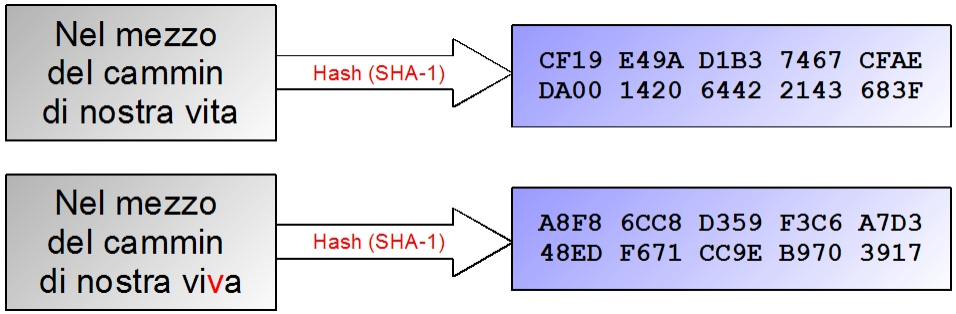
Metodologie e strumenti per la crittoanalisi della funzione di hash SHA-1 e sue implicazioni sulla sicurezza di rete

Luigi Esposito

Prima giornata di incontro con Borsisti GARR, Roma, 22.06.2010

Funzioni di hash crittografiche

- Associano a messaggi di lunghezza arbitraria brevi stringhe di lunghezza fissa, dette valori di *hash*



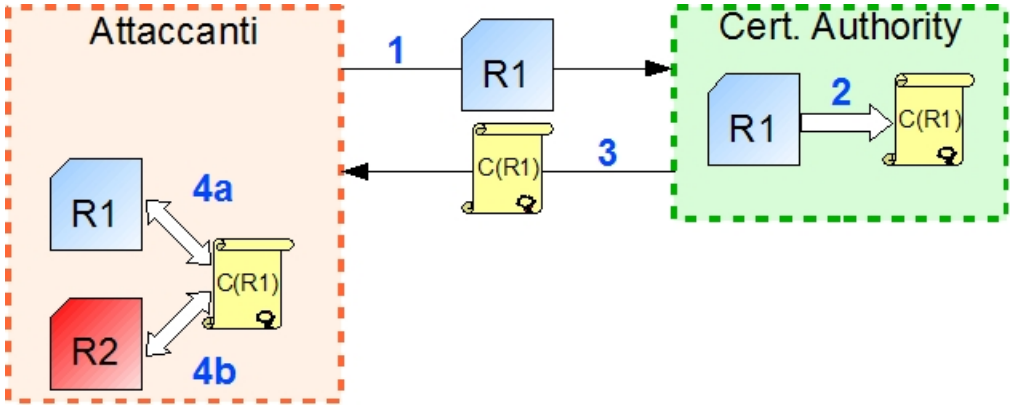
- Applicazioni:
  - Verifica dell'integrità ed autenticità dei messaggi in vari protocolli di rete
    - A livello applicazione (PGP, S/MIME, SSH)
    - A livello trasporto (TLS/SSL)
    - A livello networking (IPsec)
  - Identificazione di file, dati o software
    - Controllo versione software: Git, Mercurial, Monotone
    - Identificazione file (ad es. nei P2P e negli archivi)
  - Derivazione di sequenze di chiavi e password

Funzioni di hash crittografiche – Proprietà

- Tre proprietà rispetto alle funzioni di hash semplici
  - Resistenza all'individuazione della prima preimmagine
    - Dato H(m), impraticabile risalire ad m
  - Resistenza all'individuazione della seconda preimmagine
    - Dati m_1 ed $H(m_1)$, impraticabile trovare $m_2 \neq m_1$ tale che $H(m_2) = H(m_1)$
  - Resistenza alle collisioni
    - Impraticabile individuare m_1, m_2 con $m_1 \neq m_2$ tali che $H(m_2) = H(m_1)$
- La capacità di resistere alle collisioni è la prima proprietà ad essere attaccata
  - La funzione è considerata a rischio quando è scoperto un attacco più efficiente del *birthday attack*
  - Viene meno il ruolo principale della funzione: quello di essere una "firma" virtualmente univoca del messaggio

Attacco di Sotirov et al. contro MD5

- Alcuni ricercatori costruirono due diverse richieste di certificato digitale significative, che producevano uguale valore di hash
- Dato che una Certification Authority firma soltanto l'hash, facendo validare la prima, ottennero automaticamente validazione anche per la seconda (malevola)



- In conseguenza dell'attacco, MD5 è stata quasi completamente sostituita
  - L'alternativa principale è risultata essere proprio SHA-1

Work Plan (1/2)

