BORANG PENGESAHAN STATUS TESIS*

JUDUL: AMBIENCE WHITE NOISE AS ASOURCE OF RANDOM NUMBER

SESI PENGAJIAN: _________2007/2008_______

Saya _______RABI'ATUL-ADAWIYYAH HASUBULLAH________

(HURUF BESAR)

mengaku membenarkan tesis (PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan (/)

_______ SULIT (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972

_______ TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

____/__TIDAK TERHAD

(TANDATANGAN PENULIS)	(TANDATANGAN PENYELIA)

Alamat tetap:
No. 103,Jalan Sg Kolek,26800,
Kuala Rompin,Pahang

NUR AZMAN ABU
Nama Penyelia

Tarikh: 17 Jun 2008

Tarikh: 17 Jun 2008

CATATAN: * Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

AMBIENCE WHITE NOISE AS A SOURCE OF RANDOM NUMBER

RABI'ATUL-ADAWIYYAH HASUBULLAH

This report is submitted in partial fulfillment of the requirements for the
Bachelor of Computer Science (Computer Networking)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2007

DECLARATION

I hereby declare that this project report entitled

AMBIENCE WHITE NOISE AS A SOURCE OF RANDOM NUMBER

Is written by me and my own effort and that no part has been plagiarized without citations.

STUDENT : ______________________ Date: 23.06.2008

(RABI'ATUL-ADAWIYYAH HASUBULLAH)

SUPERVISOR : ______________________ Date: 23.06.2008

(NUR AZMAN ABU)

DEDICATION

Special thanks to my beloved family and person who always support me for complete this project to gather to achieve the Bachelor of Computer Science. Besides, I also would like to special thanks to my supervisor who is also my P.A for supervising my project all this semester.

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to thanks Universiti Teknikal Malaysia Melaka (UTeM) for this 'Projek Sarjana Muda' in fulfill the requirements of completing Bachelor of Computer Science (Computer Networking) and improves the students in all the best ways.

Special thanks to Encik Nur Azman bin Abu, my supervisor for his invaluable guidance and constructive suggestions and devices throughout this project which really help me progress throughout the project.

I would like to express my sincere to all my classmates and others colleagues for their support and help in one way or another.

Last but not least, I wish to express my deepest appreciation and heartfelt thanks to my beloved family for their understanding, motivation, support and sacrifices so that I attend and success in this project.

ABSTRACT

This project shall capture audio sound. The Least Significant Bit (LSB) of audio sound will be analyzed and tested using National Institute of Standard and Technology (NIST) Test Suite. The NIST random Test Suite shall produce P-value. The P-value will determine the randomness of the captured audio sound. The P-value will generate the key that use for random number.

ABSTRAK

Project ini merekodkan signal audio. Nilai Nombor Se Bit Terkecil (LSB) audio yang direkodkan akan dianalisis dan diuji menggunakan National Institute of Standard and Technology (NIST) Test Suite. NIST random test suite akan menghasilkan nilai-P. nilai tersebut akan menentukan kerawakan audio yang direkodkan. Nilai tersebut juga akan menghasilkan kekunci yang akan digunakan sebagai sumber nombor rawak.

TABLE OF CONTENTS

CHAPTER	SUBJECT		PAGE
	ACKNOWLEDGEMENTS		**iii**
	ABSTRACT		**iv**
	ABSTRAK		**v**
	TABLE OF CONTENTS		**vi**
	LIST OF TABLES		**vii**
	LIST OF FIGURES		**viii**
CHAPTER I	**INTRODUCTION**		
	1.1	Project Background	1
	1.2	Problem Statements	2
	1.3	Project Objective	2
	1.4	Scope	4
	1.5	Project Significance	4
	1.6	Expected Output	5
	1.7	Conclusion	6
CHAPTER II	**LITERATURE REVIEW AND PROJECT METHODOLOGY**		
	2.1	Introduction	7

