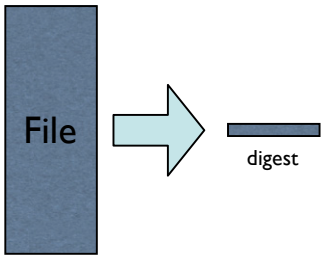
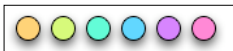
Hour #2
Crypto and
Privacy Protecting Technologies

Message Digests reduce a file to a "fingerprint."

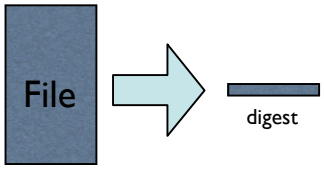
Input: 1-2^{64} bytes

Output: 128, 160, 256 or more bits





The same digest always produces the same output for an input.
Different digests produce different outputs for the same input.



Constitution of the United States of America
(In Convention, September 17, 1787)

Preamble
We the people of the United States, in order to form a more perfect union, establish justice, insure domestic tranquility, provide for the common defense, promote the general welfare, and secure the blessing of liberty to ourselves and our posterity, do ordain and establish the Constitution of the United States of America.

Article I.
Section 1. All legislative powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and a House of Representatives.
...

MD5 → bab1c005bad1ac7d58d54d0e5d0e5f3f

SHA-1 → Ff3881c932e7591e674e2d9d
772817746e8d983f

UNIX and Windows have command-line tools for computing message digests.

```
% ls -l
total 58
-rw-r--r--  1 simsong  wheel  47990 Jul 13  1990 Constitution
-rw-r--r--  1 simsong  wheel   9949 Jul 13  1990 Declaration
%
%  md5 Constitution
MD5 (Constitution) = bab1c005bad1ac7d58d54d0e5d0e5f3f
%
% sha1 Constitution
SHA1 (Constitution) = ff3881c932e7591e674e2d9d772817746e8d983f
%
```

A good message digest is impossible to predict.

Changing one input bit should change ~50% of the output bits.

message	MD5(message)
"this is a test"	ff22941336956098 ae9a564289d1bf1b
"this is c test"	c5e530b91f5f324b 1e64d3ee7a21d573
"this is a test "	6df4c47dba4b01cc f4b5e0d9a7b8d925

Message Digest Algorithms

Rivest Functions:

- MD2 (128 bits)
- MD4 (128 bits)
- MD5 (128 bits)

NIST Functions:

- SHA (160 bits) SHA-1 (160 bits)
- SHA-512, SHA-1024

Other Functions:

- Snerfu, N-Hash, RIPE-MD, HAVAL

There are two ways to "break" a message digest function.

Brute-force attack:

- Search for two messages with the same digest
  (there are many of them!)
- Create many messages until you find a specific digest.

Algorithmic attack

- Use clever math and pre-computation.

Just how big is 2^{128}?

2^{128} = 340,282,366,920,938,463,463,374,
607,431,768,211,456

If you could try a billion^2 combinations a second, it would take 10,790 billion years

- (2^{128} / 10^9 / 10^9 / (60*60*24*365) / 10^9)

MD5 “Broken”

“Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEMD,” Xiaoyun Wang and Dengguo Feng and Xuejia Lai and Hongbo Yu, August 16, 2004

http://eprint.iacr.org/2004/199/

Here is an MD5 collision:

file1.dat:

```
00000000  d1 31 dd 02 c5 e6 ee c4  69 3d 9a 06 98 af f9 5c
00000010  2f ca b5 87 12 46 7e ab  40 04 58 3e b8 fb 7f 89
00000020  55 ad 34 06 09 f4 b3 02  83 e4 88 83 25 71 41 5a
00000030  08 51 25 e8 f7 cd c9 9f  d9 1d bd f2 80 37 3c 5b
00000040  96 0b 1d d1 dc 41 7b 9c  e4 d8 97 f4 5a 65 55 d5
00000050  35 73 9a c7 f0 eb fd 0c  30 29 f1 66 d1 09 b1 8f
00000060  75 27 7f 79 30 d5 5c eb  22 e8 ad ba 79 cc 15 5c
00000070  ed 74 cb dd 5f c5 d3 6d  b1 9b 0a d8 35 cc a7 e3

MD5(file1.dat) = a4c0d35c95a63a805915367dcfe6b751
```

file2.dat:

```
00000000  d1 31 dd 02 c5 e6 ee c4  69 3d 9a 06 98 af f9 5c
00000010  2f ca b5 07 12 46 7e ab  40 04 58 3e b8 fb 7f 89
00000020  55 ad 34 06 09 f4 b3 02  83 e4 88 83 25 f1 41 5a
00000030  08 51 25 e8 f7 cd c9 9f  d9 1d bd 72 80 37 3c 5b
00000040  96 0b 1d d1 dc 41 7b 9c  e4 d8 97 f4 5a 65 55 d5
00000050  35 73 9a 47 f0 eb fd 0c  30 29 f1 66 d1 09 b1 8f
00000060  75 27 7f 79 30 d5 5c eb  22 e8 ad ba 79 4c 15 5c
00000070  ed 74 cb dd 5f c5 d3 6d  b1 9b 0a 58 35 cc a7 e3

MD5(file2.dat) = a4c0d35c95a63a805915367dcfe6b751
```

