CM0133 Internet Computing

Security in the Internet

Objectives - Security

- Security ?
- Cryptography
- En/Decryption, Symmetric, Public-Key
- Secure Communications
- Certificates, Transport-Layer-Security

Security for the WWW

- TCP and thus HTTP are clear-text protocols, which make no attempt to hide the data being transmitted. For secure data transfers, it thus is necessary to use additional technologies for providing secure data transfers.
- For the Web, the most interesting security feature are secure HTTP interactions, which are provided by HTTP over SSL (HTTPS), a protocol that layers an encryption layer (SSL or TLS) between TCP and HTTP.
- For any task involving personalization and/or trust, it is not only necessary to have a concept for providing privacy, but also to have concepts for identity and how to prove identity, which needs authentication.

http://dret.net/lectures/web-spring10/security#%282%29

Identification

- Identity is required for any non-anonymous communications
  - groups can have an identity (facebook members see more than non-members)
  - pseudonyms are "hidden identities" (the "real identity" is not visible)
  - personal identity should be tied to a person itself
- Proof of Identity is important for any privileged operation
  - signatures and seals are traditional ways
  - traditional ways are mostly protected by law (but not really safe)
  - more modern ways often include technical methods for Authentication
- Client identity on the Web can be bound in three ways:
  1. Computer (most of the time "identified" by an IP Address)
  2. Browser (in the form of a stored cookie)
  3. User (identified through some authentication method)

Authentication

- Authentication is the process of verifying an identity
  - the weakest form of authentication is simply trust
  - legal consequences can make it more risky to falsify authentication
  - technical measures should make it hard to impossible to falsify authentication
- Authentication on the Web comes in many different flavours
  - implicitly by accessing a server from some IP Address range
  - presenting a cookie from a previous formal authentication
  - presenting a password as a proof of identity
  - proving that you are owning additional authentication hardware (often PIN-enabled, see http://en.wikipedia.org/wiki/Personal_identification_number)
- Risk and potential damage should justify authentication methods

Authorization

- Authorization is the question of allowing operations
  - Identification is necessary to identify the initiator
  - Authentication is necessary to verify the initiator's identity
  - if the initiator is authorized, the operation can be performed
- Web pages often are public or restricted access
  - public web pages do not require any identification (and thus authentication)
  - restricted access Web pages can be group pages (internal company pages)
  - personal access is another popular scenario (email, facebook, online banking)
- Web servers have well-defined ways of performing authentication

Classroom Task

Pair Work: Identify Dangers in the Internet

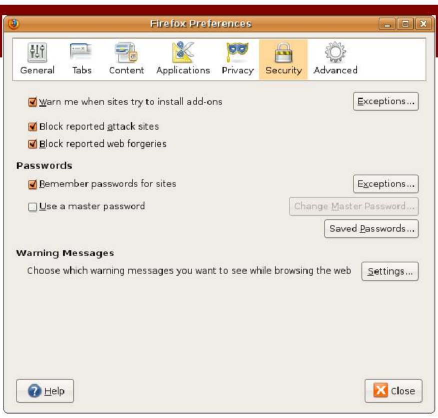
Trust and Security on the Web

- Web-based applications introduce many risks
  - do you trust your browser? (it may not safeguard your information)
  - do you trust your computer? (it may have a virus)
  - do you trust your network? (it may be monitored on various levels)
  - do you trust the server? (it may be a fake phishing [http://en.wikipedia.org/wiki/Phishing] server)
- Most of these risks are amplified by the Web's scale
  - phishing and spamming only work because the Web makes fraud more effective
- Controlling Web access is important for safe browsing
  - trusting shared browsers is risky (they may store logins and cache pages)
  - trusting the network can be risky (more and more networks are wire-tapped)
  - trusting the server is risky (phishing and poor server security)

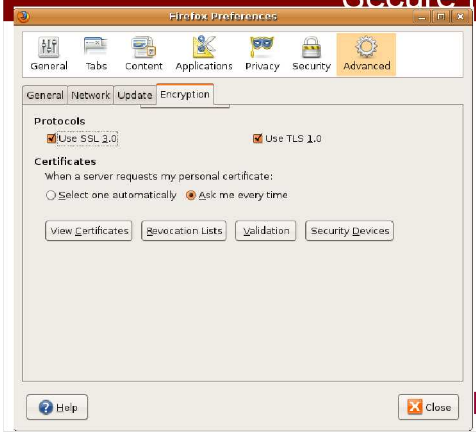
http://dret.net/lectures/web-spring10/security

