



Analysis and Design of Cryptographic Hash Functions

Pieter Retief Kasselman

Voorgelê ter vervulling van 'n deel van die vereistes vir die graad Magister in Ingenieurswese
in die Fakulteit Ingenieurswese, Universiteit van Pretoria.

18 November 1999



Summary

Analysis and Design of Cryptographic Hash Functions

by

P. R. Kasselmann and W. T. Penzhorn

Department of Electrical, Electronic and Computer Engineering

University of Pretoria

MEng Electronic Engineering

Indexing Terms: Hash Functions, Cryptanalysis, Cryptography, Message Integrity Code, Message Authentication Code, Differential Cryptanalysis, Boolean Functions, MD4, MD5, HAVAL.

Cryptographic hash functions are one of the primitive building blocks commonly used in information security. They form an important building block for authentication protocols, encryption algorithms, digital signatures and integrity checking algorithms. Two important properties of hash functions used in cryptographic applications are collision resistance and one-wayness. In this dissertation the focus is on collision resistance.

The dissertation provides a detailed overview of existing cryptographic hash functions, including definitions of fundamental properties, generic threats, and popular designs for cryptographic hash functions. Special attention is given to dedicated cryptographic hash functions related to the MD4 hash function.

Between 1990 and 1994 a number of practical cryptographic hash functions were designed and implemented, following the design principles of MD4. These cryptographic hash functions include MD4, MD5, SHA, SHA-1, HAVAL, RIPEMD-128 and RIPEMD-160. These functions were designed to exhibit the properties of collision resistance and one-wayness.

In this dissertation the attacks by Dobbertin on MD4 and MD5 are reconstructed. A novel approach is introduced that allows the execution of the attack on MD4 to be optimised. This new approach allows a reduction in computation time for a collision by a factor 64.

Based on these attacks a generalised attack is formulated. The generalised attack provides a new framework for the analysis of the collision resistant property of any cryptographic hash function.



This newly derived framework for the analysis of cryptographic hash functions is then applied to reduced versions of SHA and HAVAL. The results obtained in this investigation are the first cryptanalytical result to be published on the HAVAL hash function. The investigation shows that a collision can be found for a reduced version of HAVAL in less than a minute on a 200 MHz Pentium Pro personal computer. This result suggests that three and even four round HAVAL should not be used for security applications where message integrity and non-repudiation is required.

Based on the findings of these cryptanalytic attacks, a new set of design criteria for dedicated cryptographic hash functions is formulated. The design criteria aim to alleviate the common weaknesses identified in dedicated hash functions such as MD4, MD5, SHA, SHA-1 and HAVAL. Thereby the generalised attack developed in this dissertation can be thwarted.

Samevatting

Analise en Ontwerp van Kriptografiese Hutsfunksies

deur

P. R. Kasselmann and W. T. Penzhorn

Department Elektriese, Elektroniese en Rekenaar Ingenieurswese

Universiteit van Pretoria

M Ing Electroniese Ingenieurswese

Indekseringsterme: Hutsfunksies, Kripto-analise, Kriptografie, Boodskap Integriteit Kode, Boodskap Stawing Kode, Differensiële Kriptoanalise, Boolese Funksies, MD4, MD5, HAVAL.

Kriptografiese hutsfunksies is een van die primitiewe boublokke wat algemeen gebruik word in informasiesekerheid. Dit vorm 'n belangrike boublok vir stawingsprotokolle, enkripsiealgoritmes, digitale handtekeninge en integriteitmeganismes. Twee belangrike eienskappe van hutsfunksies is weerstand teen botsings en die eenrigtingeienskap. In hierdie verhandeling val die fokus op die botsingweerstandseienskap.

Die verhandeling bevat 'n volledige oorsig van bestaande kriptografiese hutsfunksies, insluitend definisies van fundamentele eienskappe, generies bedreigings en populêre ontwerpe vir kriptografiese hutsfunksies. Spesiale aandag word gegee aan toegewyde kriptografiese hutsfunksies wat verwant is aan die MD4 hutsfunksie.