Uses of Digest Functions

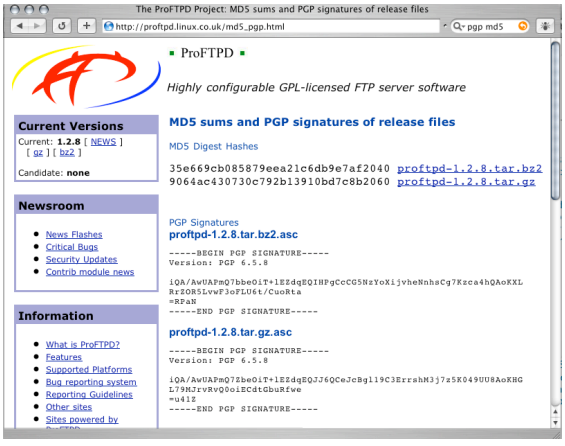
Integrity

- Verifying downloaded code
- Use Digest to determine if two files are identical
- Verifying SSL streams

Authentication

- verifying a shared secret w/o encryption

MD5 hashes are commonly used to verify downloaded code.



In this application, MD5's "break" is irrevellant.

Hashes can be used to securely store passwords.

```
gigawalt:fURfuu4.4hY0U:129:129:Walter Belgers:/home/gigawalt:/bin/csh
root:$1$zlC9.Vfl$9rXSaQqe1HWDaNNOSTJzh.:0:0::0:0:Nitroba &:/root:/bin/tcsh
```

Instead of storing the password, store the hash of the password.

"Cracking" the password requires hashing every password entry to see if it matches the hash.

Unix originally used a DES-based hash, now it uses an MD5 hash.

password file has both salt and encrypted pw

Hash ("Rfuu4.4hY0U")

gigawalt:fURfuu4.4hY0U:129:129:Walter Belgers:/home/gigawalt:/bin/csh

Salt ("fU")

The "salt" assures that the same password can encrypt many different ways.

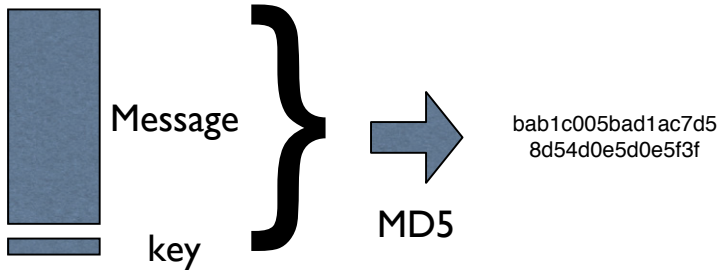
MACs and HMACs allow hash functions to be used for authentication.

MAC = “Message Authentication Code”

HMAC = “Keyed Hashing for Message Authentication” (RFC 2104)

- http://www.ietf.org/rfc/rfc2404.txt
- http://www.cs.ucsd.edu/users/mihir/papers/hmac.html

MACs: The Big Idea



RFC 2104: HMAC

$$HMAC(f,K,M) = f(K \oplus 0x5c^{64} \cdot f(K \oplus 0x36^{64} \cdot M))$$

More complicated than concatenating the key and taking the hash, but more secure!

Uses of HMACs

Data integrity and authentication

- BGP uses HMAC
- IPsec Authentication Header and Encapsulating Security Payload use HMAC as a digital "signature."

Password protocols

Other uses of MACs

Hash Trees - Shurety digital notary

S/KEY

SecureID

Password Challenge-Response

Symmetric EncryptionFunctions

Lucifer

DES

3DES

RC2

RC4

Blowfish

AES

...

"I cannot forecast to you the action of Russia," said Winston Churchil.

"It is a riddle wrapped inside a mystery inside an enigma."

Symmetric Functions: the key that seals also unseals.

M' = f(M,key) *encryption* or *sealing*

M = f'(M',key) *decryption* or *unsealing*

f=f' or f≠f' (some algorithms have a decrypt mode, some don't need it).

Germany used the "Enigma Machine" to encipher communications in WW.

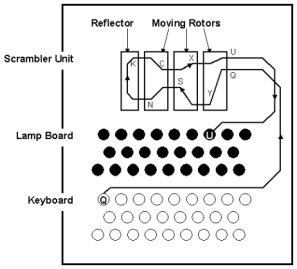
Code clerks set the "code of the day" on dials.

Later models: Set additional code with plugs and wires.

Press a button with the letter to encrypt; the encrypted letter lights up.

Each key press advances the dials

Inside the Enigma



http://www.math.miami.edu/~harald/enigma/enigma.gif

After WW2, cryptography remained a military interest.

Academia was largely disinterested

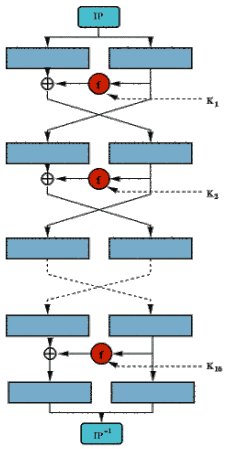
The US National Security Agency became the largest Employer of mathematicians in the world.

IBM developed the "Lucipher" cipher to encrypt data for ATM networks.

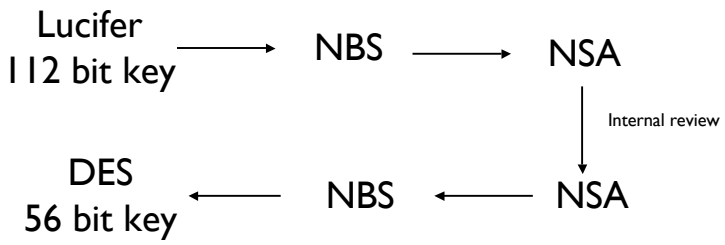
The client: Lloyds of London

The algorithm:

- Symmetric Algorithm
- Fiestel Network
- explicit encrypt/decrypt
- 112 bit key
- Substitution and transposition within 8-character blocks



IBM submitted "Lucifer" to the National Bureau of Standards.



