                    Updated Security Considerations for
              the MD5 Message-Digest and the HMAC-MD5 Algorithms

Abstract

   This document updates the security considerations for the MD5 message
   digest algorithm. It also updates the security considerations for
   HMAC-MD5.

Status of This Memo

   This document is not an Internet Standards Track specification; it is
   published for informational purposes.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Not all documents
   approved by the IESG are a candidate for any level of Internet
   Standard; see Section 2 of RFC 5741.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc6151.

1. Introduction

   MD5 [MD5] is a message digest algorithm that takes as input a message
   of arbitrary length and produces as output a 128-bit "fingerprint" or
   "message digest" of the input. The published attacks against MD5
   show that it is not prudent to use MD5 when collision resistance is
   required. This document replaces the security considerations in RFC
   1321 [MD5].

   [HMAC] defined a mechanism for message authentication using
   cryptographic hash functions. Any message digest algorithm can be
   used, but the cryptographic strength of HMAC depends on the
   properties of the underlying hash function. [HMAC-MD5] defined test
   cases for HMAC-MD5. This document updates the security
   considerations in [HMAC], which [HMAC-MD5] points to for its security
   considerations.

   [HASH-Attack] summarizes the use of hashes in many protocols and
   discusses how attacks against a message digest algorithm's one-way
   and collision-free properties affect and do not affect Internet
   protocols. Familiarity with [HASH-Attack] is assumed. One of the
   uses of message digest algorithms in [HASH-Attack] was integrity
   protection. Where the MD5 checksum is used inline with the protocol
   solely to protect against errors, an MD5 checksum is still an
   acceptable use. Applications and protocols need to clearly state in
   their security considerations what security services, if any, are
   expected from the MD5 checksum. In fact, any application and
   protocol that employs MD5 for any purpose needs to clearly state the
   expected security services from their use of MD5.

2. Security Considerations

   MD5 was published in 1992 as an Informational RFC. Since that time,
   MD5 has been extensively studied and new cryptographic attacks have
   been discovered. Message digest algorithms are designed to provide
   collision, pre-image, and second pre-image resistance. In addition,
   message digest algorithms are used with a shared secret value for
   message authentication in HMAC, and in this context, some people may
   find the guidance for key lengths and algorithm strengths in
   [SP800-57] and [SP800-131] useful.

   MD5 is no longer acceptable where collision resistance is required
   such as digital signatures. It is not urgent to stop using MD5 in
   other ways, such as HMAC-MD5; however, since MD5 must not be used for
   digital signatures, new protocol designs should not employ HMAC-MD5.
   Alternatives to HMAC-MD5 include HMAC-SHA256 [HMAC] [HMAC-SHA256] and
   [AES-CMAC] when AES is more readily available than a hash function.

2.1. Collision Resistance

Pseudo-collisions for the compress function of MD5 were first
described in 1993 [denBBO1993]. In 1996, [DOB1995] demonstrated a
collision pair for the MD5 compression function with a chosen initial
value. The first paper that demonstrated two collision pairs for MD5
was published in 2004 [WFLY2004]. The detailed attack techniques for
MD5 were published at EUROCRYPT 2005 [WAYU2005]. Since then, a lot
of research results have been published to improve collision attacks
on MD5. The attacks presented in [KLIM2006] can find MD5 collision in
about one minute on a standard notebook PC (Intel Pentium, 1.6GHz).
[STEV2007] claims that it takes 10 seconds or less on a 2.6Ghz
Pentium4 to find collisions. In [STEV2007], [SLdeW2007],
[SSALMOdeW2009], and [SLdeW2009], the collision attacks on MD5 were
successfully applied to X.509 certificates.

Notice that the collision attack on MD5 can also be applied to
password-based challenge-and-response authentication protocols such
as the APOP (Authenticated Post Office Protocol) option in POP [POP]
used in post office authentication as presented in [LEUR2007].

In fact, more delicate attacks on MD5 to improve the speed of finding
collisions have been published recently. However, the aforementioned
results have provided sufficient reason to eliminate MD5 usage in
applications where collision resistance is required such as digital
signatures.