Privacy

Security

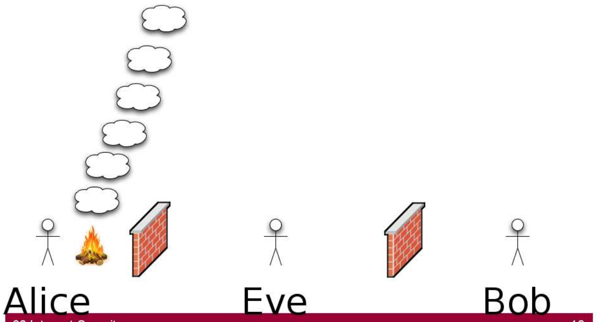


Secure Protocols



Cryptography

http://www.bbc.co.uk/dna/h2g2/alabaster/A233966



Section

Cryptography

Plain-text

Alice uses er key to encrypt the plain-text

⊕ ← Key

Cypher-text

⊕ ← Key

Plain-text

Bob uses his key to decrypt the cypher-text

http://dret.net/lectures/web-spring10/security

Reversing Thee Message

.thgin yb sklaw tunaep ehT

The cipher is called Reversing Thee Message and it works by writing the message backwards. Decrypted, the message would read:

The peanut walks by night.

http://www.bbc.co.uk/dna/h2g2/alabaster/A233966

Caesar Code / ROT-14

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n

ROT 14

The hen has laid its eggs

↓

Hvs vsb vog zowr whg suug

Zsh hvsa soh qoys

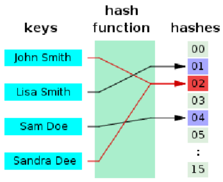
Cryptography

- Cryptography is structured into different layers
  - layering is a well-established principle for separation of concerns
- Cryptographic primitives implement very basic functionality
  - changes and advancements in this area are limited to very specialized researchers
  - it is easy to make fatal mistakes which then challenge everything built on top if it
- Cryptographic protocols assemble primitives into application-level solutions
  - primitives solve very basic security problems (fingerprints, encryption, …)
  - protocols combine these into applications (digital signatures, secure communications, …)

http://dret.net/lectures/web-spring10/security

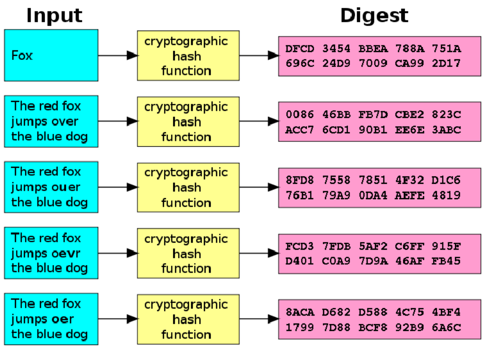
Hash Function

- A hash function is any well-defined procedure or mathematical function that converts a large amount of data into a small datum, usually a single integer that may serve as an index to an array. The values returned by a hash function are called hash values, hash codes, hash sums, or simply hashes.



A hash function that maps names to integers from 0 to 15. There is a collision between keys "John Smith" and "Sandra Dee".

Cryptographic Hash Function



Cryptographic Hash Function

- The ideal cryptographic hash function has four main or significant properties:
  - it is easy to compute the hash value for any given message,
  - it is infeasible to find a message that has a given hash,
  - it is infeasible to modify a message without changing its hash,
  - it is infeasible to find two different messages with the same hash.

One-Way Function

Variable length original data

Fixed length "digest" of data

- Hashes (or message digests) are well-known in computer science
- One-way functions are cryptographically safe hashes
  - very hard to find an input producing a given output
  - very hard to find two inputs producing the same output ("collision")

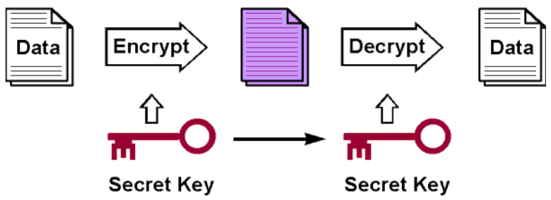
http://dret.net/lectures/web-spring10/security

Plausible Encryption

- Secret-Key is was most people think of when thinking of encryption
  - symmetric cryptography is another popular term
    - One key for encryption and decryption
    - Revealing the key makes encrypted data openly readable
  - there must be a secure channel to transport keys, such as diplomatic pouches [http://en.wikipedia.org/wiki/Diplomatic_bag].
- Good for long-term relationships with few partners
  - exchange secret keys as part of the initial setup of a relationships
  - adding partners requires a secure channel for key exchange
  - changing keys requires a secure channel for key exchange
- Almost impractical in an environment with many ad-hoc partners

http://dret.net/lectures/web-spring10/security

Notice the Arrow



http://dret.net/lectures/web-spring10/security

! Known-plain-text !

