A PRACTICAL PROPOSAL FOR ENSURING THE PROVENANCE OF HARDWARE DEVICES AND THEIR SAFE OPERATION

Y. Kovalchuk, W.G.J. Howells†, H. Hu*, D. Gu*, K.D. McDonald-Maier**

**School of Computer Science and Electronic Engineering, University of Essex, UK,*
yvkova@essex.ac.uk, hhu@essex.ac.uk, dgu@essex.ac.uk, kdm@essex.ac.uk
† School of Engineering and Digital Arts, University of Kent, W.G.J.Howells@kent.ac.uk

Keywords: ICmetrics, security, encryption, embedded systems.

Abstract

This paper presents a novel technique, termed ICmetrics (Integrated Circuit metrics), that can be used for the purposes of generating encryption keys, electronic signatures, detecting attempts of frauds, or preventing malfunction of hardware components and systems. The ICmetrics technology is based on employment of measurable features derived from characteristics of a given electronic device in order to generate an identifier that uniquely determines or describes the device. Any changes in the identifier during consequent device's operation would signal about a possible safety or security breach within the electronic system. After a detailed overview of the ICmetrics technology and comparing it to the alternative techniques commonly used for securing electronic systems, the paper discusses challenges of developing the technology and brings an example to demonstrate how these issues are being addressed.

1 Introduction

Digital devices penetrate every facet of our daily life nowadays. As we rely, and in many cases are dependable, on electronic systems and hardware devices, it is essential to ensure their safe operation. In addition, the issue of security arises as we create, store, share, and manage information in digital format.

This paper is concerned with both safety and security of exploiting electronic devices despite their design or operating conditions (e.g., this could be mobile devices used by consumers, robots assisting humans or used in manufacturing, or distributed system sensing an environment, etc.). More specifically, the paper proposes a novel technique, termed ICmetrics (Integrated Circuit metrics), that can be used for the purposes of generating encryption keys, electronic signatures, detecting attempts of frauds, or preventing malfunction of hardware components and systems.

Essentially, the ICmetrics technology is based on employment of measurable features derived from characteristics of a given electronic device in order to generate an identifier that uniquely determines or describes the device. This identifier can be subsequently used for generating a unique encryption key or signature to protect the device or messages sent from/to it. Moreover, any misuse of the device would result in change of the identifier (and therefore the encryption key associated with it), signalling thus a safety or security breach. For instance, ICmetrics may prevent unauthorised access to embedded and distributed devices that are connected wirelessly; it could prevent the fraudulent cloning or imitation of a device in order to compromise its identity and subsequent communication. It can allow for implicit detection of tampering of the software or hardware associated with the device via the inclusion of spyware or similar virus software.

In the following sections, we explain how ICmetrics works, compare it to other popular techniques used for securing electronic systems, and discuss the advantages of the proposed technology. We then outline the challenges that are associated with developing ICmetrics and bring an example from the research we are currently undertaking to address these issues. In particular, we are looking at the problem of finding suitable ICmetrics features, both from the point of view of obtaining data and the properties which such features should possess so as unique and stable identifiers can be generated. The final section concludes the paper and highlights the major findings from our current research.

2 The ICmetrics technology

The ICmetrics technology is based on the idea that electronic devices often function under unique conditions; they sense different environments, run different software, perform different tasks and interact with different users. Various features can be extracted from digital devices' operation that may be integrated together to generate unique identifiers for each of the devices or create unique profiles that describe the devices' actual behaviour.

In a sense, the ICmetrics technology can be seen analogous to a biometrics based system, but employing devices' features instead of those intrinsic to humans. While biometric features can be extracted from analysis of human characteristics, such as iris, fingerprints or voice, potential sources for ICmetrics

features include programmable structures, circuits, sensors, communication peripherals, etc. Our preliminary investigation has shown that ICmetrics features may be extracted from, for example, program sequences, contents of selected memory locations, or access frequency and the system's input/output with its environment. Such features can be used for both, generating unique encryption keys or digital signatures (e.g. to encrypt message sent from or stored on the device), and detecting cases of untypical device's behaviour (e.g., due to failure of its electronic elements or intrusion of a parasite program code).

