Name : **VivekVijayan/Raj Thakkar**
Course : **CPCS4600**
Assignment : **Final Project**
Professor : **Dr. Li Yang**
Date : **11/21/2013**

White Box Cryptography

Abstract

This paper discusses white box cryptography, which is used to protect the key from white box attack. Previously white box cryptography was applied to symmetric key encryption, which does provide protection mechanism to the key, but affects the performance and is considered difficult to update the key. The scheme proposed by the researcher mentions White Box AES and improves its low performance and key update problem by adopting a composite mode using White Box AES and Standard AES. The end result shows similar performance with AES and provides a dynamic key approach effect. [1]

Introduction

Attack contexts for cryptography module can be classified as black box, gray box, and white box attacks. Among which white box attack is considered to be the strongest attack and the adversary has all the privileges and also has complete access to the implementation of the algorithm and its dynamic execution. [1] Reverse engineering is one of the methods in white box attack where the adversary extracts the secret key in memory using the reverse engineering tool and can also observe the dynamic execution of the program which causes security vulnerabilities of cryptography modules. [1]

White box cryptography prevents key extraction by using one of its methods to hide the key in lookup tables. The adversary cannot find the secret key in physical memory as it cannot be seen in the memory directly. Thus white box cryptography is used to provide high security. The performance is affected due to the large size of the lookup table and the encoding and decoding process. The paper discussed a secure and efficient block cipher by using white box cryptography which can be used in mobile devices. [1]

Various Cryptography Modules

The cryptography module is divided into three types of attacks which are black box, gray box and white box.

Black Box Model

This module assumes that the attacker has no physical access to the key or internal workings, but can only observe the external information and behavior. The information consist of either revealing the plaintext or the cipher text of the system assuming zero visibility on code execution and dynamic encryption operations. [2]

There are three levels of attacks in the black box model. The first being passive attack which is also known as the plaintext attack which can only observe the input and output of the black box system. The second one being active attack which is also known as plain text attack which involves direct interaction. The third attack is the adaptive attack known as the plaintext-ciphertext attack. This attack gathers information by choosing a ciphertext and obtains its decryption using an unknown key. This attack may lead to knowing the ciphertext and obtaining the resulting plaintext. [1]

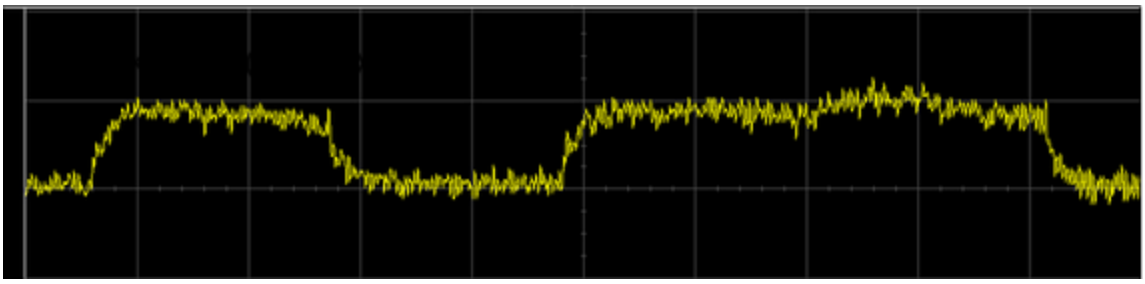
Gray box model

The grey box model is also known as the side channel attack or the partial access attack. The attacks are based on information gained from physical implementation of cryptosystem rather than a brute force attack. The adversary gains information from power consumption, timing, fault analysis and uses the information to break the system. The gray box attack states that any visibility to the inner working, side effect or execution of an algorithm can weaken the security. [1]

Side Channel Cryptanalysis

A Side channel attack is any attack based on the information gained from the physical implementation of a cryptosystem, rather than a brute force attack or theoretical weakness in algorithms. Example timing information, power consumption, electromagnetic leaks or sound can provide an extra source of information which can be exploited to break the system. Some of the attacks are:

- **Timing attack**: attacks that are based on measuring how much time various computations take to perform.
- **Power monitoring attack**: attacks that make use of varying power consumption by the hardware during computation.
- **Electromagnetic attacks**: Attacks that are based on leaked electromagnetic radiation which can directly provide plaintexts and other information.
- **Acoustic cryptanalysis**: attacks which exploit sound produced during a computation.
- **Different fault analysis**: in which secrets are discovered by introducing faults in a computation.
- **Data remanence**: in which sensitive data are read after supposedly having been deleted. [2]



In the above example we can see an attempt made to decode RSA key bits using power analysis. The left peak represents the CPU power variations during the step of the algorithm without multiplication, the right peak shows a step with multiplication, allowing to read bits 0, 1. [2]

White Box Model

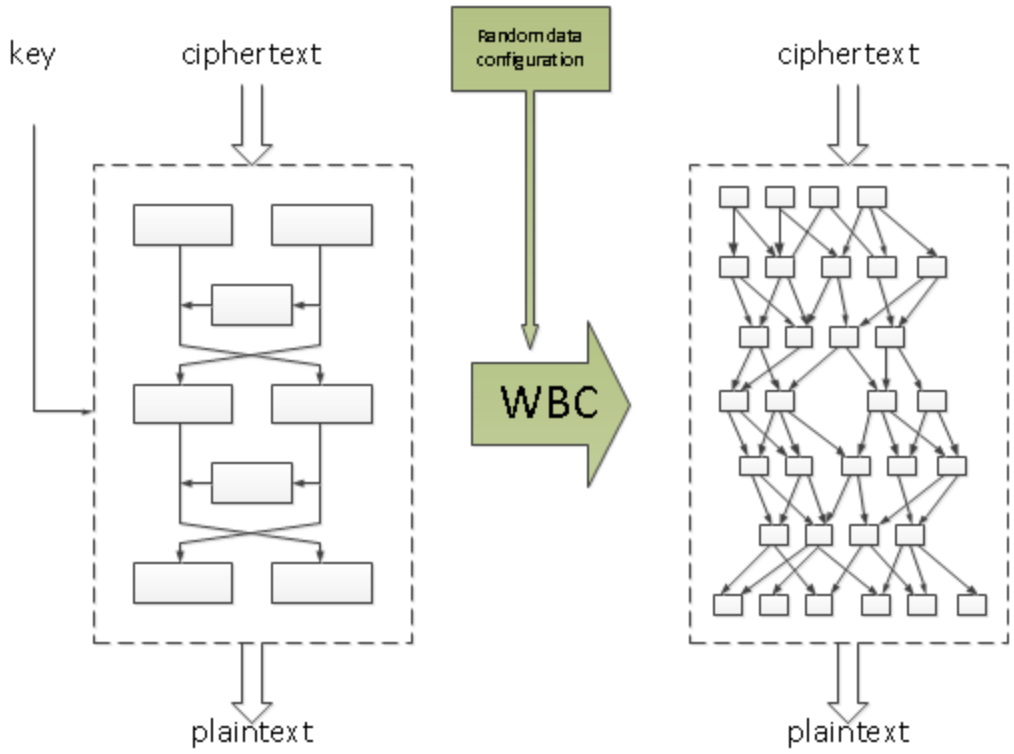
In this model the adversary has full control of the targets execution environment assuming that:

- Attacks with full privilege have complete access to the implementation algorithms.
- Dynamic execution can be observed and important data such as cryptographic keys can be seen.
- Detailed algorithms in the system are completely visible and alterable. [1]

