6.857 Homework	**Problem Set 1**	# 1-3 - The SHA-3 Competition
Elette Boyle		
David Harvison		
Rodrigo Ipince		
Benjamin Switala		March 12, 2009

Overall Analysis

Out of the four hash functions assigned to our group (ECHO, CRUNCH, SIMD, Dynamic SHA), we selected ECHO as the strongest. The ECHO paper addressed the most security issues, and was the only one to actually prove resistance in certain cases. They exhibited an ideal resistance of the function against preimage, 2nd preimage, collision, and k-multi-collision attacks. They also rigorously show resistance to differential cryptanalysis. In comparison, CRUNCH and Dynamic SHA have been found vulnerable to length-extension attacks.

CRUNCH

The CRUNCH algorithm is comprised of four stages: preprocessing, encryption permutation, compression function and hash computations. Preprocessing is just padding the message and setting initial values. The encryption permutation stage is based on an unbalanced Feistel scheme. Then the compression function is constructed by XORing two permutations from the encryption stage and selecting the number of bits depending on the message length required. The hash computation involves using the the compression function and the Merkle-Damgard construction together.

In general the author's arguments are based on previous work in testing designs similar to what CRUNCH uses but hand wave away the differences between the algorithms.

They argue for pseudorandomness based on the symmetric Feistel schemes, while utilizing an unbalanced Feistel scheme. In addition they argue that when the internal functions used in the Feistel scheme are secret they can achieve a proven level of security, though their internal functions are not secret and they admit that this changes the security bounds. They hand wave and say that doing more rounds should increase the security but do not show bounds.

In citing collision resistance they cite a paper which studied a hash functions whose compression functions use two permutations of fixed key block ciphers. Using this they claim collision resistance up to the birthday problem.

In claiming preimage resistance they cite they same paper as collision resistance. Using this they make a claim on the bound of security.

They do not mention resistance to differential cryptanalysis, linear cryptanalysis or length extension attacks. Though according to the SHA-3 Zoo, CRUNCH has been shown vulnerable to length extension attacks.

Overall, CRUNCH does not seem like a strong competitor in the SHA-3 competition because the

majority of their proofs are hand waved and they have already been shown vulnerable to length extension attacks.

SIMD

The SIMD algorithm uses a Merkle-Damgård mode of operation. Each compression step is computed as:

$$H_i = P(H_{i-1}, E_{M_i'}(H_{i-1} \oplus M_i)),$$

where

- H_i is the ith chaining variable,
- M_i is the ith message block,
- M_i' is the result of sending M_i through a Reed-Solomon-like expansion,
- $E_k()$ is a Feistel ladder block encryption with key k, and
- $P(a, b)$ is a few extra Feistel rounds, using a as a key to encrypt b.

The final compression stage is slightly different from the previous, and the final output is truncated to the desired output length.

The paper cites proofs which show the Merkle-Damgård iteration structure utilized is indifferentiable from a random oracle if the compression function is a random oracle. (These papers are [Chang, Nandi], [Lucks], and [Maurer et al].) These papers apparently prove security up to 2^n queries (where n is the length of the hash function), and conclude there are no generic collision, second-preimage, or preimage attacks on this mode of operation. It is important to note that the maintained internal state H_i in SIMD is twice as large as the output, to strengthen against attacks launched after some number of rounds. This internal-state chaining variable is only truncated in the final step.

The paper includes a discussion of MAC constructions based on SIMD. The first case is using SIMD in the HMAC construction. They state this method can be proved secure by the security proof of [Bellare]. The second case is just taking $\text{MAC}_k(M) = \text{SIMD}(k||M)$; ie, taking SIMD of the concatenation of they key together with the message M. They assert there are no generic shortcut attacks on this construction, based on the security proof in "the indifferentiability framework paper."

They also consider the application of SIMD as a key derivation function. If the compression function is a pseudorandom function, then SIMD is a PRF. In this case, they assert SIMD is a "good" randomness extractor, based on results in [Fouque, et al].