2.1 Comparison to alternative techniques

There have been many techniques developed for the purposes of securing electronic systems, both on the level of hardware (e.g. the Physical Unclonable Functions (PUF) technologies [14] and in particular, Hardware Intrinsic Security [5]) and the user (e.g., biometrics, encryption, passwords etc.). ICmetrics can be seen as a hybrid approach that exploits features derived based on the interaction of the hardware with their users and/or environment.

While PUF-based techniques may be successful in providing secure key storage mechanisms or preventing cloning of hardware devices [5], they are not able to address attacks on the software level. Any unauthorised changes to program(s) running on hardware devices (e.g., inclusion of spyware or similar virus software) may undermine the security and safety of the whole electronic system.

Passwords or encryption keys can be used to control access to software programs and information stored on hardware devices; however both can be forgotten or stolen. Similarly, the major weakness of the many existing biometric-based systems is that they rely on the explicit template storage [7]. Although some work has also been done on direct encryption of keys [1,2,8,11-13,16], these proposals apply to a restricted problem domain and do not successfully overcome the problems associated with intra-sample variation in the generic case. In addition, these authorisation techniques cannot detect improper operation of digital devices.

The ICmetrics technology addresses the above issues and can ensure both security and safety of electronic systems. In summary, the technology provides the following advantages over other approaches used for securing electronic systems:

- The removal of the need to store any form of template for validating devices. Unique identifiers for the devices are generated in real time based on the devices' current operation and predetermined feature values distributions.
- There is no back door. The security of a system will be as strong as the ICmetrics associated with it and the encryption algorithm employed. The only mechanisms to gain subsequent access are to provide another sample of the ICmetrics or to break the cipher employed by the encryption technology.
- The compromise of a system does not release sensitive ICmetrics template data which would allow unauthorised access to other systems protected by the same ICmetrics or indeed any system protected by any other ICmetrics templates present.
- The removal of the need for the storage of the private key associated with the encryption system. This is a natural consequence of the system since the key will be uniquely associated with the given ICmetrics sample and a further ICmetrics sample may be used to regenerate the required private key. As there is no physical record of the key, it is not possible to compromise the security of sensitive data via unauthorised access to the key.
- Tampering with the constitution of a hardware device will cause its behaviour to change, potentially causing the features underlying the ICmetrics to change, perhaps dramatically, thus causing the generated ICmetrics to change. Consequently, a faulty or maliciously tampered device will be autonomously prevented from decrypting its own stored data or participating in any initiated secure communications, as the regenerated keys will differ from those created before its integrity was compromised. In other words, the ICmetrics approach can be made to fail securely and safely; it also provides a very high immunity from cloning and tampering.

2.2 Current development of ICmterics

ICmetrics is a novel concept. At this early stage of our research, we explore the possibility of applying the ICmetrics technology for the purposes of generating stable encryption keys based on the features derived from embedded systems' operation (see section 4 for details).

In its current state, the ICmetrics system is designed to be employed on previously unseen devices and to faithfully reproduce encryption keys for such devices on further application to them by examining a pre-defined set of measurable features of such devices. However, the system does require detailed knowledge of the likely distribution of such features within their domain of possible values for typical devices. Therefore, a significant calibration phase is required for each application domain for which the ICmetrics system is to be used prior to its employment in the generation of encryption keys. This calibration phase operates on samples taken from typical example devices which may or may not include devices for which encryption keys will subsequently be required. Although the system is, subsequent to the calibration phase, designed to operate on previously unseen devices, this is governed by the restriction that the measured features will behave approximately as predicted by analysis of the sample devices.

ICmetrics is therefore a two phase system: (1) *Calibration* phase is applied only once (or every time the operational conditions of the given device(s) change); (2) O*peration* phase is applied each time an encryption key is desired for the given device.

