1970's by IBM
  - with inputs from the NSA (National Security Agency) of US
- The most widely used encryption standard in the world
- 64 bits per block
- Use 56-bit key (7x8)

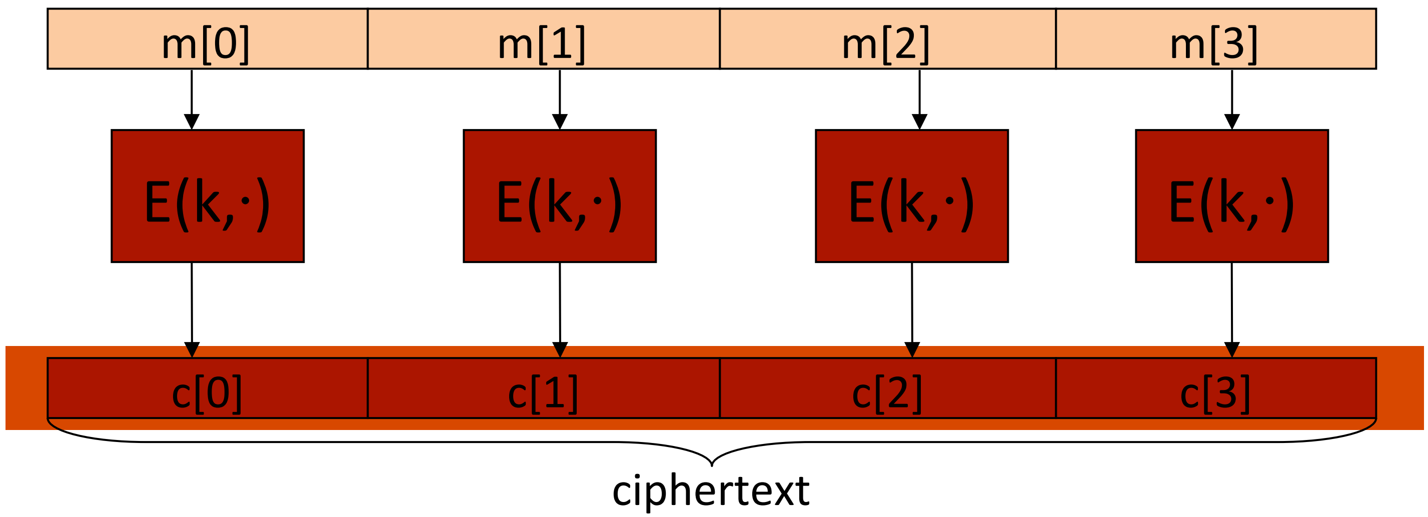
Block Cipher by Iterations



Mode of Operations

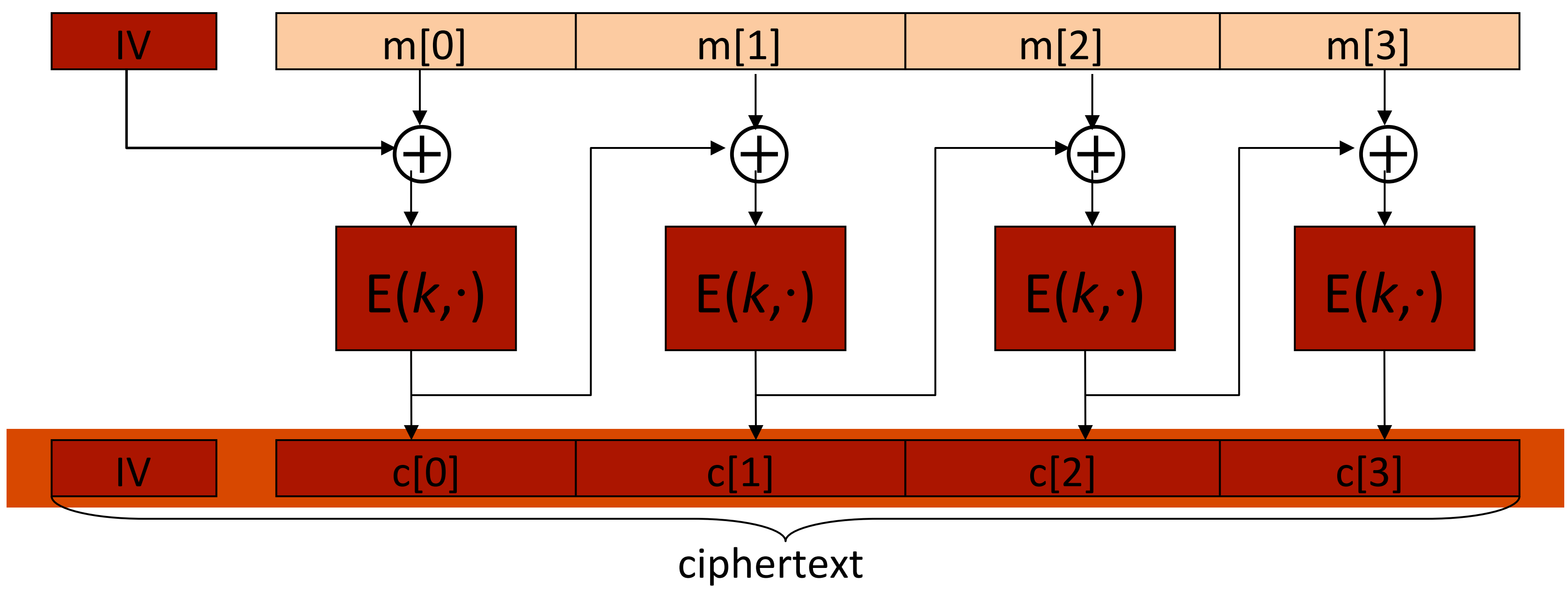
- What if I want to encrypt more than 1 block?
- Easy: just encrypt it “block-wise”
- What does “block-wise” really mean?

ECB, depicted



CBC w/ Random IV

- c[0] = E$\big($*k*, IV ⊕ m[0]$\big)$, c[1] = E$\big($*k*, c[0] ⊕ m[1]$\big)$, …

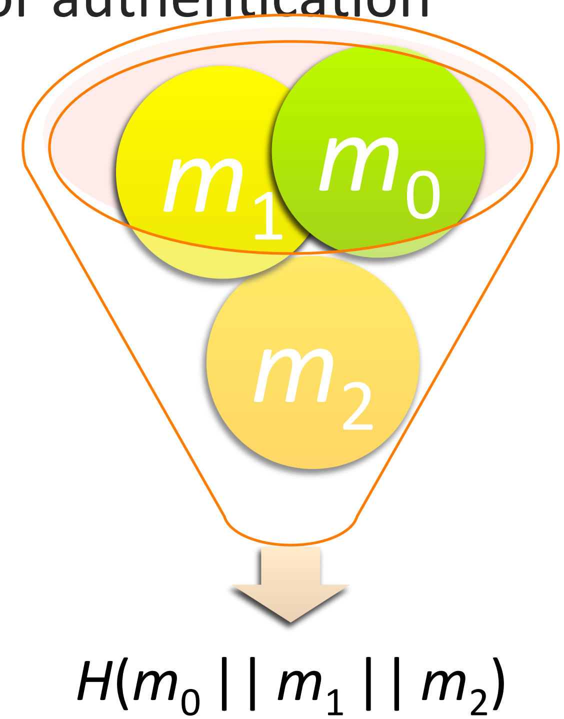


Authentication

- Using a block-cipher as a mean for authentication?
- What if I want to authenticate more than 1 block?

Hash Function for Message Digest

- Hash function accepts a *variable* size message *M* as input and produces a *fixed-size* message digest *H(M)* as output
- Message digest is sent with the message for authentication
- Produces a fingerprint of the message
- Authenticate just the hash of messages
  - instead of all the messages



Let’s play a game

- I randomly select a subset of size *n* of student from this class.
- I will pay you if no two of them share the same birthday.
- Otherwise you pay me.
- For what *n* you will enter this game?

