E-Business Security Technologies

Michael B. Spring
Department of Information Science and Telecommunications
University of Pittsburgh
spring@imap.pitt.edu
http://www.sis.pitt.edu/~spring

Overview

- Introduction to security
- Secure Transactions
  - Encryption and Secure Sockets
  - Authentication
- Private Data
  - Regulation and privacy
  - Cookies and sessions
- Secure Sites
  - Firewalls and access controls

Introduction to security(1)

- Security is never guaranteed and comes at a cost in terms of convenience and access
- Historically, information systems were made secure by isolating them
- Increasingly, the desire to use the Web has mandated that these information systems be opened in a controlled way

Introduction to security(2)

- Websites need to be secure
- In classic terms, the provider needs to assure:
  - Confidentiality: data cannot be read
  - Integrity: data cannot be written
  - Authenticity: entities are who they say they are
- In addition, providers need to assure physical security from fire, flood, etc. and logical security – safety from denial of service, Trojan horses, etc.

Security Encryption and Privacy

Concern about security is the primary reason why people do not use E-Business – (Tilson et al., 1998)

- Encryption and Authentication
  - E-commerce transactions need to be secure in transit
  - Partners need to know who they are dealing with
- Privacy
  - What information about users is collected?
  - How is that information secured?
  - What assurances about security can be provided?

Private (Single) Key Encryption

- Encryption is the masking of a message by substitution, permutation, or application of a key.
- Shannon proved that a one time random key the length of the message assures unbreakable cipher
- Single key encryption/decryption is fast and virtually unbreakable(effort versus benefit)
- One of the best known single key encryption methods is DES (Digital Encryption Standard)
- Unfortunately, management of single keys is untenable on the web. The solutions is dual key.

Public (Dual) Key Encryption

- Rivest, Shamir, and Adelman developed a public-private key encryption system known as RSA
- Very simply, RSA keys are developed as follows:
  - two large primes(p,q) are used to generate a number n=p*q
  - A number e, also a prime is calculated such that 3<e<(p-1)*(q-1)
  - A number d is determined such that e*d=1mod(p-1)*(q-1)
- The keys become (e,n) and (d,n) and practically d is not calculable without knowledge of p and q.

Definitions of Terms

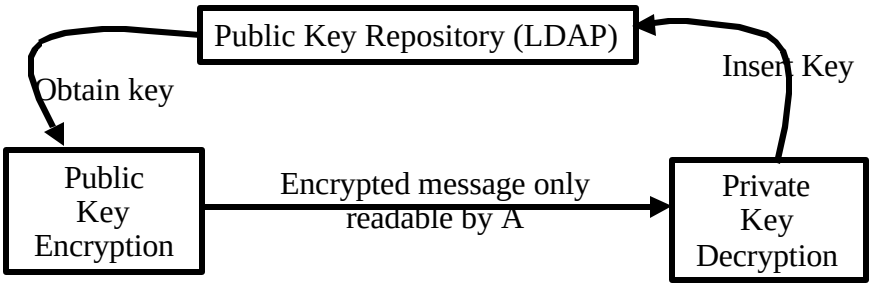
- Identification is the process of claiming to be someone who has rights
  - This represents an ID of some sort
- Authentication is the assurance that the initiator is who they say the are
  - This represents a password of some sorts
  - Passwords are of three types – knowing, having, and being
- Non-repudiation is the assurance that the initiator can not deny their participation

Web Security for
Confidentiality/Integrity & Authenitication

- The need to transfer data rapidly suggested the need to find a way to use DES
- The need to authenticate users suggested the need for a digital signature
- The need to transfer keys securely between diverse users mandated a dual-key system
- The solution is found in the use of secure sockets (SSL) and certificates (backed by certificate authorities)

Private Messages(Dual Key)

1. Consider two users, A and B.
2. B wishes to send a private message to A
3. B uses A's public key to encrypt the message
4. Only A can decrypt the message using A's private key



Digital Signatures

- User A takes their private key and encrypts a message they wish to send.
- The resulting cipher is attached to the original message.
- User B uses A's public key to decrypt the cipher attached to the message
- If the clear text message and the decrypted message match, the message has integrity and it must have been sent by A.
- The cipher is called a digital signature.

