

On the Full Cost of Collision Search for SHA-1

Christophe De Cannière and Florian Mendel
and Christian Rechberger

ECRYPT Hash Workshop, May 25, 2007

*Institute for Applied Information Processing
and Communications (IAIK) - Krypto Group*

*Faculty of Computer Science
Graz University of Technology*



Agenda

- Anatomy of a collision search attack on SHA-1
- How to speak about the cost/feasibility of collision search in a hash function?
- Some details of a **real** collision search for 70-step SHA-1
- Conclusions / future work

Anatomy of a collision search attack on SHA-1

- Message difference
- High probability characteristic (trail / path) for part of CF
- Generalized characteristic for full CF
- Speed-up methods
- Actual search (search space large enough?)

Anatomy of a collision search attack on SHA-1

- **Message difference**
- **High probability characteristic (trail / path) for part of CF**
- Generalized characteristic for full CF
- Speed-up methods
- Actual search (search space large enough?)

XOR-Approx → Search for candidate characteristic

Several proposals: [RO05], [WYY05], [PRR05]

Precise evaluation of candidates: [DR06]

(no counting of Hamming weight, no ad-hoc rules)

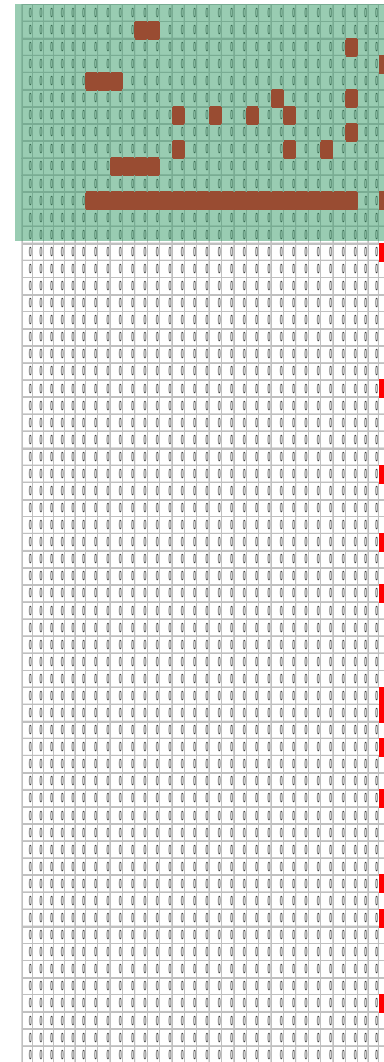
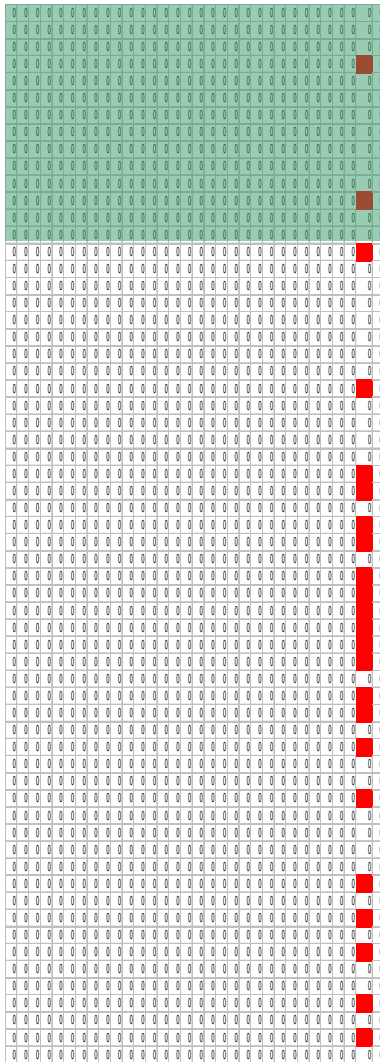
Anatomy of a collision search attack on SHA-1

- Message difference
- High probability characteristic (trail / path) for part of CF
- **Generalized characteristic for full CF**
- Speed-up methods
- Actual search (search space large enough?)

Differential attacks on hash functions

- Good characteristics for block ciphers:
 - Optimise probability
 - Minimise number of chosen plaintexts
- Good characteristics for hash functions
 - Optimise probability
 - Minimise effort to solve equations
 - Equations in first steps are always easy
 - Only a small part of the message involved
 - Inputs are known
 - Late start / Early stop

Good characteristics



Anatomy of a collision search attack on SHA-1

- Message difference
- High probability characteristic (trail / path) for part of CF
- **Generalized characteristic for full CF**
- Speed-up methods
- Actual search (search space large enough?)