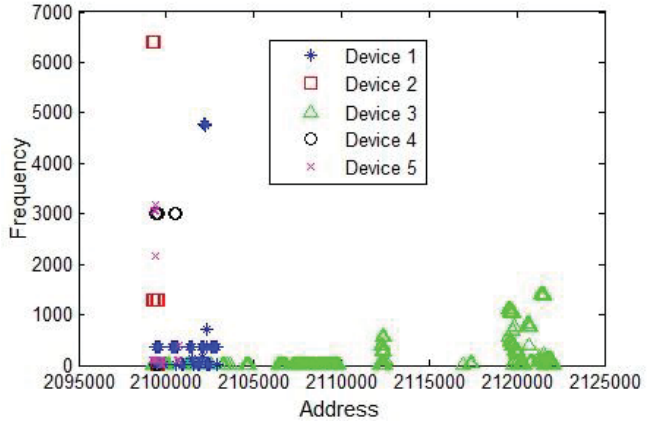

Figure 1: Original distributions of program counter values.

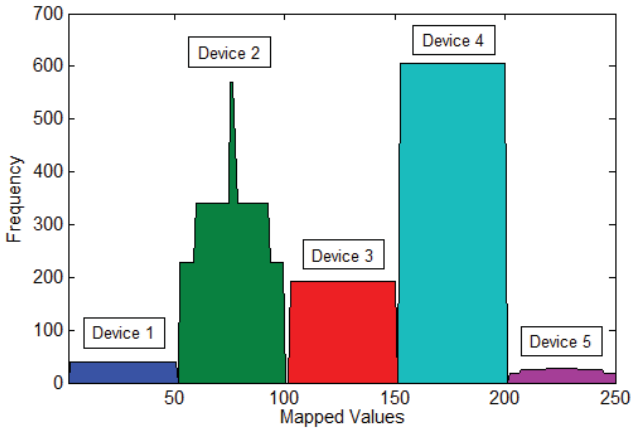

Figure 2: Mapped distributions of program counter values.

The calibration phase consists of the following three steps:

1. Record a set of desired measurements (features) associated with the devices in the given set. For example, the program counter values of a processor core can be logged during execution of various software programs on hardware devices and considered as ICmetrics features.
2. Generate feature distributions for each feature tabulating the frequency of each occurrence of each discrete value within the given value scale for each device. Figure 1 shows an example of the distributions of the program counter (address) values recorded for five different devices (see section 4.1 for configuration details). It may be noted from Figure 1, that there is a number of tightly grouped address values for each device, which makes the task of separating devices in the feature space difficult.
3. Normalise the feature distributions generating normalisation maps for each feature. The purpose of this step is to bring unusual distributions of measured features (as can be seen from Figure 1) to their normal (Gaussian) form so as values suitable for key generation could be defined. As an example, Figure 2 represents feature distributions after applying the normalisation procedure over a limited set (identical for all devices) of original feature samples. Full details on the normalisation procedure can be found in our earlier work [10].

The operational phase includes three steps:

1. Measure features (e.g. the program counter values) for the device for which an encryption key is desired.
2. Apply the normalisation maps to generate values suitable for key generation.
3. Apply the key generation algorithm.

This means that subsequently to the normalization step, component feature values for a device must be combined in such a way to form a unique number so as to identify each device. This number will form the basis for the subsequent derivation of the key required for the actual encryption process.

3 ICmetrics challenges

While the ICmetrics technology may overcome many issues that can be found in traditional security approaches, it presents many research challenges to be addressed in order to build an effective security and safety infrastructure. Among these questions are the following:

1. *How to obtain the feature measurements in real time so as the current system performance is not impacted or this impact is minimal?*

Securely collecting and measuring features drawn from the performance of electronic devices is a major challenge. Appropriate hardware and software instrumentation is required in order to gain access to device's features and properties during its performance. Such interfaces should allow for non-intrusive way of measuring and recording of feature values that does not change the system's operation which is essential for all real-time applications.