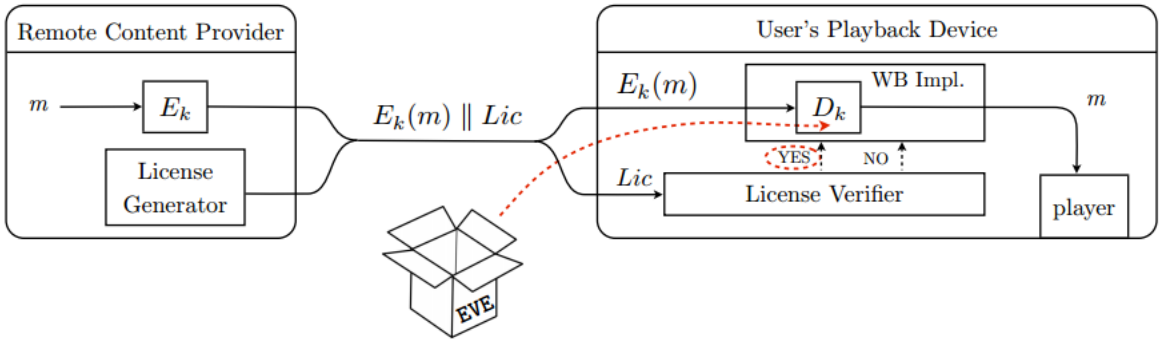
The adversary extracts the cryptographic key in this model as he knows all the algorithms and observes the dynamic execution, through which he can extract the keys from the memory. The adversary using this type of model attack has the strongest power to the cryptography system. To protect them techniques like obfuscation and temper resistance is used. [1]

White Box Cryptography

The research on cryptography in a white box model is called white box cryptography and implementation on a cryptography model is referred to as white box implementation. The adversary has full control of execution environment, white box cryptography is motivated to implement cryptographic primitives in a secure way. [1]



The above diagram shows a high level overview showing a White box cryptography key implementation. The Key is hardcoded in the code. White Box Cryptography transformations will generate code for an application where it is hard to extract they Key from the code.



The above diagram shows white box cryptography implementation in Trusted Digital media player which is applied in an untrusted environment.

The Goal of the adversary is to extract the secret decryption key in order to:

- Decrypt the encrypted content by passing the license.
- Distribute the key to non authorized users. [4]

The White Box AES uses lookup table that maps input plain text to ciphertext using a fixed key. If a block of cipher text is L bits, then the corresponding lookup table entries is 2^L. where L=64 bits or 128 bits. The amount of memory required to store this table is not possible. Thus a number of small lookup tables are used. [3]

White Box AES without protection

Chows's whitebox AES implementation without protection is used for this description.

AES 128 [3]

This is an iterated block cipher that maps a 16-byte input to a 16-byte output using a 16-byte Key. This has 10 rounds in it, where each round updates a 16-byte state variable which is treated as one-dimensional array, applied using combination of four basic transformations

- **First Round**: **Add round Key,** where each byte of the state is combined with block of round key with bitwise XOR.
- **Rounds**:
    - **Sub-bytes**: Is a non-linear substitution step where each byte is replaced with another according to the lookup table.
    - **ShiftRows**: Is a transposition step where each row of the state is shifted cyclically a certain number of steps.
    - **Mix Columns**: a mixing operation which operates on the columns of the state, combining the four bytes in each column.
- **Final Round**:
    - Sub-Bytes.
    - ShiftRows.
    - AddRoundKey.
- **T-Box**: Every round has AddRoundKey and SubBytes transorfmations combined into a series of 16 lookup tables that map bytes to bytes called as T-Boxes. There are in total 160-Tboxes.
- **T_{yi}-Tables:** For every round 1 to 9, when a byte is mapped through a T-Box, it is input into a Mix Columns transformation. In total of 144 **T_{yi}-Tables** are created to accept the outputs of the T-boxes in rounds 1 to 9.
- **XOR Tables:** For each round 1 to 9, Twelve 32-bit exclusive ors are required to determine the result of Mix-columns. To carry out the computation, 96 copies of the XOR table are created in each round (864 copies of XOR tables in each round).
- **Table Composition:** When a T-Box is fed directly into a **T_{yi}-Table** for rounds 1 to 9, the two separate tables can be replaced with their composition. Example in round 1, T_0^1 and T_{y0} could be replaced with new look-up table $(T_{y0})\ 0\ (T_0^1)$ where

$$(T_{y0})\ 0\ (T_0^1)(x) = T_{y0}(T_0^1(x)).$$

- Composing Table lookups reduces the number of individual table accesses required to carry out an encryption. Thus throughout the rounds 1 to 9, The T-boxes and T_{yi} tables are composed. [3]

