CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Resonant-Tunnelling Diodes as PUF building blocks
Authors	Bagci, Ibrahim Ethem;McGrath, Thomas;Barthelmes, Christine;Dean, Scott;Bernardo Gavito, Ramón;Young, Robert James;Roedig, Utz
Publication date	2019-01-14
Original Citation	Bagci, I. E., McGrath, T., Barthelmes, C., Dean, S., Bernardo Gavito, R., Young, R. J. and Roedig, U.(2019) 'Resonant-Tunnelling Diodes as PUF building blocks', IEEE Transactions on Emerging Topics in Computing. doi: 10.1109/TETC.2019.2893040
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://ieeexplore.ieee.org/document/8611366 - 10.1109/TETC.2019.2893040
Rights	© 2019, the Authors. This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/ - http://creativecommons.org/licenses/by/3.0/
Download date	2023-09-30 02:27:46
Item downloaded from	https://hdl.handle.net/10468/9170

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2019.2893040, IEEE Transactions on Emerging Topics in Computing

Resonant-Tunnelling Diodes as PUF Building Blocks

Ibrahim Ethem Bagci, Thomas McGrath, Christine Barthelmes, Scott Dean, Ramón Bernardo Gavito, Robert James Young, and Utz Roedig

Abstract—Resonant-Tunnelling Diodes (RTDs) have been proposed as building blocks for Physical Unclonable Functions (PUFs). In this paper we show how the unique RTD current-voltage (I-V) spectrum can be translated into a robust digital representation. We analyse 130 devices and show that RTDs are a viable PUF building block.

Index Terms—Physical Unclonable Functions, Identification, Authentication, Physical Security

1 Introduction

Physical Unclonable Functions (PUFs) provide an alternative method to generate a secret. Instead of storing the secret in digital memory or asking a user to provide it, it is derived from a physical characteristic. A PUF can be constructed in various ways, for example, using scattering patterns of an optical medium [1] or chip-specific transistor switch delay variations [2]. The assumption is that the secret cannot be duplicated, as it is bound to a physical entity, which cannot be cloned.

RTDs, simple electronic structures utilising quantum confinement, have been proposed as building blocks for PUFs [3]. The RTD encapsulates a quantum nanostructure between two electrical contacts and displays an exotic I-V characteristic not seen in classical devices. The I-V spectrum exhibits a peak which is highly dependent on the quantum confinement within its nanostructure, and the quantum confinement depends on the overall atomic arrangement of the device. The atomic arrangement is subject to random process variations during manufacture. Therefore, each manufactured device exhibits a spectrum with a uniquely positioned peak. The peak location can be translated into unique device specific data.

An RTD has a number of benefits. It represents a physical system which is extremely hard to clone due to the devices nanoscale size and complexity. An RTD can be produced together with an Integrated Circuit (IC) on the same wafer without introducing additional manufacturing steps. As an RTD is simple and small in size many can be included within a chip providing a large amount of unique data. Electronic PUFs typically suffer from stability issues when implemented and sensitivity to ambient conditions, this will make scaling the technology challenging as the feature size in CMOS transistors continues to shrink. PUFs based on quantum tunnelling, however, have the opposite relation. As the characteristic size of features within them reduces then measurements from them become more robust; variations in the resonant voltage vary proportionally to fluctuations in the well's width with an inverse square relation.

The general design of RTDs was shown in [3]. Our previous work focuses on the physical properties of the RTDs. A comprehensive discussion on how to obtain data from the RTD and an evaluation of data quality in a PUF context is missing. In this paper we investigate in detail how the RTDs can be used to form a PUF. Specifically, the contributions of this paper are:

- *Randomness and Stability:* We analyse the randomness and stability of the RTD spectra by evaluating 130 newly manufactured devices.
- *Digitisation:* We show how an RTD measured spectrum can be digitised. We propose to extract a single bit from each device by comparing spectrum peak to a threshold.
- *Entropy:* We analyse the min-entropy of RTD digital outputs. We show that the entropy depends on accuracy of the selected threshold.
- *Robustness:* We investigate robustness of the RTD digital outputs by analysis of error rates.