Introduction of “Signed-bit differences”,

First solution by [WYY05], manual

Introduction of “Generalized characteristics”,

Automated Method: [DR06]

Generalized conditions

x_i	x_i^*
0	0
0	1
1	0
1	1

Type	Possibilities
XOR	2
Signed-bit	4-6
Generalized:	16

Anatomy of a collision search attack on SHA-1

- Message difference
- High probability characteristic (trail / path) for part of CF
- Generalized characteristic for full CF
- **Speed-up methods**
- **Actual search (search space large enough?)**

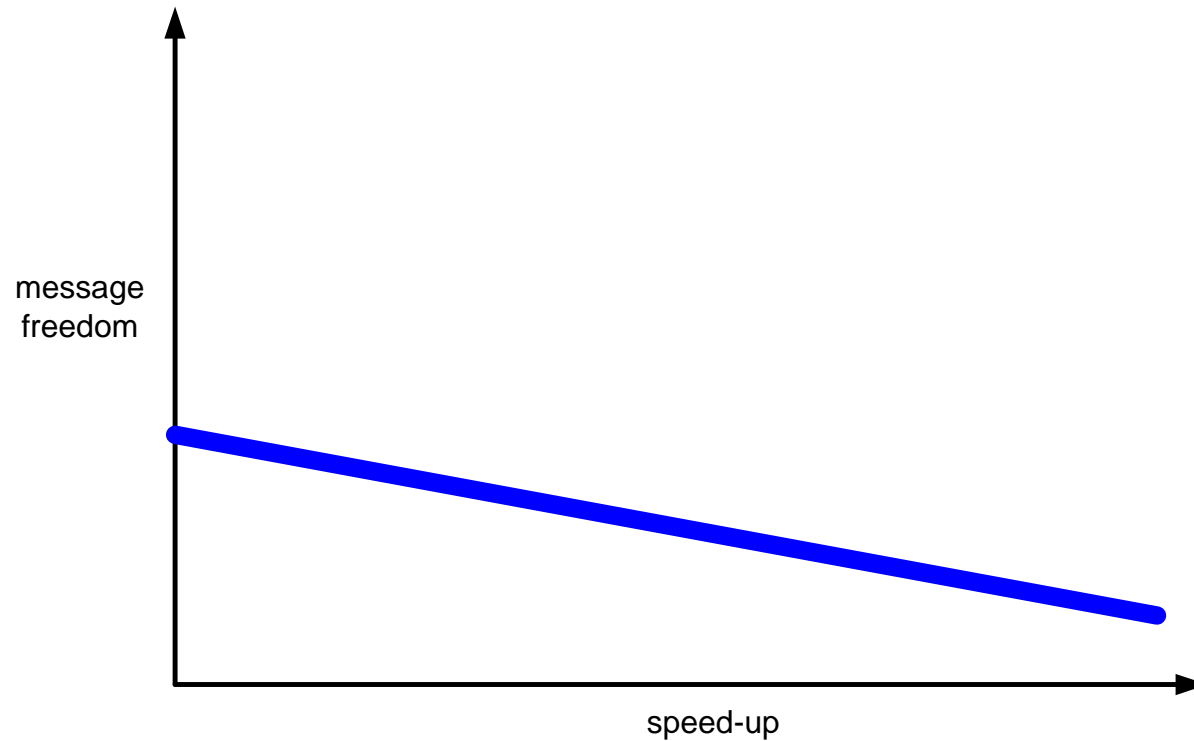
- Neutral-bit technique [Biham and Chen]
- Advanced message modification [Wang et al.]
- Greedy method [De Cannière and Rechberger]
- Symbolic computation [Sugita et al.]
- Boomerang method [Joux and Peyrin]
- anyone?

