

COMPUTER SECURITY

HASH FUNCTIONS

MESSAGE AUTHENTICATION CODES

Willy Sudiarto Raharjo

Teknik Informatika
Universitas Kristen Duta Wacana

Message Digest

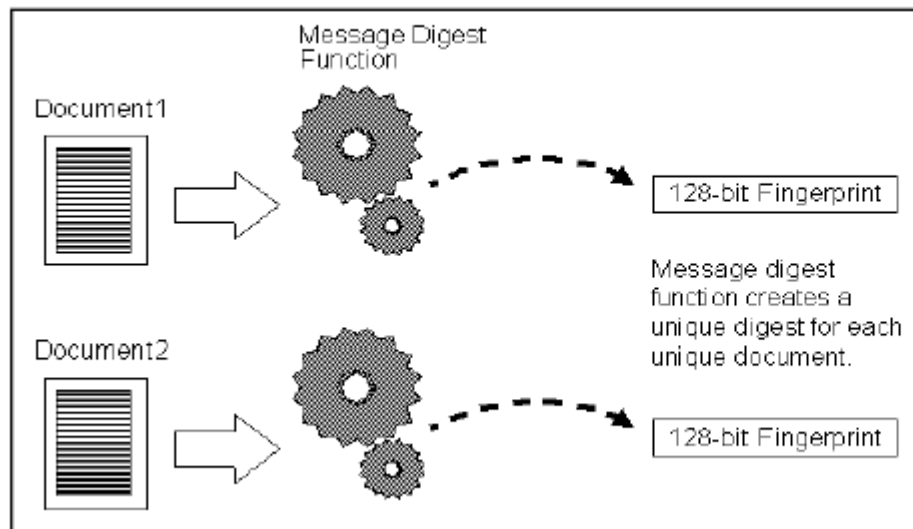
Message digest functions change the information contained in a file, (small or large) into a single large number, typically between 128 and 256 bits in length

The best message digest functions combine these mathematical properties

Every bit of the message digest function is influenced by the function's input

If any given bit of the function's input is changed, every output bit has a 50 percent chance of changing

Given an input file and its corresponding message digest, it should be computationally infeasible to find another file with the same message digest value



HASH Functions

Message digests are also called one-way hash functions because they produce values that are difficult to invert, resistant to attack, mostly unique, and are widely distributed

Message digest algorithms themselves are not used for encryption and decryption operations

They are used in the creation of digital signatures, message authentication codes (MACs), and encryption keys from passphrases

Message digest functions:

- HMAC
- MD2
- MD4
- MD5
- SHA
- SHA-1



HASH Function

- Compression
 - ◆ For any size of input x , the output of $y = h(x)$ is small and fixed
- Efficiency
 - ◆ Easy to compute $h(x)$ for any input x
- One-way
 - ◆ Difficult to invert $h(x) = y$ given any value of y
- Weak collision resistance
 - ◆ Difficult to find y , given x and $h(x)$ such that $h(y) = h(x)$ and $y \neq x$
- Strong collision resistance
 - ◆ Difficult to find x and y , such that $h(x) = h(y)$ and $x \neq y$

Pre Birthday Problem

- Suppose N people in a room
- How large must N be before the probability someone has same birthday as you is $\geq 1/2$?
- Probability of random people doesn't have the same birthday as you: $364/365$
- Probability of none of N people have the same birthday as you: $(364/365)^N$
- Probability at least 1 person has the same birthday: $1 - (364/365)^N$
- Solve: $1/2 = 1 - (364/365)^N$ for N
- We find $N = 253$

- How many people must be in a room before probability is $\geq 1/2$ that any two (or more) have same birthday?

$$1 - 365/365 \dots 364/365 \dots (365 - N + 1) / 365$$

- Set equal to $1/2$ and solve: $N = 23$
- Surprising? A paradox?