In the next section we give an introduction to PUFs and discuss related work. Section 3 describes the process of digitisation of a measured RTD spectrum. Section 4 analyses the min-entropy and robustness of the RTDs digital outputs. Section 5 discusses the results and implementation aspects of RTDs. Section 6 concludes the paper with a discussion on future work.

2 Background and Related Work

It has been shown that process variations exist in ICs [4], [5], [6], and these variations can be used to extract unique numbers to identify ICs [7]. Since the identification information can be read by an attacker easily from the IC, this technique cannot be used for authentication. To securely authenticate the device PUFs are proposed [8].

A PUF is a device that uses the physical characteristics of an IC to generate a secret. We can describe a PUF as a function, which takes an input challenge c and gives a response $r = f(c)$ in return. The response is dependant on the physical characteristic of the IC and the challenge c.

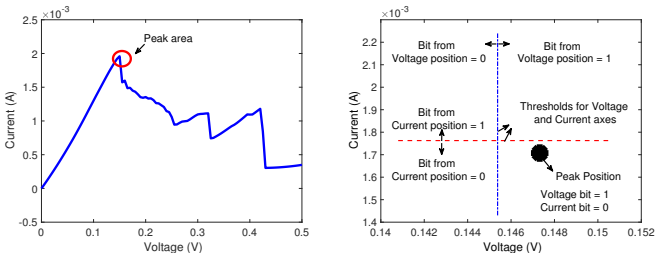
Various methods have been proposed to construct a PUF. Optical PUFs use the scattering patterns from an optical medium [1]. Arbiter PUFs use gate delays [9], and some of the PUFs fabricated on silicon use ring oscillators [10] or statistical delay variations between two identical paths [2]. SRAM-PUFs exploit the power-up state of SRAM cells [11], [12]. Rowhammer PUFs use the Rowhammer effect in DRAM modules [13]. Coating PUFs use a passive dielectric coating sprayed on the chip to explicitly introduce random elements [14]. BoardPUFs characterise the Printed Circuit Boards (PCBs) by embedding a number of capacitors in the internal layers of PCBs and analyse their variations [15].

Supporting a large number of challenge response pairs (CRPs) requires a dedicated on-line database where the CRPs can be stored. When authentication is needed, a challenge from the database is taken and sent to the PUF and then a check is performed to ensure the response agrees with the one stored in the database. The number of CRPs is important when describing the capability of a PUF [16], [17]. Some PUFs provide a CRP database which scales exponentially and polynomially with the system size, and some of them have just one CRP. The number of CRPs is important when considering PUF applications. For example, a PUF with a large number of CRPs may be used for low-cost authentication; a PUF with a single CRPs is often used for secure key generation (called Physically Obfuscated Key (POK) in this context [18]). Depending on the number of CRP, different security requirements exist [17].

- *I. E. Bagci is with the School of Computing and Communications, Lancaster University, Lancaster, UK (e-mail: i.bagci@lancaster.ac.uk).*
- *T. McGrath, C. Barthelmes, S. Dean, R. Bernardo Gavito and R. J. Young are with the Physics Department, Lancaster University, Lancaster, UK (e-mail: thomas.mcgrath@lancaster.ac.uk; c.barthelmes@lancaster.ac.uk; scott-dean@outlook.com; r.bernardogavito1@lancaster.ac.uk; r.j.young@lancaster.ac.uk).*
- *U. Roedig is with the School of Computer Science and Information Technology, University College Cork, Cork, Ireland (e-mail: u.roedig@cs.ucc.ie).*

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2019.2893040, IEEE Transactions on Emerging Topics in Computing



(a) (b)

Fig. 1: (a) I-V characteristic. (b) An example of bit generation process out of an I-V spectrum. Number of bins is selected as 2. Bit numbers of 1 and 0 are generated from voltage and current axes, respectively.