Can you trust DES?

NSA said they made it “better.”

“Better” for who?

- 56 bit key (was 112)
- new “sboxes”
  (what was wrong with old ones?)

In fact, it was more secure, but NSA couldn’t explain why because the Lucifer vulnerability was classified.

Don Coppersmith, ``The Data Encryption Standard (DES) and its strength against attacks,'' *IBM Journal of Research and Development*, 38(1994), pp. 243-250.

The only way to break a DES-encrypted message is to brute force search for a key

In the 1980s, it was hypothesized that someone could build a DES-cracking machine for $1M

In the 1990s, John Gilmore and & EFF built one for $250K. “Deep Crack.” Time to crack a key: 4-7 days. http://www.eff.org/descracker

Nevertheless, DES is still used.

Is weak crypto better than no crypto?

weak crypto	no crypto
stops casual disclosure	doesn't give people a false sense of security
gets people used to use crypto	gives people incentive to move to strong crypto
"Most people don't need crypto anyway"	"so why use it?"

Today there is a wide choice of strong ciphers.

Triple DES (3DES): 3 keys = 168 bits

RC2 & RC4: 40-2048 bits

AES: 128, 192, or 256 bits

Modes of Operation define how a block cipher is used on data longer than a block.

ECB - Electronic Code Book

CBC - Cipher Block Chaining

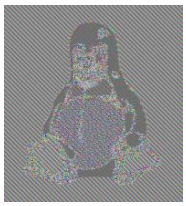
CFB - Cipher Feed Back (XOR generator)

Counter Mode

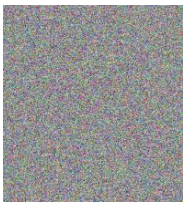
A strong cipher with the wrong mode of operation can have no effective security.

In general, Electronic Code Book (ECB) is easy to implement but not very secure.

original



ECB



CBC

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

Public Key Algorithms

DH
RSA
Digital Signatures
Certs and Certification

Public Key: One key seals, the other key unseals

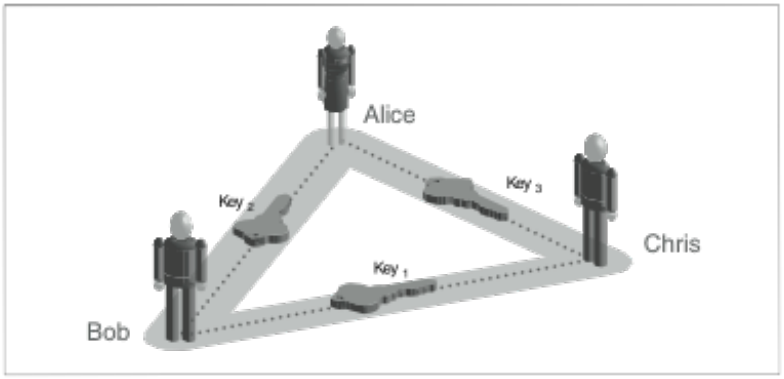
$M' = f(M,K_1)$
$M = f'(M',K_2)$

Obvious today; was revolutionary in 1974!

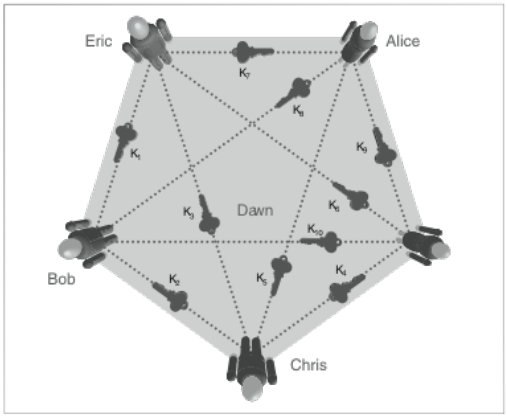
Secret Key vs. Public Key

	secret key	public key
algorithm type	symmetric	asymmetric
basis	substitution and transposition	math
speed	fast	slow
encrypts	blocks of data	numbers
uses	encrypting files	encrypting email

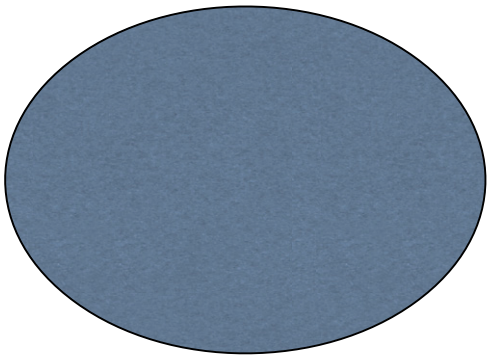
With symmetric cryptography, 3 people need 3 keys to communicate.



Five people need 10 keys to communicate.



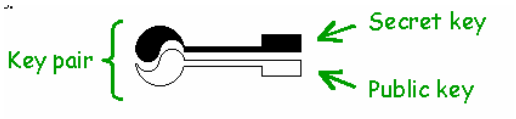
And 1000 people need 499,550 keys to communicate.



$$\#\ \text{keys} = \frac{(n)(n-1)}{2}$$

Public key cryptography uses two keys.

