Randomness of Physiological Signals in Generation Cryptographic Key for Secure Communication Between Implantable Medical Devices Inside The Body And The Outside World

Hassan Chizari*, Emil Lupu†, Paula Thomas*

July 11, 2018

Abstract

A physiological signal must have a certain level of randomness inside it to be a good source of randomness for generating cryptographic key. Dependency to the history is one of the measures to examine the strength of a randomness source. In dependency to the history, the adversary has infinite access to the history of generated random bits from the source and wants to predict the next random number based on that. Although many physiological signals have been proposed in literature as good source of randomness, no dependency to history analysis has been carried out to examine this fact. In this paper, using a large dataset of physiological signals collected from PhysioNet, the dependency to history of Interpuls Interval (IPI), QRS Complex, and EEG signals (including Alpha, Beta, Delta, Gamma and Theta waves) were examined. The results showed that despite the general assumption that the physiological signals are random, all of them are weak sources of randomness with high dependency to their history. Among them, Alpha wave of EEG signal shows a much better randomness and is a good candidate for post-processing and randomness extraction algorithm.

1 Introduction

The problem of secure communication between Implantable Medical Devices (IMDs) inside the body and the outside world is one of the emerging areas of research in cybersecurity of IoT devices. There are many real-world case studies on vulnerabilities of IMDs to attacks, as they are usually using a single permanent symmetric key for communication with outside world. Hackable insulin pumps [1], brain neural implants [2] and pacemakers [3] are a few examples of the most recent attacks reported in the news and literature. As usage of public key cryptography is not applicable in IMDs due to very limited resources [4, 5], a series of solutions have been proposed in the science community to solve the

problem. Among these solutions, proximity based algorithms are the most successful one with this assumption that if a device gets very close to the body, it is safe to consider that we can trust the device [6]. In order to start the communication, the device outside the body and the IMD need to read a physiological signal of the body at the same time and generate the communication key from that. With this solution, there would be no need to have a permanent key saved inside the IMD nor to have a key exchange process. Although, this solution received high attention from scientific community, one problem is still remained unsolved and that is whether or not the physiological signal could provide strong cryptographic key.

A physiological signal must have a certain level of randomness inside it to be a good candidate for generating cryptographic key from that. Entropy of a source is usually a tool to examine the strength of that source to generate random numbers. In addition to Entropy, dependency to history is another tool to examine the strength of randomness which is usually missed in the literature. Dependency to the history is a measure where the adversary has infinite access to the history of generated random bits from the source and wants to predict the next random number based on that. Dependency is measured using Santha-Vazirani method. Although, many physiological signals have been claimed to be a good source of randomness in the literature, no extensive study till now examined the strength of those sources against dependency to history. In this paper, using a large dataset of physiological signals collected from PhysioNet, the dependency to history of Interpuls Interval, QRS Complex and EEG signals are examined.

2 Related Works

Several physiological signals have been proposed in the literatures to be used as the source of randomness for generating secret key such as Brain waves or electroencephalograms (EEG) [7], electrocardiogram (EKG) [8] and Photoplethysmogram (PPG) [9], Electrocardiography (ECG) [10] and InterPulse Intervals (IPI) [6]. IPI is the time difference between two peaks of an ECG signal. ECG signal has three peaks for every heart beat named as Q, R and S. So, three IPI values could be extracted from the time difference between each two corresponding peaks (Q-Q, R-R, S-S). Among these, R has the highest peak and the easiest one to detect. In the rest of this paper, whenever we refer to IPI, it is the R-R time difference. QRS is another physiological measure which is the time difference between the peak of Q signal and the peak of S signal. Since the peak of R signal is between these two, this physiological feature is called QRS Complex.