Tussen 1990 en 1994 is 'n aantal kriptografies hutsfunksies ontwerp en geïmplimenteer. Hierdie ontwerpe is gegrond op die ontwerpbeginsele van MD4. Die kriptografiese hutsfunksies sluit in MD4, MD5, SHA, SHA-1, HAVAL, RIPEMD-128 en RIPEMD-160. Hierdie funksies is almal ontwerp om die eenrigting en botsingsweerstand eienskappe te vertoon.

In hierdie verhandeling word die aanvalle van Dobbertin op MD4 en MD5 gerekonstrueer. 'n Unieke benadering word voorgestel wat die aanval op MD4 optimeer. Die nuwe benadering verminder die berekeningkompleksiteit om 'n botsing te verkry met 'n faktor 64.

'n Veralgemeende aanval word geformuleer op grond van hierdie aanvalle. Die veralgemeende aanval voorsien 'n nuwe raamwerk vir die analise van die botsingsweerstand eienskap van enige toegewyde kriptografiese hutsfunksie.

Hierdie nuwe raamwerk vir die analiese van kriptografiese hutsfunksies word dan toegepas

op afgeskaalde weergawes van SHA en HAVAL. Die resultate van hierdie studie is die eerste kriptanalitiese resultate wat vir HAVAL gepubliseer is. Die studie toon dat 'n botsing vir die laaste twee rondtes van drieronde HAVAL verkry kan word in minder as 'n minuut op 'n 200 MHz Pentium Pro persoonlike rekenaar. Hierdie resultaat dui aan dat drie en selfs vier rondte HAVAL nie gebruik moet word vir sekuriteitstoepassings waar boodskapintegriteit vereis word nie.

Op grond van die kriptanalitiese resultate word 'n nuwe stel ontwerpseriteria vir toegewyde kriptografiese hutsfunksies geformuleer. Die ontwerpseriteria is daarop gemik om die gedeelde swakhede geïdentifiseer in toegewyde hutsfunksies soos MD4, MD5, SHA, SHA-1 en HAVAL te vermy. Hierdeur kan die veralgemeende aanval wat in die verhandeling ontwikkel is gefnuik word.



Acknowledgements

I would like to make use of this opportunity to thank the following individuals and organisations.

My study leader, Prof. W.T. Penzhorn for his contributions, suggestions and encouragement throughout this project.

Prof. G.J. Kühn for the many stimulating conversations on the topic of cryptology.

Dr Bart Preneel and Antoon Bosselaers from the COSIC research group at the Katholieke Universiteit Leuven, Belgium for their comments on my work and suggesting the analysis of the HAVAL hash function.

My wife, Marelize, for her continued support and companionship throughout this project.

Ciphertec cc for the opportunity to perform research on the topic of cryptographic hash functions.

The management of Nedcor Bank Ltd for the time afforded to me in order to complete this dissertation.



CONTENTS

1	Introduction	1
1.1	Information Security	1
1.2	Hash Functions and Security	2
1.2.1	Applications of Hash Functions	2
1.2.2	Properties of Hash Functions	4
1.2.3	Hash Functions Today	8
1.3	Problem Statement	9
1.4	Hypothesis	9
1.5	Scope	9
1.6	Dissertation Objectives and Methodology	10
1.7	Results	11
2	Taxonomy of Cryptographic Hash Functions	12
2.1	Introduction	12



2.1.1	MAC	13
2.1.2	MDC	13
2.2	Approaches to the Design and Analysis of Cryptographic Hash Functions	15
3	Threats Against Hash Functions	17
3.1	Introduction	17
3.2	Taxonomy of Attackers	17
3.2.1	Capabilities	17
3.2.2	Position	18
3.3	Terminology	19
3.3.1	MDC Terminology	20
3.3.2	MAC Terminology	20
3.4	Attacks on MDCs	21
3.4.1	Attacks Independent of the Algorithm	21
3.4.2	Attacks Dependant on the Chaining	23
3.5	Attacks on MACs	29
3.5.1	Key Collisions	29
3.5.2	Exhaustive Key Search	29
3.5.3	Chosen Text Attacks	30
3.5.4	Known and Chosen Text Attack	30



