Chapter 1

Introduction

1.1 Preliminaries

Cryptography is where security engineering meets mathematics. Modern communication environments create the potential for the leakage of sensitive information and copying without authorization. The issues of protecting information transmitted over open channels has been widely studied in cryptology which covers confidentiality, authentication, data integrity, non-repudiation and other related information security objectives. Historically, cryptography was developed as a tool to protect secret information, especially in the military and government communities. Cryptographic algorithms were designed and used for ensuring the secrecy of internal communications only. Interest in cryptography has greatly increased with the invention of public key cryptography in 1976. This revolutionary work opened a new research field in cryptology and public key cryptography attracted a lot of interest from the academic community resulting in the publication of numerous research papers. Cryptography can be partitioned into two fields, symmetric key cryptography and asymmetric key cryptography. Symmetric ciphers fall into two categories: block ciphers and stream ciphers. Stream ciphers and block ciphers are the two most important families of symmetric encryption algorithms. Block ciphers are stateless and given plaintext directly output ciphertext. Stream ciphers maintain state and produce keystream which is externally combined with the plaintext to make the ciphertext. While design principles and security of block ciphers
are relatively well understood, stream cipher design still requires much re-
search. A great variety of design principles and the lack of understanding
of their interrelations is another reason for the poor state of the theory
of stream ciphers. Stream ciphers are interesting for two reasons. First,
they can be designed to allow much faster keystream generation in software
than a block cipher in stream cipher mode. Second, they can be designed
to be smaller in hardware. Hence, for a stream cipher to be motivating it
must either be very fast in software or very small in hardware. To address
the lack of standards for secure stream ciphers, several competitions were
undertaken by the cryptographic community. Cryptographic Research and
Evaluation Committee (CRYPTREC) was initiated by Japan, which rec-
ommended a few stream ciphers among them: 128-bit RC4, MUGI and
MULTI-S01. eSTREAM [1] initiated in 2004 was a project to identify new
stream ciphers that might be suitable for widespread adoption. It was set
up since none of the six stream ciphers submitted to NESSIE project [2]
organized by EU in the year 2000 has been selected as the portfolio. After
three evaluation phases of eSTREAM project in four years, the project was
completed with 7 primitives as the finalist portfolios in September 2008.
Grain v1, Mickey 2.0 and Trivium are the three selected hardware oriented
stream ciphers. Due to its simple structure, Trivium [3] has been a target
for quite a lot of cryptanalysis, but so far no attacks faster than exhaustive
key search is known on the full round cipher. Most of the stream ciphers
specified recently make use of two inputs to generate a keystream sequence:
secret variables representing the key, and public variables representing the
IV called the initial vector. This gives a clear practical advantage over pre-
viously proposed stream ciphers with single input. On the other hand, the
use of an IV as input has considerable impacts on the cryptanalysis and
on the formalization of the security requirements of stream ciphers. As
for cryptanalytic implications, an adversary can compare the keystream
sequences associated with several known, related or chosen IV values, and
potentially derive information corresponding to the internal state values
that could not be derived from single keystream sequence. We assume that
the attacker can simulate the cipher during the preprocessing phase, and
can apply a chosen IV attack during the online phase. A one-time pad
or Vernam cipher having perfect secrecy relies entirely on confusion and
does not employ diffusion. Since stream ciphers are considered to be a weaker version of an one-time pad, they are said to employ only confusion. From the security perspective, several stream ciphers and block ciphers are found to be vulnerable to cryptanalysis due to their weak design causing poor diffusion [4].

1.2 Block Ciphers

Symmetric key block ciphers are the most important elements in many cryptographic systems. They are modeled as a permutation and individually provides confidentiality. As a fundamental building block, they allow construction of pseudorandom number generators, stream ciphers, Message Authentication Codes, and hash functions. The input and output of the block ciphers are of fixed size. A block cipher is a keyed family of pseudorandom permutation which outputs a ciphertext given a plaintext and a key. In block ciphers, the plaintext is broken into blocks of fixed length say $n$. Each plaintext block is mapped into a ciphertext block of the same length by means of an invertible key-dependent substitution. Each key defines a permutation on $n$-bit blocks.

Shannon in his landmark paper of 1949 [5], presents the principles of confusion and diffusion. Both these properties have become the cornerstone in the design of a secure block cipher. In an iterated block cipher, a complex round function is used repeatedly, each time taking as input the output from the previous round. The well-known example of such a cipher is DES. The iterated structure in DES has its origin from the Feistel cipher which dates back from 1974 [6]. For the case of implementation of a Feistel cipher, we need not implement two different algorithms, one for encryption and one for decryption, instead we reverse the scheduling of the subkeys and this allows using the same algorithm for encryption and decryption with the same key. DES is now considered to be insecure for many applications due to its small key size of 56-bits. There are also some analytical results which demonstrate theoretical weaknesses in the cipher, although they are infeasible to mount in practice. The algorithm is believed to be practically secure in the form of Triple DES, although there are theoretical attacks. Later the cipher has been superseded by the Advanced Encryp-
tion Standard (AES). AES is a specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001. AES is based on the Rijndael cipher [7] developed by two Belgian cryptographers, Joan Daemen and Vincent Rijmen, who submitted a proposal to NIST during the AES selection process. Rijndael is a family of ciphers with different key and block sizes. For AES, NIST selected three members of the Rijndael family, each with a block size of 128 bits, but three different key lengths: 128, 192 and 256 bits.