Signed Encrypted Messages

- A "total" solution -- message from A to B
- Because dual key encryption of long message is time consuming, the MD5 algorithm is used to calculate a message digest
- Digest to a signature using A's Private Key
- Message encrypted using B's Public Key
- B receives the secure message decrypts is using B's Private Key and authenticates it using A's Public Key

Certificates

- The missing ingredient here is some assurance that I am who I say I am
- A certificate is a parcel of digitally signed information – basically a signed message
  - The signature comes from a "Certificate Authority" who is recognized as trustable
- If what the certificate says is true matched what the decrypted signature (using the Certificate Authorities Public Key) it is considered valid.

https

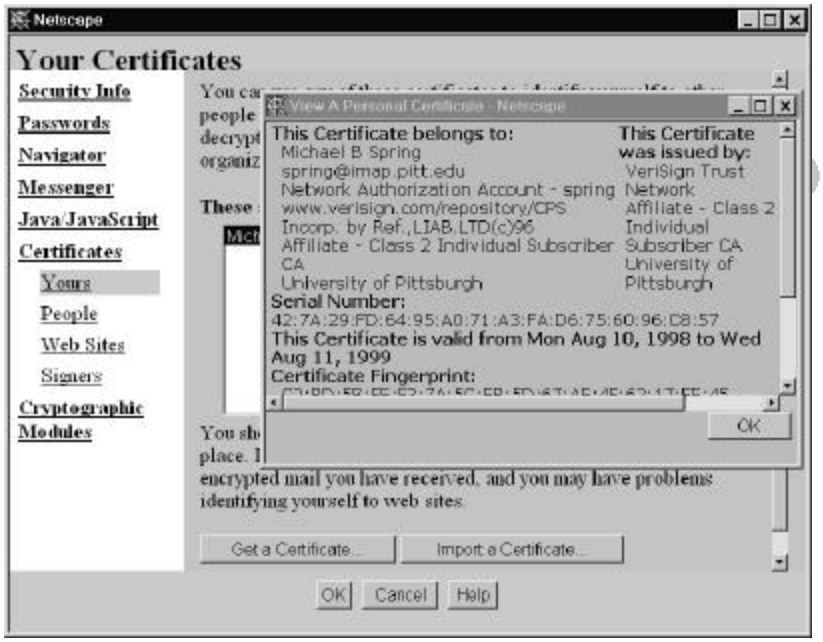
- All of this functionality, and more is provided to the user through https.
- When a URL specifies https:// as the service, it is an indication that the client and the server will communicate over port 135 rather than 80 and will use the secure sockets protocol – or layer
- https combines certificates, private and public keys in a way that is transparent to the user

Secure Sockets Layer

- Client sends SSL info including generated data
- Server sends SSL info including generated data and certificate
- Client authenticates the server or exits (see next slide)
- Client generates premaster secret and sends it using the servers public key – also generates the session key
- Server authenticates client(optional) and uses the premaster secret to generate the session key
- Client and server signal use of the session key
- Messaging begins using the session key

How authentication works

- Assuming either a client or a sever requests authentication, they submit a certificate
  - Certificate is decrypted using the public key
  - The validity period of the certificate is checked
  - The Certificate Authority's(CA) distinguished name is checked against the checkers list
  - Does the CA's public key (held locally) decrypt the CA signature that was in the certificate
  - Does the domain name of the certificate owner match


Netscape
Your Certificates
Security Info
Passwords
Navigator
Messenger
Java/JavaScript
Certificates
Yours
People
Web Sites
Signers
Cryptographic Modules
View A Personal Certificate - Netscape
This Certificate belongs to:
Michael B Spring
spring@imap.pitt.edu
Network Authorization Account - spring
www.verisign.com/repository/CPS
Incorp. by Ref.,LIAB.LTD(c)96
Affiliate - Class 2 Individual Subscriber CA
University of Pittsburgh
This Certificate was issued by:
VeriSign Trust Network
Affiliate - Class 2 Individual Subscriber CA
University of Pittsburgh
Serial Number:
42:7A:29:FD:64:95:A0:71:A3:FA:D6:75:60:96:C8:57
This Certificate is valid from Mon Aug 10, 1998 to Wed Aug 11, 1999
Certificate Fingerprint:
OK
encrypted mail you have received, and you may have problems identifying yourself to web sites.
Get a Certificate...
Import a Certificate...
OK
Cancel
Help