The encryption portion of the compression function is very similar to that of MD5 (and other members of the MD/SHA family). Unfortunately, instead of proving or citing proofs of security strengths of the function, the paper just quickly mentions the fixes they introduced in attempt to protect against particular attacks faced by MD5. For instance, they avoided a particular differential pattern attack by adding an extra rotation in the function design.

They briefly address general differential cryptanalysis in the paper, arguing that the expansion stage protects against differential attacks by making small changes in inputs yield significant changes in the outputs. The message expansion algorithm acts like an error-correcting code with high minimum distance: in their case, any two nonidentical messages will have expansions that vary in at least 520 bits in SIMD-256 (or 1032 in SIMD-512). These expanded values are used as the keys to the encryption steps, and differences between keys will propagate in the nonlinear portion of the encryption. The paper does not, however, provide a rigorous analysis or proof of resistance.

This is the extent of the security discussion in the paper. There is no mention of other vulnerabilities, such as linear cryptanalytic attacks on potential weaknesses in the encryption nonlinearity, or potential side-channel attacks on the substitution tables in the Feistel encryption.

Dynamic SHA

Dynamic SHA is based on SHA-2, with a twist: the compression function on a so-called "data dependent" function. In the author's words, "when the message is changed, different rotate right operations may be done." This is unlike old frameworks.

Collision resistance and and preimage resistance are claimed, but the proofs are hand-waved. Indifferentiabiliity from a Random Oracle is not directly mentioned, but the author admits that some hashes are more probable than others.

Resistance to differential cryptoanalysis is claimed, however resistance to linear cryptanalysis is not mentioned. Although the paper claims resistance to length extension attacks, such attacks have already been found. Dynamic SHA does not appear to be a strong competitor for the SHA-3 competition.

ECHO

The ECHO hash function builds on top of the Advanced Encryption Standard (AES) and the Merkle-Damgård paradigm. It reuses as many aspects of AES as possible and in turn, the authors claim it inherits many of its security properties. ECHO also adopts features from the HAIFA model and uses a double-pipe strategy (described in [Lucks]) to tackle some deficiencies of the models they build on top of. ECHO supports any hash output length from 128 to 512 bits, but the paper only focuses on standard lengths (224, 256, 384, and 512 bits).

In terms of security, the paper claims a resistance to preimage, 2nd preimage, collision, and k-multi-collision attacks that is as good as that of an ideal hash function. They support their claim with the combined usage of HAIFA and a double-pipe strategy (in which they use a chaining variable that is twice the size of the hash output). ECHO obtains preimage, 2nd preimage, and collision resistance directly from HAIFA, and incorporates the large chaining variable to achieve resistance to multi-collision and herd attacks. They do not, however, offer formal proofs for these claims. In the same section (Section 9), they also claim resistance to fixed-point and message extension attacks, which carry over from HAIFA.

The paper dedicates a large portion of its security analysis to differential attacks. It covers characteristic, differentials, fixed-key characteristic, and truncated differentials. The authors state (with

proof) upper bounds on the expected probabilities of success on some of these attacks. For example, the expected probability of the best characteristic over four or more operations in ECHO, averaged over salt and counter, is upper bounded by 2^{-750}. Proofs are presented in Section 8.2.3.

The paper also investigates ECHO's resistance against various attack methods that have been proposed against AES (since ECHO heavily relies on AES). Structural, algebraic, related-key, and known-key distinguisher attacks are covered. The treatment of these attacks is somewhat less formal, and most of the arguments rely on the security of AES itself. However, the treatment would not qualify as hand-wavy, since they do offer upper bounds for most of the attacks under certain assumptions (Section 8.3).

In general, the authors provided a reasonably formal treatment of the security of ECHO and even pointed out the places where they believe ECHO might have some vulnerabilities. ECHO seems to be a relatively strong candidate for the SHA-3 competition.