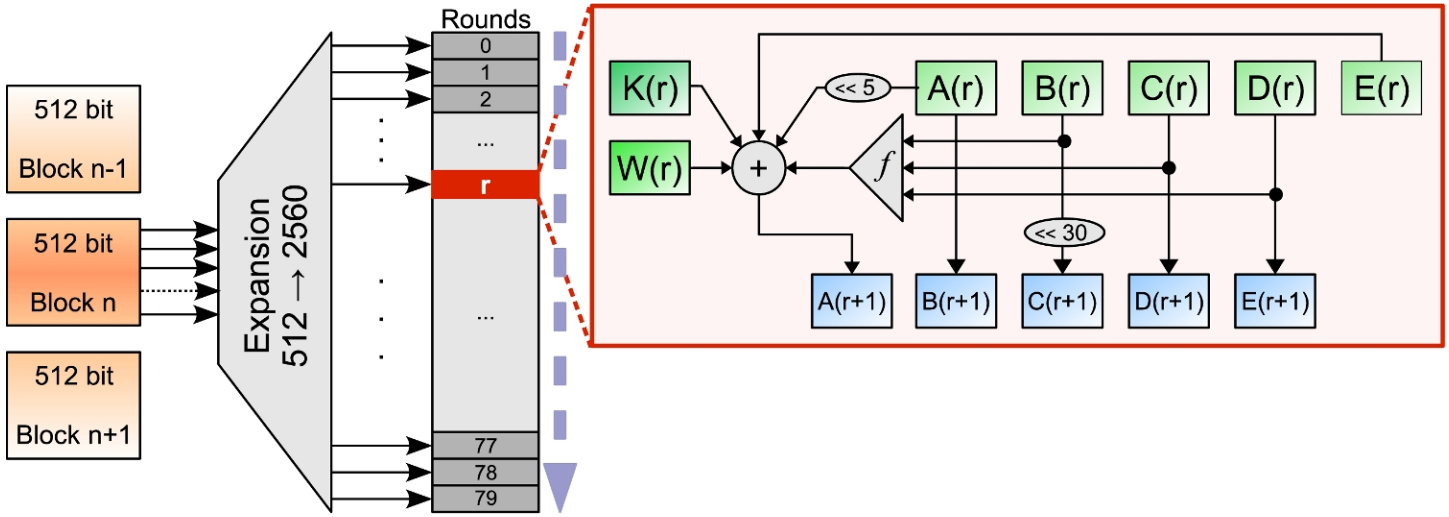
Attività	Avvio (mesi)	Durata (mesi)	Stato
Studio approfondito di funzioni di hash crittografiche e tecniche di crittoanalisi allo stato dell'arte	0	2	✓
Metodi e strumenti per la ricerca automatizzata di collisioni			
Sviluppo applicazione altamente ottimizzata per la ricerca di collisioni in SHA-1, che inglobi le più efficienti tecniche note	2	2	✓
Progettazione e sviluppo di versioni parallele dell'applicazione, per l'adattamento a diverse architetture (HPC, CELL B.E., FPGA)	4	8	⌛
Individuazione di collisioni per versioni di SHA-1 progressivamente più vicine a quella standard	4	12	⌛
Perfezionamento algoritmico anche mediante nuove tecniche e metodologie	6	8	⌛

Work Plan (2/2)

Attività		Avvio (mesi)	Durata (mesi)	Stato
Interscambio con altri gruppi di ricerca attivi in questo settore				
	Per la crittografia differenziale, in particolare su SHA-1	4	12	⏳
	Per l'implementazione ad alta efficienza di algoritmi HPC	4	12	⏳
Valutazione dell'impatto dei risultati				
	Sulle applicazioni che correntemente utilizzano SHA-1	12	3	...
	Su altre funzioni di hash esistenti o in fase di progettazione (es. concorso SHA-3)	12	6	...
Produzione di documentazione tecnica specifica e pubblicazioni scientifiche		4	12	⏳

SHA-1 – Funzionamento

- SHA-1 è stata introdotta nel 1995 dal National Institute of Standards and Technologies, USA
- Soprattutto dopo la dimostrazione di Sotirov et al., è di fatto la funzione di hash maggiormente utilizzata
- L'algoritmo ha natura iterativa
  - 80 round per ciascun blocco da 512 bit processato



SHA-1 – Ricerca di collisioni

- Di base, la ricerca di collisioni per SHA-1 è fondata sulle *caratteristiche differenziali*
  - Sono schemi strutturati di vincoli che intercorrono tra bit omologhi di due messaggi
- Due messaggi *conformi* ad una caratteristica costruita opportunamente hanno maggiore probabilità di collidere rispetto a due messaggi qualunque
- SHA-1 risulta ancora inviolata
  - All'inizio dell'attività erano tuttavia note collisioni per versioni ridotte a 70 round

Caratteristiche differenziali

(x,x*)	(0,0)	(1,0)	(0,1)	(1,1)
#				
0	✓			
u		✓		
3	✓	✓		
n			✓	
5	✓		✓	
x		✓	✓	
7	✓	✓	✓	
1				✓
–	✓			✓
A		✓		✓
B	✓	✓		✓
C			✓	✓
D	✓		✓	✓
E		✓	✓	✓
?	✓	✓	✓	✓