- Known-plain-text attack
  - Key is unknown, but when given plain-text the crypto-system produces encrypted text
  - Careful choice of the plain-text allows the retrieval of the key
- Defences
  - Longer keys and padding with random data
  - Use a different key for each encryption

! Key-exchange !

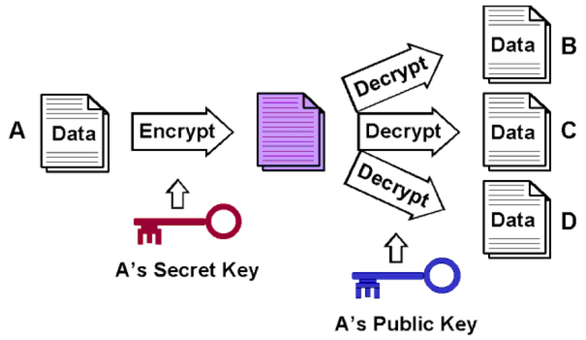
- Key-exchange
  - The secret key must be transferred from Alice to Bob
- Solutions
  - Out-of-band transfer - Disk, CD, hand-written Note,...
  - Key-exchange algorithms

Implausible Encryption

- Public-Key intuitively is hard to accept as a concept
  - asymmetric cryptography is another popular term
- Key pairs of one public and one secret key
  - key generation is the process of generating these key pairs
- The public key can be made available to the public
  - only the secret key can do the inverse operation of the public key
- Good for short-term relationships with many partners
  - publish your public key so that it can be used worldwide
  - everybody can encrypt data using the public key
  - only the owner of the secret can can decrypt the message and read it
- Computationally expensive and not good for a large amounts of data

http://dret.net/lectures/web-spring10/security

No arrow here ...



http://dret.net/lectures/web-spring10/security

Public Key

- Use different keys for encryption and decryption
  - Alice publishes her public-key and keeps her private-key secret
  - Bob uses Alice's public-key to encrypt the plain-text
    - The plain-text cannot be decrypted using the public-key
  - Alice uses her private-key to decrypt the cypher-text

http://en.wikipedia.org/wiki/RSA

Public Key

- Alice generates a large random number
- This number is split into a public and a private component using a “trapdoor” function
  - Allows for easy splitting of the random number
  - Makes it hard to guess the private component from the public component

Building Secure Applications

- Cryptographic primitives in most cases are not sufficient
  - they provide basic functionality for fundamental tasks
  - they must by combined to provide solutions for real-world problems
- Typical problem #1: How to ensure key authenticity
  - with insecure keys, the majority of cryptographic methods is worthless
- Typical problem #2: How to communicate securely without shared keys
  - many interesting scenarios are based on ad-hoc interactions
  - secret-key does not work, public-key needs to verify the peer
- Typical problem #3: How to check authenticity and integrity of data
  - integrity can be done with checksums, but these could be forged
  - authenticity needs a cryptographically secure way of combining identity and data

http://dret.net/lectures/web-spring10/security

Digital Signature

- also digital signature scheme
- is a mathematical scheme for demonstrating the authenticity of a digital message or document.
- A valid digital signature gives a recipient reason to believe that the message was created by a known sender, and that it was not altered in transit.
- Digital signatures are commonly used for software distribution, financial transactions, and in other cases where it is important to detect forgery and tampering.