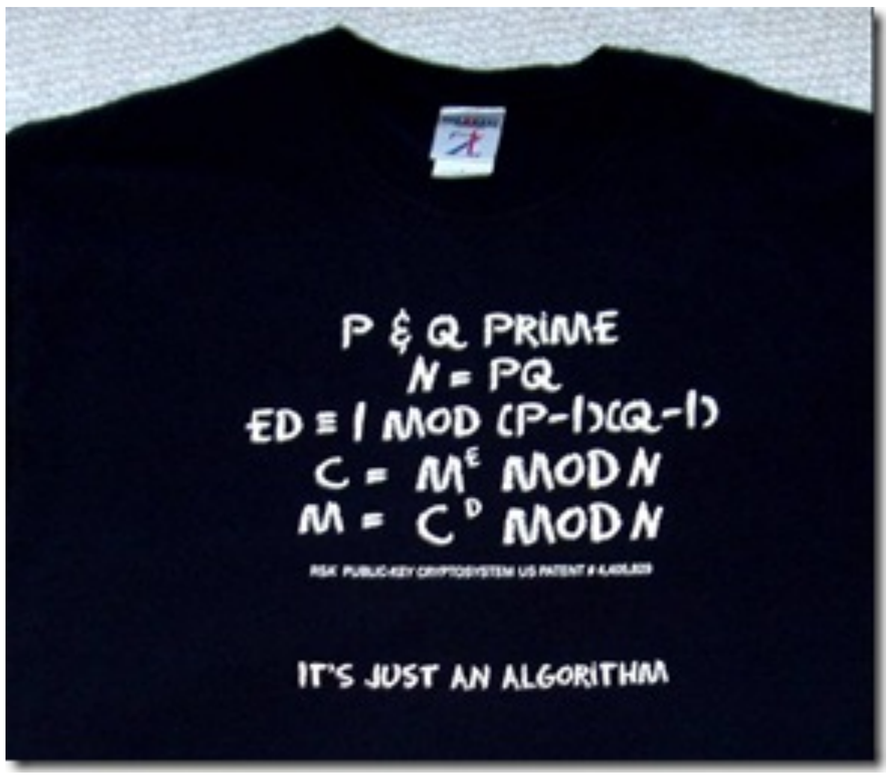
- Wait, why I suddenly talk about this??

Symmetric-Key vs. Asymmetric-Key

- How Alice and Bob have a shared secret at the first place?
- Bob has “secret knowledge” (*Decryption Key)* that Eve does not have.
- But Alice does not have Bob’s “secret knowledge”
- Wait… Alice only needs to perform Encryption but not Decryption
- Can Bob *publish* an encryption key, s.t. Alice (and Eve) will know, while only Bob can do the decryption?
- Answer: Public-Key Encryption (PKE)

RSA Algorithm

- Most widely accepted and implemented approach to PKE
- “Block cipher” where $0 \leq m, c \leq N - 1$ for some N



RSA Key Generation

- Choose two large prime numbers p, q. (e.g., 1024 bits each)
- Compute $N = pq$, $\Phi = (p-1)(q-1)$
- Choose e ($< N$) that has no common factors with Φ
  - i.e., e, Φ are "relatively prime"
- Choose d such that ed-1 is exactly divisible by Φ
  - i.e., $ed - 1 = K\Phi$ for some integer K
  - which means $ed \bmod \Phi = 1$ // remainder of ed divided by Φ is 1
- Public key is (N, e)
- Private key is (N, d) // or just d