Public key = seals/encrypts data
Private key = unseals/decrypts data



Whitten's "Metaphor Tailoring."

Ralph Merkle figured this out in 1974, but nobody understood it!

Reviewers at ACM didn't understand the project!

- "Too far out of the mainstream of cryptography."
- "Bad science: everybody knows that it is important to keep cryptography keys secret."

Communications finally published the paper in 1978, with an editorial note.

Whitfield Diffie & Matin Hellman

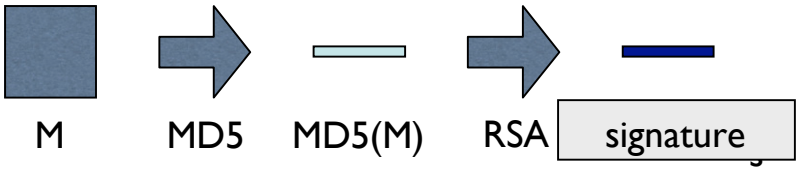
"Multi-User Cryptographic Techniques," written in fall 1975 for the 1976 National Computer Conference

Proposed the idea of Public Key Cryptography.

May 1976 - Diffie Hellman algorithm invented.

Interactive protocol for 2 participants.

Digital Signatures

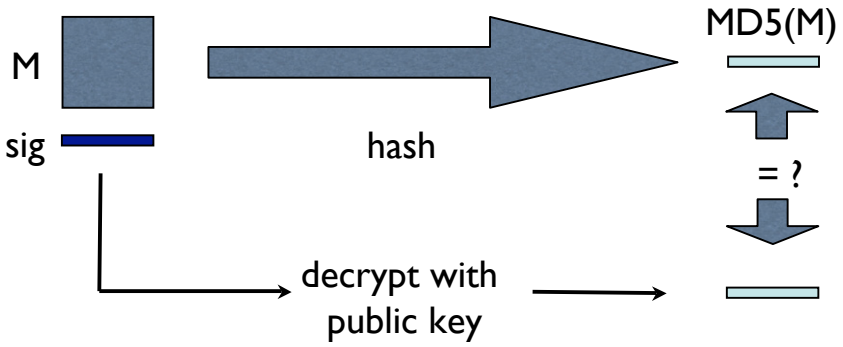


Encrypt with the *secret* key, decrypt with the *public* key.

Used for verifying that the signer had the private key.

Instead of encrypting the entire Message, we usually encrypt a hash

Verifying a Digital Signature



If the hash matches the decrypted signature, the signature verifies!

Using Digital Signatures

To sign a digital signature, you need...

- your private key.

To verify a digital signature, you need...

- the other person's public key...
- the name of the algorithm the person has used for the digital signature.

Certificates bind public keys to identities. [Kohnfelder '78]

"Simson Garfinkel"
KeyID 9c309

Signed by KeyCertCo

Digital Certificates

Certificates "register" public keys

Certificates are signed with digital signatures!

Certificates signed by a "Certificate Authority"

X.509:

Name
Organization
Public Key
Valid from
Valid to
Algorithms
Other info
...

Signature from Certificate Authority

Certificate Authorities

Issue Certificates, not keys

Process:

- User creates public/private keypair
- User sends Certificate Signing Request (CSR) to the CA.
- CA verifies the sender's identity.
- CA sends the certificate back to the

What good is a Certificate from a CA?

In Theory:

- Allows you to "prove your identity" on the Internet. (Age, Sex, Name)
- Allows you to digitally sign documents.
- Allows users to prove "membership" without having to distribute a membership list.

In practice:

- Allows you to run an SSL server without a warning

Certificate Revocation Lists (CRLs)

List of “mistakes.”

- –User lost their Private key.
- –CA signed the wrong key.

Technically, should be checked whenever a CA cert is trusted.

Most application do not check CRLs.

Most public key systems are actually hybrid systems.

- –Use Diffie-Hellman or RSA to exchange a 128-bit session key
- –Use RC2/RC4/AES to encrypt bulk information
- –Use certificates to vouch for public keys.

Certs and Keys with OpenSSL

OpenSSL command-line interface:

- Useful for making keys, certs and CSRs.
- Useful for simple testing
- Useful for converting one format to another (handles PKCS, PEM, and others)
- Useful for testing SSL servers

OpenSSL Commands

ca - Certificate Authority Management
ciphers - lists ciphers in your implementation
crl - Manage Certificate Revocation Lists
dgst - calculation of md digests
dsa - Manages DSA algorithm
dsparam - Generate and manage DH keys

Random Numbers are *Very Important* for public key cryptography:

Random Numbers

- –Use them to pick your initial public/private key pair.
- –Use them for picking session keys

Come to think of it, they are important for symmetric key cryptography too!