2.2 Literature Review 8
2.2.1 Domain 8
2.2.2 Keyword 8
2.2.3 Previous Research 9
2.3 Proposed Solution 17
2.3.1 Project Methodology 17
2.4 Project Requirements 21
2.4.1 Software Requirements 22
2.4.2 Hardware Requirements 22
2.5 Project Schedule and Milestone 22
2.6 Conclusion 24

CHAPTER III ANALYSIS
3.1 Introduction 25
3.2 Problem Analysis 26
3.2.1 Analysis of Current System 26
3.2.1.1 Random Number Generator Using MSP430 26
3.2.2 Problem Statement 28
3.3 Requirement Analysis 29
3.3.1 Data Requirement 29
3.3.2 Functional Requirement 30
3.3.2.1 Significant ordering bit 30
3.3.2.2 Random number NIST test suite 31
3.3.2.2.1 Random or pseudorandom number generator 33
3.3.3 Non-Functional Requirement 33
3.3.3.1 Audio Standardization 33
3.3.3.2 Result of Experiment 34

3.3.4 Others Requirement 35

3.3.4.1 Software Requirement 35

3.3.4.2 Hardware Requirement 36

3.4 Conclusion 37

CHAPTER IV DESIGN

4.1 Introduction 38

4.2 High-Level Design 39

4.2.1 User Interface Design 39

4.2.2.1 Navigation Design 40

4.2.2.2 Input Design 40

4.2.2.3 Output Design 40

4.3 Detailed Design 41

4.4 Conclusion 41

CHAPTER V IMPLEMENTATION

5.1 Introduction 43

5.2 Software development and environment setup 43

5.3 Software Configuration management 44

5.3.1 Configuration environment setup 44

5.3.2 Version Control Procedure 44

5.4 Implementation Status 49

5.5 Conclusion 50

CHAPTER VI TESTING

6.1 Introduction 51

6.2 Test Plan 52

6.2.1 Test Organization 52

6.2.2 Test Environment 53

6.2.3 Test Schedule 53

6.3 Test Strategy 54

6.3.3 Classes of Test 54

6.4 Test Design 55

6.4.1 Test Description 55

6.4.2 Test Data 55

6.5 Test Result and Analysis 55

6.5.1 Test Aspect Result 57

6.6 Conclusion 67

CHAPTER VII PROJECT CONCLUSION

7.1 Introduction 68

7.1.1 Strength 68

7.1.2 Weaknesses 69

7.2 Propositions for Improvements 69

7.3 Contribution 69

7.4 Conclusion 70

REFERENCE 71

APPENDICES A – Coding of selected test 73

APPENDICES B – Proposal Form 91

APPENDICES C – Gantt Chart 96

LIST OF TABLES

TABLE	TITLE	PAGE
2.1	PSM 1 Milestone	23
3.1	Data model of the system	29
3.2	Indicating of random bit	31
3.3	Sample of *P-value*	34
3.4	Software Requirement	35
3.5	Hardware Requirement	36
4.1	Input Design Interface	40
4.2	Output Design	40
5.1	Conclusion of Hypothesis	45
5.2	Description of Selected NIST Test Suite	46
5.3	User Interface implementation Status	49
5.4	User Interface implementation Status	50
6.1	System Tester	53
6.2	Test Schedule	53

LIST OF FIGURES

DIAGRAM	TITLE	PAGE
2.1	Waterfall Model	20
3.1	Overview of the current system	27
3.2	Flow of the functioning of system	30
3.3	Pseudo-Random Number Generator requires random seed in order to be secure.	32
4.1	User Interface	39
6.1	Result of Entropy test	58
6.2	Result of Frequency Test	59
6.3	Result of LongestRunsOfTest	60
6.4	Result of Monobit Test	61
6.5	Result of Runs Test	62
6.6	Result of Serial Test	63
6.7	Result of Serial Test	64
6.8	Result of Spectral Test	65
6.9	Result of Cumulative Test	66

CHAPTER 1

INTRODUCTION

1.1 Project Background

The white room noise has a potential to be a source of random numbers. The random numbers plays an important role in cryptosystems. Randomness and random number primarily used for several purposes in many applications such as statistical sampling, experimental simulation, cryptography and etc. randomness will be introduce to pseudo-random number generators by Computer Algorithm. Pseudorandom numbers have the characteristics that they are predictable which can be predicted if we know where the first sequence number is taken from.