It has been shown that existing PUFs are often clonable. A device can be produced which has identical characteristics, defeating the purpose of a PUF. For example, an SRAM-PUF was cloned by Helfmeier et al. [19], and it has been shown that some PUFs have vulnerabilities to side-channel attacks [20]. However, recent work by Goorden et al. [21] on random-scattering PUFs makes use of the no-cloning theorem in quantum mechanics, and prevents the initial challenge to be cloned. It has also be shown that simulation attacks are possible where the PUF is replaced by a device producing the desired output. For example, Arbiter PUFs, Ring Oscillator PUFs, XOR Arbiter PUFs, Lightweight Secure PUFs and Feed-Forward Arbiter PUFs have all been attacked by Rührmair et al. using machine learning techniques [22].

The works by Rührmair et al. [23] and Jaeger et al. [24] are the closest to ours in terms of the material. As opposed to previous work, using an RTD as a source of a PUF uses much less resources and operate with lower power, have a smaller device geometry and rely on atomic structure, which is the most difficult system to attempt to clone.

3 Extracting Unique Data from RTDs

In this work we use RTDs, simple electronic structures exhibiting quantum confinement [3]. The RTD encapsulates a quantum nanostructure between two electrical contacts and displays an exotic I-V characteristic not seen in classical devices. The I-V spectrum exhibits a peak which is highly dependent on the quantum confinement within its nanostructure, and the quantum confinement is subject to the overall atomic arrangement of the device. The atomic arrangement is subject to random process variations during manufacture. Therefore, each manufactured device exhibits a spectrum which can be used to uniquely identify the device.

3.1 Digitisation

The I-V spectrum as shown in Figure 1a has a number of characteristics which can be considered when transforming the signal into a digital representation. The obvious choice is the location and the height of the peak position. However, it is also possible to consider a number of other elements, these include the position of the valley (where the current drops to before rising again), the slope of the curve after the peak (the negative differential resistance (NDR) region) or the width of the peak (the full-width half maximum). Furthermore, unlike a conventional diode, all these features appear in both bias directions giving us double the number of elements to explore [25].

We decided in this work to only consider peak position and height as they are straightforward to characterise. It would be possible to integrate the RTD with a simple electronic circuit for peak finding. This would be more efficient than sweeping over the entire spectrum and then subsequently applying a computer algorithm for peak finding. However, in this work we followed this inefficient approach as we used prototype devices not equipped with dedicated circuitry for peak finding. The peak finding method used will be explained in Section 3.2.

We can generate a digital output using I or V axes (considering peak height or location). The axis is divided into 2 bins within the range in which we expect peaks to be positioned. Peak ranges are known and depend on the device specifics. The bin order is used as the bit number that is extracted from the device. Hence, we extract only one bit from a device. Increasing the number of bins helps to generate a larger unique bit sequence, i.e., we can extract more than one bit numbers from a peak position. However, at the same time this causes a less robust bit sequence as the peak location/height is then more likely to fluctuate between quantisation steps for each measurement as these are subject to noise. Another drawback of increasing the number of bins is that it may cause bias in the output. Some of the peak positions may be clustered in some positions, and the bit sequences from those peaks would be similar to each other. Furthermore, 2 bins approach is easier to implement in hardware. Using only 2 bins would divide the spectrum into 2, and by choosing a proper *threshold* we can get a bit number with equal probability. Notice that RTDs are tiny devices. Therefore, even if we extract only one bit from a device, many of them can be used together to obtain a bit sequence of required size while keeping the total size of the interconnected devices still small.

When we use 2 bins per axis, if the peak position falls into left half of the voltage axis, we extract bit number 0; or if it falls into right half of the voltage axis, we extract bit number 1. Similarly, if the peak position is in lower half of the current axis, we extract bit number 0; or if it is in the upper half of the current axis, we extract bit number 1. Figure 1b shows an example of the bit extraction (digitisation) process for a device. Peak position is shown as a black dot. In this example, we generate the bits 1 and 0 from the voltage and current axes, respectively.