Certificate

- Certificates are digital signatures issued by a trusted party
  - most digital signatures are created with certified public keys
  - this means the digital signature is created based on a digitally signed key
- Who can you trust on the Web?
  - trust can only start to grow based on initial trust in something
  - many systems come with pre-installed trust (root certificates)
  - certificates from other issuers will cause browsers to complain [https://katapultmedia.com/]
- Certificates (like domain names) are a very easy way to make money
  - in theory there are different levels of certificates with different levels of identity checking
  - in practice most sites choose the cheapest one that does not give an error message

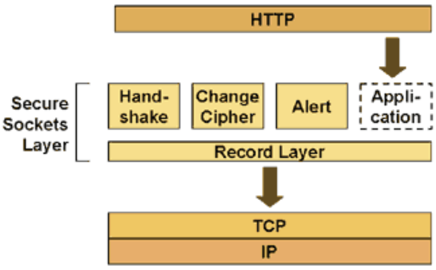
Secure Communications

- Public-Key cryptography is computationally expensive
  - it is possible to encrypt all traffic using asymmetric key pairs
  - this generates considerably more load on the server side
- Combining public-key and secret-key cryptography
  1. check the public key for authenticity (using a Certificate)
  2. generate a key for a secret-key encryption scheme
  3. use the public key to securely transmit the secret key
  4. use the secret key for securely transmitting the payload
- Combines the advantages of both methods
  - the lower complexity of secret-key algorithms
  - the ability of public-key algorithms to work without a secure channel

HTTP and Security

- HTTP sends clear-text messages
- Making HTTP secure requires additional mechanisms
- Encryption is done by a layer on top of TCP
  - Secure Sockets Layer (SSL) is the protocol layer invented by Netscape
  - Transport Layer Security (TLS) is the standardized Internet version
  - TLS adds more encryption schemes and more flexibility
- Lower-level methods may also provide encryption
  - Virtual Private Networks (VPN) provide IP-based encryption
  - WEP and WPA provide network interface encryption

HTTP and SSL



http://dret.net/lectures/web-spring10/security

SSL – Secure Socket Layer

- SSL is a cryptographic protocol or encryption protocol used to for secure application-level data transport.
- SSL implements algorithms for secure communication on the internet.

TSL – Transport Layer Security

- TSL stands for Transport Layer Security
- A browser requesting a secure page adds an "s" to the "http" when sending out the public key and certificate.
- TSL carries out three checks
  - Certificate comes from trusted party
  - Certificate is currently valid
  - Certificate has a relationship with the site from which it's coming

Transport Layer Security

- First version introduced 1995 by Netscape as SSL 2.0
  - SSL 2.0 suffers from serious cryptographic defects
- SSL 3.0
  - Fixed the major defects of SSL 2.0
- TLS 1.0
  - Can be used to secure any TCP based connection
  - IETF standard

Transport Layer Security

Client Hello
↓
Server Hello
↓
Server Certificate
↓
Key Exchange
↓
Start Secure Connection
↓
Finish Secure Connection

ATTACKS !

Dictionary Attack

- Brute-Force Attack with a dictionary or list of known words to “guess” a password.
  - Qwerk, 123456, admin, admin123
- Every request takes a couple of ms to s → a few hundred requests per minute possible
- Store IP address of the attacker / user – if too many requests per minute are detected block the account
  - Banking Passwords often allow no more than three attempts then account is blocked and direct personal interaction with the provider required for additional security checks.

! Null Certificates !

- Certificates where the Common-Name includes a null character (\0)
  - www.lloyds.co.uk\0www.evilsite.com
- Certificate provider checks that the attacker controls evilsite.com
- Browser only displays the Common-Name up to the null character
  - www.lloyds.co.uk

Man in the Middle Attack (MITM)

- A MITM is achieved by an attacker who intercepts communication between victims making them believe that they are talking directly to each other over a private connection.
- Communication is controlled by the attacker without the victims knowing.
- The attacker has the capability to intercept all messages going between victims and inject new ones.

Means to prevent MITM

- Public key infrastructures
- Mutual authentication, via Secret keys
- Mutual authentication via Passwords
- Latency examination / using long Cryptographic hash function calculations
  - E.g. if hash function takes 10s and both parties take 20 seconds normally but the actual examined calculation takes 60 seconds to reach each party, this can indicate a third party
- Second (secure) channel verification
- One time pads
  - encryption, which has been proven to be impossible to crack (http://en.wikipedia.org/wiki/One-time_pad#Example)

Summary - Internet Security

- Security is hard – there is NO 100 % Security
- Certificates are used to guarantee a party's authenticity
- Certificates are digital signatures issued by trusted parties
- Once authenticated, public keys can be used to securely communicate
- Encryption on the Web is based on SSL and TSL - HTTPS

Literature

- http://en.wikipedia.org/wiki/Internet_security
- http://dret.net/lectures/web-spring10/security