Cost of speed-up

1) Computational cost

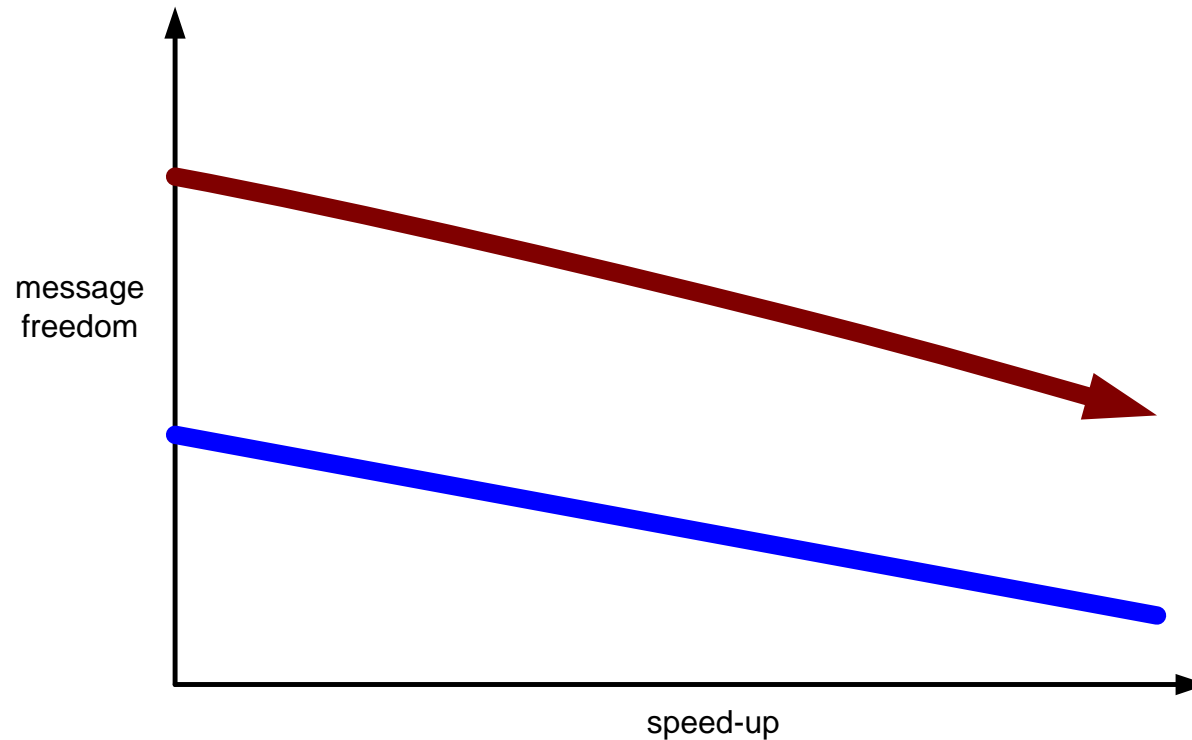
Cost of speed-up

2) Loss of degrees of freedom



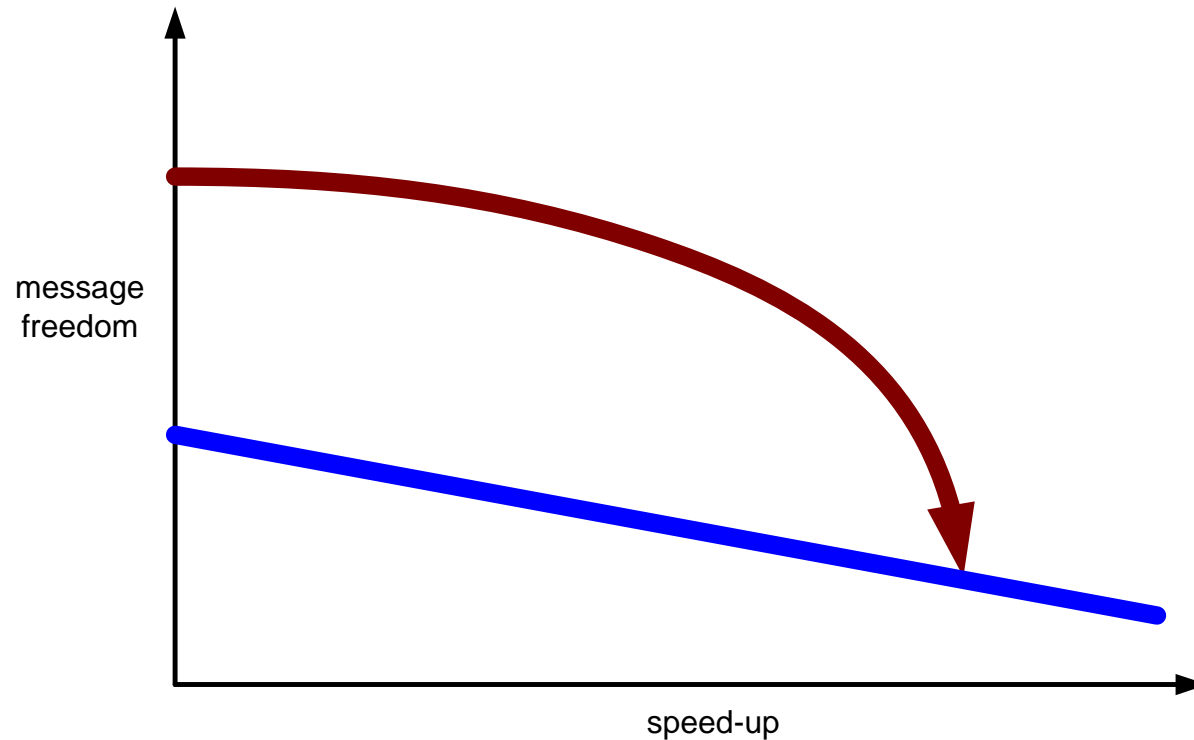
Cost of speed-up

2) Loss of degrees of freedom case MD4/MD5/SHA



Cost of speed-up