It should be noted that there is a dependency between the peak position on the voltage and current axes. Hence, it might be advisable to only use one axis at a time to extract digital information.

3.2 Peak Finding

In this work, RTD spectra are fit to an analytical expression for the current density of a tunnel diode as a function of voltage [26]. The fit was found to accurately follow the experimental results that were obtained, and so the local maximum value of the equation shown in [26] was used. This corresponds to the tunnelling region maximum used for the voltage position of the peak current value as opposed to the raw, recorded voltage value. This is to take advantage of analysis of all the data points in the spectrum, as opposed to just the maximum point, leading to a more accurate representation of actual peak value. It is worth noting this model does not account for the plateauing of the current shortly after the peak voltage is reached. These occurrences are likely due to charge trapping effects of electrons within the quantum well of the diode. The least squares fit was therefore applied to the initial incline and only small portion tunnelling decline, before the current starts to plateau.

The analytic expression in [26] describes the shape of the I-V curve. However, the curve parameters are determined by the specifics of the RTD in question. An attacker will therefore be aware of the shape of a curve (function) but does not have any information about peak position (function parameters).

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2019.2893040, IEEE Transactions on Emerging Topics in Computing

3.3 Threshold Selection

In this work we are dividing the voltage and current axes into 2 bins by selecting a threshold to extract a bit number from an RTD. Ideally, the threshold should divide the axes into two where the number of peaks on the both sides of it are equal. Device specifics and its physical characteristics can give an idea about choosing the threshold. Another way to decide the threshold is to measure a subset of device and get the mean values of the peak positions. This may not give the true threshold but we should get an estimation about its location.

The application scenario will determine how many of the devices can be used to calculate the threshold. If the PUFs will be used for identification purposes, all the devices must be profiled during the manufacturing phase so that we can know the true threshold. This because the identification application would require comparing the PUF results that are constructed from RTDs against a database. In this case the threshold would divide the axes into 2 where in both sides of the threshold there will be equal amount of peak positions. Therefore, probabilities of extracting bit numbers 0 and 1 will be 0.5 for each RTD.

In some scenarios, RTDs may be manufactured and shipped in batches at different times. And sometimes profiling all manufactured RTDs may not be feasible due to their large number. Therefore, the threshold must be estimated using the first shipped batch or some portion of the RTDs. Then the outputs of the remaining devices would be extracted by comparing their peak positions against the estimated threshold. This will not guarantee that the number of peaks in both side of the threshold will be equal.

3.4 Min-Entropy of RTD Outputs

In information theory, the min-entropy corresponds to the most conservative way of measuring the unpredictability of a set of outcomes. It allows measurement of the quality of the results. The goal is to get unpredictable outputs as often as possible, and in best case it means having min-entropy equal to the size of the outputs.

Let p_{max} denote the most likely outcome of random variable X, then min-entropy of X is defined as:

$$H_\infty(X) = H_{min}(X) = -log_2 p_{max}$$

In our work, we extract 1 bit from an RTD. In an ideal case, the probabilities of bit numbers 0 and 1 should be equal, $p = 0.5$. Therefore, most likely outcome of any bits would be $p_{max} = 0.5$, and min-entropy of $H_\infty = 1$ can be obtained from a device in the best case.

The threshold selection has a crucial impact on the amount of min-entropy that can be obtained from a device. The highest min-entropy can be achieved only if the threshold can divide the axis where the each side of the threshold have equal size of peaks. With poor threshold selections, the number of peaks will be different at each side of the threshold, and it will decrease the min-entropy.