In total we have (144 composed T-Box and T-yi tables, 864 xor tables and 16-T-boxes) tables for our implantation. The above steps can be summarized as shown below

```
state ← plaintext
for r = 1...9
    ShiftRows
    TBoxesTyiTables
    XORTables
ShiftRows
TBoxes
ciphertext ← state
```

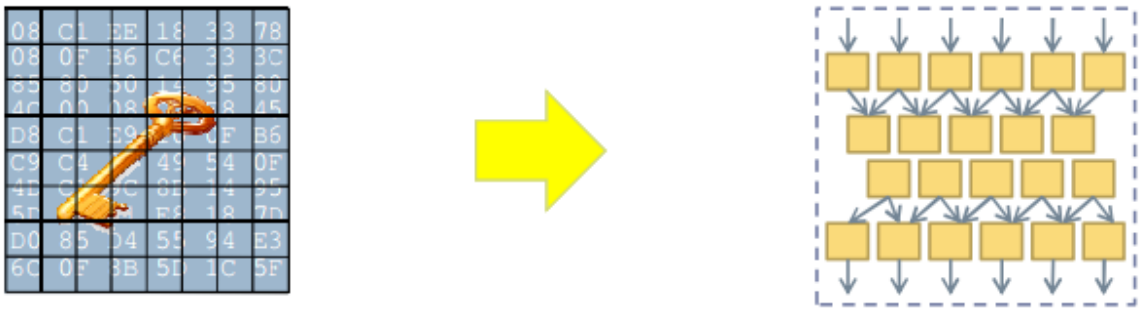
[3]

Protected Implementation

The White Box implementation AES- 128 encryption for a particular key executes in an environment which is under the control of an attacker. The use of a disassembler/debugger makes it easy for the attacker to learn the contents of the various lookup tables including the composed T-box/Tyi tables incorporating the round keys. This leads to steps to implement protection:

- Encoding.
- Mixed Bijections.
- External Encoding.

Obfuscation Strategy



The idea in Obfuscation Strategy is to spread the key information on entire network and to make every building block independent of the key. The Objective is to force the adversary to analyze

the complete network in order to obtain key information. The techniques used are: partial evaluation, by-pass encoding, matrix decomposition etc. [5]

Encoding

If the attacker gets to know the composed T-box/Tyi Tables for round 1, then the adversary can easily recover the AES-Key. Something must be done to protect the contents of the composed tables in round 10.

Encoding is simply a bijection which is used to protect the composed table T, we choose bijections f and g and form a new table T^1 where

$$T^1 = g_0 \; T_0 \; f^{-1}$$

Where f is input encoding and g is output encoding. [3]

Mixed Bijections

Application of concatenated input and output encodings help achieve confusion. To achieve diffusion linear transformations are composed to their inputs and outputs. An invertible linear transformation is referred to as mixing bijection. [3]

External Encoding

Questions arise considering white-box attack context as to why the adversary wants to extract the cipher key when software is available to decrypt the cipher text. Chow says the goal was to design an implementation so that it does not map raw ciphertext to raw plain text, but rather map encoded ciphertext to encoded plain text.

Thus to protect the table based AES implementation, mixing bijections, internal encodings and then external encoding is applied. Once mixing bijection is applied, the number of lookup-tables in round 1 to 9 doubles. Thus the table counts are as follows:

228 8- bit to 32-bit tables (1024 bytes each).

1728 8-bit to 4-bit tables (128 bytes each).

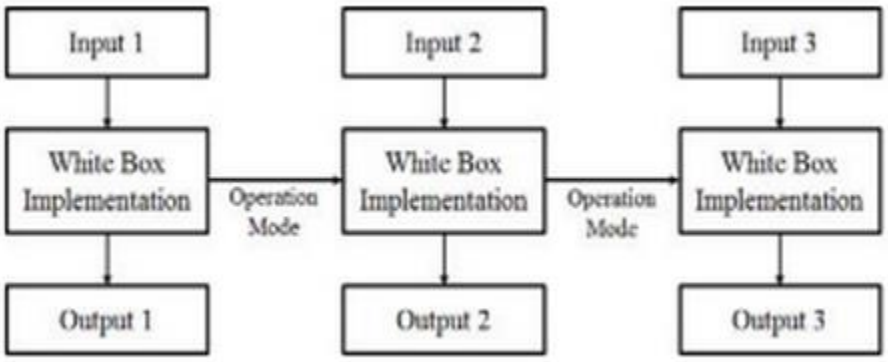
16 8-bit to 8-bit tables (256 bytes each).