2. *Which characteristics of devices' behaviour could provide suitable ICmetrics features?*

The specific characteristics that can be used in ICmetrics may take the form of internal signal distributions or metrics derived from the highly changing signals. Real examples investigated so far include: the address and value from the data transactions of a processor, its program counter, and metrics for the effectiveness of the program and data caches derived from performance counters. The key observation made is that any modifications taking place to either the software executing on a given hardware configuration (in the form of the addition of Spyware or similar) or to the hardware (by tampering with its available memory, configuration or external sensors) manifest themselves in the form of variations in the measurable characteristics. This would have the significant effect of modifying any value derived from the data used for encryption or validation key production by the system, prohibiting its continued participation in any further secure communications or terminating its safe operation. Thus, these are all good examples of features to be considered for generating encryption keys. In general however, various application domains may require different features depending on the context and environment of their operation. It is important to determine the most suitable ones that would allow for generating a stable and secure encryption key for each particular case.

3. *Which techniques can be used for analysing multimodal features that possess non-standard distributions and are different in nature?*

In the context of ICmetrics, analysis of the feature distributions associated with the embedded circuit features presents some novel challenges as compared to many traditional pattern recognition tasks. This is because many of the observed features possess highly non-standard multi-modal distributions (see Figures 1 for example), often a product of the device under investigation operating in a number of distinct states. The problem of incorporating pattern features with unusual distributions is well known within pattern recognition problems, even if not easily addressed. The problem is, however, more acute when features are derived from characteristics of electronic devices, and appropriate techniques should be developed. This includes finding right techniques for classification and normalisation of data in multi-dimensional space.

In our previous work [10], the target space was linear in nature. We used enhanced Peak-Trough detection [6,15] and kernel estimation algorithms [3] to determine the various modal clusters taking one feature at a time. Our current research, however, is focused on investigation of multi-dimensional spaces combining various features where each circuit mode is equi-distant from every other. This would allow the system to be applied to devices which have not formed part of the calibration sample within any enrolment of known samples from the devices. Such a generalization provides an improved mapping onto the key generation space and allows the multi-modal nature of the feature distributions to be effectively integrated within the overall system. Considering multiple features that are different in nature has also another advantage of designing hybrid ICmetrics that can include features derived not only from hardware characteristics, but also from signals employed in human-computer interfacing (e.g. voice, gestures, brain signals, etc.).

4. *Which algorithms should be employed for integrating ICmetrics feature values in order to achieve generation of stable encryption keys?*

The generation of encryption keys requires developing suitable methods for combining selected features so as to produce a unique basis number – an initial number unique to the electronic system from which actual encryption keys may be derived. The main requirement for such methods is that they should allow for generating basis numbers with low intra-sample variance (that is, the values produced for the same device) but high inter-sample variance (that is, the values produced for different devices) with the ideal case being no inter-sample overlap of potential basis numbers.

5. *How to build an evaluation and calibration platform?*

The stability of encryption keys generated by the ICmetrics system depends on several factors as discussed above, including: (i) number of devices available for training and the environment of their operation; (ii) features employed; (iii) mathematical algorithms and their settings used for feature processing (e.g., clustering, normalization, and de-correlation techniques); (iv) methods applied for combining selected features to generate a basis number. Therefore, it is important to ensure the evaluation platform is in place to allow for controlled experiments to test and compare performance of various algorithms for feature extraction and processing, both in isolation and their various combinations. It may well appear that certain algorithms work better for certain types of devices, and there should be a way to find right techniques for each case quickly and reliably.

In addition to experimenting with various algorithms during the calibration phase, it is equally important to evaluate the system's performance during its operational phase. In other words, once the system is trained on test devices, its ability to generate stable encryption keys for unseen devices of similar type has to be estimated. Furthermore, taking into account that the conditions of operation of the same electronic device may change (e.g., a computer running a different software), the evaluation and calibration platform should also include interfaces for fast recalibration of the ICmetrics system. The interfaces should be convenient to use by non-experts to allow for the system to be deployed outside the laboratory.