While there are only few works using EEG as the source of randomness (e.g. [7]), IPI is the most popular one. The idea of using IPI as a source of randomness has been proposed by [6], where a random extraction algorithm is needed to convert IPI value to a random number. Several randomness extractor algorithms from IPI have been proposed in the literature. Proposed methods are

including using XOR function [11], gray-coding [12] and using frequency domain [13, 14, 15]. In some studies a combinations of algorithms are used for randomness extraction. For instance, [16, 17, 18] used accumulation, modulo, contract mapping and gray-coding for the extractor and [19] proposed a combination of concatenation, quantization and gray-coding as the randomness extractor from IPI.

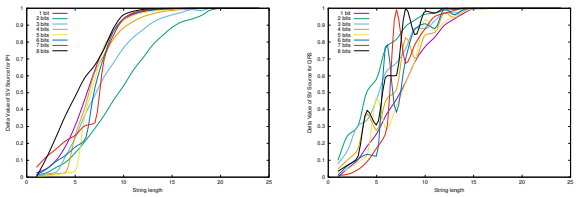
Whatever the extraction method is, it needs to be evaluated for the quality of generated randomness. There are several physiological randomness extraction methods which did not perform any randomness test to examine the quality of proposed algorithm (e.g. [20, 21, 13, 14, 22, 23, 15]). To evaluate the quality of a randomness extractor, there are two aspects to consider: the dataset and the methodology of evaluation. Till now for all proposed randomness extraction methods, these two aspects are somehow ignored. For instance, [12] evaluated the randomness property of the output by 5 minutes ECG data of 10 subject by NIST Statistical Test Suite (STS) [24] randomness test. [16, 17, 18] tested the algorithm using 5 minutes ECG signal of 40 subjects with NIST STS. In another work, [10] used histogram analysis on 1500 consecutive IPI values. [25] tested their proposed method with 100 subjects ECG data with Entropy test. [19], to evaluate the randomness of proposed algorithm, used Temporal Ratio [26] method over 5 minutes ECG data of 50 subjects. Nevertheless, none of aforementioned works, examined the dependency to the history of the proposed physiological signal for randomness extraction.

3 Data Collection

In this work, using PhysioNet [27], we have created a dataset of IPI, QRS and EEG signals. The dataset contains 202,569,491 QRS values, 895,621,566 IPI values and 597,931,520 EEG values. We used 8 bit coding for both IPI and QRS. For EEG signals, as proposed by [7], we extracted the five frequencies of brain (Alpha, Beta, Delta, Gamma and Theta) using the EEGLab developed by [28]. In order to extract the brain waves, EEGLab transfers the signal from time domain to frequency domain. Then, using Table , based on the level of the frequency of the wave they will be extracted from the main signal. Finally, the value of signal is multiplied by 1000 and we used the three least significant digits of it as a 10 bit coded number.

Table 1: Characteristics of EEG rhythms (from [29])

State	Unconscious		Conscious		
Rhythm	Delta	Theta	Alpha	Beta	Gamma
Frequency (Hz)	0.5-4	4-8	8-13	13-30	>30
Amplitude (μV)	20-200	10	20-200	5-10	5-10



(a) IPI Data (b) QRS Data

Figure 1: Dependency analysis for all ECG signals

4 Statistical Analysis

A randomness source is called Santha-Vazirani (SV-source) [30] where the outcome of last generated bit is related to the previous outcomes. In another word, the source is not independent. Let consider for source X and $\delta \in [0, 1]$, then we have:

$$\forall i \in n, \forall x_i \in \{0,1\} \rightarrow \\ \frac{1-\delta}{2} \leq Pr[X_i = x_i | \forall x_{i-1}] \leq \frac{1+\delta}{2} \tag{1}$$

δ is the bias for the new bit x_i, which it has some dependencies to the previous bits in the source $\{0,1\}^{i-1}$. In a simpler form we have:

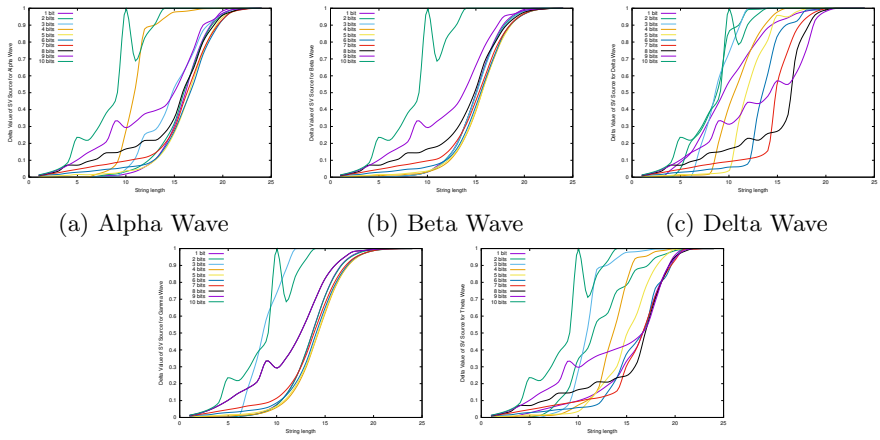
$$\forall x, y \in \{0,1\}^n \rightarrow \frac{Pr[X = x]}{Pr[Y = y]} \leq \frac{1+\delta}{1-\delta} \tag{2}$$

The best possible δ value is zero for any string length which demonstrates that the source is not Santha-Vazirani. If δ is equal to zero, the probability of having zero or one is always 0.5, no matter how much of the history of data is available. More importantly, it has been shown [31, 32] that even slightly biased SV-sources, (i.e. sources with low δ), are not suitable for many cryptographic purposes. To evaluate the predictiveness of source X from Eq. 2, we calculated the maximum ($Pr[X = x]$) and the minimum ($Pr[X = x]$) distribution value. Then, δ can be calculated as:

$$\delta = \frac{Pr[X = x] - Pr[X = x]}{Pr[X = x] + Pr[X = x]} \tag{3}$$

for $\forall x, y \in \{0,1\}^n$ where $n = 1..24$.

To examine the dependency to history, from each signal, we selected a series of bits. For instance, in QRS, firstly we selected the first least significant bit and applied the analysis on it. Then, we selected the first two bits of every QRS value and applied the analysis on it. This goes until the 8 bits of QRS, since the QRS has originally 8 bits. For IPI also, since it consists of 8 bits, we examined the dependency to history analysis on 8 series of data, starting from only one



(a) Alpha Wave (b) Beta Wave (c) Delta Wave

(d) Gamma Wave (e) Theta Wave

Figure 2: Dependency analysis for all EEG signals

least significant bit as the first series till the 8 least significant bit (or complete IPI) as the 8^{th} series. For all EEG waves, since in coding we used 10 bits, 10 series of values are extractable for each wave.

The next concept in this methodology is the string size. In order to measure the dependency to history, we must calculate the distribution histogram of the bits in each data series. This distribution could be based on the strings with size one ($s_1 = \{0, 1\}$). This histogram is simply counting the number of zeros and ones in the data series. For string size 2, we have $s_2 = \{00, 01, 10, 11\}$. In theory the analysis should be applied to all possible string size (s_∞), however, in practical, due to the limitations of database sizes, there is a limit to s.

5 Results and Discussion

We applied the dependency analysis for all featured physiological signals for string sizes up to 24 ($1 \leq s \leq 24$). Figures 1 and 2 show the dependency analysis results over EEG signals of Alpha, Beta, Delta, Gamma, Theta and ECG signals of IPI and QRS. As shown, despite as what has been advised in current works, IPI and QRS have the highest dependency to the history.