3.5 Concatenation

In this work we extract only 1 bit from an RTD. Therefore, RTD by itself is not suitable to be used as a PUF. Multiple RTDs have to be connected to construct a PUF and to obtain desired amount of bit sequences. The number of RTDs required for a PUF depends on the min-entropy of a device and on the desired security level. For example, if min-entropy of $H_\infty = 0.8$ can be obtained from a device, $128 \div 0.8 = 160$ RTDs should be concatenated to get 128 bit level security. Moreover, if an error correction mechanism is used alongside the PUF, we need more RTDs as helper data reduces the overall entropy from a PUF.

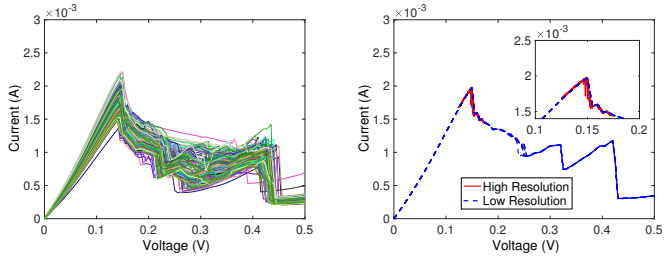


Fig. 2: I-V spectrum of 130 devices. (a) One measurement for 130 devices in low resolution. (b) 10 and 50 measurements for one device in low and high resolution, respectively.

4 Experimental Analysis

In the previous section we described how an RTD spectrum can be digitised. In this section we aim to analyse the quality of extracted bits. For this purpose we analyse the min-entropy and the robustness of the outputs.

4.1 Experimental Data

In this work we manufactured and measured 130 RTDs. For each of the devices, 10 independent measurements are obtained by sweeping from 0 V to 0.5 V with 5 mV resolution, which corresponds to 101 voltage points. We will refer these measurements as measurements taken with *low resolution*. After figuring out where the peak area is, 50 more independent measurements are obtained in a 0.05 V range around the peak with 0.5 mV resolution, which again corresponds to 101 voltage points. We will refer these measurements as measurements taken with *high resolution*. We use multiple measurements of the same device to analyse the robustness of the digital output generation process. It has to be noted that this measurement was carried out in an experimental setup such that peak position and height can be evaluated. A practical device would contain electronic circuitry to directly find peak location and peak height, eliminating some of the more complex spectrum evaluation steps.

4.2 Spectra Randomness and Stability

In this section we investigate the quality of the RTD spectra with statistical analysis.

Figure 2a shows the I-V spectrum of first measurement of each of the 130 devices. As shown in the figure, peak locations and height in each spectrum differ. Figure 2b shows the I-V spectrum of 10 and 50 individual measurements of one of the devices in low and high resolution, respectively. This figure shows that the spectrum is stable when considering a single device.

The peak location of an I-V spectrum is found by peak finding method explained in Section 3.2. Peak locations from all the measurements of each device are shown in Figure 3a. The figure shows error bars on both voltage and current axes using the mean and 99% confidence level of the 50 measurements in high resolution. We can see that even if some of the peak locations are close to each other, they are still clearly distinguishable.

Next we perform statistical tests to investigate more thoroughly how device measurements deviate from each other. To carry out such a test in a meaningful way it is first necessary to determine the distribution type of the measurements. We carry out a Kolmogorov-Smirnov test with a significance level of $\alpha = 0.01$ to check the normality of the peak locations of 130 devices with their 50 measurements in high resolution. This test concludes that peak locations and height do not fit a normal distribution. Therefore, we apply a pair-wise Wilcoxon-Mann-Whitney test to the peak locations to see how they deviate from each other. The Wilcoxon-Mann-Whitney test is a non-

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2019.2893040, IEEE Transactions on Emerging Topics in Computing

