Chapter 6
Cryptographic Hash Function: A High Level View

Imad Fakhri Alshaikhli
International Islamic University of Malaysia, Malaysia

Mohammad Abdulateef AlAhmad
PAAET, Kuwait

ABSTRACT

Cryptographic hash function verifies data integrity and sender identity or source of information. The task is accomplished by taking a variable bit patterns as an input then produces a fixed bit patterns of output. This chapter provides a detailed overview to include classification, properties, constructions, attacks, applications and an overview of a selected dedicated cryptographic hash function.

INTRODUCTION

A message digest \(\{0,1\}^*\) is an algorithm \(H\) that uses a un-constant size message as input \(\{0,1\}^*\) to produce a constant size output (sometimes called a digital fingerprint, imprint, hash result, hash code, hash value, or purely hash). All above alterative names were actually functions which made to play a fundamental role in modern cryptography practical applications, for example, a digital signature (Goldwasser et al., 1988), digital time stamp (Haber & Stornetta, 1991), message authentication code (or MAC) (Krawczyk et al., 1997), public key encryption (Cramer & Shoup, 2003), tamper detection of files and many more. Due to the application versatility it deserved the nickname “Swiss army knife of cryptography”.

Background

Cryptology is the science of hiding and revealing information. This term is subdivided into two categories; cryptography and cryptanalysis. Cryptography is the science of studying the information security. Where, its opposite term is cryptanalysis, which refers to the science of violating the information security. While, cryptographer studies the techniques needed to achieve and protect the information between any communication channels, cryptanalyst studies the techniques needed to understand the meaning of secret information and attempts to break them. More precisely, Whitfield Diffie and Martin Hellman, in their seminal paper on “New Directions in Cryptography” (Diffie
& Hellman, 1976), they have given birth to many concepts that are widely used these days by clearly distinguishing between privacy and authentication. Their definition of cryptography is "Cryptography is the study of mathematical systems for solving two kinds of security problems: privacy and authentication." Where, privacy of information ensures the secrecy of the information, and authenticity proves sender's identity. Another new definition proposed by Ronald Rivest (Rivest, 1991), he noted in chapter “Cryptography” in the “Handbook of Theoretical Computer Science” that cryptography is about communication in the presence of adversaries. The adversary is the enemy or opponent who aims to prevent the cryptographers of the cryptosystems to achieve their protection security goals. The adversaries can be in form of humans, computers operated or programmed by humans. The modern cryptography concept presented new directions of security measurements that should be considered when designing a cryptosystem. More precisely, modern cryptography is all about constructing and analyzing protocols that overcome adversaries’ threats that are related to information security. Where, information security falls into four measurements: data confidentiality, data integrity, authenticity and non-repudiation. To achieve the information security measurements, modern cryptography can be divided into several area of study. Cryptographic hash function is one branch of the cryptography family tree. Cryptographic hash function is a function that takes an arbitrary length as an input and produces a fixed size of an output. The viability of using keyed cryptographic hash function is to verify data integrity and sender identity or source of information. Consequently, the digital Holy Quran requires a keyed cryptographic hash function as protection method from any alteration or modification. Next sections will overview cryptographic hash functions in more details.

CLASSIFICATION, PROPERTIES, CONSTRUCTIONS, AND ATTACKS OF HASH FUNCTIONS

1. Classification of Hash Functions

Figure 1 classifies our research model function into two main categories unkeyed and keyed. The first one takes a un-constant span message as a single input and outputs a constant hash digest, \( Y:\{0,1\}^* \rightarrow \{0,1\}^n \). It is also known as a Modification Detection Code (MDC). Where the later accepts an random span message and a set span secret key as two inputs then outputs a set span hash digest., \( Y_K: \{0,1\}^{k}\times\{0,1\}^* \rightarrow \{0,1\}^n \). It is also known as Message Authentication Codes (MACs). Looking down at the unkeyed route stems out a subclass of three algorithms as shown in Figure 1. Below is the MDC super class hierarchy (OWHF), then (CRHF), and finally (UOWHF) (Naor & Yung, 1989).

CRHF commonly works with longer segment hash values. An unkeyed hash function is used with a set of absolute integer \( n \) that contain at least the below characteristics:

Figure 1. Classification of cryptographic hash function
(Naor & Yung, 1989).