2) Loss of degrees of freedom case SHA-1



Too optimistic estimates → New Model

- Model of search algorithm
- Counting #step computations
 - Perfect early stop strategy
- Incorporate all speed-up methods into model

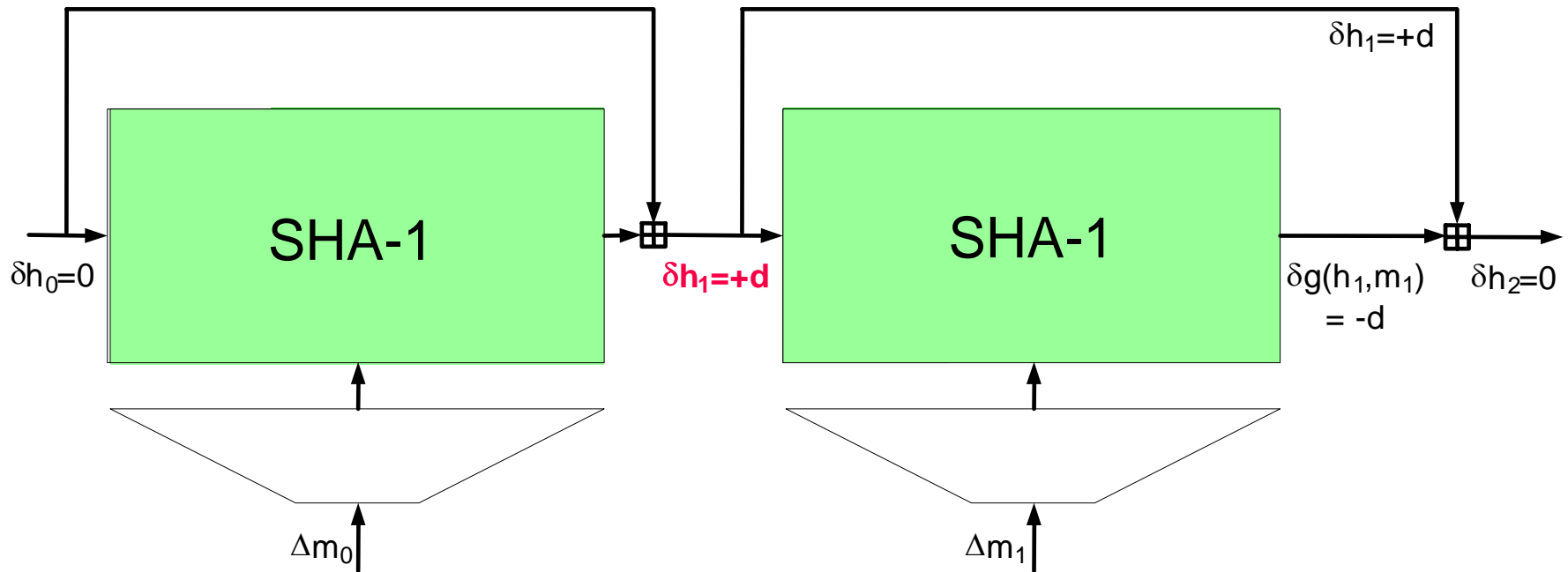
Result:

Even under ideal conditions

(requires many degrees of freedom!)

speed-ups are not as large as expected

Use of near-collisions



- Two related near-collisions give a 2-block collision
- Work effort of two blocks is about the double of one block

Actual collisions for step-reduced SHA-1

- 2005: 40 steps with ? [BCJ+05]
- 2005: 58 steps with 2^{33} [WYY05]
- 2006: 64 steps with 2^{35} [DR06]
- 2006: 64 steps with 2^{35} partially meaningful [DR06b]
- 2007: 70 steps with 2^{44} [DMR07] ← NEW

The 70-step Collision

i	Message 1 (m_0), first block				Message 1 (m_1), second block			
1–4	3BB33AAE	85AECBBB	57A88417	8137CB9C	ABDDBEE2	42A20AC7	A915E04D	5063B027
5–8	4DE99220	5B6F12C7	726BD948	E3F6E9B8	4DDF989A	E0020CF7	7FFDC0F4	EFEFE0A7
9–12	23607799	239B2F1D	AAC76B94	E8009A1E	0FFBC2F0	C8DE16BF	81BBE675	254429CB
13–16	C24DE871	5B7C30D8	000359F5	90F9ED31	5F37A2C6	CD1963D3	FFCA1CB9	9642CB56
i	Message 2 (m_0^*), first block				Message 2 (m_1^*), second block			
1–4	ABB33ADE	35AECBE8	67A8841F	8137CBDF	3BDDBE92	F2A20A94	9915E045	5063B064
5–8	9DE99252	EB6F12D7	826BD92A	23F6E9FA	9DDF98E8	50020CE7	8FFDC096	2FEFE0E5
9–12	236077A9	C39B2F5F	8AC76BF4	08009A5F	0FFBC2C0	28DE16FD	A1BBE615	C544298A
13–16	E24DE821	9B7C3099	E0035987	30F9ED32	7F37A296	0D196392	1FCA1CCB	3642CB55
i	XOR-difference are the same for both blocks							
1–4	90000070	B0000053	30000008	00000043	90000070	B0000053	30000008	00000043
5–8	D0000072	B0000010	F0000062	C0000042	D0000072	B0000010	F0000062	C0000042
9–12	00000030	E0000042	20000060	E0000041	00000030	E0000042	20000060	E0000041
13–16	20000050	C0000041	E0000072	A0000003	20000050	C0000041	E0000072	A0000003
i	The colliding hash values							
1–5	151866D5	F7940D84	28E73685	C4D97E18	97DA712B			

The 70-step Collision

i	Message 1 (m_0), first block				Message 1 (m_1), second block			
1–4	3BB33AAE	85AECBBB	57A88417	8137CB9C	ABDDBEE2	42A20AC7	A915E04D	5063B027
5–8	4DE99220	5B6F12C7	726BD948	E3F6E9B8	4DDF989A	E0020CF7	7FFDC0F4	EFEFE0A7
9–12	23607799	239B2F1D	AAC76B94	E8009A1E	0FFBC2F0	C8DE16BF	81BBE675	254429CB
13–16	C24DE871	5B7C30D8	000359F5	90F9ED31	5F37A2C6	CD1963D3	FFCA1CB9	9642CB56
i	Message 2 (m_0^*), first block				Message 2 (m_1^*), second block			
1–4	ABB33ADE	35AECBE8	67A8841F	8137CBDF	3BDDBE92	F2A20A94	9915E045	5063B064
5–8	9DE99252	EB6F12D7	826BD92A	23F6E9FA	9DDF98E8	50020CE7	8FFDC096	2FEFE0E5
9–12	236077A9	C39B2F5F	8AC76BF4	08009A5F	0FFBC2C0	28DE16FD	A1BBE615	C544298A
13–16	E24DE821	9B7C3099	E0035987	30F9ED32	7F37A296	0D196392	1FCA1CCB	3642CB55
i	XOR-difference are the same for both blocks							
1–4	90000070	B0000053	30000008	00000043	90000070	B0000053	30000008	00000043
5–8	D0000072	B0000010	F0000062	C0000042	D0000072	B0000010	F0000062	C0000042
9–12	00000030	E0000042	20000060	E0000041	00000030	E0000042	20000060	E0000041
13–16	20000050	C0000041	E0000072	A0000003	20000050	C0000041	E0000072	A0000003
i	The colliding hash values							
1–5	151866D5	F7940D84	28E73685	C4D97E18	97DA712B			