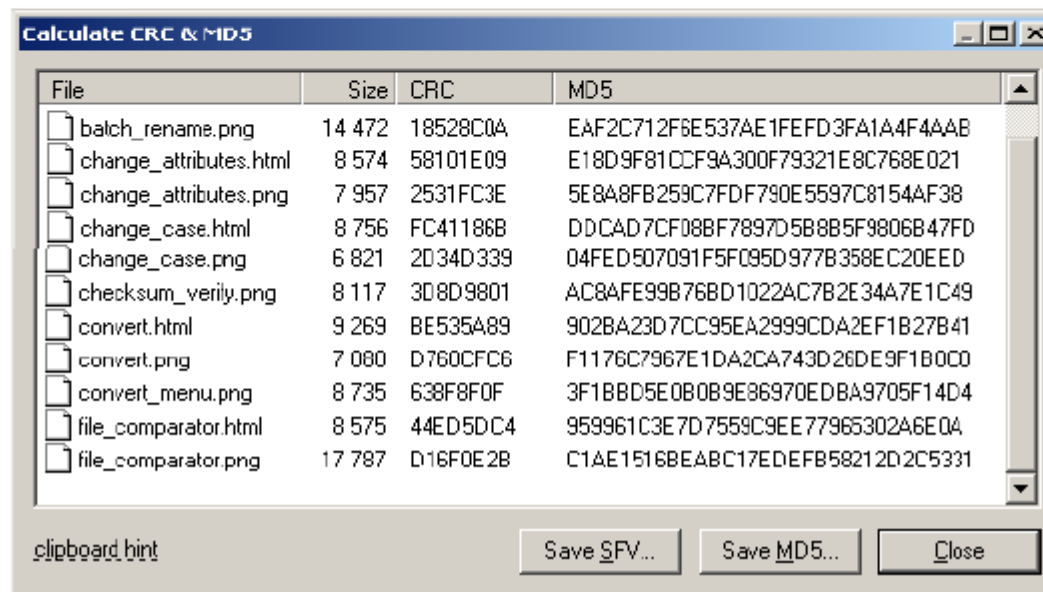
Implication:

secure N bit symmetric key requires 2^{N-1} work to “break”
secure N bit hash requires $2^{N/2}$ work to “break”