Sources of Random Numbers

good	bad
keystroke timing	time of day
packet timing (*)	process ID
radiation, lava lamp	rand(), random()
FM radio	ethernet address
microphone	blocks of CDROMs

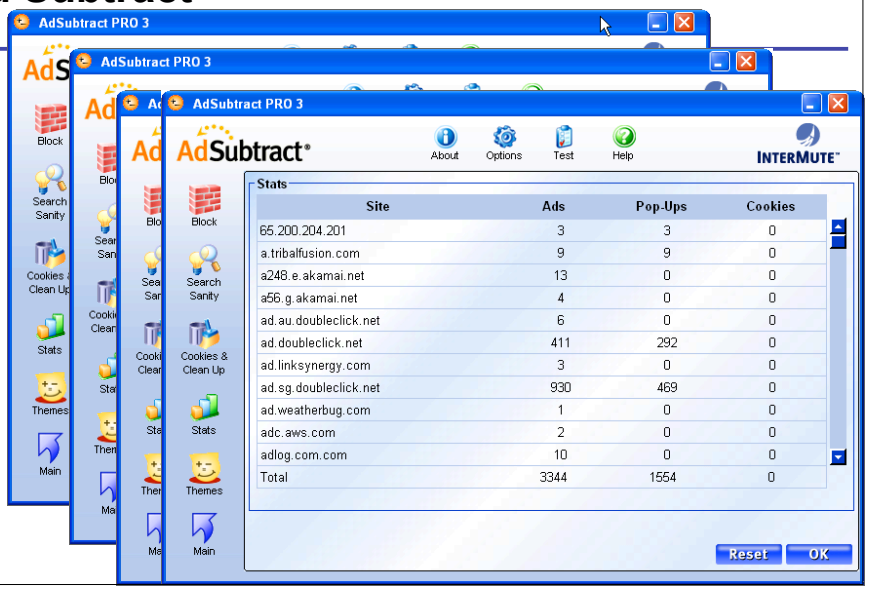
The OpenSSL system supports many ciphers and MAC functions

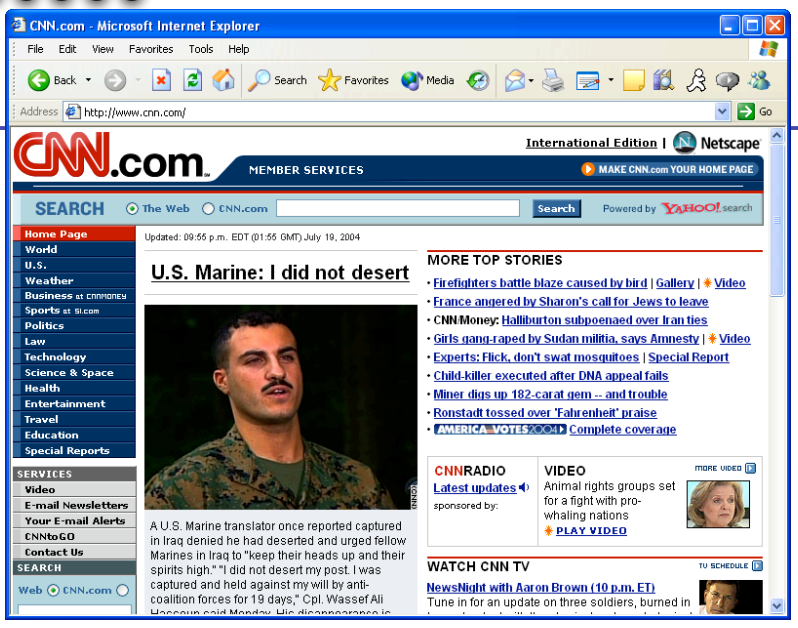
% **openssl ciphers**
EDH-RSA-DES-CBC3-SHA:EDH-DSS-DES-CBC3-SHA:DES-CBC3-SHA:DES-CBC3-MD5:DHE-DSS-RC4-SHA:IDEA-CBC-SHA:RC4-SHA:RC4-MD5:IDEA-CBC-MD5:RC2-CBC-MD5:RC4-MD5:RC4-64-MD5:EXP1024-DHE-DSS-RC4-SHA:EXP1024-RC4-SHA:EXP1024-DHE-DSS-DES-CBC-SHA:EXP1024-DES-CBC-SHA:EXP1024-RC2-CBC-MD5:EXP1024-RC4-MD5:EDH-RSA-DES-CBC-SHA:EDH-DSS-DES-CBC-SHA:DES-CBC-SHA:DES-CBC-MD5:EXP-EDH-RSA-DES-CBC-SHA:EXP-EDH-DSS-DES-CBC-SHA:EXP-DES-CBC-SHA:EXP-RC2-CBC-MD5:EXP-RC4-MD5:EXP-RC2-CBC-MD5:EXP-RC4-MD5

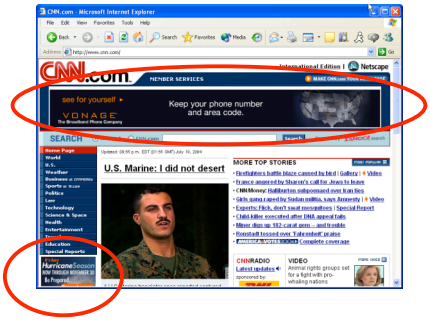
Privacy Protecting Technologies

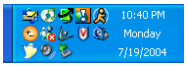
Using technology to improve privacy.