Cost of the first 70-step collision

Equivalent to

2^{41} (first block)

2^{44} (second block)

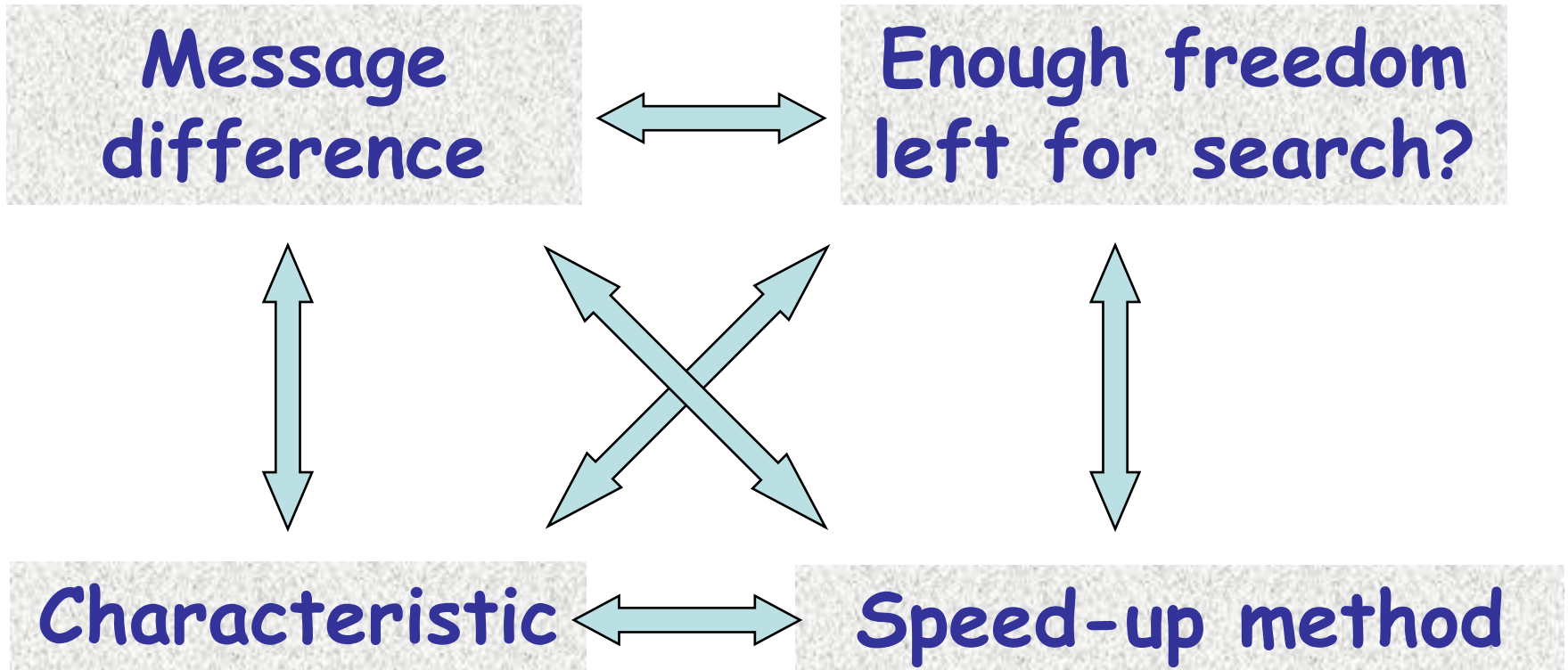
compression function calls to an

efficient implementation of 70-step SHA-1