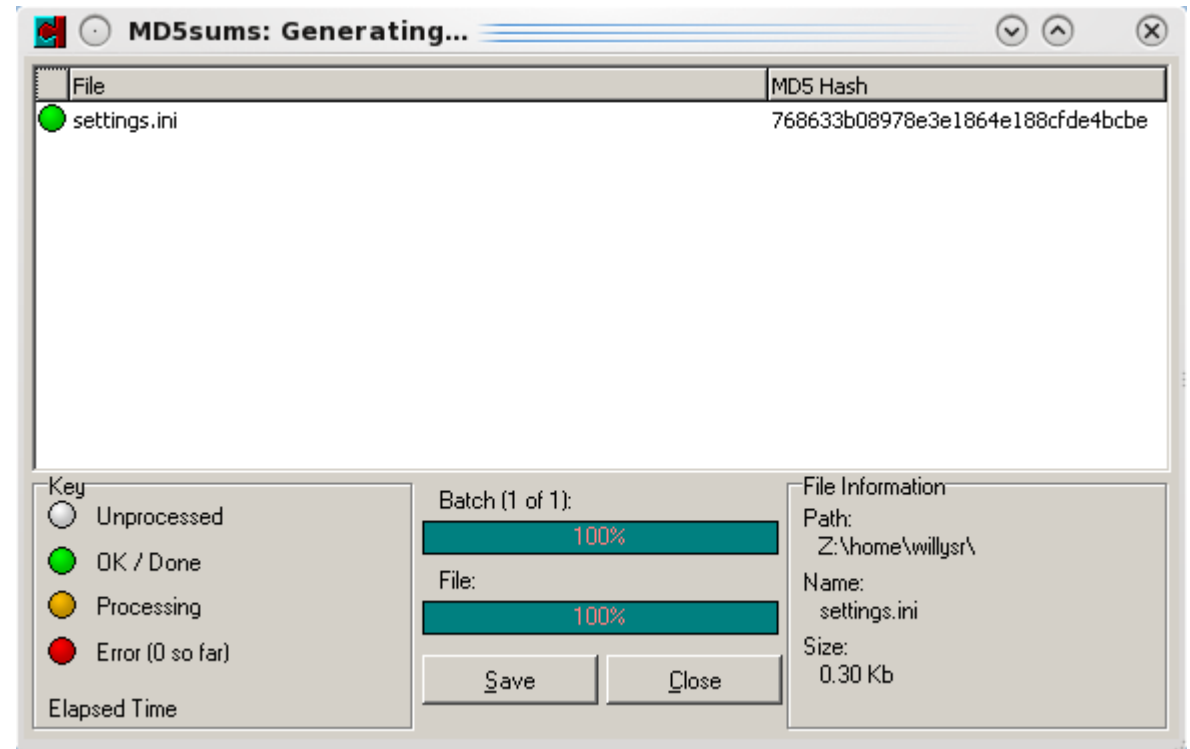
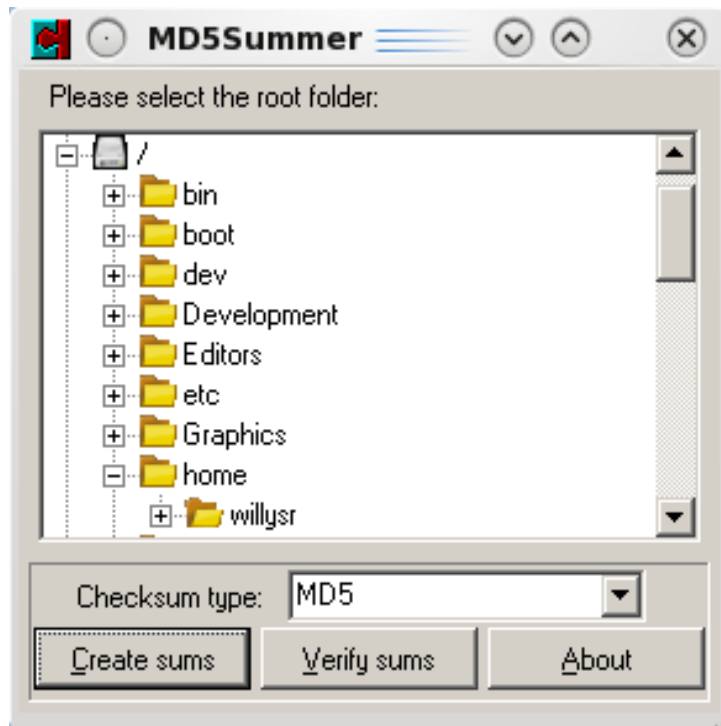
MD5

The MD5 algorithm takes as input, a message of arbitrary length, and outputs a 128-bit fingerprint or message digest of the input

The MD5 algorithm is intended for digital signature applications, where a large file is compressed in a secure manner before being encrypted with a private (secret) key under a public-key cryptosystem, such as RSA