4 Example of identifying suitable ICmetrics features

Among the challenges outlined in the previous sections, we focus our current research on finding suitable ICmetrics features and methods for obtaining them [9]. ICmetrics features are used for generating unique identifiers (basis numbers) for each of the electronic devices in the considered set based on circuits' metrics so as stable encryption keys could be further generated from these identifiers. The main requirement for the basis number is that it should remain constant for a given device on each attempt of its generation, but distinct from the basis numbers generated for other devices employed in the operational set. Furthermore, it should not be possible to derive or estimate the encryption keys generated for other devices based on the basis number of a given device. In order to achieve this, it is important to find such features associated with electronic devices, which allow for separation of the devices in the feature space.

In our previous work [9], we have explored the program counter of a processor core as a potential source for ICmetrics features. We have chosen this parameter since the set of the program counter distinct values is finite and is the same for a certain device (assuming the full program profile is taken), but is likely to vary from one device to another. In that study, we have found that separate values of the program counter do not always allow for unique identification of hardware devices. However, suitable ICmetrics features may be derived from frequencies of the program counter occurrences and their sequences observed during programs' execution flow.

Another alternative to use the program counter as the ICmetrics feature is to combine it with other measurements extracted from electronic devices' operation. In the example below, we investigate this option by adding values of the

status register and stack pointer traced during program execution flow to the correspondent series of the program counter values. Having the requirement for the ICmetrics features that they should allow for separation of the considered devices in the feature space, we test if this can be achieved by combining the proposed three measurements (features): the program counter (PC), status register (SR), and stack pointer (SP). We first describe our hardware and software platform for logging feature values and then discuss the analysis we performed over the logs.

4.1 Experimental platform

For this study, we have employed a low resource embedded system based around an ARM7 processor core, in particular an Atmel AT91SAM7S256 microcontroller [17] and 64Kbytes SRAM memory. We have used the combination of Eclipse [18], Open On-Chip Debugger (OOCD) [19], and JTAG programming port for programming the microcontroller, as well as tracing programs' execution.

We have used an intrusive single stepping tracing method to log feature values. While this method affects programs' execution times, it does not change the execution flow, meaning that the proposed method provide the same PC, SR, and SP values as would have been obtained with non-intrusive methods. To register feature values, our tracing program halts the CPU by issuing OOCD commands [19] via a telnet port and logs every single CPU instruction, meaning that the complete profile of the program execution is obtained.

Since at this early stage we have only been interested to see if the PC, SR, and SP could be used as the ICmetrics features, we have employed basic low complexity software routines to serve as a source of data so as to achieve visually representative and easily interpretable analysis results. More specifically, we have chosen several algorithms from the automotive package of the MiBench suite of benchmark algorithms [4] to design our programs representing five different devices, namely: angle conversion (D1); bit count (D2); cubic function (D3); and square roots (D5). In addition, we have included a program generating random numbers (D4).

4.2 Data analysis

Table 1 details a summary of statistics performed over the PC, SR and SP values logged during the considered devices' operation (see section 4.1). In particular, "total steps" indicates the total number of feature values recorded during the entire sessions of tracing. Since a program may use the same feature values several times during its execution flow, we have calculated how many distinct values of PC are present in the devices' profiles and compare this to the number of distinct combinations of the three features (PC, SR, and SP) recorded at each step of logging. Our aim is to compare the case of using one feature (e.g. PC) with the case when a combination of three features (PC, SR, and SP in this case) is employed.

Param.\Program	D1	D2	D3	D4	D5
total steps	263565	104971	205482	138011	102131
distinct pc	429	44	1695	52	94
distinct pc-sr-sp	1056	60	4404	103	169
unique pc	132	0	1376	7	16
unique pc-sr-sp	1056	55	4404	98	169

Table 1: Statistics of the devices' performance and metrics.