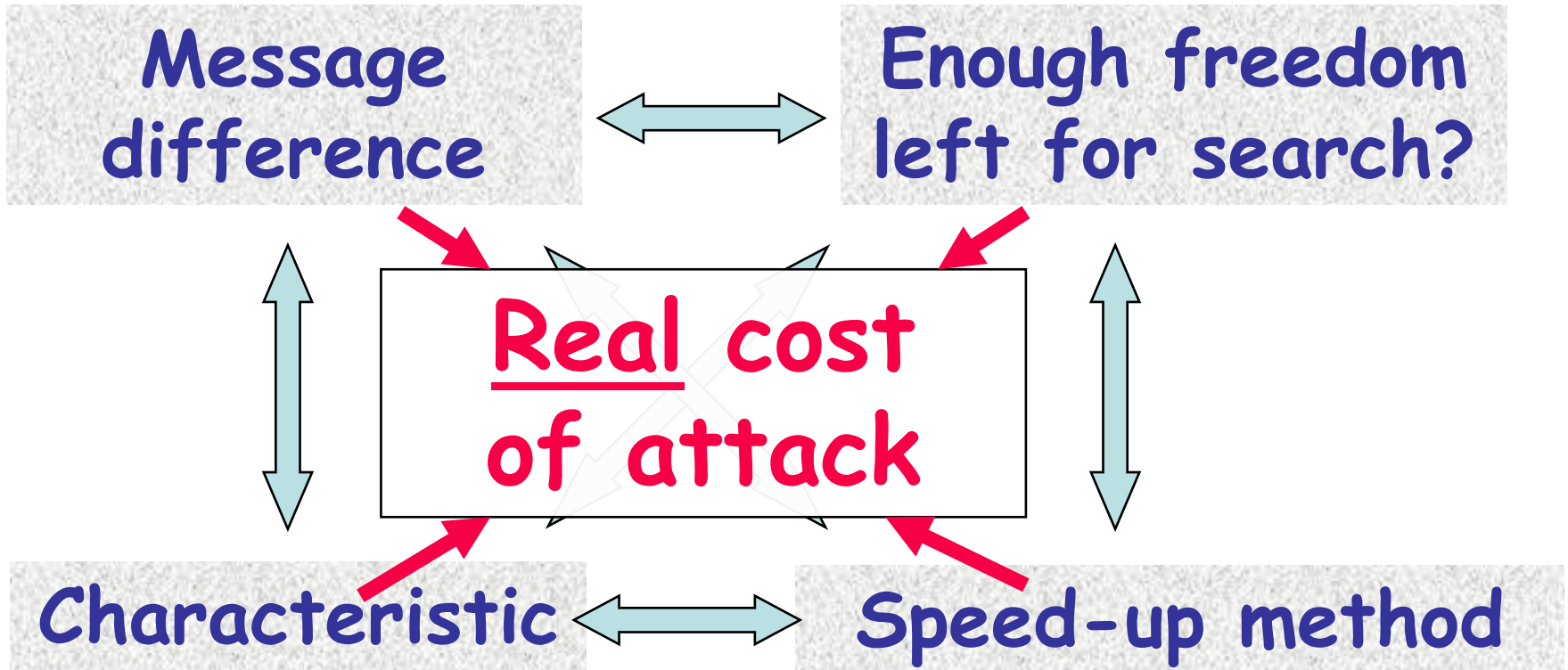
(includes a 2^3 slowdown of a less-than-optimal collision search program)

Comparison with estimates	[WYY05]	Ours
80-step	2^{69}	?
70-step	2^{50}	2^{44} [this paper]
64-step	2^{36}	2^{35} [DR06]