In this project, a recording-system will be built to capture audio or sounds using microphone, and from that recorded audio will then analyzed. The recorded audio will be analyzed and examined to determine whether they are high quality source and suitable for used as cryptographic keys such as Session Key, Master Key, Public Key or Private Key. These bits of audio will be tested based on National Institute of Standards and Technology (NIST) Test Suite. The problems that will be face in this project are uses of random number in audio. To solve this problem, I will develop a system that will generate a random number using MATLAB software. In this project, I will apply the knowledge of pseudorandom number in cryptography.

1.2 Problem Statement

In this project, I would like to solve how random number interacts in audio.

i- Randomness of noise that will contribute by Least Significant Bits.

ii- Predictability of the bits of noise as a source of random key.

1.3 Objective

i. To write a program using MATLAB programming language that is able to record the audio by using the microphone.

   a. MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in easy-to-use environment where problems and solutions are expresses in familiar mathematical notation.

   b. MATLAB features a family of add-on application-specific solutions called toolboxes. Very important to most users of MATLAB, toolboxes allow user to learn and apply specialized technology. Toolboxes are comprehensive collection of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. Areas in which toolboxes are available include signal processing, control systems, neutral networks, fuzzy logic, wavelets, simulations and many others.

ii. To analyze the bits of recorded audio using selected test.

   a. Each digital recording sample should contribute only one bit and preferably the least significant bit. At least the first one thousand samples must be discarded since the recording will take several micro seconds before actually getting nonzero input.

   b. The microphone should be set at the highest sensitivity. The common threshold value used to identify silence period should be set to zero from the regular voltage 0.1 or disabled.

iii. To write a selected Statistical Test for Random and Pseudorandom Number Generators, NIST Special Publication 800-22 using MATLAB (Matrix Laboratory) programming language that consists of sixteen different tests.

   a. This test was developed to test the randomness of binary sequences produced by either hardware or software based cryptographic random or pseudorandom number generators.

   b. These test are:

      i. The Frequency (Monobit) Test
      ii. Frequency Test within a Block
      iii. The Runs Test
      iv. Test for the Longest-Run-of-Ones in a Block
      v. The Binary Matrix Rank Test
      vi. The Discrete Fourier Transform (Spectral) Test
      vii. The Non-overlapping Template Matching Test
      viii. The Overlapping Template Matching Test
      ix. Maurer's "Universal Statistical" Test
      x. The Lempel-Ziv Compression Test
      xi. The Linear Complexity Test
      xii. The Serial Test
      xiii. The Approximate Entropy Test
      xiv. The Cumulative Sums Test
      xv. The Random Excursions Test
      xvi. The Random Excursions Variant Test

iv. To identify whether the bits of the recording audio is random, unique, and suitable to use as cryptographic keys.

   a. There are two ways for generating random numbers; by using a physical device or by using algorithm. By using a physical device it takes various factors such as noise and time of day into account.

   b. Physical hardware random number generator has a greater advantage, since it can produce completely unpredictable random sequences. .

Pseudorandom number generators are usually combined with a physical generator to obtain both the speed and some of the unpredictability.

1.4 Scope

i. The system will record the audio signal for two secoonds.

ii. The system may be implemented in Windows operating system.

iii. The audio that has been recorded will be analyzed using suitable test.

iv. The target users are who want to use the system that analyzed audio as a source of random number.

1.5 Project Significance

This project is important because I will

1. Recording the audio using microphone from a system that has been build by MATLAB programming.
2. Analyzing the 16-bit mono of 44100 Hertz audio recording using suitable test.
3. The Statistical Test for Random and Pseudorandom Number Generators will be written.
4. The 16-bit mono of 44100 Hertz bits of recording audio will be identified whether it is random, unique, and suitable to use as cryptographic keys.

