International Journal of Cryptology Research 2(1): 73-87 (2010)

One Megabit Random Ambience

[1]Nur Azman Abu and [2]Shahrin Sahib
Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka (UTeM),
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
Email: [1]nura@utem.edu.my, [2]shahrinsahib@utem.edu.my

ABSTRACT

Every cryptographic operation nowadays mandates random key as an input. These operations are mostly designed and taken care of by the developers of the cryptosystem. Due to the nature of random numbers, pseudorandom numbers are preferred instead due to its efficiency and reliability. As the name suggests, pseudorandom numbers are not truly random. Rather, they are generated from a mathematical formula. The outputs of pseudorandom number generators may exhibit most of the theoretical properties of random numbers. However, pseudorandom numbers are predictable, periodic and reproducible by the developer of the cryptosystem. True random numbers are believed to be generated only using hardware random number generators. Careful statistical analysis is still required to have any confidence the process and apparatus generates numbers that are sufficiently random to suit the cryptographic use. In this underlying research study, each moment in life is considered unique in itself. The random key is unique for the given moment generated by the user whenever he or she needs the random bits in practical cryptographic applications. An ambience of high fidelity digital image shall be tested for its randomness according to the NIST Statistical Test Suite. Recommendation on generating one megabit random cryptographic keys live shall be reported.

INTRODUCTION

Modern cryptographic system should be open and its security is no longer based on the secrecy of the algorithm and system design but rather on the randomness of the key being used. There are mainly two separate ways for generating random keys. First, the random bit can be captured from random phenomena by using a physical device. This strategy takes various factors, such as noise and time of day, into account. Extremely complex hardware random number generators are based on essentially random atomic phenomena, such as radioactive decay and thermal noise.

Second, random keys can also be generated computationally, by using algorithms. The second method generates so called pseudorandom keys which are sequences of numbers which approximate many of the properties of random keys. Pseudorandom keys can easily be regenerated, thus they are not truly random in nature. Nevertheless, the latter method is still preferred especially on large input or plaintext which requires long keys.

Furthermore, most of the cryptographic operations nowadays mandates random key as an input. These operations are mostly designed and taken care of by the developers of the cryptosystem. The main reason behind the original idea of this study is to check on the validity of having white room noise as a source of random keys. A developer of stream cipher system may easily make a claim that his/her system is highly secure since the system use room noise to capture the random key. The security of such an OTK(One-Time Key) crypto-system relies heavily on the design and the trustworthiness of the developer oneself.

In this paper, a new technique is proposed using simple apparatus which is capable of generating large random key live on demand. The technique has been designed based on the air ambience principle. The rest of the paper shall be organized as follows. Section 2 will give an overview of physical random number generators with their limitations. Section 3 will introduce the concept of air ambience as a possible source of randomness. In Section 4, strict requirements have been imposed on the random key generation apparatus. Section 5 and 6 explores the issue of user's physical environment and discusses the modus operandi of capturing high fidelity digital images as a source of random input. Section 7 explains the selected random tests for one-mega bit key. Section 8 gives a sample experimental result. Lastly, Section 9 concludes the paper.

PHYSICAL RANDOM NUMBER GENERATOR

In principle, true random number generators must capture randomness from physical phenomena. The physical phenomenon can be very simple such as the little variations in user's mouse movements or in the time difference between keystrokes from the user's typing speed. These techniques, though friendly, produce limited random bits with low entropy.

Due to the nature of random numbers, pseudorandom number generators are preferred instead due to their efficiency and reliability (Barker and Kelsey, 2007). As the name suggests, pseudorandom numbers are not truly random. Rather, they are generated from a mathematical formula. The outputs of pseudorandom number generators may exhibit most of the theoretical properties of random numbers. However, pseudorandom numbers are predictable, periodic and repeatable by the developer of the cryptosystem. The use of pseudo-random processes to generate secret quantities can result in pseudo-security (Eastlake *et. al.*, 2005).

The use of pseudo random number generator is quite misleading. It still remains a common problem today. It can be observed in the latest textbook on how to generate a simple random number in iPhone application development (Drake, 2009). The author suggests on how to generate random number via the random function from a regular old C function by setting the seed to the system clock. System clock is a 32-bit number. It carries only few bits entropy.