2.2. Pre-Image and Second Pre-Image Resistance

Even though the best result can find a pre-image attack of MD5 faster
than exhaustive search, as presented in [SAAO2009], the complexity
$2^{123.4}$ is still pretty high.

2.3. HMAC

The cryptanalysis of HMAC-MD5 is usually conducted together with NMAC
(Nested MAC) since they are closely related. NMAC uses two
independent keys K1 and K2 such that NMAC(K1, K2, M) = H(K1, H(K2,
M), where K1 and K2 are used as secret initialization vectors (IVs)
for hash function H(IV, M). If we re-write the HMAC equation using
two secret IVs such that IV2 = H(K Xor ipad) and IV1 = H(K Xor opad),
then HMAC(K, M) = NMAC(IV1, IV2, M). Here it is very important to
notice that IV1 and IV2 are not independently selected.

The first analysis was explored on NMAC-MD5 using related keys in
[COYI2006]. The partial key recovery attack cannot be extended to
HMAC-MD5, since for HMAC, recovering partial secret IVs can hardly
lead to recovering (partial) key K. Another paper presented at

   Crypto 2007 [FLN2007] extended results of [COYI2006] to a full key
   recovery attack on NMAC-MD5. Since it also uses related key attack,
   it does not seem applicable to HMAC-MD5.

   A EUROCRYPT 2009 paper presented a distinguishing attack on HMAC-MD5
   [WYWZZ2009] without using related keys. It can distinguish an
   instantiation of HMAC with MD5 from an instantiation with a random
   function with 2^97 queries with probability 0.87. This is called
   distinguishing-H. Using the distinguishing attack, it can recover
   some bits of the intermediate status of the second block. However,
   as it is pointed out in [WYWZZ2009], it cannot be used to recover the
   (partial) inner key H(K Xor ipad). It is not obvious how the attack
   can be used to form a forgery attack either.

   The attacks on HMAC-MD5 do not seem to indicate a practical
   vulnerability when used as a message authentication code.
   Considering that the distinguishing-H attack is different from a
   distinguishing-R attack, which distinguishes an HMAC from a random
   function, the practical impact on HMAC usage as a pseudorandom
   function (PRF) such as in a key derivation function is not well
   understood.

   Therefore, it may not be urgent to remove HMAC-MD5 from the existing
   protocols. However, since MD5 must not be used for digital
   signatures, for a new protocol design, a ciphersuite with HMAC-MD5
   should not be included. Options include HMAC-SHA256 [HMAC]
   [HMAC-SHA256] and [AES-CMAC] when AES is more readily available than
   a hash function.

3. Acknowledgements

   Obviously, we have to thank all the cryptographers who produced the
   results we refer to in this document. We'd also like to thank Wesley
   Eddy, Sam Hartman, Alfred Hoenes, Martin Rex, Benne de Weger, and
   Lloyd Wood for their comments.

4. Informative References

   [AES-CMAC]   Song, JH., Poovendran, R., Lee, J., and T. Iwata, "The
                 AES-CMAC Algorithm", RFC 4493, June 2006.

   [COYI2006]   S. Contini, Y.L. Yin. Forgery and partial key-recovery
                 attacks on HMAC and NMAC using hash collisions.
                 ASIACRYPT 2006. LNCS 4284, Springer, 2006.

   [denBBO1993] den Boer, B. and A. Bosselaers, "Collisions for the
                 compression function of MD5", Eurocrypt 1993.

   [DOB1995]   Dobbertin, H., "Cryptanalysis of MD5 Compress",
                  Eurocrypt 1996.

   [FLN2007]   Fouque, P.-A., Leurent, G., Nguyen, P.Q.: Full key-
                  recovery attacks on HMAC/NMAC-MD4 and NMAC-MD5. CRYPTO
                  2007. LNCS 4622, Springer, 2007.

   [HASH-Attack] Hoffman, P. and B. Schneier, "Attacks on Cryptographic
                  Hashes in Internet Protocols", RFC 4270, November 2005.