1.6 Expected Output

The proposed ambience white noise as a source of random number is:

1. Recorded the audio using microphone from a system that has been build by MATLAB programming.
2. Analyzed the 16-bit mono of 44100 Hertz audio recording using suitable test.
3. Write the Statistical Test for Random and Pseudorandom Number Generators.
4. Identified the 16-bit mono of 44100 Hertz bits of recording audio whether it is random, unique, and suitable to use as cryptographic keys.

1.7 Conclusion

As a conclusion, this is a critical project that needs to be completed by six months as the NIST Statistical Test Suite is very hard to be written in M-lint code file. Hopefully, after finishing this project, I will gain more knowledge and information about cryptography field. And I also hope the future undergraduate student will be able to understand the essential aspect for secure communication is that cryptography.

For the next the next chapter, I will write about literature review and project methodology. Literature review will provides a background of the topic, is a summary and an evaluation of previous research on a topic, and it should be selective. While the project methodology is a justification and description of the selected approach or methodology used in a project. It also provides explanation of the detail activities in the project.

CHAPTER II

LITERATURE REVIEW AND PROJECT METHOLOGY

2.1 Introduction

The literature review provides the backgrounds to the topic. It is a summary and an evaluation of previous research on a topic. It represents the method of searching, collecting, analyzing, and compute conclusion from book writers or other open sources about certain topics. Most of the literature reviews are done based upon the facts and findings from writers of IT books. For this project, the literature reviews are collected from books and journals.

Project methodology is the justification and description of the selected approach or methodology used in project. It will explain about Least Significant Bits and National Institute Statistical Test Suite (NIST).

After defining the project methodology, the next section will covered on project requirements in term of software requirements, hardware requirements and other requirements if needed. Then, it will discuss the next section which is project schedule and milestones. It is a proper guideline that will help the developer to complete the project.

2.2 Literature Review

2.2.1 Domain

The recording system using MATLAB software is in the domain and application of cryptography in Network and IT Security. The project will focus on cryptography or random number as a subject of study. This cryptography field is very wide and it is mostly focus on developing a system for cryptography domain. The domain that related in this project is cryptography field. Cryptography is the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. The goals of cryptography are privacy or confidentiality, data integrity, authentication, and non-repudiation.

2.2.2 Keyword

- **White noise** – is a sound that contains the frequency within the range of human hearing in equal amounts. People will perceive this sound as having more high-frequency content than low. This perception occurs because each successive octave has twice as many frequencies as the one preceding it. For this project, the microphone should be automatically preset at the highest sensitivity in order to capture the full audio sound range and the common threshold value used to identify silence period should be set to zero from the regular voltage 0.1 or disabled.
- **Random number** – number that occur in a sequence. This number will be described in two conditions. These conditions are the values are uniformly distributed over a defined interval or set and the other is it is impossible to predict future values base on past or present ones.

2.2.3 Previous Research

i- Research 1: Multi-Speaker voice cryptographic key generation

By: L. Paola Garcia-Perera, J.Carlos Mex-Perera and Juan A, Nolazco-Flores

This article will present the procedure for generating binary vector that be used as keys for cryptographic purpose. The research is based on Automatic Speech Recognition technology and Support Vector Machines. The drawback of the current system that used cryptographic is problem in authentication. It cause of the simplest algorithm the authorized user should deny across to the information. For the biometric key, it goal is to produce a password using the intrinsic characteristics of some specific user distinguishing quality such as voice, iris, fingerprint and face. Although the voice is a natural process and produces different utterances in short time, it gives advantages among other biometrics which allowing to assure the user's identity.

The researchers will focus on voice for this research, it characteristics which is to properly obtain reliable cryptographic key. The type of key generation has several applications in security which are access control to the network either to a computer system or to a restricted area and used for remote control by telephone. The cryptographic scheme of this research is based on the knowledge of the phonemes, which it is important to know the start and ends of the phonemes. The challenge for this research is to produce a cryptographic key that should repeatedly be the same for one user under certain conditions.