3.6	Attacks on Underlying Block Ciphers	32
3.6.1	Complementation Property	33
3.6.2	Weak Keys	33
3.7	High Level Attacks	34
3.7.1	Differential Fault Analysis	34
3.7.2	Differential Power Analysis	34
3.7.3	Attack on the Interaction with the Signature Scheme	35
3.7.4	Attacks on the Protocol	35
3.7.5	Attacks Dependant on the Algorithm	35
3.8	Attackers and Attacks	36
3.9	Feasibility	37
3.10	Conclusion	37
4	Requirements for Cryptographic Hash Functions	39
4.1	Introduction	39
4.2	Functional Requirements	39
4.2.1	Message Reduction	39
4.2.2	Repeatability	40
4.2.3	Data Type Independence	40
4.2.4	Fast Calculation	41



4.2.5	One Pass per Message	41
4.2.6	Minimum of Secret Information	41
4.2.7	Modular Design	41
4.2.8	Ease of Implementation	42
4.2.9	Machine Independence	42
4.2.10	Distribution and Obtainability	42
4.3	Security Requirements	42
4.3.1	Confusion and Diffusion	43
4.3.2	Message and Hash Value Independence	43
4.3.3	Computational Feasibility	44
4.3.4	Interaction with other Algorithms	44
4.3.5	MDC Hash Space	45
4.3.6	MAC Key and Hash Space	45
4.3.7	Message Dependence	46
4.3.8	One-Wayness	46
4.3.9	Error Extension	46
4.3.10	Distribution of Preimages	46
4.3.11	Decomposable Algorithms	47
4.3.12	Conditions on Chaining	47



4.3.13	Redundancy	47
4.4	Functional vs. Security Requirements	48
4.4.1	Repeatability and Security	49
4.4.2	Chaining and Security	49
4.4.3	Speed and Security	49
4.4.4	Speed and Machine Independence	50
4.4.5	Decomposability and Ease of Implementation	50
4.4.6	Security and Bandwidth	51
4.5	Conclusion	51
5	General Dedicated Hash Function Constructions	52
5.1	Introduction	52
5.2	Ideal Cryptographic Hash Function	52
5.3	Iterated Hash Functions	54
5.3.1	The Segmentation and Padding Rule	54
5.3.2	The Compress Function	56
5.3.3	The Chaining Rule	59
5.3.4	Construction	59
5.4	Round Function Constructions	61
5.4.1	MD4-Family Construction	61

5.5	Round Function Building Blocks	62
5.5.1	Bit Permutations	63
5.5.2	Bitwise Boolean Operations	64
5.5.3	Substitution Boxes	64
5.5.4	Modular Arithmetic Operations	65
5.6	Conclusion	66
6	Analysis of the MD4 Hash Algorithm	67
6.1	Introduction	67
6.2	Introduction to MD4	67
6.3	Notation	68
6.4	The MD4 Algorithm	68
6.4.1	Message Padding	69
6.4.2	Initial Values	69
6.4.3	Iterative Rounds	70
6.5	Cryptanalysis of MD4	76
6.6	Notation	76
6.7	Dobbertin's attack: A review	77
6.8	Alternative algorithm for establishing inner almost-collisions	80
6.9	Results	82