MD5Summer

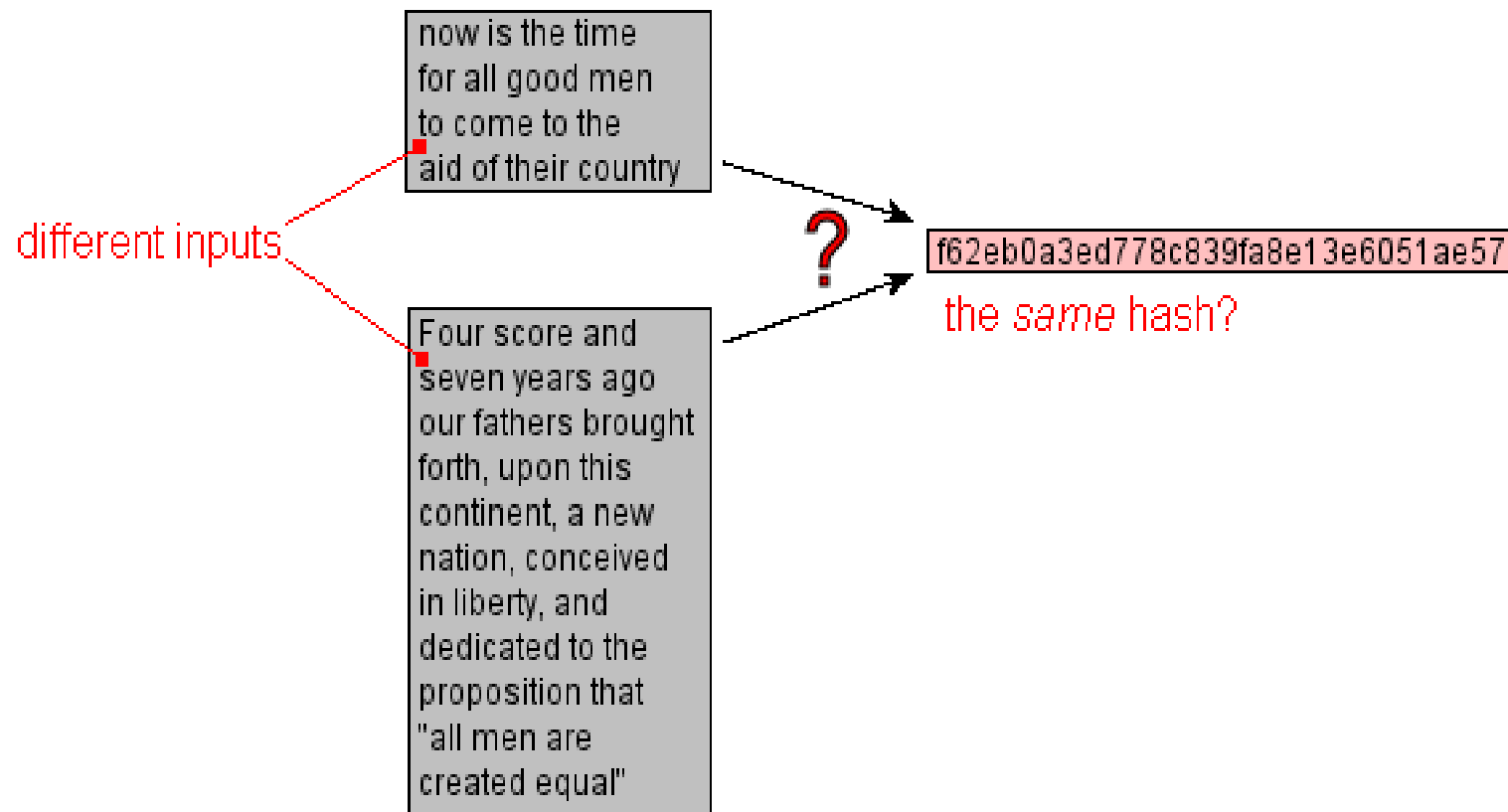


<http://md5summer.org/>

Collision

- Different contents but results in same hash values
- Input space is much larger than output space
- Example:
 - ◆ Hash function generates 128-bit output
 - ◆ Possible input values: 150-bit
 - ◆ On average, 2^{22} or more than 4.000.000 of these input values hash to each possible output value

