Security and secure systems

COS 518: *Advanced Computer Systems*
Lecture 18

Michael Freedman

Intro to crypto in 10 minutes

2

What is Cryptography?

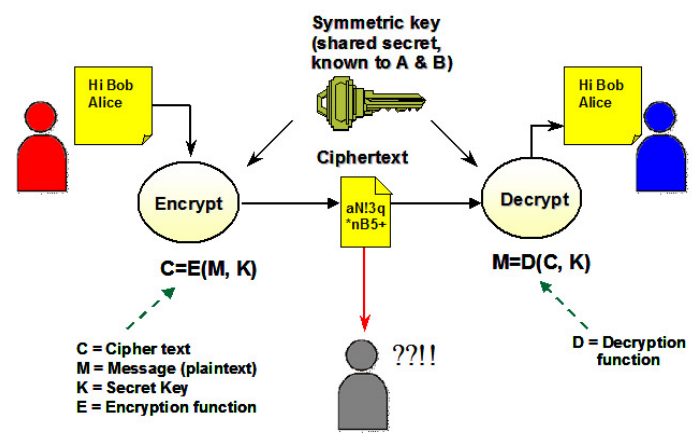
- From Greek, meaning "secret writing"
- Confidentiality: encrypt data to hide content
- Include "signature" or "message authentication code"
  – Integrity: Message has not been modified
  – Authentication: Identify source of message

plaintext —encryption→ ciphertext —decryption→ plaintext

- Modern encryption:
  – *Algorithm* public, *key* secret and provides security
  – Symmetric (shared secret) or asymmetric (public-private key)

3

Symmetric Cipher Model



4

Symmetric (Secret Key) Crypto

- Sender and recipient share common key
  - **Main challenge: How to distribute the key?**
- Provides dual use:
  - Confidentiality (encryption)
  - Message authentication + integrity (MAC)
- 1000x more computationally efficient than asymmetric

5

Public-Key Cryptography

- **Each party has (public key, private key)**
- **Alice's public key PK**
  - Bob uses PK to encrypt messages *to* Alice
    - ciphertext = encrypt (message, PK)
  - Bob uses PK to verify signatures *from* Alice
    - isValid = verify (signature, message, PK)
- **Alice's private/secret key: sk**
  - Alice uses sk to decrypt ciphertexts sent to her
    - message = decrypt (ciphertext, sk)
  - Alice uses sk to generate new signatures on messages
    - signature = sign (message, sk)

6

(Simple) RSA Algorithm

- Generating a key:
  - Generate composite **n = p * q**, where p and q are secret primes
  - Pick public exponent **e**
  - Solve for secret exponent **d** in $d \cdot e \equiv 1 \pmod{(p-1)(q-1)}$
  - Public key = (e, n), private key = d
- Encrypting message m: $c = m^e \bmod n$
- Decrypting ciphertext c: $m = c^d \bmod n$
- **Security** due to cost of factoring large numbers
  - Finding **(p,q)** given **n** takes $O(e^{\log n \log \log n})$ operations
  - n chosen to be 2048 or 4096 bits long

7

Cryptographic hash function

(and using them in systems)

8

Cryptography Hash Functions I

- Take message *m* of arbitrary length and produces fixed-size (short) number *H(m)*

- One-way function
  - Efficient: Easy to compute *H(m)*
  - **Hiding property:** Hard to find an *m*, given *H(m)*
    - Assumes "m" has sufficient entropy, not just {"heads", "tails"}
  - **Random:** Often assumes for output to "look" random

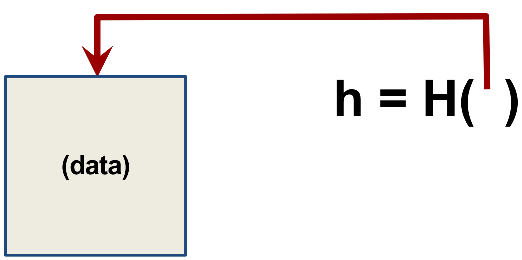
Cryptography Hash Functions II

- Collisions exist: | possible inputs | >> | possible outputs |

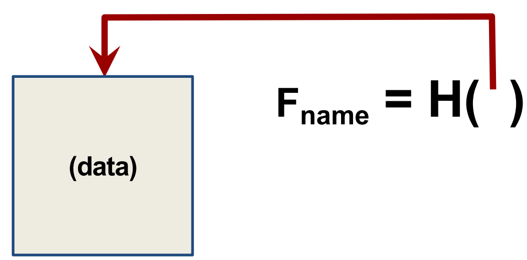
  … but hard to find

- Collision resistance:
  - Strong resistance: Find any m != m' such that H(m) == H(m')
  - Weak resistance: Given m, find m' such that H(m) == H(m')
  - For 160-bit hash (SHA-1)
    - Finding any collision is birthday paradox: 2^{160/2} = 2^80
    - Finding specific collision requires 2^160