Ad Subtract


AdSubtract PRO 3
AdSubtract PRO 3
AdSubtract PRO 3
AdSubtract
About
Options
Test
Help
INTERMUTE
Block
Search Sanity
Cookies & Clean Up
Stats
Themes
Main
Stats
Site
Ads
Pop-Ups
Cookies
65.200.204.201 3 3 0
a.tribalfusion.com 9 9 0
a248.e.akamai.net 13 0 0
a56.g.akamai.net 4 0 0
ad.au.doubleclick.net 6 0 0
ad.doubleclick.net 411 292 0
ad.linksynergy.com 3 0 0
ad.sg.doubleclick.net 930 469 0
ad.weatherbug.com 1 0 0
adc.aws.com 2 0 0
adlog.com.com 10 0 0
Total 3344 1554 0
Reset
OK


CNN.com - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://www.cnn.com/
Go
CNN.com
International Edition | Netscape
MEMBER SERVICES
MAKE CNN.com YOUR HOME PAGE
SEARCH
The Web
CNN.com
Search
Powered by YAHOO! search
Home Page
World
U.S.
Weather
Business at CNNMONEY
Sports at SI.com
Politics
Law
Technology
Science & Space
Health
Entertainment
Travel
Education
Special Reports
SERVICES
Video
E-mail Newsletters
Your E-mail Alerts
CNNtoGO
Contact Us
SEARCH
Web
CNN.com
Updated: 09:55 p.m. EDT (01:55 GMT) July 19, 2004
U.S. Marine: I did not desert
A U.S. Marine translator once reported captured in Iraq denied he had deserted and urged fellow Marines in Iraq to "keep their heads up and their spirits high." "I did not desert my post. I was captured and held against my will by anti-coalition forces for 19 days," Cpl. Wassef Ali
MORE TOP STORIES
Firefighters battle blaze caused by bird | Gallery | Video
France angered by Sharon's call for Jews to leave
CNN/Money: Halliburton subpoenaed over Iran ties
Girls gang-raped by Sudan militia, says Amnesty | Video
Experts: Flick, don't swat mosquitoes | Special Report
Child-killer executed after DNA appeal fails
Miner digs up 182-carat gem -- and trouble
Ronstadt tossed over 'Fahrenheit' praise
AMERICA VOTES 2004 Complete coverage
CNNRADIO
Latest updates
sponsored by:
VIDEO
Animal rights groups set for a fight with pro-whaling nations
PLAY VIDEO
MORE VIDEO
WATCH CNN TV
TV SCHEDULE
NewsNight with Aaron Brown (10 p.m. ET)
Tune in for an update on three soldiers, burned in



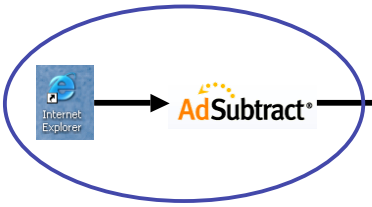


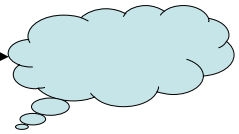
"Search Sanity"

Ad Subtract: Client-Side Java Proxy

Advantages:

- Multiplatform
- Easy to debug
- Client/server

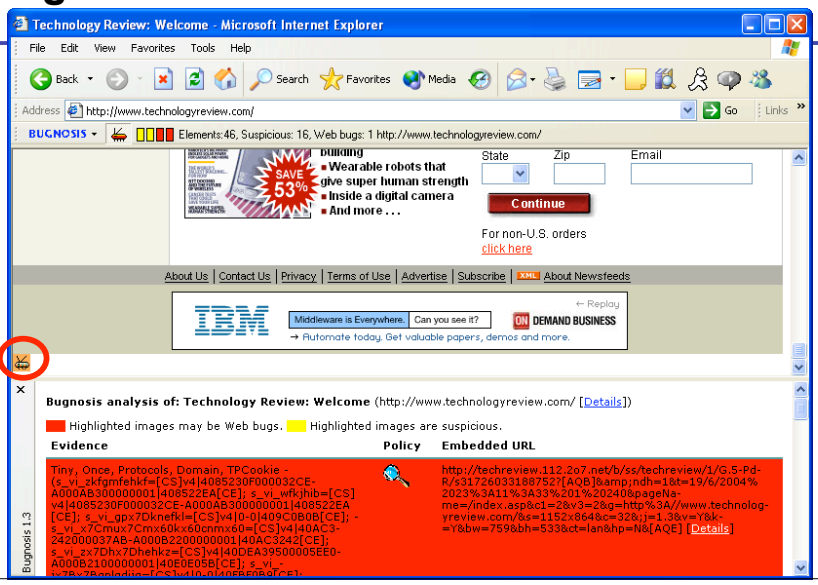




Disadvantages:

- Doesn't work with SSL
- Install footprint
- Need to parse HTML

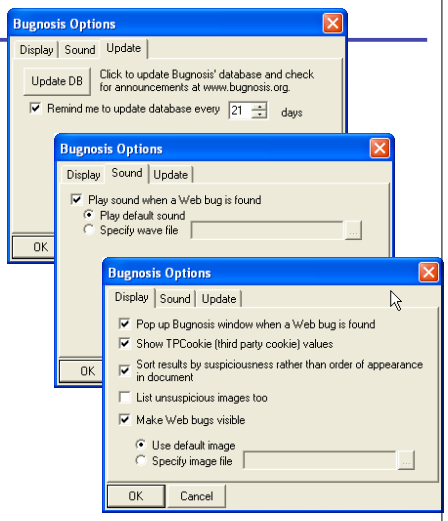
Bugnosis



Bugnosis

Features:

- Browser helper object
- Accesses HTTP & HTTPS
- Downloads updates
- Designed for journalists