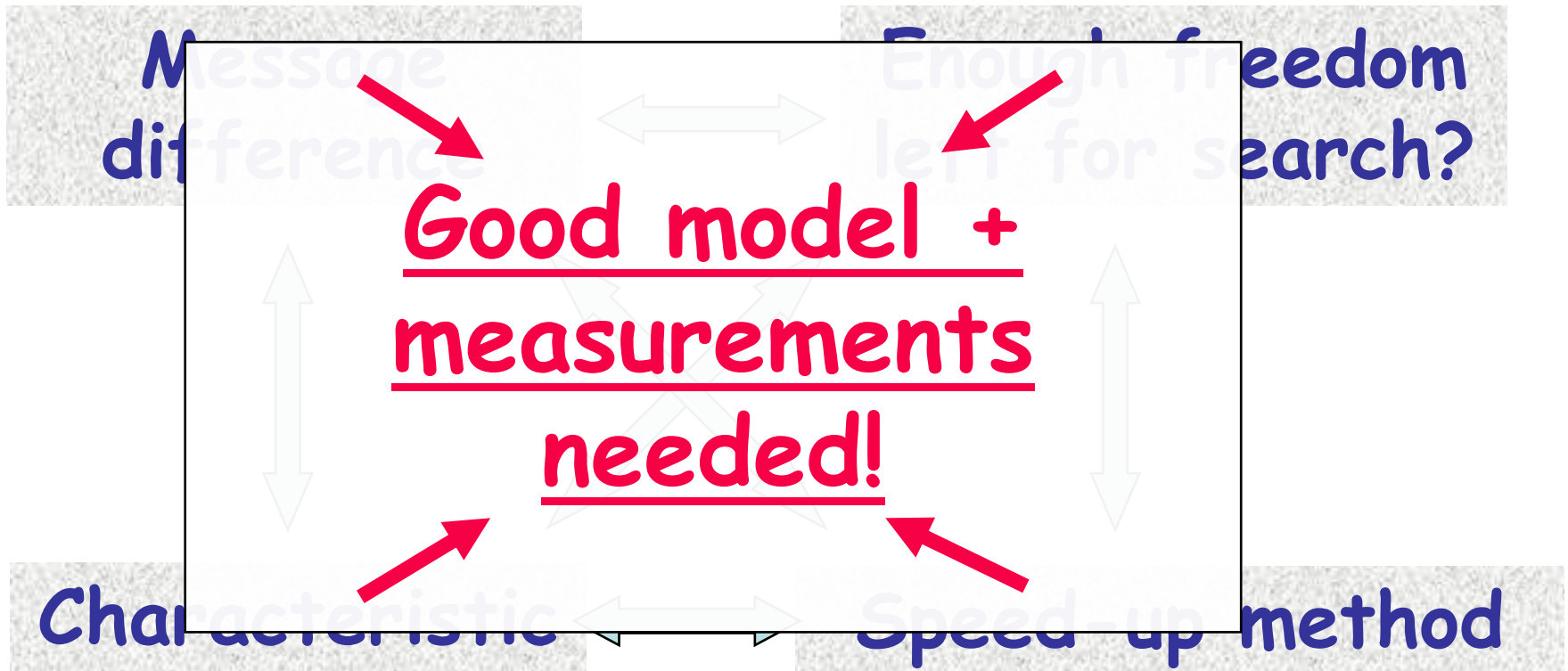
Problem of optimization



Problem of optimization



Problem of optimization



Conclusions

- Little memory, trivially parallelizable, negligible communication
 - Full cost reduces to runtime on single machine
- It's NOT about probability, but effort to solve equations
 - Some are hard: → strong relation to concept of probability
 - Some are easier → ...
- Generalized characteristics → Corrective factors → Real Cost
- Demonstration: first collision for 70-step SHA-1

Future Work

- Collision for full (80-step) SHA-1 is getting closer
- Actual cost of Wang's latest announcement ($2^{62} * x$) unknown, but seems optimized, details unpublished
- No single bottleneck for search → multidimensional optimization (degrees of freedom!)

Future Work

- Apply new insights to other hash functions like RIPEMD-160, SHA-2, new proposals?
- Improved Attacks on NMAC/HMAC?
See [Rechberger and Rijmen, FC2007]
- Results on (2nd-)preimage resistance?
- Will there ever be that highly developed dedicated cryptanalytic tools for other hash functions?

On the Full Cost of Collision Search for SHA-1

Q&A

Christophe De Cannière and Florian Mendel
and Christian Rechberger

***Institute for Applied Information Processing
and Communications (IAIK) - Krypto Group***

***Faculty of Computer Science
Graz University of Technology***



Used References

- [BCJ+05] Biham, Chen, Joux et al.: “Collisions of SHA-0 and Reduced SHA-1”, EUROCRYPT 2005
- [DR06] De Cannière, Rechberger: “Finding SHA-1 Characteristics: General Results and Applications”, ASIACRYPT 2006
- [DR06b] De Cannière, Rechberger: “Meaningful Collisions for SHA-1 at no extra cost?”, CRYPTO 2006 Rump Session
- {PRR05] Pramstaller, Rechberger, Rijmen: „Exploiting Coding Theory for Collision Search Attacks on SHA-1“, IMA C&C, 2005
- [RR07] Rechberger, Rijmen: “On Authentication with HMAC and Non-Random Properties”, Financial Cryptography 2007
- [RO05] Rijmen, Oswald: “Update on SHA-1”, CT-RSA 2005
- [WYY05] Wang, Yin, Yu: “Finding Collisions in the Full SHA-1”, CRYPTO 2005