From: Praveen Gauravaram [p.gauravaram@mat.dtu.dk]
Sent: Thursday, August 13, 2009 12:12 PM
To: hash-function@nist.gov
Cc: hash-forum@nist.gov
Subject: OFFICIAL COMMENT: Skein (Round 2)
Attachments: Skein-observation.pdf

An observation on Skein hash function is attached.

Best wishes,
Praveen

An Observation in Skein and its Application

Praveen Gauravaram

Technical University of Denmark
Matematiktorvet, Building S303
DK-2800 Kgs. Lyngby
Denmark

Abstract. In this note, an observation is made regarding the Skein hash function [1] which has been selected for the second round of NIST's SHA-3 hash function competition. This observation is applied to re-asses the security of Skein MAC. If n is the state size and m is the minimum of state and output sizes, then the stated security of Skein MAC is up to $\min(2^{n/2}, 2^m)$. We note that whenever the output size is more than n bits, the security of Skein MAC is always upto $2^{n/2}$.

1 Skein as a cascaded hash function

To produce large size hash values, Skein hash function repeats the output transformation as much as needed with the same chaining input state but with different counter as the input to the output transformation [1, Figure 7]. The desired length of the output is encoded as one of the inputs to the Skein hash function[1]. This can be seen as a cascaded hash function construction where each hash function in the cascade differs only in one input to the output transformation.

It is known from the attacks of Joux [2] that the cascade of two hash functions is not really more secure than that of the single best function in the cascade. This result can be generalised to the cascade of more than two hash functions. This attack also holds for Skein hash function when it is used to generate, for example, $2n$-bit output following the description in section 2.4 of [1].

The output transformation uses the same input chaining value to the output transformation with different index inputs in order to generate a variable length output. Hence, instead of doing Joux multicollision, we can just find one collision on the first hash function which will also be a collision for the second one. In this sense, Skein hash function with more than n-bit output can be visualised as a weak cascade hash function. It seems that designers of Skein implicitly observes this attack by noting that "Producing large outputs is often convenient, but of course the security of Skein is limited by the internal state size."

[1] The 32-byte configuration string encodes the desired output length (see page 6 of [1]).

2 Application of the above observation on Skein MAC

A dedicated MAC mode has been proposed for the Skein hash function (see sections 2.6, 4.3 of [1]) in the style of secret prefix MAC $[H(k\|(.))]$ [2] where a secret key is processed as a message block with 0 input state first followed by the steps involved in the hashing process. This construction is a provably secure PRF if the underlying block cipher, namely three-fish, in Skein is a PRP and hence it is also a provably secure MAC. It has been observed by the designers that the PRF/MAC modes generate a variable length output which is encoded as one of the inputs to the hash function.

If n is the state size and m is the minimum of state and output size, then it has been claimed that [1, p.28] it takes up to $\min(2^{n/2}, 2^m)$ to forge the Skein MAC. Let output size be $2n$ bits, then $m = n$ bits. Now we can ask MAC oracle on $2^{n/2}$ equal length messages hoping for two messages to collide producing the same $2n$-bit output. That is with a good probability, we get a collision before the application of the two output transformations after $2^{n/2}$ MAC oracle queries. Let x and x^* be the colliding messages. Now we ask the MAC oracle for the $2n$-bit tag for the message $x\|y$. Note that $H(k\|x\|y) = H(k\|x^*\|y)$ and hence we output $x^*\|y$ as the forged message. Designers have made a similar observation from the point of view of the hash function in [1, p.38]. In the sense of this generic analysis, the complexity of the forgery attack is always upto $2^{n/2}$ instead of $\min(2^{n/2}, 2^m)$ whenever the output size is made more than n.

Note that for a single lane Skein-MAC (without cascade of multiple outputs), the security against the forgery attacks with the complexity of $\min(2^{n/2}, 2^m)$ still holds.

Acknowledgments: Many thanks to Stefan Lucks for reviewing this note and recommending some corrections.

References

1. N. Ferguson, S. Lucks, B. Schneier, D. Whiting, M. Bellare, T. Kohno, J. Callas, and J. Walker. The Skein Hash Function Family. First Round of NIST's SHA-3 Competition, 2008. Available at `http://www.skein-hash.info/` (Accessed on 2/4/2009).
2. A. Joux. Multicollisions in Iterated Hash Functions. Application to Cascaded Constructions. In M. Franklin, editor, *Advances in Cryptology-CRYPTO 2004*, volume 3152 of *LNCS*, pages 306–316, Santa Barbara, California, USA, Aug. 15–19 2004. Springer.

[2] The UBI chaining mode used for Skein hashing prevents the trivial forgery attacks that work on the secret prefix MAC following the Merkle-Damgård hash structure as noted in [1].