Private Messaging

PGP – first generation

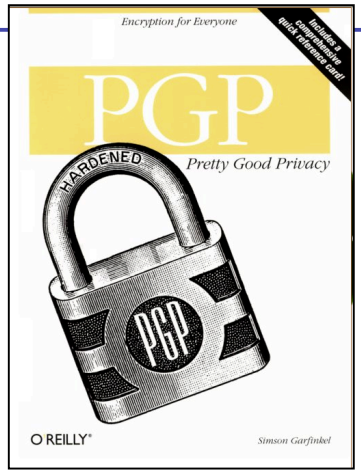
Hush Mail – web based

The Martus Project – application specific

Disappearing Ink (Omniva) - Deletion

PGP

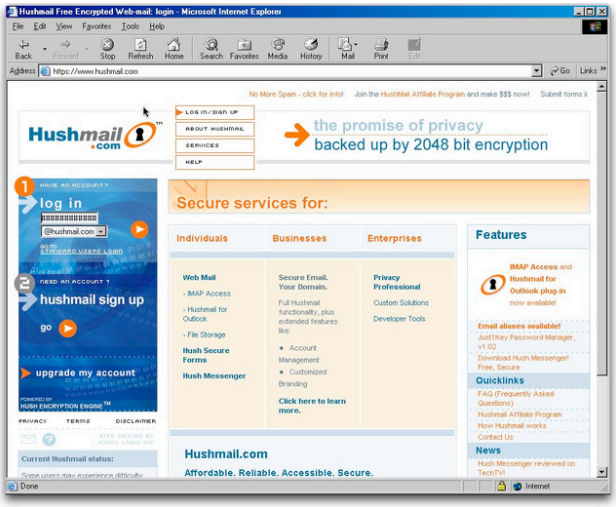
Add-on
Plug-in
S/MIME vs. OpenPGP
Political Baggage



Hush Mail

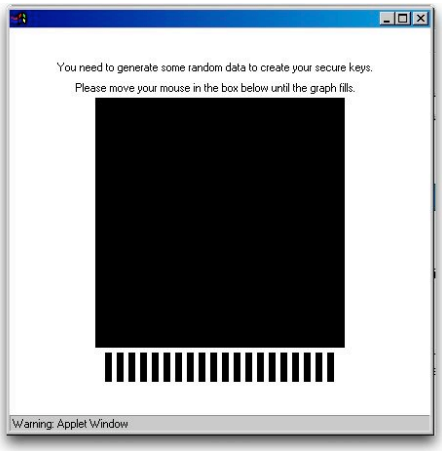
Second-generation
Web-based
Java Crypto Client

Hush Mail

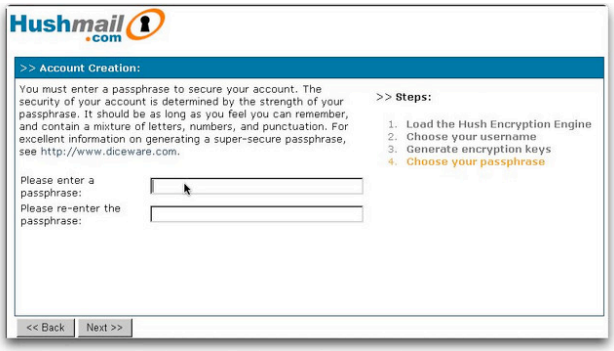


Hush Mail

Hush Mail

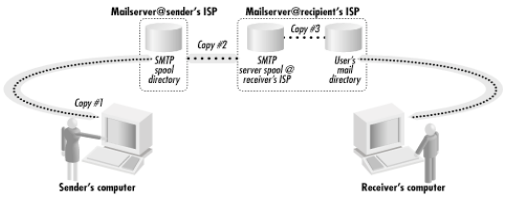


Hush Mail

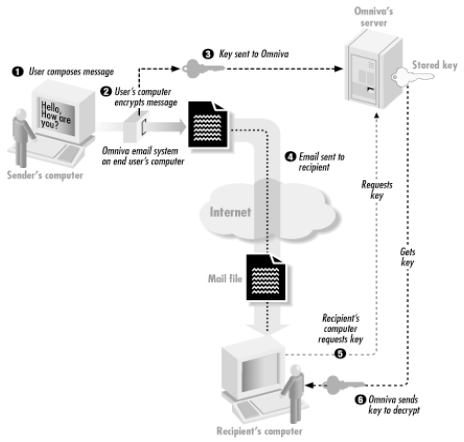


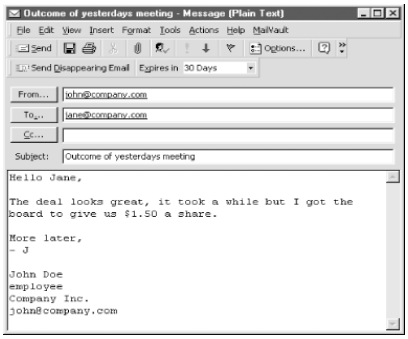
Hush Mail Interface

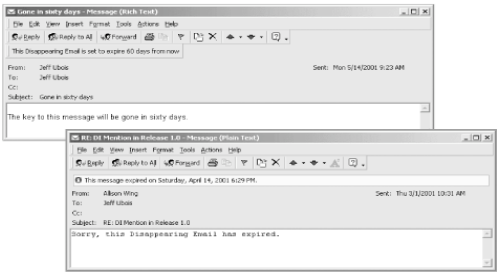
Email gets copies a lot; each copy can violate a person's privacy.



Omniva encrypts email messages with a key that is deleted at a predetermined time.