A physical true random number generator uses a natural phenomenon as an entropy source to produce random bits. It captures and measures unpredictable natural processes by using a dedicated hardware device. The measurement usually follows certain principle of physics such as radiation (Tsoulfanidis and Landsberger, 2010), quantum theory (Gottfried and Yan, 2003) (Garrison and Chiao, 2008), laser pulse (Gibbon, 2005) (Diels and Rudolph, 2006), gas theory (Sone, 2007), chaos theory (Ott, 2002) (Gleick, 2008), thermal/electrical noise (Gupta, 1997) (Davenport and Root, 1987) (Ritter, 2004) and so on. However, most of them require apparatus and equipment of especially dedicated expensive devices.

RANDOM AMBIENCE

Physical hardware random number generator has a greater advantage, since it can produce completely unpredictable random sequences. Air ambience is the natural choice here. The random movement of microscopic particles suspended in the air is caused by collisions among particles and the gas molecules of air. Density of air is about 1.29 kg m–3 or 0.0129 g cm–3. Average mass of air (a mixture of different gases) is 28.9 g/mol as shown in Table 1 One mole is 6.02 x 1023 particles, namely, Avogadro's number.

The particles suspended in the air caused by collisions between these particles and the gas molecules of air move in random directions, at high speed. Average velocity of a single air particle is around 500 m/s at room temperature 27^{o}C or 300 Kelvin. For oxygen O_2, the mass M = 32 g/mol at the room temperature of 298 Kelvin = 25^{o}C, its velocity is approximately 444 m/s. Assuming the molecular collisions driving the motion are completely random, the motions in the three directions are uncorrelated.

TABLE 1: The velocity of major air molecules at 20° C

Molecule	Molar Mass (g/mol)	rms Velocity at 20° C (m/s)
Hidrogen H_2	2	1902
Helium He	4	1352
Water H_2O	18	637
Nitrogen N_2	28	511
Oxygen O_2	32	476
Carbon Dioxide CO_2	44	408

Based on the basic principles of physics, the molecules in the atmosphere obey Newton's law of motion. Nevertheless, as a whole they move randomly, meaning, any molecule may move in any direction with any speed. At any given moment, a certain percentage of molecules move at high speeds and a certain percentage move at low speeds (Serway and Jewett (2003). The movement of these molecules in the air generates the random ambience.

CHALLENGES IN GENERATING RANDOM KEYS

Most physical generator can only generate short keys that are of insufficient length for modern cryptographic protocols. Others requires expensive special device such as radioactive scanner. The users and owners of cryptosystem are mostly nontechnical. It is important to have a generation process and apparatus which are physical, economics, convenient, efficient to use and secure.

The system should use only requires simple hardware and utilize only few basic computer algorithms. The randomness should come from the source rather than the algorithm itself. The measuring device should work automatically on demand from user's physical environment. In order to make sure the randomness of the numbers being generated, they should pass a full set of statistical tests once generated.

The criteria of the modern cryptographic key generation process and apparatus should be as follows:

i. Ability to capture physical source of randomness
ii. Use of minimum hardware
iii. Economically viable to operate
iv. Use of minimum mathematical formula
v. Involve only basic computer algorithm
vi. Pass Statistical tests
vii. Convenient to non-technical users
viii. Computationally efficient
ix. Automate operation on measuring device
x. Capable of generation on demand from user's physical environment.

The last criterion requires the keys to be generated on demand from user's physical environment or somehow related to the user. In a way, the apparatus must be mobile which can be carried around without much hassle. At the same time, it must be robust enough to produce random bits consistently under various practical climates.

USER'S PHYSICAL ENVIRONMENT

There are several studies have been done on capturing random numbers on demand from user's physical environment which satisfies the last criterion. Previously, the first author has studied on raw white room noise a source of cryptographic keys (Abu and Muslim, 2008). In the study, the key shall be generated based on user's surrounding image. The key is not really coming from the user visible image itself rather it is coming the surrounding user's ambience shall be taken into consideration.

In this underlying study, each moment in life is considered unique in itself. Based on earlier study, each random key is unique for the given moment generated by the user whenever he or she needs the random bits in a practical cryptographic application (Abu and Muslim, 2007).

DIGITAL IMAGES AS SOURCE OF RANDOM INPUT

In this research study, the randomness of air ambience is tested to its full potential. The objective is to generate one megabit per snapshot. A typical 14-bit medium range digital camera has been selected as the source of random input. The digital camera usually comes with its own software package driver. It is crucial the compression mode to be switch off or disable for this particular use. Each pixel should contribute only one bit, namely, the least significant bit. The idea is to capture the air ambience of the moment in life as a unique source of random ambience. The least significant bit of the center square of a digital image is then converted into a binary sequence via circular reading (Lang *et. al.*, 2009).