In QRS (Fig. 1b, the best combination of bits are when the 5 least significant bits (out of 8 bits) are selected. In this QRS series of Data up to string size of seven the SV-Delta value is less than 0.5 ($s = 7$ and $\delta = 0.41522364$). Moreover, for string sizes up to 18 the SV-Delta is less than one ($s = 18$ and $\delta = 0.999947422$). If SV-Delta is equal to one, it means that at one of the bars in the distribution histogram of the series is zero. In IPI series (Fig. 1a), less dependency to history has been observed compared to QRS when only the two least significant bits of IPI are selected as the source of randomness. In IPI, the

SV-Delta is less than 0.5 for string sizes up to 9 ($s = 9$ and $\delta = 0.460185251$). Moreover, no zero distribution has been found in IPI distribution until the string size of 21 ($s = 20$ and $\delta = 0.99923426$). Although IPI showed lower dependency to history compared to QRS, both of them suffers from this problem.

EEG signals in overall show better quality compared to ECG signals. Among EEG signals, Delta wave has the highest dependency to history (Fig. 2c). In Delta wave itself, the best combination of bits is when 9 least significant bits of Delta Wave value is selected. The value of SV-Delta, in Delta wave, is lower than 0.5 up to string size of 15 ($s = 15$ and $\delta = 0.47886351$). The SV-Delta is lower than one up to string size of 22 ($s = 22$ and $\delta = 0.99851962$). Next is Theta wave where the 7 least significant bits of its coding provide the least dependability to history for it. Alpha, Beta and Gamma showed more randomness compared to Delta and Theta. Interestingly, all these three waves belong to conscious operation of the brain, while Delta and Theta are waves of unconscious mind.

Among the conscious mind waves, Alpha provides the least dependency to the history. The best combination of bits to achieve this is the seven least significant bits. In Alpha wave, the SV-Delta value is less than 0.5 for string sizes up to 17 ($s = 17$ and $\delta = 0.450936539$). Moreover, the SV-Delta value is less than 1 for string sizes up to 23 ($s = 23$ and $\delta = 0.998452611$). Another interesting point regarding the Alpha wave is that for string size up to 15 the SV-Delta value is less than 0.287290776. The reason for this very big jump from 0.287 in string size of 16 to 0.754 in string size of 19 and afterwards could be interpreted as the sample size error. This is related to the problem of sample size.

The necessary sample size to measure the dependency to history is increasing by the 2 with the power of string size. In order to have 95% confident interval, considering than the sample size is uniform (best case scenario), for string size of $s = n$, the number values needed in the series is $1000 * 2^n$. For instance, in string size of $s = 24$, the sample size should consists of at least 16,777,216,000 values. But, as represented above, the number of samples for each physiological signal is much lower than this. A meaningful boundary for string size could be $s = 18$, where the sample size should consists of at least 262,144,000 values. Considering the maximum value of $s = 18$, still ECG physiological signals (IPI and QRS) show high dependency to history. Meanwhile, in EEG signals, Alpha Wave has the lowest dependency to history, but still even that small amount of dependency makes it not a very good candidate for randomness extraction.

6 Conclusion

A physiological signal must have a certain level of randomness inside it to be a good candidate for generating cryptographic key from that. A strong random source should have many features including independency to history. Dependency to the history is a measure where the adversary has infinite access to the history of generated random bits from the source and wants to predict the next random number based on that. In this paper, using a large dataset of

physiological signals collected from PhysioNet, the dependency to history of Interpuls Interval (IPI), QRS Complex, and EEG signals were examined. The results showed that despite the general assumption that the physiological signals are random, all of them are weak sources of randomness with high dependency to their history. Among them, Alpha wave of EEG signal shows a much better randomness and is a good candidate for post-processing and randomness extraction algorithm. The results of this paper have a great impact on the security of IMDs, where till now the quality of randomness of physiological signals considered to be high. With this result, a new set of researches is needed to investigate the randomness extraction algorithms from body physiological signals for secure communication between IMD inside the body and its outside world.

References

[1] Finkle, Jim. (2016, Oct.) J&J warns diabetic patients: Insulin pump vulnerable to hacking. [Online]. Available: http://www.reuters.com/article/us-johnson-johnson-cyber-insulin-pumps-e-idUSKCN12411L

[2] L. Pycroft, S. G. Boccard, S. L. F. Owen, J. F. Stein, J. J. Fitzgerald, A. L. Green, and T. Z. Aziz, "Brainjacking: Implant Security Issues in Invasive Neuromodulation," *World Neurosurgery*, vol. 92, no. C, pp. 454–462, Aug. 2016.