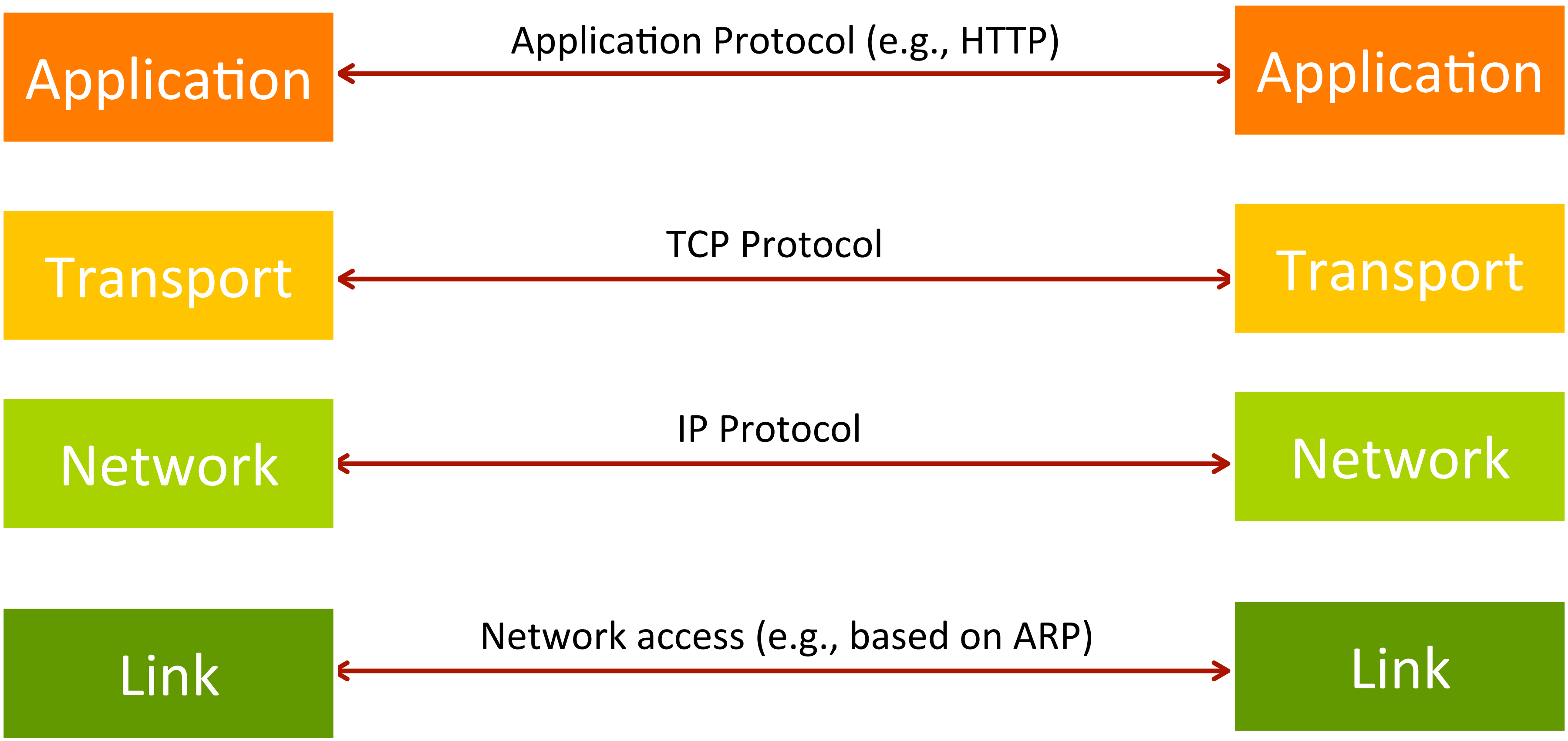
RSA Encryption and Decryption

- Enc(m) → c --- $c := m^e \bmod N$
- Dec(c) → m --- $m := c^d \bmod N$
- Fermat's Little Theorem: For any prime p, and any x such that p does not divide x (e.g., $1 \le x \le p$-1), we have $x^{p-1} = 1 \bmod p$
- Euler's Theorem: For any integer N, and x in $\mathbf{Z}_N$*, $x^{\Phi(N)} = 1 \bmod N$
  - Treat "x in $\mathbf{Z}_N$*" as "x is relatively prime to N"
  - $\Phi(N) = (p - 1)(q - 1)$
- Correctness: $c^d \bmod N = m^{(K\Phi+1)} \bmod N = m^{K\Phi} m \bmod N = m \bmod N$

