Generalization of the WHIRLPOOL Hash Function

David Albertson [1], Liljana Babinkostova[1], Alexander Hegedus[2], Henry-Louis de Kergorlay[3], Desislava Nikolov[4], and Laura Wells[5]

[1]Boise State University, [2]Alma College, [3]Wesleyan University, [4]Stellenbosch University, [5]Providence College

Introduction

The hash function WHIRLPOOL was developed in 2003 and accepted by NESSIE, an international organization that identifies quality cryptographic functions. Hash functions are used for password storage, message integrity verification, pseudorandom number generation, and non-repudiation in digital security. A hash function takes a string of arbitrary length and returns a string of a fixed length. Understanding the algebraic structure of a cryptosystem is essential for improving its security. In particular, it is important to maximize the order of the group generated by the hash function.

Objectives

- The current WHIRLPOOL function works in $GF(2^8)$. Generalize it to an arbitrary $GF(p^r)$ in order to anticipate future innovations.
- Find the group generated by the generalized WHIRLPOOL function in terms of parameters p, r, and the dimensions of the state matrix.

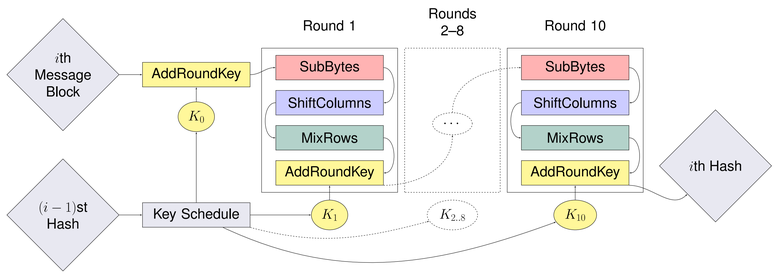

Figure 1 : WHIRLPOOL Hash Function [2]

Each round function can be thought of as a permutation on the space of state matrices. We analyze the parity of each of the round functions separately.

ShiftColumns (π)

ShiftColumns is a linear transformation that diffuses information among the rows of the state matrix.

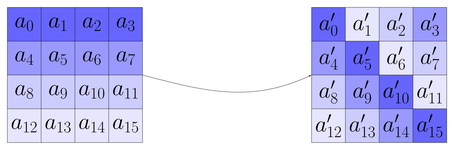

Figure 2 : The ShiftColumns Function

Theorem

Let $n(c)$ denote the number of configurations in a c-cycle; and $N(c)$ denote the number of configurations in a c-cycle, overcounting every configuration in a d-cycle, where $d|c$.
Then

$$N(c) = (p^r)^{c \cdot gcd(m, c(i))}$$

Therefore,

$$n(c) = N(c) - \sum_{d||c} n(d) = \sum_{d|c} \mu(d) N\left(\frac{c}{d}\right)$$

where μ is the Möbius function.

Counting the cycles in this manner is similar to counting other combinatorial objects, such as monic irreducible polynomials, necklace polynomials, and the number of Lyndon words.

MixRows (ρ)

MixRows is a linear diffusion layer which ensures that a small change in input results in a large change in the final output.

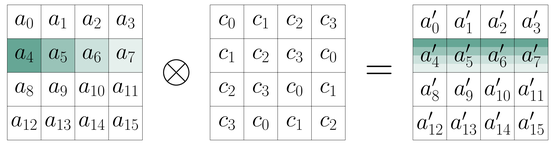

Figure 3 : The MixRows Function

If $\mathcal{M}$ is the algebra of $n \times n$ circulant matrices, then

$$\mathcal{M} \cong \frac{F[x]}{\langle x^n - 1 \rangle} \cong \frac{F[x]}{\langle f_1^{m_1} \rangle} \times \frac{F[x]}{\langle f_2^{m_2} \rangle} \times \cdots \times \frac{F[x]}{\langle f_r^{m_r} \rangle}. \qquad [4]$$

Theorem

Let C be a circulant matrix where $C \leftrightarrow (a_1, a_2, \cdots, a_r)$ and $a_i \in \frac{F[x]}{\langle f_i^{m_i} \rangle}$.
If $\sum_{i=1}^{n} c_i \not\equiv 1$, then MixRows is an odd permutation if and only if p, m and $\frac{p^{rn} - 1}{\text{lcm}(\{\text{ord}(a_i)\}_1^r)}$ are odd. If $\sum_{i=1}^{n} c_i \equiv 1$, then MixRows is an odd permutation if and only if p, m and $\frac{p^{rn} - p^r}{\text{lcm}(\{\text{ord}(a_i)\}_1^r)}$ are odd.