[3] A. Hern. (2017, Aug.) Hacking risk leads to recall of 500,000 pacemakers due to patient death fears.

[4] C. Camara, P. Peris-Lopez, and J. E. Tapiador, "Security and privacy issues in implantable medical devices: A comprehensive survey," *JOURNAL OF BIOMEDICAL INFORMATICS*, vol. 55, no. C, pp. 272–289, Jun. 2015.

[5] M. R. Kanjee and H. Liu, "Authentication and key relay in medical cyber-physical systems," *Security and Communication Networks*, vol. 9, no. 9, pp. 874–885, May 2014.

[6] C. Poon, Y. T. Zhang, and S. D. Bao, "A novel biometrics method to secure wireless body area sensor networks for telemedicine and M-health," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 73–81, Apr. 2006.

[7] G. Bajwa and R. Dantu, "Neurokey: Towards a new paradigm of cancelable biometrics-based key generation using electroencephalograms," *Computers & Security*, vol. 62, pp. 95–113, Sep. 2016.

[8] A. Ali and F. A. Khan, "An Improved EKG-Based Key Agreement Scheme for Body Area Networks." *ISA*, vol. 76 CCIS, no. Chapter 29, pp. 298–308, 2010.

[9] K. K. Venkatasubramanian, A. Banerjee, and S. K. S. Gupta, "Plethysmogram-based secure inter-sensor communication in body area

networks," in *Proceedings - IEEE Military Communications Conference MILCOM*, Arizona State University, Tempe, United States. IEEE, Dec. 2008, pp. 1–7.

[10] G. Zheng, G. Fang, M. A. Orgun, R. Shankaran, and E. Dutkiewicz, "Securing wireless medical implants using an ECG-based secret data sharing scheme." *ISCIT*, 2014.

[11] W. Wang, K. Hua, M. Hempel, D. Peng, H. Sharif, and H.-H. Chen, "A Stochastic Biometric Authentication Scheme Using Uniformed GMM in Wireless Body Area Sensor Networks," in *st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, South Dakota State University, Brookings, United States. IEEE, 2010, pp. 1620–1624.

[12] T. Hong, S.-D. Bao, Y.-T. Zhang, Y. Li, and P. Yang, "An improved scheme of IPI-based entity identifier generation for securing body sensor networks," in *2011 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 1519–1522.

[13] S. N. Ramli, R. Ahmad, and M. F. Abdollah, "Electrocardiogram (ECG) signals as biometrics in securing wireless body area network," in *2013 8th International Conference for Internet Technology and Secured Transactions, ICITST 2013*, Universiti Teknikal Malaysia Melaka, Ayer Keroh, Malaysia. IEEE, Jan. 2013, pp. 536–541.

[14] S. N. Ramli, R. Ahmad, M. F. Abdollah, and E. Dutkiewicz, "A biometric-based security for data authentication in Wireless Body Area Network (WBAN)," in *International Conference on Advanced Communication Technology, ICACT*. Universiti Teknikal Malaysia Melaka, Ayer Keroh, Malaysia, Apr. 2013, pp. 998–1001.

[15] K. Kalaivani, V. Anjalipriya, R. Sivakumar, and R. Srimeena, "An efficient Bio-key Management scheme for telemedicine applications," in *Proceedings - 2015 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development, TIAR 2015*, SRM Group Of Educational Institutions, Chennai, India. IEEE, Dec. 2015, pp. 122–126.

[16] S.-D. Bao, "A matching performance study on IPI-based entity identifiers for body sensor network security," in *2012 5th International Conference on Biomedical Engineering and Informatics, BMEI 2012*, Ningbo University of Technology, Ningbo, China. IEEE, Dec. 2012, pp. 808–811.