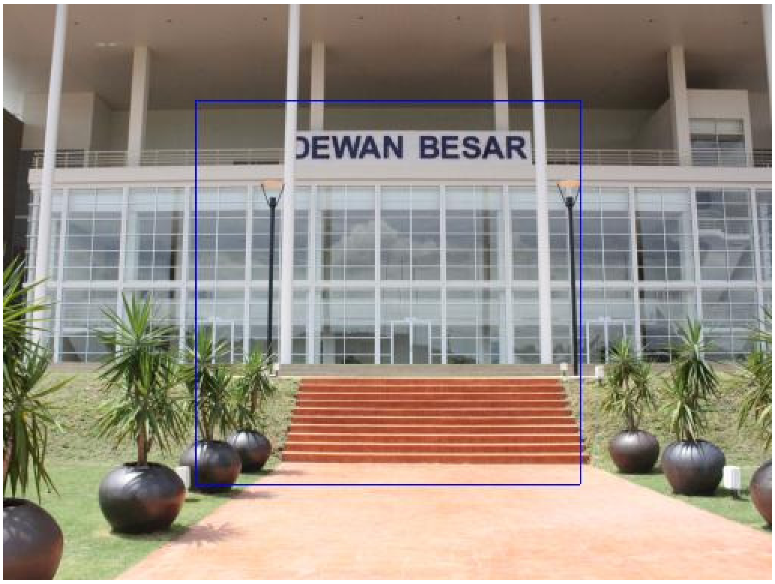

Figure 1: A sample of 14-bit digital image

In Figure 1, a sample of digital image is taken. In Figure 2, the least significant bit of the 1024 by 1024 bits in center image appears to be noise bits. The binary sequence intended cryptographic keys shall then be tested for their randomness using selected random tests suitable for large number.

Each image capture consists of 3 layers. The exclusive-OR of the 3 RGB layers has ben chosen to get the unbiased random bit sequence. This approach is simply the result of successively XORing at least three bit streams. This is considered a trivial mixing function. This simple technique is useful and efficient particularly since the camera captures random bits in parallel.

In general, this XOR corrector represents a simple linear function, which applies an exclusive-or operation on blocks of *n* bits in order to generate one output bit. It can dramatically reduce the bias on the generator output at the cost of reducing *n*-times its bit-rate. However, the bias of the output bit-stream is reduced only if the original bits are independent. The main advantages of the XOR corrector are its simplicity and the possibility to maintain a constant output bit-rate (Fischer *et. al.*, 2009).

Use of stronger mixing functions to extract more of the randomness in a stream of skewed bits is examined in Section 5.2 of (Eastlake *et al*, 2005). Further reading should refer to (Naslund and Russell, 2000) on this issue. This research study shall make use of this particular technique since it is efficient and convenience to use.

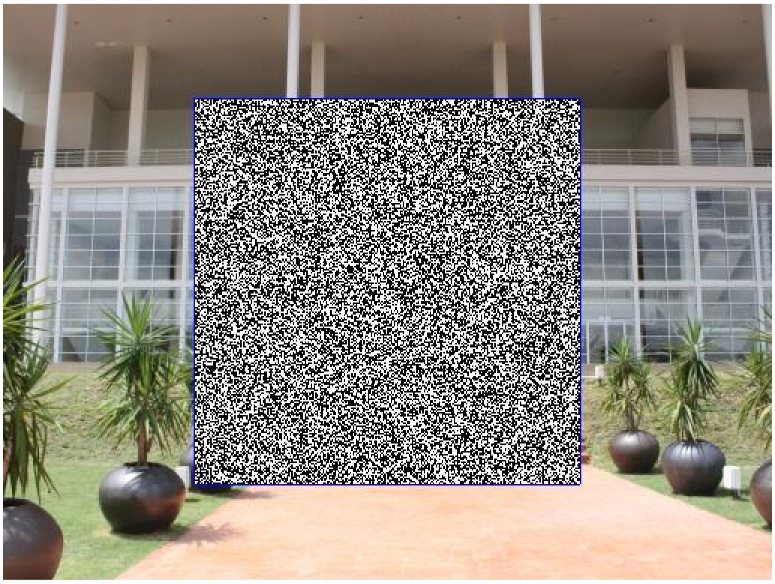

Figure 2: The random ambience is visualized as a plain noise on the center image

ONE MEGABIT RANDOM STATISTICAL TESTS

This study requires dedicated random statistical tests for large input. The candidate of random key here is a block of 2^{20} = 1,048,576 bits. A statistical test is formulated to test a specific null hypothesis (H_0). For the purpose of this study, the null hypothesis under test is that the binary sequence $\boldsymbol{\varepsilon}$ being tested is random against the alternative hypothesis (H_1) for which the binary sequence $\boldsymbol{\varepsilon}$ is not random.