MD5 Collision Example



Real MD5 Collision Example

```
#!/usr/bin/perl -w

use strict;

my $v1=<<END_V1;
d1 31 dd 02 c5 e6 ee c4 69 3d 9a 06 98 af f9 5c
2f ca b5 87 12 46 7e ab 40 04 58 3e b8 fb 7f 89
55 ad 34 06 09 f4 b3 02 83 e4 88 83 25 71 41 5a
08 51 25 e8 f7 cd c9 9f d9 1d bd f2 80 37 3c 5b
d8 82 3e 31 56 34 8f 5b ae 6d ac d4 36 c9 19 c6
dd 53 e2 b4 87 da 03 fd 02 39 63 06 d2 48 cd a0
e9 9f 33 42 0f 57 7e e8 ce 54 b6 70 80 a8 0d 1e
c6 98 21 bc b6 a8 83 93 96 f9 65 2b 6f f7 2a 70
END_V1

my $v2=<<END_V2;
d1 31 dd 02 c5 e6 ee c4 69 3d 9a 06 98 af f9 5c
2f ca b5 07 12 46 7e ab 40 04 58 3e b8 fb 7f 89
55 ad 34 06 09 f4 b3 02 83 e4 88 83 25 f1 41 5a
08 51 25 e8 f7 cd c9 9f d9 1d bd 72 80 37 3c 5b
d8 82 3e 31 56 34 8f 5b ae 6d ac d4 36 c9 19 c6
dd 53 e2 34 87 da 03 fd 02 39 63 06 d2 48 cd a0
e9 9f 33 42 0f 57 7e e8 ce 54 b6 70 80 28 0d 1e
c6 98 21 bc b6 a8 83 93 96 f9 65 ab 6f f7 2a 70
END_V2

my $p=join("",map {chr(hex($_))} split /\s+/, $v1);
my $q=join("",map {chr(hex($_))} split /\s+/, $v2);

print `echo -n \"'$p'\"|md5sum`;
print `echo -n \"'$q'\"|md5sum`;
```

SHA

The SHA algorithm takes a message of the arbitrary length as input and outputs a 160-bit fingerprint or message digest of the input

The algorithm is slightly slower than MD5, but the larger message digest makes it more secure against brute-force collision and inversion attacks

Key Strength

40-bit key algorithms are of no use

56-bit key algorithms offer privacy, but are vulnerable

64-bit key algorithms are safe today but will be soon threatened as the technology evolves

