**CHAPTER II  
BASIC OF THEORY**

1. **Definition of Cryptography**

Cryptography is the conversion of data into a secret code for transmission over a public network. Today, most cryptography is digital, and the original text called plain text turned into a coded equivalent called cipher text via an encryption algorithm. The cipher text is decrypted at the receiving end and turned back into plain text [1].

Cryptanalysis is the study of analyzing [information systems](https://en.wikipedia.org/wiki/Information_system) in order to study the hidden aspects of the systems. Cryptanalysis is used to breach [cryptographic](https://en.wikipedia.org/wiki/Cryptographic) security systems and gain access to the contents of encrypted messages, even if the [cryptographic key](https://en.wikipedia.org/wiki/Key_(cryptography)) is unknown.



**Figure 2.1 Enigma Machine**

**REF: (http://static.bbc.co.uk/history/img/ic/640/images/resources/topics/enigma.jpg)**

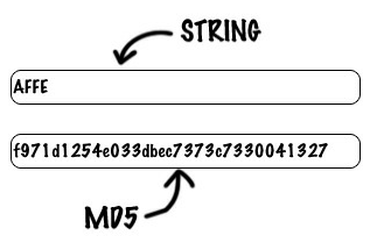
1. **Definition of Encryption**

Encryption is the process of encoding a message so that it can be read only by the sender and the intended recipient. Any procedure used in cryptography to convert plain text ito chipter text (encryipted message ) in order to prevent any but the inteded recipient from reading data. Encryption systems often use two keys, a public keys and a private keys. The public key available to anyone, and a private key that aloows only the recipient to decode the message [2].

1. **Definition of MD5 Algorithm**

MD5 is an [algorithm](http://whatis.techtarget.com/definition/algorithm) that is used to verify [data integrity](http://searchdatacenter.techtarget.com/definition/integrity) through the creation of a 128-bit message digest from data that is claimed to be as unique to that specific data as a fingerprint is to the specific individual. MD5 which was developed by Professor Ronald L. Rivest of MIT is intended for use with [digital signature](http://searchsecurity.techtarget.com/definition/digital-signature) applications, which require that large files must be compressed by a secure method before being encrypted with a [secret key](http://searchsecurity.techtarget.com/definition/private-key) under a [public key](http://searchsecurity.techtarget.com/definition/public-key) crypto system.

MD5 is currently a standard, Internet Engineering Task Force request for comments 1321. According to the standard, it is computationally infeasible that any two messages that have been input to the MD5 algorithm could have as the output the same message digest, or a false message could be created through apprehension of the message digest [3].



**Figure 2.3 MD5 Algorithm in Common Use**

**REF: (http://zdnet3.cbsistatic.com/hub.png)**

The MD5 hash consists of a small amount of binary data, typically no more than 128 bits. All hash values share the following properties:

1. The length of the hash value is determined by the type of the used algorithm, and its length does not depend on the size of the file. The most common hash value lengths are either 128 or 160 bits.
2. Every pair of non identical files will translate into a completely different hash value, even if the two files differ only by a single bit. Using today's technology, it is not possible to discover a pair of files that translate to the same hash value.
3. Each time a particular file is hashed using the same algorithm the exact same hash value will be produced.
4. All hashing algorithms are one way. Given a checksum value, it is infeasible to discover the password. In fact, none of the properties of the original message can be determined given the checksum value alone [4].

Put simply, MD5 is hash value typically expressed in text format as a 32 digit hexadecimal number, MD5 has been utilized in a wide variety of cryptographic applications, and is also commonly used to verify [data integrity](https://en.wikipedia.org/wiki/Data_integrity).

1. **History of MD5 Algorithm**

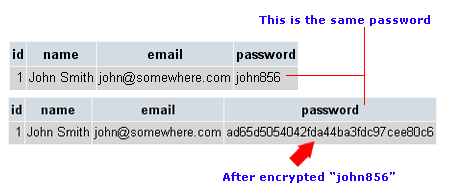
In 1990, Ron Rivest invented the hash function MD4. In 1992, he improved on MD4 and developed another hash function: MD5. In 1993, the National Security Agency published a hash function very similar to MD5, called the Secure Hash Algorithm (SHA). Then in 1995, citing a newly discovered weakness that it refused to elaborate on, the NSA made a change to SHA. The new algorithm was called SHA-1. Today, the most popular hash function is SHA-1, with MD5 still being used in older applications.

One way hash functions are supposed to have two properties. One, they're one-way. This means that it's easy to take a message and compute the hash value, but it's impossible to take a hash value and re-create the original message. (By "impossible," I mean "can't be done in any reasonable amount of time.") Two, they're collision-free. This means that it's impossible to find two messages that hash to the same hash value. The cryptographic reasoning behind these two properties is subtle, and I invite curious readers to learn more in my book Applied Cryptography.

Breaking a hash function means showing that either -- or both -- of those properties aren't true. Cryptanalysis of the MD4 family of hash functions has proceeded in fits and starts over the past decade or so, with results against simplified versions of the algorithms and partial results against the whole algorithms [4].

Simply, MD5 developed by Professor Ronald L. Rivest of MIT, is intended for use with digital signature applications, which require that large files must be compressed by a secure method before being encrypted with a secret key, under a public key cryptosystem. MD5 is currently a standard, Internet Engineering Task Force (IETF) Request for Comments (RFC) 1321.

SHA-1 (Secure Hash Algorithm) is a most commonly used from SHA series of cryptographic hash functions, designed by the National Security Agency of USA and published as their government standard. SHA-1 produced the 160-bit hash value. Original SHA (or SHA-0) also produce 160-bit hash value, but SHA-0 has been withdrawn by the NSA shortly after publication and was superseded by the revised version commonly referred to as SHA-1. The other functions of SHA series produce 224-, 256-, 384- and 512-bit hash values. SHA-0 published in 1993 as the Secure Hash Standard, FIPS PUB 180 by National Institute of Standards and Technology. SHA-1 published in 1995 in FIPS PUB 180-1. SHA-256, SHA-384 and SHA-512 first published in 2001 as draft FIPS PUB 180-2 and released as official standard in 2002. SHA-224 published in 2004 as change notice for FIPS PUB 180-2 [5].



**Figure 2.2 MD5 Algorithm Data Used**

**REF: (http://www.phpeasystep.com/imgs/encrypt\_password.gif)**