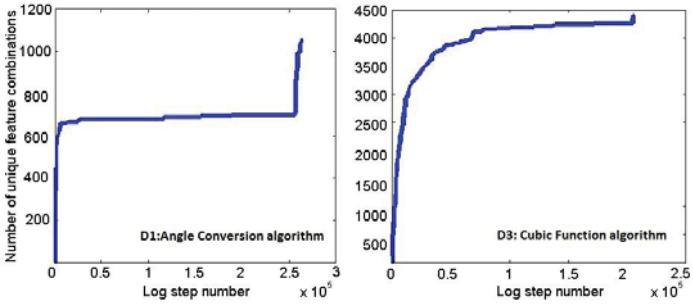

Figure 3: Information gain graphs for D1 – the device running the angle conversion algorithm (left) and D3 – the device running the cubic function algorithm (right).

For ICmetrics research, we are also interested in how the devices' profiles differ from each other. Therefore, we have further refined the number of distinct PC values, and the combinations of PC, SR, and SP values, by finding the number of values/combinations that occurred in the profile of a certain device, but not in the profiles of the rest of the devices. This is reflected in "unique pc", in case when only PC is considered as a feature, and "unique pc-sr-sp", in case three features are employed.

It can be noted from Table 1 that while taking only the PC as the ICmetrics feature does not allow to separate the five devices in the feature space (D2 has no unique PC values as compared to the rest of the devices), there exist combinations of the three features (PC, SR, and SP) that describe each of the devices uniquely. These unique combinations could be used for generating unique basis numbers for each device as discussed at the beginning of section 4. It is interesting to note that D2 and D4 share only five combinations of the PC-SR-SP values (60 distinct as opposed to 55 unique combinations for D2, and 103 distinct as opposed to 98 unique combinations for D4), and the remainder of the devices have these combinations all unique.

Another interesting problem to explore for the ICmetrics research is to find optimal logging times so as to achieve maximum information gain within a minimum time. From our experiments, we have found that this problem should be addressed for each problem domain separately, depending on the specification of the devices employed and their environment of operation. To illustrate this, Figure 3 provides graphs of information gain based on tracing of the three considered features (PC, SR, and SP) for two different devices, D1 (left) and D3 (right). The graphs show how many unique combinations of the three features (axis Y) are obtained

at each logging step (axis X). It can be noticed from Figure 3 that, depending on the complexity of the program running on a device, information gain graphs take different shape; it may therefore take different logging times to obtain the same number of unique combinations of ICmetrics features.

5 Conclusion

This paper has introduced the ICmetrics technology – a novel technique that can be used for the purposes of generating encryption keys, electronic signatures, detecting attempts of frauds, or preventing malfunction of hardware components and systems. The technology is based on employment of measurable features derived from characteristics of electronic devices in order to generate identifiers that uniquely determine the devices. ICmetrics can be seen analogous to a biometrics based system, but employing devices' features instead of those intrinsic to humans. The paper has provided the comparison of the proposed technology to alternative techniques widely used to secure electronic systems (e.g., PUFs, encryption, biometrics, etc.) and demonstrated the advantages of the ICmetrics system. We have also discussed the challenges associated with the implementation of ICmetrics and brought an example to demonstrate how some of these problems are being addressed in our current research. In particular, we have explored possible ICmetrics features that allow for unique identification of the devices in a given set. From our experimental results, we conclude that suitable ICmetrics features can be extracted from, for example, frequencies of the program counter occurrences and their sequences observed during programs' execution flow. Furthermore, the combination of several features, such as the program counter, status register and stack pointer, may be used to separate devices in multi-dimensional space.

Acknowledgements

The authors gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council under grant EP/K004638/1 and the EU Interreg IV A 2 Mers Seas Zeeën Cross-border Cooperation Programme – SYSIASS project: Autonomous and Intelligent Healthcare System (project's website http://www.sysiass.eu/).