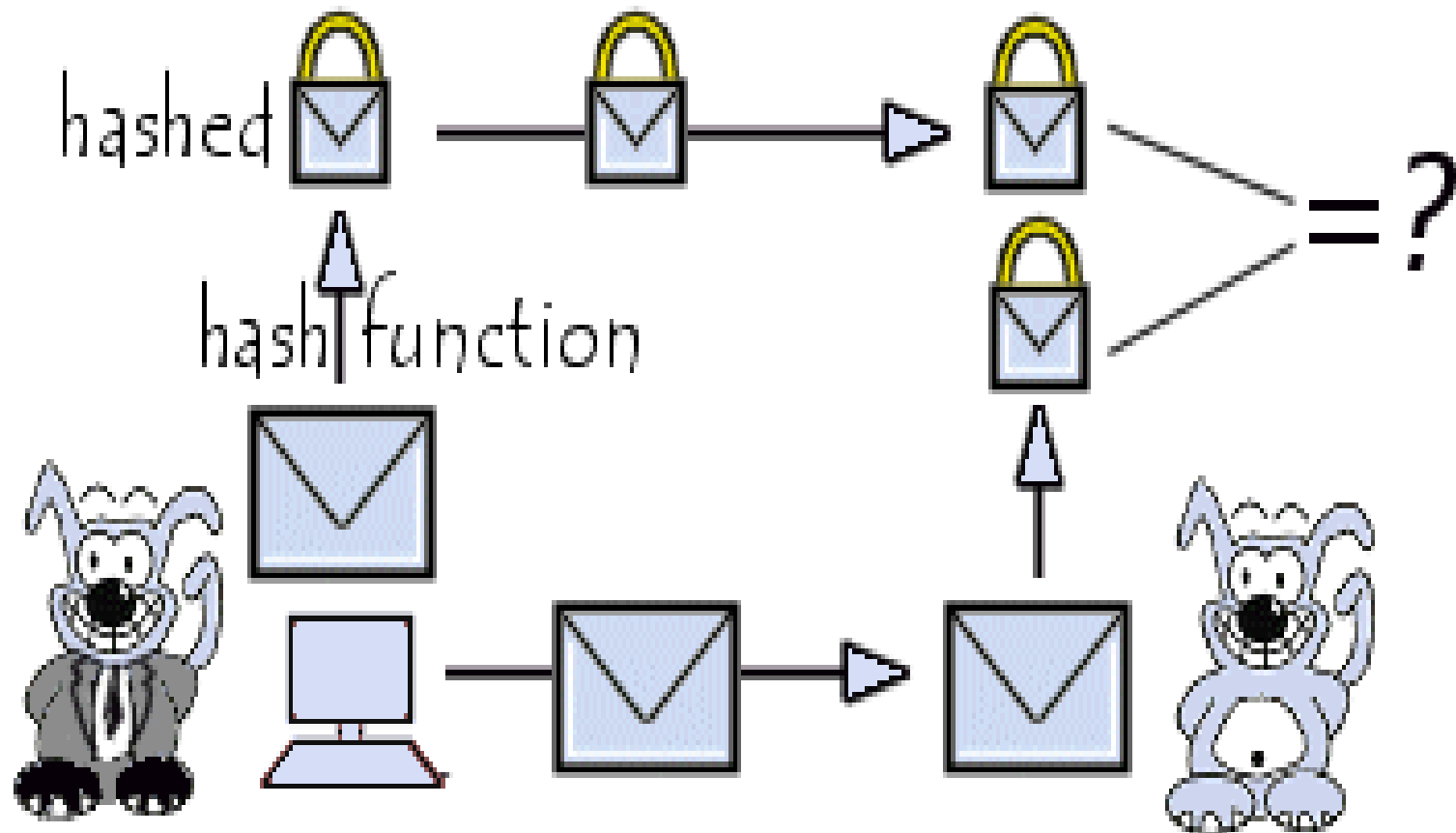
128-bit and over algorithms are almost unbreakable