For the previous research by Monrose, it showed that method to produce the key which an acoustic space was formed derived from the utterances of a significant group of people. A vectorial quantization was applied and a set of 20 centroids was obtained. Afterwards, these sequences were segmented. An iterative process using Viterbi algorithm was used. A partition plane for the vector space was suggested and a binary biometric key was conformed. At last, Montorel proposed the hardening of the key using shared schemes.

The purpose of this research by these researchers is to produce cryptographic key given the voice signal and the spoken user passphrase. The speech signal is pre-processed and the Automatic Speech Recognition (ASR) technique is used to extract the phoneme model and the phoneme starts and ends. This set of features is the input of the Support Vector Machine (SVM) model the key is generated.

The primary task of ASR is to obtain the transcription of a given speech signal and to obtain the starts and ends of the phonemes if properly configured. The speech signal is pre-processed by calculating it Mel Frequency Cepstral Coefficients (MFCC). MFCC is a common transformation that has shown to be very robust. To obtain the MFCC, each speech signal is divided into short frames. For each frames, the amplitude spectrum is obtained using the Discrete Fourier Transform (DFT). Then, the spectrum is converted to the log scale. After that, the Filter bank is used to smooth the scaled spectrum. At last, the discrete cosine transform is applied to eliminate the correlation between components. As a result, the 13-dimension vector, each dimensions corresponding to one parameter. To emphasize the dynamical features of the speech in-time, the time derivative (Δ) and the time acceleration (Δ^2) of each parameter are calculated. Then, the final representation is a 39 dimension vector formed by 12-dimension MFCC, followed by 1 energy coefficient, 13Δ and $13\Delta^2$.

ii- Research 2: Physical Random Number Generators for Cryptographic Application in Mobile Devices

By: Shinichi Yasuda, Tetsufumi Tanamoto,Ryuji Ohba, Keiko Abe, Hanae Nozaki, and Shinobu Fujita.

This article present the small random number generators using silicon devices that generate large fluctuating signal as noise source device. The noise signal is directly input to an oscillator without preamplifier because the noise of this signal is large. The researcher present the small physical random number generator using an astable multivibrator and post-processing circuits, which can generate excellent quality random number s that suitable for cryptographic application.

The difficulty of predicting secret data relies on the quality of random numbers used as identification numbers, encryption keys, and data blinding. For the mobile device, random number generators (RNGs) must be small for implementation in a tiny chip. For solution, pseudo random number generators (PNRGs) are generally used for mobile devices because PRNGs can be produced using only a digital circuit such as a linear feedback shift register (LFSR). The random number generators using white noise, such as thermal noise or the shot noise of Si devices, are already commercially available. Even it can generate high quality random numbers; the size of the RNGs is relatively large because such noise signals are so small that they need to be greatly amplified. One effective solution to achieve both downsizing and acceptable performance of RNGs is to use current fluctuations in specific noise source devices that the researchers have developed based on metal oxide semiconductor (MOS) devices, such as MOS capacitor after soft breakdown. This paper also describes a small RNG that can generate high-quality random numbers using a circuit composed of an astable multivibrator and a one bit counter to convert analog signals to digital signals, which can eliminate 1/f-like properties. Furthermore the maximum of generation rate of this circuit is analyzed. We also present the other small RNG having much larger generation rate.

The researcher has study various noise source devices, such as MOS capacitor after SBD, Silicon dot tunnel diode and Silicon dot FET. Silicon dot tunnel diode is a MOS capacitor like device which has Si nanocrystals in the oxide film. Silicon dot FET is a specific FET which has Si-dot in the gate insulator. The characteristics of Silicon dot tunnel diode is observed that the current changes largely even in the case of constant applied voltage. For example current fluctuation is caused by the electron-electron interaction between tunnelling electrons and trapped electrons in dots. Electron trapping and detrapping occurs very frequently. The normalization power spectrum density of the