Netscape
Certificate Signers' Certificates
Security Info
Passwords
Navigator
Messenger
Java/JavaScript
Certificates
Yours
People
Web Sites
Signers
Cryptographic Modules
These certificates identify the certificate signers that you accept:
ValiCert Class 2 VA
ValiCert Class 3 VA
VeriSign Class 1 Primary CA
VeriSign Class 2 Primary CA
VeriSign Class 3 Primary CA
VeriSign Class 4 Primary CA
Verisign Class 1 Public Primary Certification Authority - G2
Verisign Class 1 Public Primary Certification Authority - G3
Verisign Class 2 Public Primary Certification Authority - G2
Verisign Class 2 Public Primary Certification Authority - G3
Verisign Class 3 Public Primary Certification Authority - G2
Verisign Class 3 Public Primary Certification Authority - G3
Verisign Class 4 Public Primary Certification Authority - G2
Verisign Class 4 Public Primary Certification Authority - G3
Edit
Verify
Delete
OK
Cancel
Help

Data Privacy and Regulation

- There is increasing attention to regulation of e-business
  - Process for taxing e-business transactions
  - Process for controlling privacy and use of transaction gathered data
  - Process of controlling access to websites (PICS, robot exclusion)

Cookies

- A cookie is a code delivered by the server to the client in the header of a message
- A cookie header specifies:
  - The name and value of the cookie
  - The expiration date(default session)
  - The domain name(default single machine – cannot be a top level domain)
  - The path (default the path of the resource)
  - Security – if secure only send across https (default not)

Browsers and Cookies

- Browsers can disable cookies or request user approval
- Generally, the cookie name an value are encrypted and interpretable only by the server
- Browsers may have policies related to cookie retention. For example, Netscape:
  - Limits all cookies to under 4KB in size
  - Limits the total number of cookies to 300
  - Limits the cookies per domain to 20

Wallets

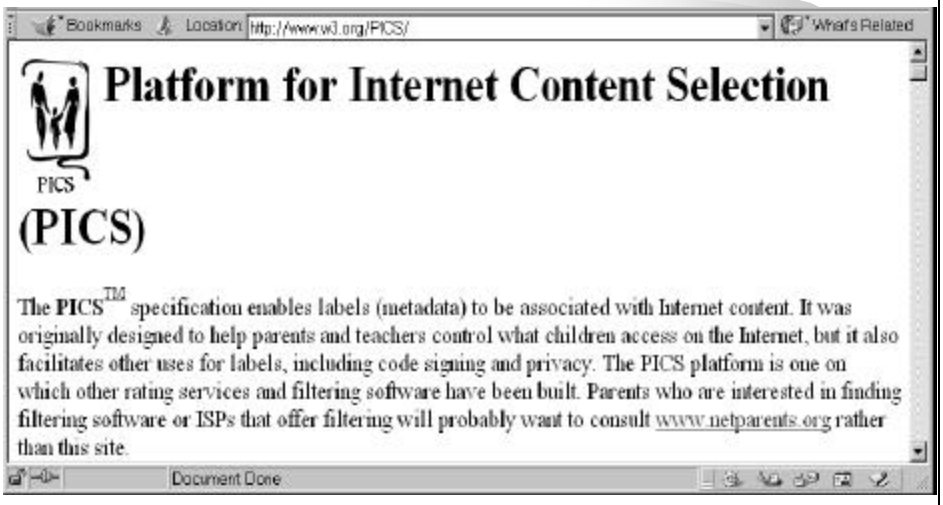
- A cookie might be used as a mechanism to allow a server to access information stored in a DBMS about a user – the cookie is a key.
- User information might include credit card information, shipping preferences, mailing address, etc. In an appropriate context, this information constitutes a wallet.