A block cipher can be used in one of several modes of operation. Let $K$ be the key, $E_K$ be the encryption function, $D_K$ be the decryption function, $m = (m_1, m_2...)$ be the message and $c = (c_1, c_2...)$ be the ciphertext. Let IV be the public initialization vector. Then the five common modes of operation for block ciphers are given as follows.

**Electronic Code Book Mode (ECB).** This is the most obvious mode in which all plaintext blocks are encrypted and decrypted independently. Thus we have $c_i = E_K(m_i)$ and $m_i = D_K(c_i)$. A weakness of ECB Mode is that redundancy in the plaintext blocks will be preserved in the ciphertext blocks, since same plaintext blocks are encrypted to the same ciphertext blocks.

**Cipher Block Chaining Mode (CBC).** This mode chains the previous ciphertext with the current plaintext. Encryption and decryption is given by $c_i = E_K(m_i \oplus c_{i-1})$ and respectively, where $c_0 = IV$.

**Output Feedback Mode (OFB).** This mode will turn the block cipher into a stream cipher. A keystream $z = (z_1, z_2...)$ is generated as $z_i = E_K(z_{i-1})$ with $z_0 = IV$. The encryption and decryption is given as $c_i = m_i \oplus z_i$ and $m_i = c_i \oplus z_i$ respectively.

**Cipher Feedback Mode (CFB).** This mode of operation also behaves like a stream cipher. A keystream $z = (z_1, z_2...)$ is generated as $z_i = E_K(c_{i-1})$ with $c_0 = IV$. Encryption and decryption is given as $c_i = m_i \oplus z_i$ and $m_i = c_i \oplus z_i$ respectively.
Counter Mode (CTR). This mode of operation is also a stream cipher mode of block cipher. Here the keystream $z = (z_1, z_2, ...)$ is generated as $z_i = E_K(IV \parallel ctr)$ where $ctr$ is a counter which is incremented for each encryption.

A block cipher in CFB mode defines a self-synchronizing stream cipher while OFB mode and Counter Mode define a synchronous stream cipher.

### 1.3 Stream Ciphers

Stream ciphers are an important class of cryptographic algorithms that are used in practice. Unfortunately, their theoretical foundations have not been systematically investigated and the theory lags behind the practice. Stream ciphers have been adapted as the most popular and practical primitives for use in secure communication.

Design and analysis of stream ciphers was kept confidential for a long time and was made public in the 1970s, when several research papers on the design of LFSR-based stream ciphers appeared. Cryptanalysis discovered during the NESSIE and eSTREAM projects [1] [2] have made it possible to strengthen cipher designs to a large extent, and attacking new algorithms has become more difficult. Till the end of 1990s there were no standards for stream ciphers and the advent of these projects made an attempt to standardize the design of stream ciphers.

They are categorized into synchronous and self-synchronous stream ciphers. In general stream ciphers were considered to be fast in hardware and slow in software, and block ciphers were slow in hardware and fast in software. Stream ciphers were also classified based upon their output unit. Ciphers with output unit a single bit are known as bit based stream ciphers like Grain v1, Trivium, Mickey 2.0, WG-29, WG-8, WG-7, A5/1 etc. Those with output unit of more than a bit (atleast one byte) are called word based stream ciphers like RC4, Turing, Sober, Dragon, Snow, HC-128, Serpent, Rabbit, Salsa20 etc. Block ciphers have a major influence on the design of word-based stream ciphers. By integrating the components used in building block ciphers, word-based stream ciphers have experienced much improvement in speed over bit-based stream ciphers without any perceptible security degradation.
The classical example of a stream cipher is the one-time pad in which keystream bits are generated randomly and independently. Each keystream bit $k_i$ is combined with a corresponding plaintext bit $p_i$ to form the ciphertext $c_i = p_i \oplus k_i$ for $0 \leq i < n$, where $n$ is the length of the message. When used properly, the one-time pad is unconditionally secure, though it suffers from the requirement that the keystream needs to be as long as the message it encrypts. The goal of a stream cipher is to emulate a one-time pad in producing a keystream that appears unpredictable and random, as if it were the product of a non-deterministic process. Hence stream ciphers deterministically generate lengthy keystream from very small input key and the retrieval of the key should involve a process that is computationally infeasible.