Mix-Nets, Web & IP transport

Chaum's mix-net scheme

The big idea: anonymity needs company

1 mix: you trust the mixer

More mixes -> Less Trust

Mixing needs to be in space and time

Key features of an anonymous remailer

Strips identity from messages passing through

Provides mapping of nyms to “true names”

- But only if replies are important

Optional:

- Mixing - only if traffic in and out is observable
- Encryption -

Anonymous Web Browsing

Web Caches

Anonymizer

Anonymous Transport Services:

- Freedom
- Onion Routing

Web Caches provided low-cost privacy protection.

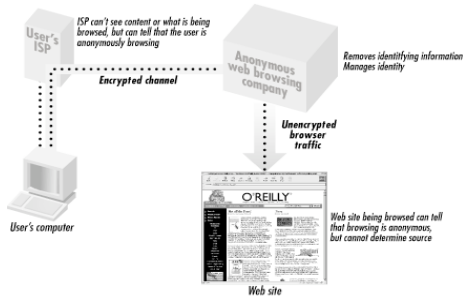
cache-ntc-ah12.proxy.aol.com - - [10/May/2003:22:47:31 -0400] "GET /clips/1999.TR.LCS35-FountainOfIdeas.pdf HTTP/1.0" 200 65536 "http://aolsearch.aol.com/aol/search?query=fountain+ideas&page=2" "Mozilla/4.0 (compatible; MSIE 6.0; AOL 7.0; Windows NT 5.1; .NET CLR 1.0.3705)“

cache-ntc-ah12.proxy.aol.com - - [10/May/2003:22:47:39 -0400] "GET /clips/1999.TR.LCS35-FountainOfIdeas.pdf HTTP/1.1" 206 688128 "-" "Mozilla/4.0 (compatible; MSIE 6.0; AOL 7.0; Windows NT 5.1; .NET CLR 1.0.3705)“

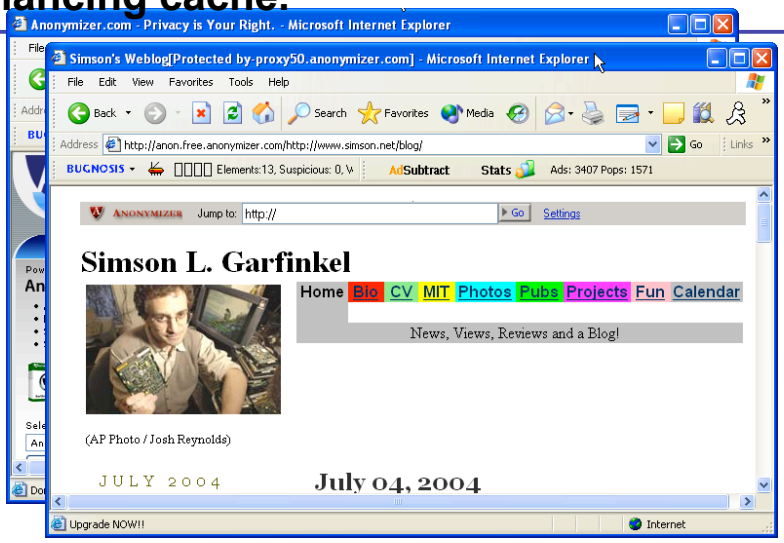
cache-ntc-ah12.proxy.aol.com - - [10/May/2003:22:47:44 -0400] "GET /clips/1999.TR.LCS35-FountainOfIdeas.pdf HTTP/1.1" 206 1024 "-" "Mozilla/4.0 (compatible; MSIE 6.0; AOL 7.0; Windows NT 5.1; .NET CLR 1.0.3705)“

cache-ntc-ah12.proxy.aol.com - - [10/May/2003:22:47:47 -0400] "GET /clips/1999.TR.LCS35-FountainOfIdeas.pdf HTTP/1.1" 206 75 "-" "Mozilla/4.0

Third-party caches can make this technology widely available.



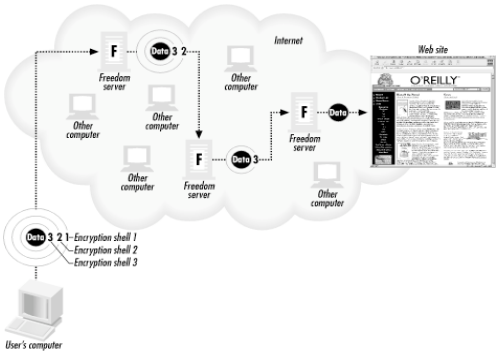
Anonymizer is a commercial privacy-enhancing cache.



Anonymizer.com rewrites URLs

```
<td width=90 style='background:aqua;
                text-align:center; font:bold; font-family:Arial'>
    <a href='http://anon.free.anonymizer.com/http://www.simson.net/photos.php'
title='Photos by and of Simson Garfinkel'> Photos </a>
  </td>
  <td width=90 style='background:lime;
                text-align:center; font:bold; font-family:Arial'>
    <a href='http://anon.free.anonymizer.com/http://www.simson.net/pubs.php'
title='Publications, both academic and journalistic.'> Pubs </a>
  </td>
   <td width=90 style='background:magenta;
                text-align:center; font:bold; font-family:Arial'>
    <a href='http://anon.free.anonymizer.com/http://www.simson.net/projects.php'
title='Current projects'> Projects </a>
  </td>
```

Onion Routing (Freedom Network & others) provide anonymity at the IP level.



[Anonymous] Publication Systems can provide a different kind of anonymity.

Napster (1999 – 2001)

KaZaA ?

Freenet

Freehaven

For further information:

EPIC Online Guide to Privacy Protecting tools:
http://www.epic.org/privacy/tools.html