References

[1] A. Bodo, "Method for Producing a Digital Signature with Aid of a Biometric Feature", *German patent DE 42 43 908 A1* (1994).

[2] J. Daugman, "Biometric Decision Landscapes", *Tech Report TR482, University of Cambridge Computer Laboratory, UK* (2000).

[3] G. H. Givens, "Consistency of the Local Kernel Density Estimator", *Department of Statistics, Colorado State University, Fort Collins, Colorado* (1994).

[4] M.R. Guthaus, J.S. Ringenberg, D. Ernst, T.M. Austin, T. Mudge, and R.B. Brown, "MiBench: A free, commercially representative embedded benchmark suite", *Proceedings of the International Workshop on Workload Characterization*, pp. 3-14 (2001).

[5] H. Handschuh, G.-J. Schrijen, and P. Tuyls, "Hardware Intrinsic Security from Physically Unclonable Functions", *Towards Hardware-Intinsic Security. Foundations and Practice,* (Eds.) A.-R. Sadeghi, D. Naccache, Springer, pp. 41-53 (2010).

[6] K. Harmer, G. Howells, W. Sheng, M. C. Fairhurst, and F. Deravi, "A Peak-Trough Detection Algorithm Based on Momentum", *International Congress on Image and Signal Processing. Sanya, Hainan, China,* (2008).

[7] S. Hoque, M. Fairhurst, G. Howells, and F. Deravi, "Feasibility of generating biometric encryption keys", *Electronics Letters*, **vol. 41**, pp. 309-311 (2005).

[8] P. K. Janbandhu, M.Y. Siyal, "Novel Biometric Digital Signatures for Internet-based Applications", *Information Management and Computer Security*, **Vol.9, No.5**, pp. 205-212 (2001).

[9] Y. Kovalchuk, W.G.J Howells, H. Hu, D. Gu, K.D. McDonald-Maier, "ICmetrics for Low Resource Embedded Systems", *the third International Conference on Emerging Security Technologies* (2012).

[10] E. Papoutsis, G. Howells, A.B.T. Hopkins, K.D. McDonald-Maier, "Integrating Multi-Modal Circuit Features within an Efficient Encryption System", *Journal of Information Assurance and Security*, **Vol. 2 No. 2**, pp. 117-126 (2007).

[11] W. Sheng, G. Howells, M. C. Fairhurst, F. Deravi, K. Harmer, "Consensus Fingerprint Matching with Genetically Optimised Approach" *Pattern Recognition,* **Vol. 42, No. 7**, pp. 1399-1407 (2009).

[12] W. Sheng, G. Howells, M. C. Fairhurst, F. Deravi, "Template-free Biometric Key Generation by means of Fuzzy Genetic Clustering", *IEEE Transactions on Information Forensics and Security*, **Vol. 3 No. 2**, pp. 183-191 (2008).

[13] W. Sheng, G. Howells, M. C. Fairhurst, F. Deravi, "A Memetic Fingerprint Matching Algorithm", *IEEE Transactions on Information Forensics and Security*, **Vol. 2, No 3, Part 1**, pp. 402-412 (2007).

[14] G. E. Suh, S. Devadas, "Physical unclonable functions for device authentication and secret key generation", *In Proceedings of the 44th Design Automation Conference, IEEE*, pp. 9-14 (2007).

[15] B. S. Todd, "An algorithm for the detection of peaks and troughs in physiological signals", *Oxford University Computing Laboratory Programming research Group*, (1997).

[16] Y. Yamazaki, N. Komatsu, "A Secure Communication System Using Biometric Identity Verification", *IEICE Transaction on Inf. and Syst.,* **E84-D (7)**, pp. 879-884 (2001).

[17] Atmel's SAM7S datasheet: http://www.atmel.com/Images/doc6175.pdf

[18] Eclipse official web-site: http://www.eclipse.org/

[19] Online OpenOCD User's Guide: http://openocd.sourceforge.net/doc/html/index.html#Top