For each statistical test, a set of *p*-values (corresponding to the set of sequences) is produced. For a fixed significance level, a certain percentage of *p*-values are expected to indicate failure. For example, if the significance level is chosen to be 0.01 (i.e., $\alpha = 0.01$), then about 1% of the sequences are expected to fail. A sequence passes a statistical test whenever the *p*-value $\geq \alpha$ and fails otherwise. The parameter α denotes the significance level that determines the critical region of acceptance and rejection. Even though NIST recommends that α be in the range [0.001, 0.01], for this large cryptographic keys it is still practical to use smaller significant level α such as 0.001.

Only 8 tests are particularly suitable for large cryptographic keys size. The selected random tests (Rukhin *et. al.,* 2001) for one-megabit key are listed in the Table 2 below.

TABLE 2: The list of suitable random tests for large 1M-bit numbers

0	Statistical Test	Min practical key size
1	Frequency Monobit Test	128 bits
2	Binary Matrix Rank Test	38,912 bits
3	Non-overlapping Template Matching Test	1,000,000 bits
4	Overlapping Template Matching Test	1,000,000 bits
5	Maurer's Universal Test	387,840 bits
6	Excursions Test	1,000,000 bits
7	Excursions Variant Test	1,000,000 bits
8	Linear Complexity Test	1,000,000 bits

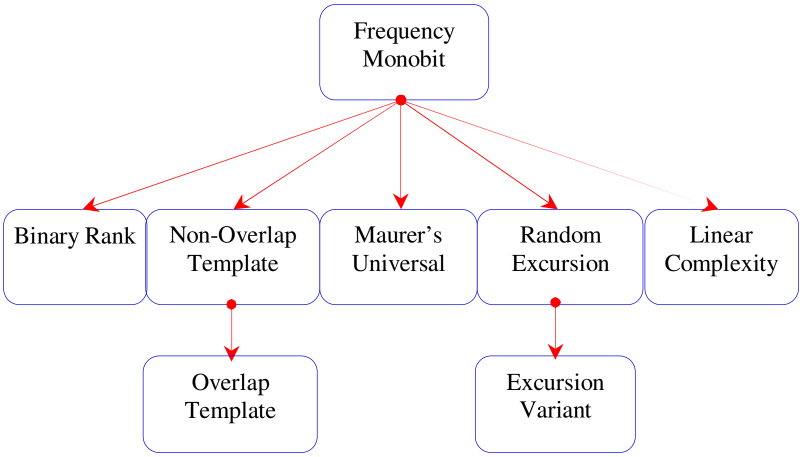

Figure 3: The hierarchy of random tests: Failing higher test will ensure the failure of the lower test.

The 8 selected tests including the first Frequency Monobit Test are basically relies heavily on the randomness of the binary sequence. Figure 3 shows the hierarchy of the tests. Once a particular block key set fails one test it is considered non-random and will certainly fail the next test in lower hierarchy.

Following similar general statistical hypothesis testing procedure, basic 8 steps shall be utilized. In the statistical hypothesis testing for randomness, the step-by-step procedure shall be as follows:

Step 0: Define the random variables in hypothesis testing.
Step 1: Set the null hypothesis H_0 : $\boldsymbol{\varepsilon}$ is random bit sequence versus an alternative hypothesis H_1: $\boldsymbol{\varepsilon}$ is not random bit sequence.
Step 2: Set the critical significant level α.
Step 3: Identify the statistical distribution to refer to.
Step 4: Define critical region as the rejection rule.
Step 5: Compute the test statistic and *p*-value.
Step 6: Decide on the sample case.
Step 7: Draw the conclusion for the test.

All 8 random tests has been encoded using 32-bit processing style in order to achieve real time performance.