[17] F. Miao, S.-D. Bao, and Y. Li, "Physiological Signal Based Biometrics for Securing Body Sensor Network," in *New Trends and Developments in Biometrics*. InTech, Nov. 2012.

[18] S.-D. Bao, F. Miao, and Y. Li, "Biometric key distribution solution with energy distribution information of physiological signals for body sensor network security," *IET Information Security*, vol. 7, no. 2, pp. 87–96, Jun. 2013.

[19] D. K. Altop, A. Levi, and V. Tuzcu, "Towards using physiological signals as cryptographic keys in Body Area Networks," in *Proceedings of the 2015 9th International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth 2015*, Sabancl University, Tuzla, Turkey. ICST, Dec. 2015, pp. 92–99.

[20] K. Cho and D. H. Lee, "Biometric based secure communications without pre-deployed key for biosensor implanted in body sensor networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Korea University, Seoul, South Korea. Berlin, Heidelberg: Springer Berlin Heidelberg, Mar. 2012, pp. 203–218.

[21] C. Hu, X. Cheng, F. Zhang, D. Wu, X. Liao, and D. Chen, "OPFKA - Secure and Efficient Ordered-Physiological-Feature-Based Key Agreement for Wireless Body Area Networks," in *Proceedings - IEEE INFOCOM*, George Washington University, Washington, United States. IEEE, 2013, pp. 2274–2282.

[22] N. Jammali and L. C. Fourati, "PFKA: A physiological feature based key agreement for wireless body area network," in *2015 International Conference on Wireless Networks and Mobile Communications (WINCOM)*. IEEE, 2015, pp. 1–8.

[23] G. Zheng, G. Fang, M. A. Orgun, and R. Shankaran, "A Comparison of Key Distribution Schemes Using Fuzzy Commitment and Fuzzy Vault Within Wireless Body Area Networks," in *IEEE th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications PIMRC*, Macquarie University, North Ryde, Australia. IEEE, 2015, pp. 2120–2125.

[24] A. Rukhin, J. Sota, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, and S. Vo, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," National Institute of Standards and Technology, Computer Security Division., Gaithersburg, MD, Gaithersburg, MD, Tech. Rep., 2000.

[25] R. M. Seepers, C. Strydis, I. Sourdis, and C. I. De Zeeuw, "Adaptive Entity-Identifier Generation for IMD Emergency Access," in *First Workshop on Cryptography and Security in Computing Systems*, Erasmus University Medical Center, Rotterdam, Netherlands. New York, New York, USA: ACM Press, 2014, pp. 41–44.

[26] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern Recognition*, vol. 38, no. 1, pp. 133–142, Jan. 2005.

[27] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[28] S. Makeig, "Dynamic Brain Sources of Visual Evoked Responses," *Science*, vol. 295, no. 5555, pp. 690–694, Jan. 2002.

[29] S. Valipour, A. D. Shaligram, and G. R. Kulkarni, "Detection of an alpha rhythm of EEG signal based on EEGLAB ," *Int. Journal of Engineering Research and Applications*, vol. 4, no. 1, pp. 154–159, 2014.

[30] M. Santha and U. V. Vazirani, "Generating quasi-random sequences from semi-random sources," *Journal of Computer and System Sciences*, vol. 33, no. 1, pp. 75–87, Jan. 1986.

[31] J. L. McInnes and B. Pinkas, "On the Impossibility of Private Key Cryptography with Weakly Random Keys," in *Advances in Cryptology-CRYPT0' 90*. Berlin, Heidelberg: Springer Berlin Heidelberg, May 2001, pp. 421–435.

[32] P. Austrin, K.-M. Chung, M. Mahmoody, R. Pass, and K. Seth, "On the Impossibility of Cryptography with Tamperable Randomness." *CRYPTO*, vol. 8616, no. Chapter 26, pp. 462–479, 2014.