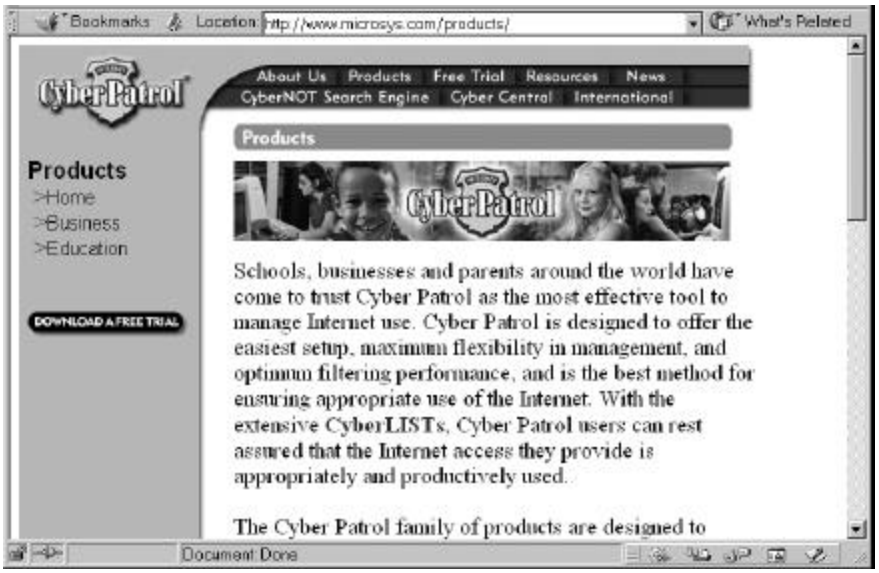
PICS

- PICS or Platform for Internet Content Selection was developed to allow client side control of content
- PICS information is placed in a <META> element in the head of each document
- Browsers implemnting PICS look for these tags to determine if the page is suitable
- Some systems use automated tagging of documents, others use human reviewers

PICS Example

```
<head>
 <META http-equiv="PICS-Label" content='
 (PICS-1.1 "http://www.gcf.org/v2.5"
   labels on "1994.11.05T08:15-0500"
        until "1995.12.31T23:59-0000"
        for
  "http://w3.org/PICS/Overview.html"
   ratings (suds 0.5 density 0 color/hue 1))
 '> </head>
```



Secure Sites

- E-business sites need to be secure in three primary ways:
  - Intranet sites that provide access to mission critical corporate data need to be protected
  - Public sites that collect user information need to secure that information
  - Public sites that are critical to the business mission must be secure from corruption or isolation

Basics

- Web security includes the same basics as an IS system, including such things as flood control and fire suppression, and physical access
- Security management also involves backup of data and duplication of facilities
- Security management also involves monitoring of notices of vulnerability and careful control of all code generated, but particularly that code that processes user inputs

Firewalls

- Firewalls operate at three basic levels
  - Routers that filter on IP's and ports
  - Dual homed bastion servers making internal addresses invisible
  - Proxy servers processing specific protocols
- Firewall configurations can be nested and can combine all three components

Access controls

- Web servers allow for a variety of different access control mechanisms
- The Unix/Apache model is the clearest
  - Each directory may contain an .htaccess file
  - The .htaccess file allows or disallows either individuals or domains
  - In the case of individuals, the users id and password are stored in a .htpasswd file above the web root
- Note that passwords passed from the client to the server based on an authorize challenge are not encrypted – they are simply encoded

Sample .htaccess file

```
<Limit GET POST>
order deny,allow
deny from all
allow from yourdomain.dom
allow user spring jones
</Limit>
```

"Intrusion" Detection

- At its simplest, e-businesses should take care to protect three kinds of intrusion
  - Intrusion based on exploiting software vulnerabilities introduced by CGI programs
  - Intrusion based on exploitation of known vulnerabilities on standard services
  - Intrusion consisting of denial of service based on requests for service

Agents, Spiders, & Robot Exclusion

- Agents and spiders will play an increasingly large role in e-commerce?
  - Agents will be focused on particular tasks
    - Searching for the lowest price for inputs (highest bidder for outputs)
    - Buying (selling) when triggers are met
  - Spiders will be focused on finding information
- Both types of "bots" should be compliant with the robot exclusion rules

Robot Exclusion

- Disallow all robots from the site
  User-agent: *
  Disallow: /
- Disallow a specific robot
  User-agent: cybermapper
  Disallow: /
- Disallow all robots from a given branch (/notes)
  User-agent: *
  Disallow: /notes