TABLE 3: The set of random tests on one megabit key

index	Statistical Test
1	Frequency Monobit Test
2	Binary Matrix Rank Test
3	Non-overlapping Template Matching Test
4	Overlapping Template Matching Test
5	Maurer's Universal Test
6	Excursions Test for positive states $x = +1, +2, +3, +4$
7	Excursions Test for negative states $x = -1, -2, -3, -4$
8	Excursions Variant Test for positive state $x = +1, +2, \ldots, +9$
9	Excursions Variant Test for negative state $x = -1, -2, \ldots , -9$
10	Linear Complexity Test

EXPERIMENTAL RESULTS

In this research study, a sample image has been captured using Cannon EOS D50. Preferably, the least significant bits are converted into one-dimensional bits via circular reading instead of popular zigzag scan. The circular scan should start from the center of the image as proposed in (Lang *et. al.*, 2009). The one megabit binary sequence has been tested using the 8 NIST Statistical Tests suitable only for long binary sequence. The *p*-value indices are shown in Table 3 for easy reference.

The result shown in the Table 4 below as the *p*-values of the tests is larger than required $\alpha = 0.01$ in order to reject the null hypothesis as random sequence. The *p*-values on 2^{20}-bit of the least significant bit 14-bit image in Figure 4 after going through circular reading passes all 8 NIST Statistical Tests.

TABLE 4: The 32 *p*-values of 8 one-megabit random tests on 2^{20}-bit coming from center image after going through circular reading

Test #	*p*-value	Test #	*p*-value
1	0.517966	8($x = +4$)	0.462637
2	0.856962	8($x = +5$)	0.435930
3	0.187608	8($x = +6$)	0.485934
4	0.988402	8($x = +7$)	0.398170
5	0.516056	8($x = +8$)	0.246257
6($x = +1$)	0.433662	8($x = +9$)	0.207266
6($x = +2$)	0.965296	8(AveP)	0.537566
6($x = +3$)	0.927847	9($x = -1$)	0.599426
6($x = +4$)	0.690607	9($x = -2$)	0.879487
6(AveP)	0.754353	9($x = -3$)	0.689665
7($x = -1$)	0.849145	9($x = -4$)	0.928818
7($x = -2$)	0.811385	9($x = -5$)	0.467494
7($x = -3$)	0.850378	9($x = -6$)	0.423870
7($x = -4$)	0.930865	9($x = -7$)	0.457528
7(AveN)	0.860443	9($x = -8$)	0.568967
8($x = +1$)	0.854148	9($x = -9$)	0.637407
8($x = +2$)	0.951640	9(AveN)	0.628073
8($x = +3$)	0.796116	10	0.564300

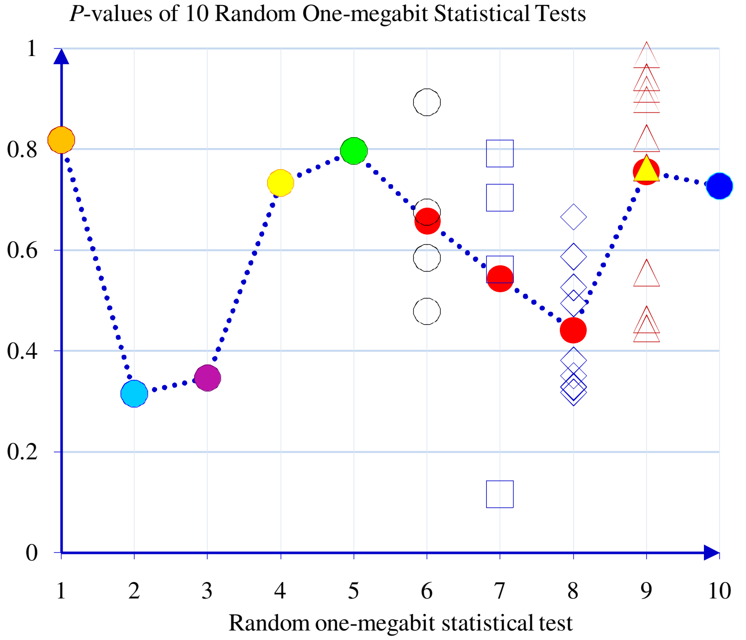

Figure 4: The 32 *p*-values are plotted with their average values

In certain practical cryptographic application it is crucial to produce large cryptographic key live on demand (Abu and Sahib, 2010). The results show that 14-bit high fidelity digital camera is capable of generating one megabit random key based on ambience principle live on demand.

CONCLUSION

The aim of this paper is to show the usage of air ambience for generating true random megabits. Due to difficulty of proving unpredictability in a theoretical way, the proposed true random bit is subjected to set of statistical tests. Due to the air ambience, the collections of least significant bits in the digital image have been shown to easily pass selected statistical tests.