256-bit and above are impossible

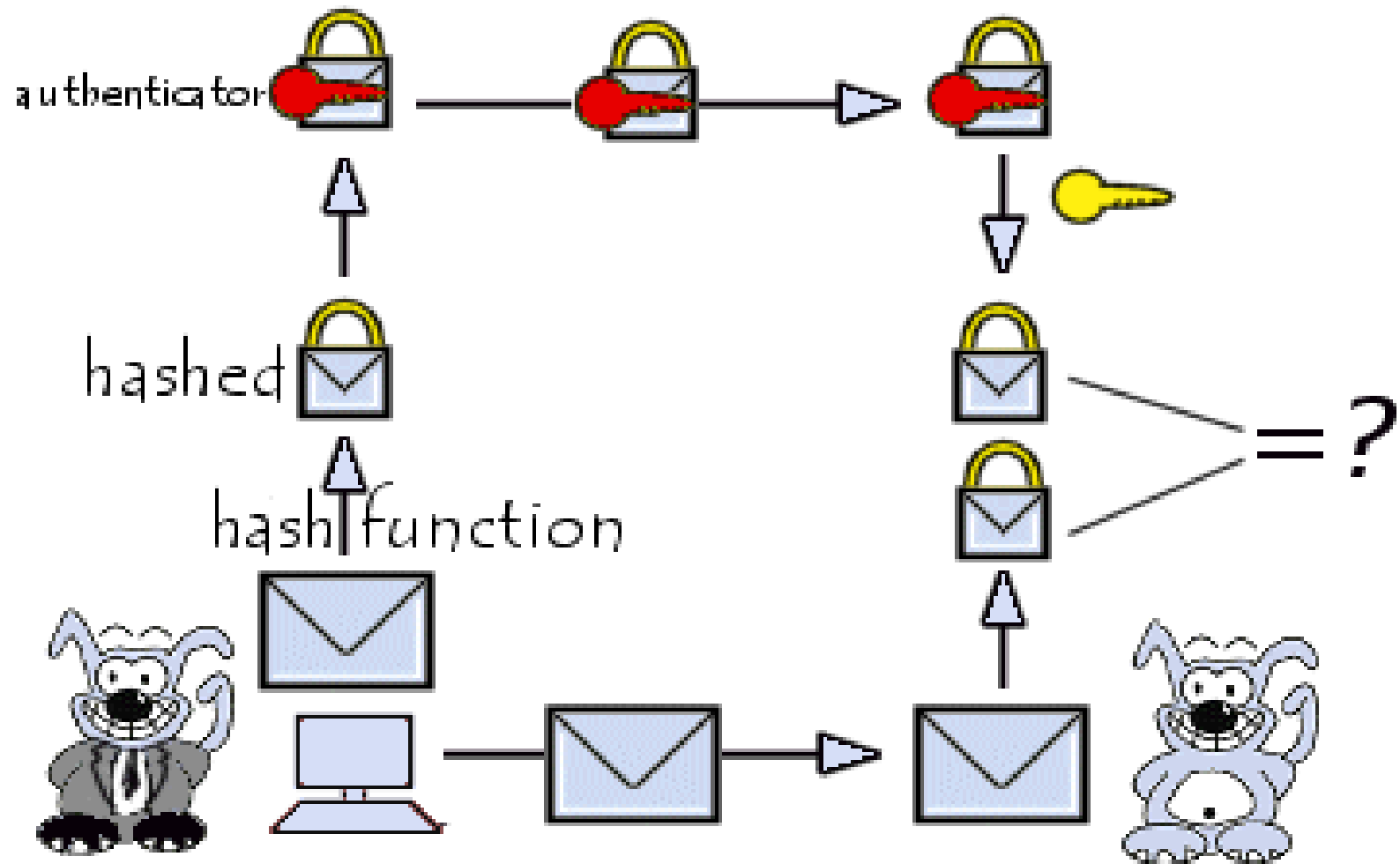


Algorithm	Output size (bits)	Internal state size	Block size	Length size	Word size	Collision attacks (complexity)	Preimage attacks (complexity)
GOST	256	256	256	256	32	No	
HAVAL	256/224/192/160/128	256	1024	64	32	Yes	
MD2	128	384	128	No	32	Almost	
MD4	128	128	512	64	32	Yes (2^8) ^[10]	With flaws (2^{102}) ^[11]
MD5	128	128	512	64	32	Yes (2^{32}) ^[12]	
PANAMA	256	8736	256	No	32	Yes	
RadioGatún	Arbitrarily long	58 words	3 words	No	1-64	With flaws (2^{352} or 2^{704}) ^[13]	
RIPEMD	128	128	512	64	32	Yes	
RIPEMD-128/256	128/256	128/256	512	64	32	No	
RIPEMD-160/320	160/320	160/320	512	64	32	No	
SHA-0	160	160	512	64	32	Yes (2^{39}) ^[14]	
SHA-1	160	160	512	64	40	Proven (2^{63}), With flaws (2^{52}) ^[15]	No
SHA-256/224	256/224	256	512	64	32	No	No
SHA-512/384	512/384	512	1024	128	64	No	No
Tiger(2)-192/160/128	192/160/128	192	512	64	64	"Pseudo-near collision" ^[16]	
WHIRLPOOL	512	512	512	256	8	No	

Integrity verification



Data sealing



MAC

- Message Authentication Codes
- Information used to authenticate a message
- Input: Secret Key + Message
- Output: MAC
- Provides
 - ◆ Data integrity
 - ◆ Authenticity
- But does not provide non repudiation. Why??

MAC