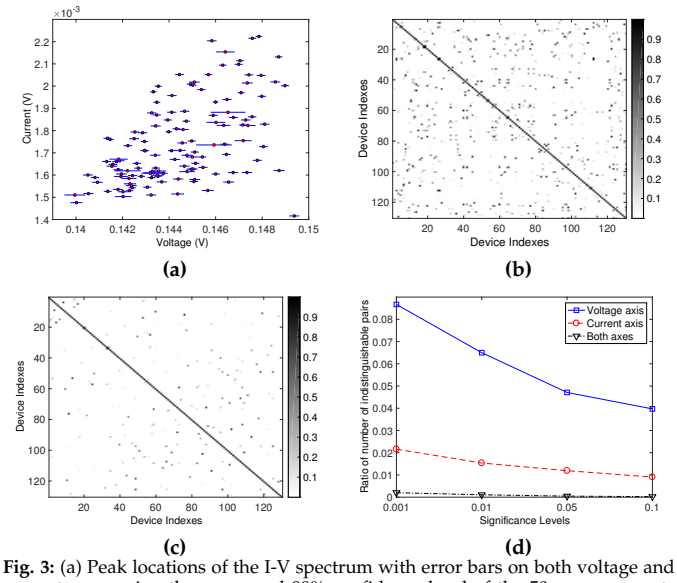


Fig. 3: (a) Peak locations of the I-V spectrum with error bars on both voltage and current axes, using the mean and 99% confidence level of the 50 measurements in high resolution. Individual peaks are clearly discernible. (b), (c) p-values when 130 devices are compared to each other with the peak positions in voltage and current axes, respectively. (d) Ratio of the indistinguishable pairs over the total pairs with varying significance levels on voltage and current axes, and on both axes at the same time.

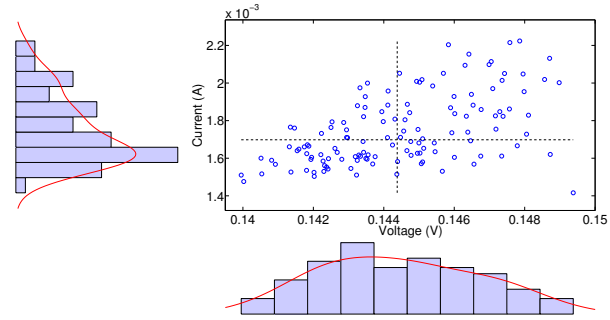


Fig. 4: Peak positions averaged from 50 measurements in high resolution of 130 devices.

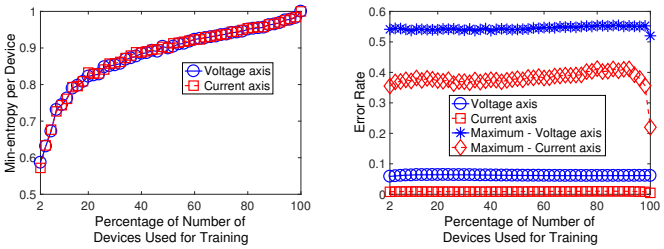


Fig. 5: (a) Min-entropy amount when using different numbers of devices for training. (b) Error rates when using different numbers of devices for training.

parametric test, which is an alternative to a t-test for non-normally distributed sets.

We are applying the Wilcoxon-Mann-Whitney test with varying significance levels. The null hypothesis for the test is that two samples come from the same population, and the alternative hypothesis is that they do not. The null hypothesis is rejected if the resulting p-value of the test is less than the significance level. Figures 3b and 3c show the p-values when 130 devices are compared to each other with the peak positions in voltage and current axes, respectively. Here we perform 8385 pairwise comparisons. A small p-value indicates strong evidence against the null hypothesis. Therefore, we should obtain smaller p-values from the statistical test. In figures 3b and 3c, small p-values are shown with lighter colours. We can see that the current axis performs better than the voltage axis.

Figure 3d shows ratio of number of indistinguishable pairs over the total pairs with significance levels of $0.001, 0.01, 0, 05$ and 0.1. The figure shows the indistinguishable pairs for either voltage and current axes, and when they are indistinguishable on both axes at the same time. We have 17, 9, 4, and 2 pairs when they are indistinguishable on both axes with significance levels $0.001, 0.01, 0, 05$ and 0.1, respectively. This means that majority of the peak positions are unique.