SubBytes (λ)

By introducing nonlinearity, the S-box used in the SubBytes function increases resistance against differential and linear cryptanalysis.

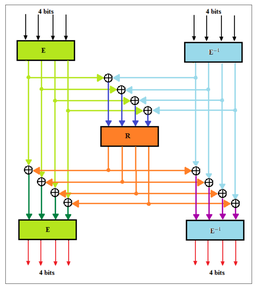

Figure 4 : Generating the S-Box

Definition

$$E(u) = \begin{cases} 0 & \text{if } u = p^r - 1 \\ g^u & \text{otherwise} \end{cases},$$

where $u \in GF(p^r)$ and g is a generator of $GF^*(p^r)$ such that neither E nor E^2 has fixed points.

Theorem

Let α be a generator $\pmod p$, σ_α be its induced permutation, and $T = \{\tau_j = \{j, (p-1)-j\} | 1 \le j < p-1, j \ne \frac{p-1}{2}\}$.

- $(\sigma_{\alpha^{-1}}$ has $1-$ cycle $(a))$ iff $(\sigma_\alpha((p-1)-a) = a)$
- $(\sigma_{\alpha^{-1}}$ has $2-$ cycle $(a, b))$ iff $(a \in \tau_i, b \in \tau_j, i \ne j$, and $\sigma_\alpha((p-1)-a) = b$, and $\sigma_\alpha((p-1)-b) = a)$

Theorem

Let p be an odd prime and $r = 1$. When $p \equiv \pm 3 \pmod 8$ and 2 is a primitive root $\pmod p$, then σ_2 contains the 2-cycle $\left(\frac{p+1}{2}, p-2\right)$. When $p \equiv -1 \pmod 8$ and $(p-4)$ is a primitive root $\pmod p$, $\left(\frac{3p-1}{4}, p-2\right)$ is a 2-cycle of σ_{p-4}.

AddRoundKey (σ)

AddRoundKey adds entropy to the algorithm by introducing outside influence from the previous hash value.

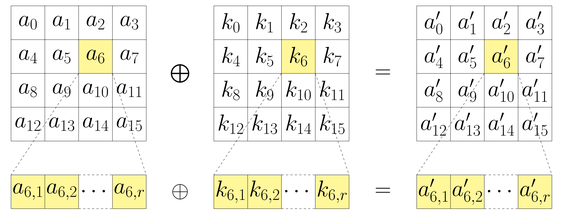

Figure 5 : AddRoundKey Function

Theorem [1]

For all $k \in M_{m,n}(GF(p^r))$ and $r, m, n > 1$, the function $\sigma[k]$ is an even permutation.

Future Work

- Determine the parity of the generalized WHIRLPOOL block cipher.
- For SubBytes: Determine when we can define the E function for the generalized WHIRLPOOL function in $GF(p^r)$.
- Find the group generated by the generalized WHIRLPOOL function.

References

[1] L. Babinkostova, K. Bombardier, M. Cole, T. Morrell, and C. Scott, *Algebraic Structure of generalized Rijndael-like SP networks*, (2012) (arXiv: 1210.7942)

[2] P. Barreto, V. Rijmen, *The Whirlpool Hashing Function*, First Open NESSIE Workshop, Leuven, Belgium, (2003), 320-335.

[3] R. Sparr and R. Wernsdorf, *Group theoretic properties of Rijndael-like ciphers*, Discrete Applied Mathematics, Vol. 156 (2008), 3139–3149.

[4] Z. Zhang, *Construction of the Orthogonal Groups of $n \times n$ Circulant Matrices Over Finite Fields*, Masters thesis, Concordia University, National Library of Canada, (1997).

Acknowledgments

Funding for this project was provided by the National Science Foundation under Award DMS 1062857 and by Boise State University.