   [HMAC]   Krawczyk, H., Bellare, M., and R. Canetti, "HMAC:
                  Keyed-Hashing for Message Authentication", RFC 2104,
                  February 1997.

   [HMAC-MD5]   Cheng, P. and R. Glenn, "Test Cases for HMAC-MD5 and
                  HMAC-SHA-1", RFC 2202, September 1997.

   [HMAC-SHA256] Nystrom, M., "Identifiers and Test Vectors for HMAC-
                  SHA-224, HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512",
                  RFC 4231, December 2005.

   [KLIM2006]   V. Klima. Tunnels in Hash Functions: MD5 Collisions
                  within a Minute. Cryptology ePrint Archive, Report
                  2006/105 (2006), http://eprint.iacr.org/2006/105.

   [LEUR2007]   G. Leurent, Message freedom in MD4 and MD5 collisions:
                  Application to APOP. Proceedings of FSE 2007. Lecture
                  Notes in Computer Science 4715. Springer, 2007.

   [MD5]   Rivest, R., "The MD5 Message-Digest Algorithm", RFC
                  1321, April 1992.

   [POP]   Myers, J. and M. Rose, "Post Office Protocol - Version
                  3", STD 53, RFC 1939, May 1996.

   [SAAO2009]   Y. Sasaki and K. Aoki. Finding preimages in full MD5
                  faster than exhaustive search. Advances in Cryptology
                  - EUROCRYPT 2009, LNCS 5479 of Lecture Notes in
                  Computer Science, Springer, 2009.

   [SLdeW2007]   Stevens, M., Lenstra, A., de Weger, B., Chosen-prefix
                  Collisions for MD5 and Colliding X.509 Certificates for
                  Different Identities. EuroCrypt 2007.

   [SLdeW2009]   Stevens, M., Lenstra, A., de Weger, B., "Chosen-prefix
                  Collisions for MD5 and Applications", Journal of
                  Cryptology, 2009.

   [SSALMOdeW2009]
               Stevens, M., Sotirov, A., Appelbaum, J., Lenstra, A.,
               Molnar, D., Osvik, D., and B. de Weger. Short chosen-
               prefix collisions for MD5 and the creation of a rogue
               CA certificate, Crypto 2009.

   [SP800-57]   National Institute of Standards and Technology (NIST),
               Special Publication 800-57: Recommendation for Key
               Management - Part 1 (Revised), March 2007.

   [SP800-131] National Institute of Standards and Technology (NIST),
               Special Publication 800-131: DRAFT Recommendation for
               the Transitioning of Cryptographic Algorithms and Key
               Sizes, June 2010.

   [STEV2007]   Stevens, M., "On Collisions for MD5", Master's Thesis,
               Eindhoven University of Technology,
               http://www.win.tue.nl/hashclash/
               On%20Collisions%20for%20MD5%20-%20M.M.J.%20Stevens.pdf.

   [WAYU2005]   X. Wang and H. Yu. How to Break MD5 and other Hash
               Functions. LNCS 3494. Advances in Cryptology -
               EUROCRYPT2005, Springer, 2005.

   [WFLY2004]   X. Wang, D. Feng, X. Lai, H. Yu, Collisions for Hash
               Functions MD4, MD5, HAVAL-128 and RIPEMD, 2004,
               http://eprint.iacr.org/2004/199.pdf

   [WYWZZ2009] X. Wang, H. Yu, W. Wang, H. Zhang, and T. Zhan.
               Cryptanalysis of HMAC/NMAC-MD5 and MD5-MAC. LNCS 5479.
               Advances in Cryptology - EUROCRYPT2009, Springer, 2009.

Authors' Addresses

   Sean Turner
   IECA, Inc.
   3057 Nutley Street, Suite 106
   Fairfax, VA 22031
   USA

   EMail: turners@ieca.com

   Lily Chen
   National Institute of Standards and Technology
   100 Bureau Drive, Mail Stop 8930
   Gaithersburg, MD 20899-8930
   USA

   EMail: lily.chen@nist.gov