The overall storage requirement for the tables is 508kb. [3]

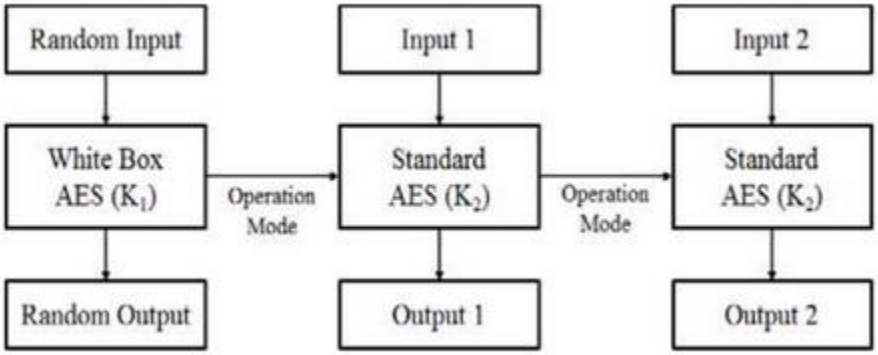
White Box cryptography provides high security that makes it difficult to extract secret key from cryptographic module. The cryptography module also has some disadvantages such as slow performance and key update problem of white box cryptography.

Proposed scheme

White Box AES as desgined by Chow is slower than the standard AES. An Example of White Box Chain is shown below



The White Box implementation is not efficient incase of large block cipher encryption or decryption. A more efficient method is needed for practical usage as shown below.



The above proposed method shows combining White Box AES and Standard AES to solve the problem of slow performance and key update in White Box AES. [1]

The White Box AES is used only once, an input is fed to the WB AES and it figures out a random output. The random input in WB AES is used as a secret value in the Standard AES which affects the results the remaining outputs through the block cipher. The above design is considered to be more effective than the white box AES chaining and the performance calculated is better than white box chains. [1] The adversary doesn't know the random input entered in the White Box AES scheme proposed. Even though the adversary comes to know the random output he would not be able to reuse it due to random property. As the user encrypts the message he may choose any random values for input in white BOX AES. The receiver who has the same white Box AES can only know the random output. [1]

To guarantee security CBC (Cipher Block Chaining) or OFB(Output Feedback) modes of operation has to be used to guarantee security.

Performance Analysis

15 messages of different length were made and were given as input in four types of AES which are White-Box AES, Standard AES, White Box AES Composition and the composed scheme. Four schemes were executed several times and the average time was figured out. The environment used to conduct the test is as shown below.

OS	Microsoft Windows 7
IDE	Microsoft Visual Studio 2005
CPU	Intel® Core™2 Duo CPU 2.4GHz
RAM	4GB
Program Language	C++

