Small Size Hashes with Enhanced Security

Lars R. Knudsen

Department of Mathematics, Tech. University of Denmark
Building 303, DK-2800 Lyngby, Denmark (Email: Lars.R.Knudsen@mat.dtu.dk)

(Received July 1, 2005; revised and accepted Aug. 9 and 15, 2005)

Abstract

This paper contains techniques for enhancing the strength of any cryptographic hash function. For an "ideal", traditional hash function with an m-bit result, the complexity of a collision attack is approximately $2^{m/2}$. Here constructions are presented where collisions are harder to find.

Keywords: Cryptology, hash functions

1 Introduction

A *cryptographic hash function* takes as input a binary string of arbitrary length and returns a binary string of a fixed length. Hash functions which satisfy some security properties are widely used in cryptographic applications such as digital signatures, password protection schemes, and conventional message authentication [1]. In the following let

$$H : \{0,1\}^* \to \{0,1\}^n$$

denote a hash function which returns a string of length n.

Definition 1.1 *Given a hash function H, a* collision attack *finds $x' \neq x$, such that $H(x) = H(x')$.*

Definition 1.2 *Given a hash function H, an x and a $y = H(x)$, a* 2nd preimage attack *finds $x' \neq x$, such that $H(x') = y$. Given a $y = H(x)$, a* preimage attack *finds x', such that $H(x') = y$.*

Definition 1.3 *Given a hash function H, a* collision attack *finds $x' \neq x$, such that $H(x) = H(x')$.*

Definition 1.4 *Given a hash function H, an x and a $y = H(x)$, a* 2nd preimage attack *finds $x' \neq x$, such that $H(x') = y$. Given a $y = H(x)$, a* preimage attack *finds x', such that $H(x') = y$.*

Clearly the existence of a 2nd preimage attack implies the existence of a collision attack. Also, it can be shown (under suitable assumptions) that a preimage attack implies a collision attack.

For an ideal hash function, the complexities of the preimage attacks are 2^m, that is, the best approach is a brute-force search with per trial has a probability of success of 2^{-m}. With an n-bit hash result it is well known that in a collection of $2^{n/2}$ arbitrary, different messages one has a good probability of finding two messages, which hash to the same value, e.g., a collision. This is called the birthday attack, named after the birthday paradox [1].

2 The Proposal

For most hash functions there is an initial value which needs to be specified. This can either fixed for all users, or be chosen by the communicating parties. Let now $H : \{0,1\}^s \times \{0,1\}^* \to \{0,1\}^n$ denote a hash function. If IV denotes an s-bit initial value of H, and M an arbitrary message, then $H(IV, M)$ denotes the hashed value of the message M using the initial value IV.

Let Y be some fixed m-bit value, where $m < n$ and $m \leq s$. Y is either part of the description of the hash function or is a value mutually agreed upon between the involved parties. Let M be the message to be hashed. Then find the smallest positive value $X = IV' - IV \bmod 2^s$ such that $H(IV', M)$ truncated to the m rightmost bits equals Y, that is, $H(IV', M) = (Z \mid Y)$. Then define $\text{Hash}(IV, M) = Z$. If it is assumed that H behaves like a random function, then the rightmost m bits of a hash result will equal some predetermined value Y with probability 2^{-m}. Thus, if one hashing operation is one unit, then the expected number of iterations one has to perform to compute the hash of a given message is 2^m. Note that it is required that $s \geq m$, in fact, it is recommended that $s > m$. The probability that in t operations one misses the target m-bit value Y is $(1 - 2^{-m})^t$. With $t = 2^m$ one gets $(1 - 2^{-m})^{2^m} \simeq 0.36$. With $s = u + m$ the probability of a miss is $(1 - 2^{-m})^{2^{m+u}} \simeq 0.36^{2^u}$. E.g., with $s \geq m + 8$ the probability of a miss is at most 2^{-377}.

Thus, one gets an $(n - m)$ bit hash value at the cost of around 2^m computations, but the security level against collision attacks is more than $2^{(n-m)/2}$ in the traditional birthday attack. In fact we can regard the hash function construction as a function $J : \{0,1\}^* \to \{0,1\}^{n-m}$ where the expected number of iterations to compute one function value of J corresponds to 2^m iterations of H. Then

Table 1: Comparison of the proposed construction and the traditional approach. One unit is one hashing operation.

	# ops compute hash	# ops to verify hash	hash size	security (collisions)	security preimages
Traditional	1	1	n	$2^{n/2}$	2^n
Proposal	2^m	2^m	$n-m$	$2^{n/2+m/2}$	2^n

a birthday attack requires around $2^{(n-m)/2}$ iterations of J, which amounts to $2^{(n+m)/2}$ iterations of H.

The target value in a preimage attack is of m bits just as in the traditional hashing approach. Assuming that the m-bit value Y above is randomly chosen, then if the best approach for finding preimages for the hash function H is a brute-force attack, then so is the best approach for finding preimages for the proposed construction based on H. Table 1 lists the properties of the proposed construction.

3 Example

A popular hash function is Secure Hash Algorithm (SHA) [2]. It produces a 160-bit hash result using a 160-bit initial value, so using the above notation $n = s = 160$. With $m = 16$ the proposed construction yields a hash function with a 144-bit result for which the expected complexity of finding a collision is 2^{88} compared to 2^{72} for a traditional 144-bit hash function and compared to 2^{80} for SHA. On a typical personal computer today, 2^{16} SHA computations of a $512q$-bit message take no more than $q/4$ seconds using a good implementation of SHA.

References

[1] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1997.

[2] NIST, *Secure Hash Standard*, FIPS 180-1, US Department of Commerce, Washington D.C., April 1995.

Lars R. Knudsen is a professor in the Department of Mathematics at the Technical University of Denmark and an adjunct professor in the Department of Informatics at the University of Bergen in Norway. He is an elected director of the International Association of Cryptologic Research (http://www.iacr.org) and an associate editor of the Journal of Cryptology.