ElGamal Encryption (not Factoring-based)

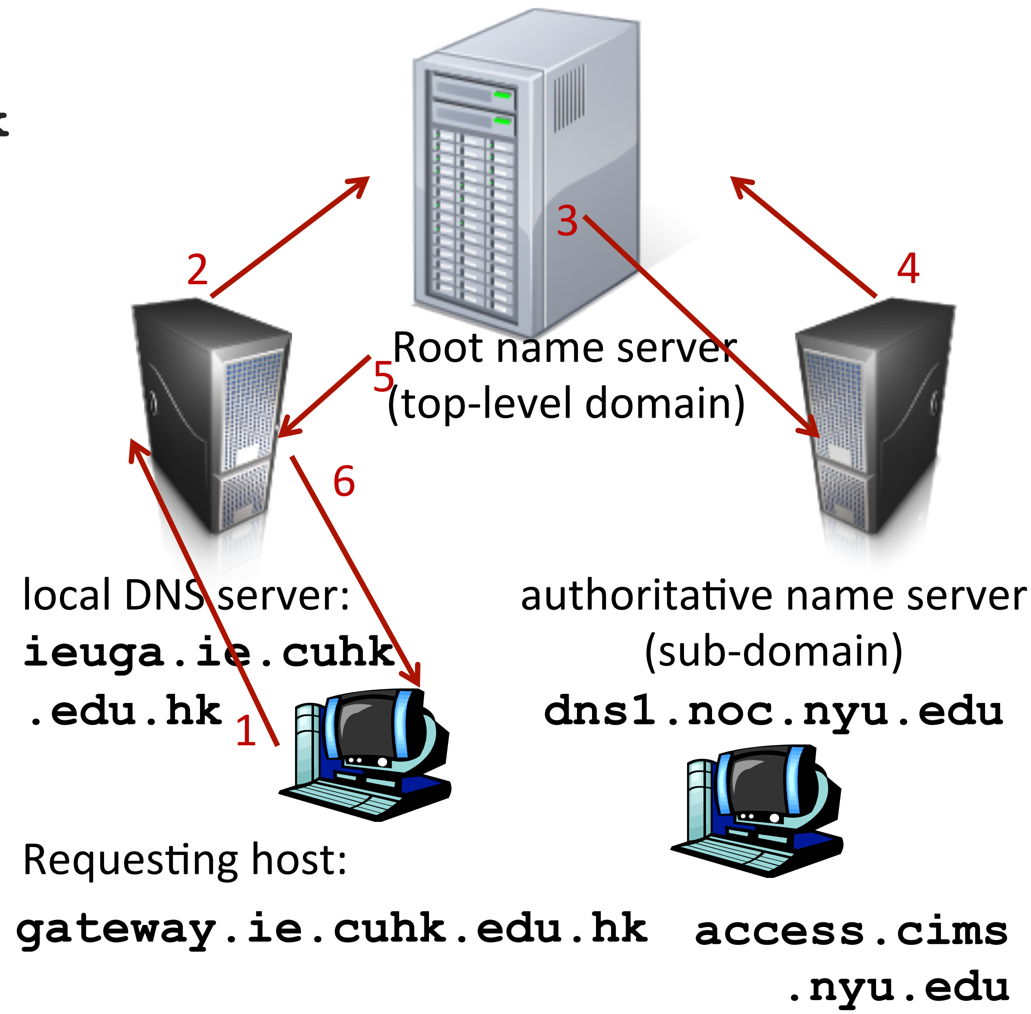
- ($sk = y$, $pk = Y = g^y$) $\leftarrow$ KeyGen(q)
- Pick $1 \leq r \leq q - 2$
- Set C_1 to be g^r mod q
- Compute $K = Y^r$ mod q
- Set C_0 to be MK mod q
- i.e., ($C_0 = M(\text{pk})^r$, $C_1 = g^r$) $\leftarrow$ Enc(pk, M)