We also look at the inter-device spectrum stability. We compare half of the measurements of an RTD to the other half of the measurements of the same RTD with significance levels of $0.001, 0.01, 0, 05$ and 0.1. If the statistical test rejects the null hypothesis, it means that the two halves of the same device are considered as coming from different populations. Out of 130 devices, 2 of them rejected the null hypothesis on voltage axis and 7 of them rejected on current axis, with each significance level. However, none of them rejected the null hypothesis on both axes at the same time.

These results show that RTD spectra are mostly unique to each device, and remain mostly stable between multiple measurements.

4.3 Entropy of the RTD Outputs

Section 4.2 shows that the ratio of non-unique peak positions is very low. However, to be able to use RTDs as a PUF source, e.g., for key generation, we need to extract bits from the devices. Here, we investigate the min-entropy of the outputs of RTDs.

Ideally, the distribution of the bits extracted from RTDs should be uniform. As a result, there won't be any bias on an output or group of outputs. Uniform distribution of the outputs will help to get highest entropy from a device, which will determine the number of RTDs to be used in a PUF for a specific security level. In this work we generate only one bit from an RTD, therefore we aim to obtain a min-entropy of 1.

Figure 4 shows the distributions of the peak positions averaged from 50 measurements in high resolution of 130 devices. Vertical dashed line on the voltage axis and horizontal dashed line on current axis show the thresholds on the axes that are calculated from the measured 130 devices. Thresholds divide the axes into 2 bins, where in each bin there are equal number of peaks. We can see that the peak positions on both axes follow close to normal distribution, and the ones on the current axis are slightly biased to the lower side of the spectrum. Here, thresholds on the voltage and current axes affect the distribution of bit numbers 0 and 1 that are extracted from the devices.

Figure 5a shows the min-entropy result when using different number of devices for training, i.e., for finding the threshold for an axis. We run the test 1000 times to randomly select a specific percentage of devices for training and average the final results. Then 50 measurements in high resolution of all devices are averaged and compared against the estimated threshold.

We can see that increasing the number of devices for training increases the entropy. Clearly, we obtain a min-entropy $H_\infty = 1$ if we can use all the devices for training to estimate the threshold.

4.4 Robustness of the RTD Outputs

In this section, we investigate the robustness of the outputs of RTDs. Each I-V spectrum of RTDs is subject to noise, therefore different measurements of an RTD may have slight differences.

We evaluate the robustness by calculating the rate of obtaining different bit numbers from a device when it is measured

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2019.2893040, IEEE Transactions on Emerging Topics in Computing

at different times, which we call *error rate*. Devices that give peak positions far away from the thresholds are less likely to give different bit numbers from their different measurements. However, if the peak position of a device is close to the threshold, each measurement would cause the peak position to fall either side of the threshold.

We first find the threshold by using the training devices, and average the peak positions of the measurement of all devices. Then we record which side of the threshold the averaged peak positions fall. Then we check whether the peak positions of individual measurements of each device fall to the same side of the threshold where the corresponding device's averaged peak positions fall. To get the error rate, number of mismatches are summed for all test devices and their measurements; and divided by [Number of test devices]X[Number of measurements].

Figure 5b shows the error rates when using different number of devices for training. Again, we used 50 measurements of 130 devices. We repeat the test 1000 times to randomly select training devices and average the results. Error rate performance of current axis is better than voltage axis. The number of training devices does not have much effect on error rates. This is because the number of peaks around different thresholds, even if they are estimated with small number of training devices, are similar to each other (peak positions are not clustered around the thresholds). The figure also shows the maximum error rate of all devices and their measurements. We obtain less maximum error rates when the number of training devices are close to the number of all devices.

4.5 Improving Error Rates

In Figure 5b, out of 130 devices, 47 devices recorded errors on voltage axis, 7 devices on the current axis, and 3 devices on both axes when training with all the devices. One can disable these RTDs from the final product batch at the time of provisioning, so that more robust PUFs can be created.