A synchronous stream cipher takes as input a $k$-bit secret key $K$ and an $n$-bit public initial vector IV. The key and IV loading function generate the loaded state by filling both initial vector and key in a fixed pre-defined manner. Initialization function mixes the inputs to generate a random looking initial state, the state which is obtained after updating the loaded state of the cipher a fixed pre-defined number of times. Thereafter, the keystream is given as output and the state is continuously updated. Formally, a modern stream cipher is a combination of the following three functions $(L, U, H)$.

Key and IV Loading Function $L : \{0, 1\}^k \times \{0, 1\}^n \rightarrow \{0, 1\}^m$

State update Function $U : \{0, 1\}^m \rightarrow \{0, 1\}^m$

Output Function $H : \{0, 1\}^m \rightarrow \{0, 1\}$

$S_t : \text{State at time instant } t \text{ and } S_{t+1} = U(S_t, K)$ and keystream bit $z_t = H(S_t, K)$.

Synchronous stream ciphers also have the property that the keystream can be generated before plaintext is available. This can also be seen as an advantage over block ciphers. It is easy to design a stream cipher and the difficult task is to make it secure and at the same time provide good software or hardware performance.
1.4 Authenticated Stream Ciphers

There are many possible ways to design authenticated encryption schemes: generic compositions, single-pass modes, two-pass combined modes. Generic composition is a traditional way of achieving both authenticity and confidentiality, where we find two algorithms yielding each of these properties and then use their combination on a message with the help of two independent keys. Generic compositions are relatively slow, as they are two-pass constructions. Single-pass modes [8] [9] provides authenticated encryption after one time processing of a message. Since all of the single-pass modes have been protected by patents, two-pass combined modes have been appeared. They process data in two steps but only use a single key for both encryption and authentication. CCM, EAX, CWC and GCM [10] are few examples of two-pass combined modes. The term Authenticated Encryption (AE) scheme is used to refer to a shared-key based transform, whose goal is to provide both confidentiality and authenticity of the encapsulated data [11]. This approach is more efficient than applying a two-step process of providing confidentiality for a message by encrypting the message, and in a separate pass providing integrity protection by generating a Message Authentication Code (MAC). Authenticated encryption using symmetric ciphers can be provided by either stream ciphers with build in authentication mechanisms or block ciphers using appropriate modes of operation. Dedicated stream ciphers have the potential for higher performance and smaller footprint in hardware than block ciphers, which makes stream ciphers suitable for resource constrained environments, where storage and computational power are limited. If properly used dedicated stream ciphers can be used for designing primitives which provides both confidentiality and integrity with cost of $n$ operations for $n$ blocks of plaintext message. So there is significant interest in the development of stream ciphers which provide authenticated encryption. Hence the goal of authenticated encryption is to provide confidentiality and integrity for messages simultaneously, with much less effort than a two-pass process.
1.5 Objectives of the Thesis

The aim of the thesis is to investigate initial vector (IV) dependent stream ciphers. Specifically, we analyzed lightweight stream ciphers having LFSR, NFSR, nonlinear Boolean functions and FCSRs as building blocks. Our analysis shows the relevance of bit diffusion property in modern stream ciphers. We show this by proposing two statistical tests for diffusion known as $SAC-r$ and $SAC-c$ defined for determining the level of diffusion in the keystream of stream ciphers for a bit flip in the Key-IV bit sequence. Some results on cryptanalysis of a lightweight stream cipher Trivium using cube attack is presented. We proposed a linearity test for testing the superpolys in cube attack. Further to this, we designed a secure lightweight FCSR based stream cipher and a single pass authenticated stream cipher.

1.6 Outline of the Thesis

The thesis is organized as follows: Chapter-1 gives an introduction to the thesis. Chapters 2 and 3 describes the preliminaries required for the contributory chapters 4, 5, 6 and 7.

In chapter-4, we propose a modifications to the BEAN cipher by suggesting an optimal output Boolean function in terms of good cryptographic properties and also propose a change in its initialization function for randomizing its internal state. The modified cipher is known to resist the linear distinguishing attack.

Chapter-5 describes two statistical methods $SAC-r$ and $SAC-c$ diffusion tests for determining the level of diffusion in the keystream of any stream cipher for a bit flip in the Key-IV bit sequence. These tests are applied on stream ciphers Grain and Trivium and the results are detailed here.

A modified test for detecting the linearity of polynomials over $\mathbb{F}_2$ is proposed in Chapter-6, where we show 69 extremely sparse linearly independent linear equations obtained from very small cubes for Trivium reduced to 576 rounds. These equations are different from those obtained by Vielhaber in 2007.

Chapter-7 presents an authenticated encryption algorithm which is resistant to most of the attacks on stream ciphers such as linear, differential,
key recovery, and algebraic cryptanalysis. We use a cryptographically secure 8 bit rotation symmetric S-box for reducing the amount of space in table lookups, when implemented in hardware.