TCP Protocol Stack



DNS Example

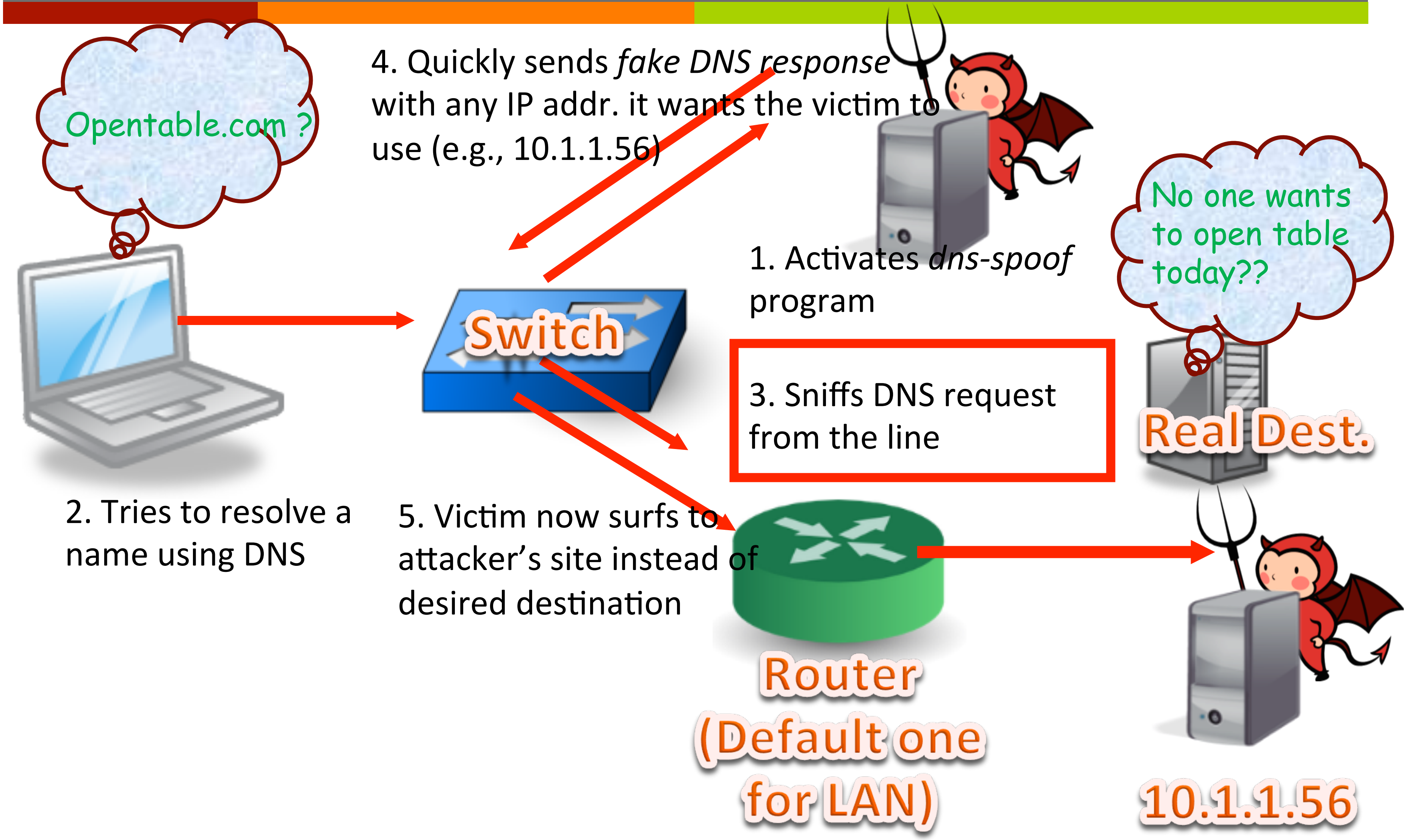
- **`gateway.ie.cuhk.edu.hk`** wants IP address of **`access.cims.nyu.edu`**
- Contacts its local DNS server **`ieuga.ie.cuhk.edu.hk`**
- **`ieuga.ie.cuhk.edu.hk`** contacts root name server
  - if necessary
- Root name server contacts authoritative name server, **`dns1.noc.nyu.edu`**
  - if necessary