Hash Pointers

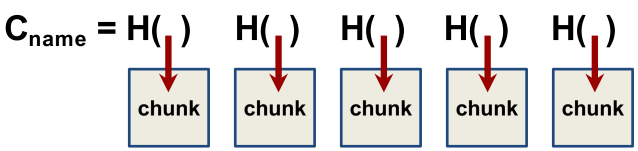


Self-certifying names



- P2P file sharing software (e.g., Limewire)
  - File named by F_{name} = H (data)
  - Participants verify that H (downloaded) == F_{name}

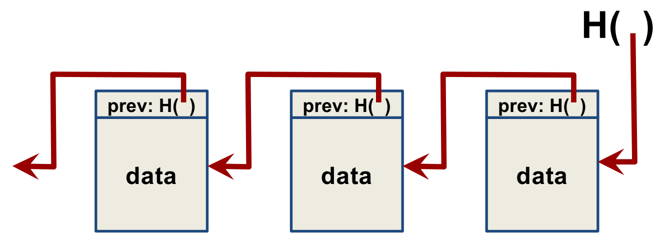
Self-certifying names



- BitTorrent
  - Large file split into smaller chunks (~256KB each)
  - Torrent file specifies the name/hash of each chunk
  - Participants verify that H (downloaded) == C_{name}
  - Security relies on getting torrent file from trustworthy source

13

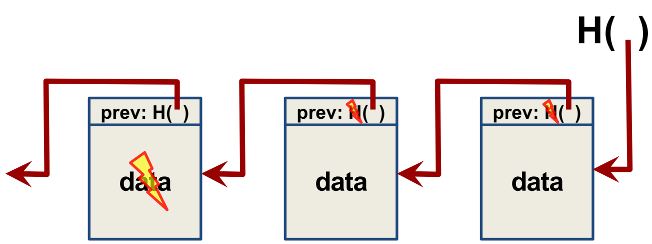
Hash chains



Creates a "tamper-evident" log of data

14

Hash chains



If data changes, all subsequent hash pointers change

Otherwise, found a hash collision!

15

Examples: HTTP and DNS security

16

“Securing” HTTP

- Threat model
  - Eavesdropper listening on conversation (confidentiality)
  - Man-in-the-middle modifying content (integrity)
  - Adversary impersonating desired website (authentication, and confidentiality)
- Enter HTTP-S
  - HTTP sits on top of secure channel (SSL/TLS)
  - All (HTTP) bytes written to secure channel are encrypted and authenticated
  - Problem: What is actually authenticated to prevent impersonation? Which keys used for crypto protocols?

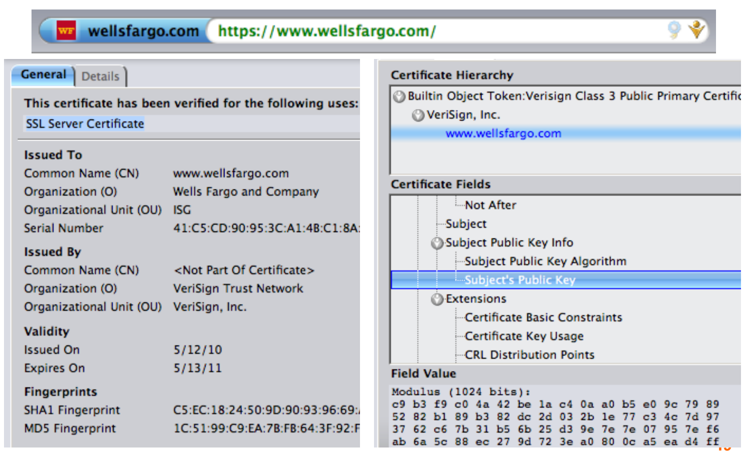
17

Learning a valid public key

- What is that lock?
  - Securely binds domain name to public key (PK)
    - Believable only if you trust the attesting body
    - Bootstrapping problem: Who to trust, and how to tell if this message is actually from them?
  - If PK is authenticated, then any message signed by PK cannot be forged by non-authorized party