Current digital camera is capable of capturing high quality image. Thus, image from user's environment can be a good source of random cryptographic key. In this paper, the least significant bit of the ambience digital image has been statistically proven to be random. In certain practical cryptographic application it is crucial to produce large cryptographic key live on demand. This technique provides a way to generate large random key within short period of time.

This research study has been designed to support practical cryptographic operations. A step-by-step procedure has been developed to produce large 2^{20} bit keys. Selected NIST random tests have been used to check on the binary sequence. The one megabit cryptographic key has been generated per digital image and tested for randomness. All 32 *p*-values are larger than $\alpha = 0.01$. Air ambience is a good source for generating true random megabits.

REFERENCES

Barker, Elaine and Kelsey, John. 2007. Recommendation for Random Number Generation Using Deterministic Random Bit Generators, *NIST Special Publication* 800-90 (Revised), March 2007.

Eastlake, 3rd D., Schiller, J. and Crocker, S. 2005. Randomness Requirements for Security, *BCP 106, RFC 4086*, June 2005.

Drake, Matt J. 2009. How to Use Random Numbers in Your iPhone App, *Learn How to develop an iPhone Apps*, 6 May 2009.

Tsoulfanidis, Nicholas and Landsberger, Sheldon. 2010. *Measurement and Detection of Radiation*, 3rd ed., CRC Press.

Gottfried, Kurt and Yan, Tung-mow. 2003. *Quantum Mechanics: Fundamentals*, 2nd ed., Springer Science and Business.

Garrison, John C. and Chiao, Raymond Y. 2008. *Quantum Optics*, Oxford University Press.

Gibbon, Paul. 2005. *Short Pulse Laser Interactions with Matter: An Introduction*, Imperial College Press.

Diels, Jean-Claude and Rudolph, Wolfgang. 2006. *Ultrashort Laser Pulse Phenomena: Fundamentals, Techniques and Applications on A Femtosecond Time Scale*, 2nd ed., Academic Press.

Sone, Yoshio. 2007. *Molecular Gas Dynamics: Theory, Techniques and Applications,* A Birkhäuser Boston.

Ott, Edward. 2002. *Chaos in Dynamical Systems*, 2nd ed., Cambridge University Press.

Gleick, James. 2008. *Chaos: Making a New Science*, 20th ed., Penguin.

Gupta, Madhu S. 1997. Electrical Noise Fundamentals and Sources, IEEE Press, June 1977.

Davenport, Wilbur B. Jr. and Root, William L. 1987. *An Introduction to the Theory of Random Signals and Noise*, Wiley-IEEE Press.

Ritter, Terry. 2004. Random Electrical Noise: A Literature Survey, Last updated: 14 January 2004, *http://www.ciphersbyritter.com/ RES/NOISE.HTM*

Serway, Raymond A. and Jewett, John W. 2003. *Principles of Physics*, 3rd Edition, Brooks Cole: 564–569.

Nur Azman Abu and Zulkiflee Muslim. 2008. Random Room Noise for Cryptographic Key, *Proceedings 2nd IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2008),* Phitsanulok, Thailand, 27–29 February 2008: 381–387.

Nur Azman Abu and Zulkiflee Muslim. 2007. Random Number Generation for Cryptographic Key, *Proceedings International Conference on Engineering and ICT(ICEI2007)*, Melaka, Malaysia, 27–28 November 2007, **1**: 255–260.

Wong Siaw Lang, Nur Azman Abu and Shahrin Sahib. 2009. Cryptographic Key from Webcam Image, *International Journal Cryptology Research*, **1**(1): 115–127.

Fischer, V., Aubert, A., Bernard, F., Valtchanov, B., Danger, J.-L. and Bochard, N. 2009. *True Random Number Generators in Configurable Logic Devices*, Projet Report ANR- ICTeR 05- BLAN-0110-01, 12 February 2009.

Eastlake, D. 3rd, Schiller, J. and Crocker, S. 2005. Randomness Requirements for Security, *BCP 106, RFC 4086*, June 2005.

Naslund, M. and Russell, A. 2000. Extraction of Optimally Unbiased Bits from a Biased Source, *IEEE Transactions on Information Theory*, **46**(3), May 2000.

Andrew Rukhin, Juan Soto, James Nechvatal, Miles Smid, Elaine Barker, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks, Alan Heckert, James Dray, and San Vo. 2001. *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*, NIST Special Publication 800-22, 15 May 2001.

Nur Azman Abu and Shahrin Sahib. 2010. Random Ambience Key Generation Live on Demand, *Proceedings 2nd International Conference on Signal Processing Systems (ICSPS 2010)*, **1**: 110-114.