Spoofing

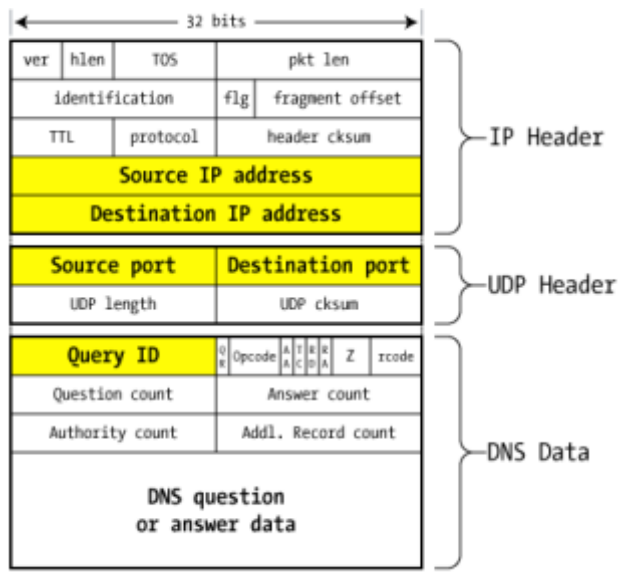
- Exploiting/altering mapping btw. domain name → IP → MAC address
  - IP Spoofing (e.g. attacker as sender)
  - DNS Spoofing (e.g. attacker as receiver)
  - ARP Spoofing (e.g. attacker as receiver)

DNS Spoofing



DNS Cache Poisoning

- As an attacker, I sit there and sniff DNS request from the line??
- To be more active, you need to know some details of DNS…