18

How to authenticate PK



The trouble with CAs

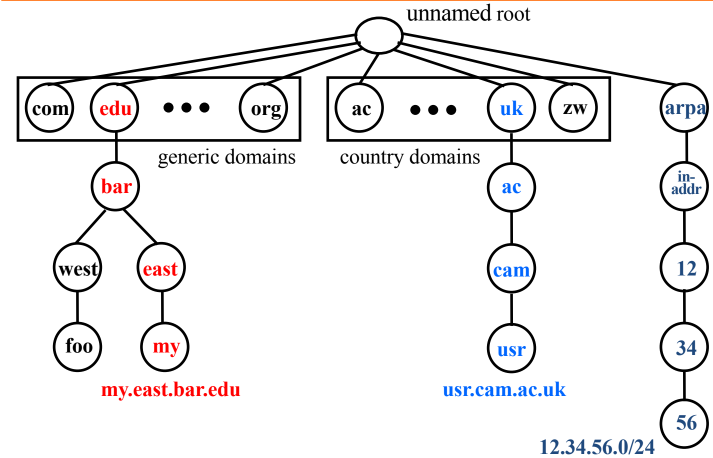
- Browse/OS vendors pick which CAs to trust
  - Sometimes they revoke this trust – e.g. DigiNotar.
- No notion of CAs having authority over only given TLD
- Trust the {Iranian, Chinese, US} national authorities?
- What standards does Apple use to pick root certs? Google? MSFT?
  - There's a restraint-of-trade issue here. Can't enter the CA business without vendor support…

22

DNS Security

23

Hierarchical naming in DNS



24

DNS Integrity: Trust the TLD operators?

- If domain name doesn't exist, DNS should return NXDOMAIN (non-existant domain) msg
- Verisign instead creates wildcard DNS record for all .com and .net domain names not yet registered
  - September 15 – October 4, 2003
- Redirection for these domain names to Verisign web portal: "to help you search"
  - and serve you ads…and get "sponsored" search
  - Verisign and online advertising companies make money…

25

DNS Integrity: Answer from authoritative server?

- DNS cache poisoning
  - Client asks for www.evil.com
  - Nameserver authoritative for www.evil.com returns additional section for (www.cnn.com, 1.2.3.4, A)
  - Thanks! I won't bother check what I asked for

26

DNS Integrity: Answer from authoritative server?

- To prevent cache poisoning, client remembers domain and 16-bit request ID (used to demux UDP response)
- But…DNS hijacking attack:
  - 16 bits: 65K possible IDs
    - What rate to enumerate all in 1 sec? 64B/packet
    - 64*65536*8 / 1024 / 1024 = 32 Mbps
  - Prevention: Also randomize the DNS source port
    - Windows DNS alloc's 2500 DNS ports: ~164M possible IDs
    - Would require 80 Gbps
    - Kaminsky attack: this source port…wasn't random after all

27

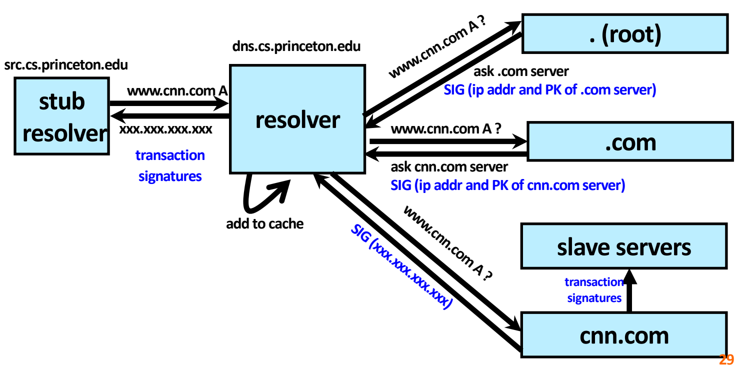
Let's strongly believe the answer! Enter DNSSEC

- The DNS servers sign the hash of resource record set with its private (signature) keys
- Public keys can be used to verify the SIGs
- Leverages hierarchy:
  - Authenticity of nameserver's public keys is established by a signature over the keys by the parent's private key
  - In ideal case, only roots' public keys need to be distributed out-of-band

28

Verifying the tree

Question: www.cnn.com ?



29