Alfabeto dei simboli che definiscono le tipologie di vincoli

i	M_i
00:	u1nn011001101001-0-010001u1n01nu
01:	10uu1000110-0----1-100-11010n010
02:	00010011011-1-01---0--111n1001nn
03:	nu0n1100011----------1111nun01n1
04:	u0nn011101100---------01101n1011
05:	nnnu111001000011-------0-un000u1
06:	un001111011110------000--n1010u0
07:	0111111010-001011111-0--01uu0101
08:	nun110000010111--00--0-10n0100u0
09:	10n00000001-------100111-nn00101
10:	unu110100-1-0------01----n00100u
11:	10n0101010----------0101u1n0111
12:	un1010110-----------0--1u00101n
13:	nnu0101111-0--1---------1nuu01n1
14:	n1n00001-0---------------01010un
15:	un1001-1--1-1--------------n00n0
16:	0un011011----0---------11u00110u
17:	nu10101-0---------------01nu00n0
18:	00n10-1-10-0--------------00110u
19:	nn101010--------------0-100110n1
20:	unn100-0----------------0n1101u1
21:	1nu1-0-10-0--------------n0111n0
22:	n111100-----------------000101u1
23:	10010-1-------------------001111
24:	001-1-01----------------10010100
25:	n01100------------------01001010
26:	0010-0-------------------00000u1
27:	01-0-10----------------01n111111
28:	10010-------------------00111010
29:	n01-1--------------------x010100
30:	n-1-00-------------------10011110
...	...

Esempio: frammento di relazione caratteristica differenziale

www.garr.it

Risultati già ottenuti (1/2)

- Studio approfondito delle funzioni di hash
  - Particolare riferimento ad MD5, SHA-0, SHA-1
- Collaborazione con T. Peyrin (Ingenico), tra i principali esperti di crittoanalisi delle funzioni di hash
- Definizione di metodi innovativi per l'identificazione di collisioni
- Sviluppo applicazione per la ricerca di collisioni in due fasi
  1. Strumento per la costruzione di caratteristiche differenziali
  2. Strumento per la ricerca di collisioni a partire da una caratteristica

Risultati già ottenuti (2/2)

- Estensione parallela dell'applicazione
  - In particolare per l'adattamento ai processori multicore di nuova generazione (CELL B.E.)
- Test di utilizzo grazie all'opportunità di accesso al cluster HPC *MariCel* di Barcellona
- Redazione e sottomissione di un primo articolo per l'illustrazione dei risultati
  - A. Cilardo, L. Esposito, A. Veniero, A. Mazzeo, V. Beltran, E. Ayugadé, *A CellBE-based HPC application for the analysis of vulnerabilities in cryptographic hash functions*, submitted to ***HPCC 10***, 2010

Collisione per SHA-1 ridotta a 71 round

- Come menzionato, il miglior risultato ottenuto finora corrispondeva a 70 round
- Complessità computazionale ridotta notevolmente

Messaggio 1				Messaggio 2			
Word	Valore (Hex)	Word	Valore (Hex)	Word	Valore (Hex)	Word	Valore (Hex)
0	A031284A	16	C66928E5	0	10312819	16	766928B6
1	8E0B07E7	17	B8D273A2	1	BE0B07EF	17	88D273AA
2	259E60AA	18	136947A4	2	259E60E9	18	136947E7
3	26865A7F	19	4C7277A5	3	F6865A0D	19	9C7277D7
4	3F7A9945	20	87640DAB	4	8F7A9955	20	37640DBB
5	1F3B9AF1	21	1E439843	5	EF3B9A93	21	EE439821
6	B79EC755	22	8F78C12A	6	779EC717	22	4F78C168
7	41CB1152	23	7EA5FA75	7	41CB1162	23	7EA5FA45
8	311807D8	24	582F0B12	8	D118079A	24	B82F0B50
9	EA18241F	25	80286705	9	CA18247F	25	A0286765
10	C126D406	26	BA240B89	10	2126D447	26	5A240BC8
11	AAF12C3D	27	8A955AE7	11	8AF12C6D	27	AA955AB7
12	7F82700C	28	AB5CB1CA	12	BF82704D	28	6B5CB18B
13	464034BF	29	2BC7E7B5	13	A64034CD	29	CBC7E7C7
14	EF55783A	30	4132CCAA	14	4F557839	30	E132CCA9
15	B805685B	31	A57DD240	15	78056849	31	657DD252

Hash	89EE5B219C39AAB795FEED4483361F39D9B52E69

Prossime attività

- Ulteriore approfondimento teorico delle vulnerabilità individuate
  - In particolare, sviluppo e perfezionamento di tecniche specifiche per la costruzione di vincoli più potenti (es., che coinvolgono gruppi di bit)
- Possibilità di estensione dell'approccio ad altre funzioni di hash (es., concorrenti SHA-3)
- Approfondimento rapporti con interlocutori nazionali e comunitari
  - Valutazione della possibilità di un'attività congiunta