Clipart from http://www.colsterworth.lincs.sch.uk

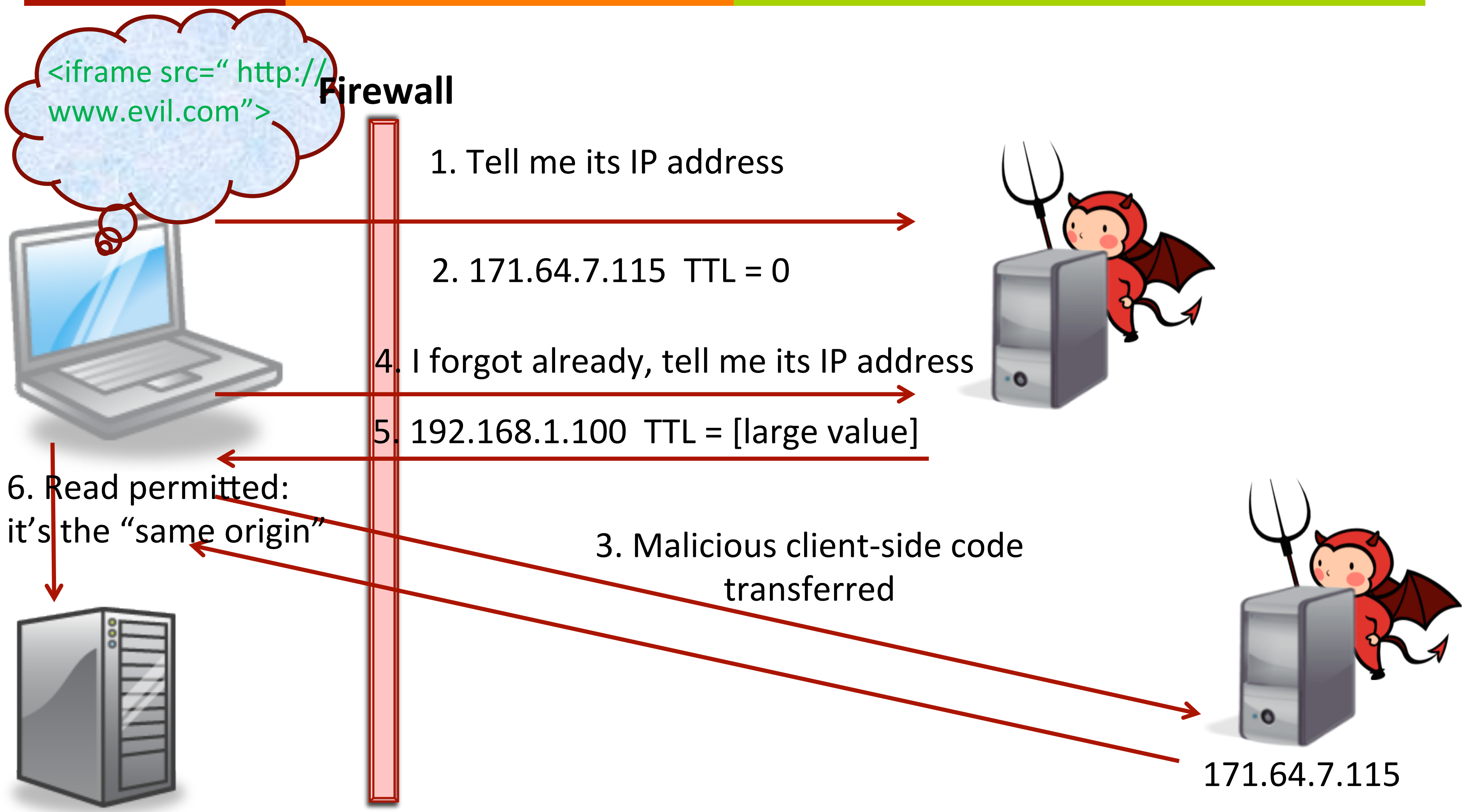
JavaScript Security

- There are no JavaScript methods that can directly access the files on the client computer
- There are no JavaScript basic methods that can directly access the network, although JavaScript programs can load URLs and submit HTML forms
- Protection via the "Same-Origin Policy"

Same Origin Policy (SOP)

- the policy permits scripts running on pages originating from the *same origin* to access each other's methods and properties with no specific restrictions
- but prevents access to most methods and properties across pages on different sites

DNS Rebinding Attack



Privilege of a running program

- A running program/process "typically" inherits the access rights of the login-account through which a program is run
- Instead of inheriting the rights of the program's runner, Unix is based on "Setuid" which may "inherit" the rights of the program's owner
  - E.g., mkdir command in UNIX changes file-system data-structure
    - i.e., need "root" or superuser privilege,
  - Thus, mkdir is owned by root but executable by users.
  - If a user runs mkdir, its "effective userid" is switched to "root"
  - Setuid-programs are especially "dangerous" because if there is a flaw in such programs, attacker can exploit it to gain superuser privilege!
    - E.g., sendmail

Buffer Overflow

- Memory errors in C and C++ programs are among the oldest classes of software vulnerabilities.
- Single biggest software security threat
- Even if we consider only classic buffer overflows, it has been lodged in the top-3 most dangerous software errors for years
  - CWE/SANS TOP 25 Most Dangerous Software Errors
- Buffer overflow vulnerabilities dominate in the area of remote network penetration vulnerabilities

Buffer Overflow in C/C++

- Store more data in the buffer (heap or stack) than it can hold
- The next contiguous chunk of memory is overwritten
- C/C++ language is inherently unsafe
  - i.e. it allows programs to overflow buffers at will
  - No runtime checks that prevent writing past the end of a buffer
  - strcpy(buf, “this string takes 27 bytes”); // buf only has 12 bytes

Stack depicted