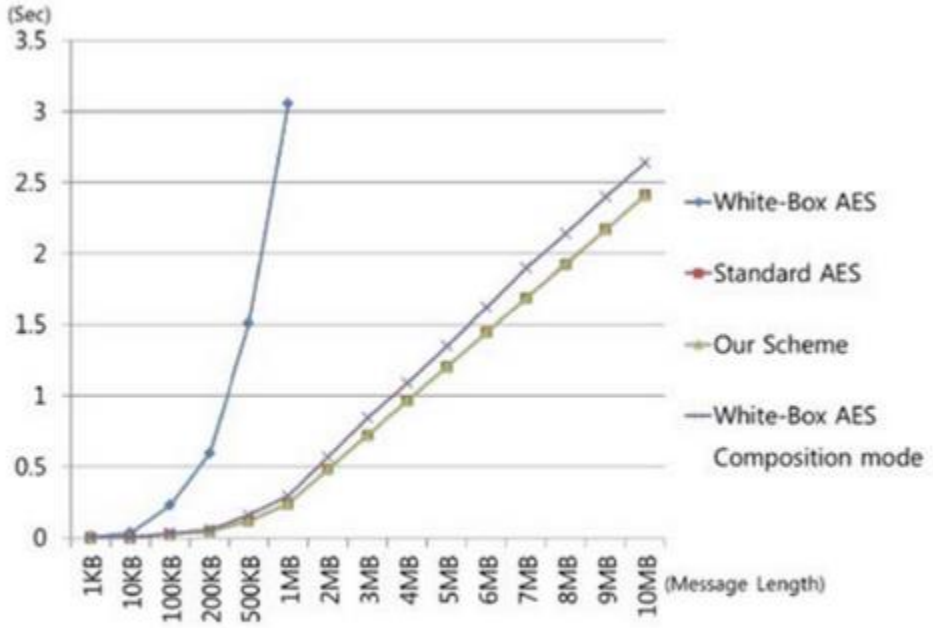


Figure 12. Encryption performance

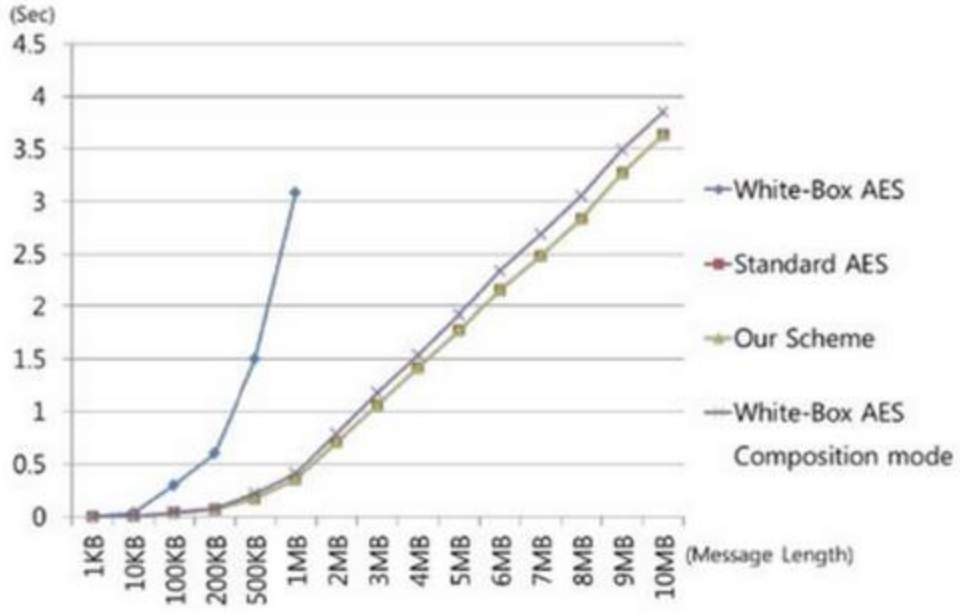


Figure 13. Decryption performance

From the above figure we can see White BOX AES has increased rapidly depending on the length of the message as it performs encode/decode process every message block. The Scheme proposed and the White Box AES Composition mode and the Standard AES scheme have alomost little overhead time. The scheme proposed uses White BOX AES only once. To encrypt and decrypt the ephermal key which is essential to encrypt and decrypt messages using Standard AES. Thus as the length of the message grows longer the scheme proposed has almost the same performance as the Standard AES. [1] When the length of the message is less than 10kb all the schemes have similar performance, but as the length of the messages increase the performance White BOX AES rapidly increases. As the message length is 10mb the encryption time is 7.9% and 11.8% decryption time when compared to white Box AES. [1]

References

[1] Understanding White Box Cryptography Whitepaper

[2] http://en.wikipedia.org/wiki/Side_channel_attack

[3] http://sac2013.irmacs.sfu.ca/slides/s15.pdf

[4] http://www.dagstuhl.de/Materials/Files/08/08253/08253.PreneelBart.Slides.pdf

[5] http://www.whiteboxcrypto.com/

[6] http://www.whiteboxcrypto.com/files/phdPresentation.pdf