6.9.1	Number of Collisions	82
6.9.2	Speedup Factor	83
6.9.3	Example	83
6.10	Conclusion	84
7	Analysis of the MD5 Hash Algorithm	86
7.1	Introduction	86
7.2	Introduction to MD5	86
7.3	Notation	87
7.4	The MD5 Algorithm	87
7.4.1	Message Padding	88
7.4.2	Initial Values	88
7.4.3	Iterative Rounds	89
7.5	Analysis of MD5	95
7.5.1	Notation	95
7.5.2	Outline of the Attack	96
7.5.3	Phase I: Inner Collisions for First Two Rounds	97
7.5.4	Phase II: Inner Collisions for Last Two Rounds	103
7.5.5	Phase III: Establishing a Connection	113
7.5.6	Determining if Solutions Exist	118

7.5.7	Conclusion	124
7.6	Acknowledgments	125
8	Generalised Analysis of the MD4 Family of Dedicated Hash Functions	126
8.1	Introduction	126
8.2	Generalised Attacks	127
8.2.1	Difference Equations	128
8.2.2	Solution of Difference Equations	130
8.3	Application of Generalised Attacks	131
8.4	Conclusion	131
9	Analysis of the SHA and SHA-1 Hash Algorithms	132
9.1	Introduction	132
9.2	Introduction to SHA	132
9.3	Notation	132
9.4	SHA	133
9.4.1	Message Padding	133
9.4.2	Initialise Chaining Variables	134
9.4.3	Message Expansion	134
9.4.4	Compress Function	134
9.4.5	Update Chaining Variables	136



9.5	SHA-1	136
9.5.1	Message Expansion	136
9.6	Analysis of SHA and SHA-1	136
9.7	SHA	137
9.7.1	Message Expansion Algorithm	137
9.7.2	Difference Equations	142
9.7.3	Extended Attack	147
9.7.4	Proposed Attack	148
9.8	SHA-1	148
9.8.1	Message Expansion Algorithm	149
9.9	Conclusion	150
10	Analysis of the HAVAL Hash Algorithm	151
10.1	Introduction	151
10.2	Introduction to HAVAL	151
10.3	Notation	151
10.4	HAVAL	152
10.4.1	Message Padding	153
10.4.2	Initialise Chaining Variables	154
10.4.3	Word Processing Order	154



10.4.4	Compress Function	155
10.4.5	Tailoring the output	159
10.5	Analysis of HAVAL	159
10.5.1	Difference Equations	160
10.5.2	Solution to Differential Equations	165
10.5.3	Collision Example	175
10.6	Conclusion	176
11	Design Criteria for Dedicated Hash Functions	178
11.1	Introduction	178
11.2	Basic Structure	178
11.3	Building Blocks	179
11.3.1	Boolean Mappings	179
11.3.2	Rotation	182
11.3.3	Message Word Reuse	182
11.3.4	Addition mod 2^{32}	185
11.3.5	Additive Constants	185
11.3.6	Composition	186
11.4	Conclusion	186
12	Conclusion	188



12.1 Discussion	188
12.2 Results	189
12.3 Summary and Future Work	189
Bibliography	197
A Additional Hash Function Constructions	198
A.1 Introduction	198
A.2 Tree Constructions	198
A.2.1 Construction	199
A.2.2 Practicality	200
A.3 Cascading of Hash Functions	200
A.4 Round Function Constructions	201
A.4.1 Block Ciphers	201
A.4.2 Stream Ciphers	209
A.5 MAC Constructions Based on MDCs	209
A.5.1 Affix Construction	210
A.5.2 IPsec recommendations	213
A.5.3 NMAC Construction	215
A.5.4 HMAC Construction	215
A.5.5 MDx -MAC Construction	217



CHAPTER 1. INTRODUCTION	
A.5.6 XOR-MAC Constructions	219
A.6 International Standards	220
A.7 Conclusion	221
B Source Code: Implementation of MD4	223
C Source Code: Attack on all three rounds of MD4	232
D Implementation: MD5	244
E Source Code: Analysis of MD5	253
E.1 Source Code: First Phase of the Attack on MD5	253
E.2 Source Code: Second Phase of the Attack on MD5	266
E.3 Third Phase of the Attack on MD5	277
F Source Code: Collisions for First Round of SHA	289
G Source Code: Implementation of HAVAL Attack	296