5 Discussion

5.1 RTD Performance

We have shown that RTD spectrum is unique to each device, and remains stable between multiple measurements.

The quality of the digital outputs of RTDs depends on the accuracy of threshold selection. The threshold is estimated using a number of training devices. Increasing the number of training devices while estimating the threshold leads to higher min-entropy results. The best min-entropy can be obtained when all the devices are used for threshold estimation. However, the number of training devices does not have much effect on the average error rates. This is because the error rate depends on the variance of the peak positions around the threshold, and these variances are similar over the axes.

We have found that current axis is better option than the voltage axis. Although the min-entropy results from each axis are almost the same, peak positions on current axis are more distinguishable among each device, and more stable between multiple measurements of a device. Moreover, the error rates are lower on current axis.

5.2 Attacks

Unclonability: The technology necessary to clone an RTD requires that the internal atomic structure be recorded with high precision before being re-assembled, atom by atom, on a separate chip. State-of-the-art technologies for measuring and remapping a three-dimensional structure's atomic make-up are not sophisticated enough to achieve this [27]. Furthermore, the attacker would have to destroy the honest party's RTD to probe the internal structure and since they cannot make a clone this would leave the legitimate user aware of some malpractice. This two-stage process necessary for cloning a device presents a higher degree of security as even if it became feasible to achieve one step, the other step would hinder a clone being produced. Moreover, it is unlikely that the technology required to complete either stage will not be advanced enough in the near future. An attacker could measure, approximate and recreate an RTD's response, such that the produced signature agrees with the initial RTD's within errors. This vulnerability can be mitigated by using a large number of RTDs in a single PUF.

Predictablity: Some PUFs providing multiple CRPs have been attacked using Machine Learning (ML) [22], which can help to predict a response for a challenge by analysing previously observed CRPs. This work describes individual RTDs as PUF elements. ML attacks cannot be run on the individual RTDs used in this study as the devices are independent. ML attacks may be possible when integrating many of the described RTDs into a full PUF device as there might be a hidden dependency between RTDs due to the manufacturing process. However, as a full PUF device incorporating many RTDs was not yet produced for this study, ML analysis is reserved for future work.

Simulation: An attacker can observe the response(s) of a PUF and replace it with another device that has the same response(s). PUFs with a single CRP can prevent this form of attack by not revealing the response during operation (Internal processing, potentially in a tamper proof environment). A PUF with multiple CRPs does not have the same restriction, as the attacker cannot enumerate all CRPs within a feasible time frame. However, in this work we have investigated a PUF building block and not a full device, therefore, this aspect has not been studied in detail.

5.3 Implementation Aspects

Integration: The RTD devices presented here use standard III-V semiconductor structures, which are the materials used for common optoelectronic devices such as laser diodes and LEDs. However, RTDs can also be made from silicon, the material of choice in the electronics industry for fabricating components such as transistors and ICs [28], [29], [30]. The facilities required for the fabrication of III-V structures and silicon is different, but these devices can be integrated into any procedure depending on the architecture wanted on chip. There is also the possibility of combining both these processes [31]. Furthermore, RTDs can be fabricated on a large scale due to the ability to integrate a huge number of devices on one chip. Typically, we can make approximately 50,000 devices on one $8''$ wafer, and advanced fabrication facilities have the possibility of making 50 of these $8''$ wafers simultaneously.

Measurement: There is a need for measurement simplicity; peak location and height should be determined quickly and without complex measurement procedures. In this study the spectrum of a device was measured and peaks were determined later using software. However, in a practical application scenario circuitry for automatic searches would be integrated with the device.

Temperature Stability: The effects of temperature upon the resonant tunnel diode within the bounds of any expected operating temperatures (-50 to 70°C [32]) are negligible. This is because the temperature of the quantum well (the active region) of the RTD becomes practically independent from the temperature of the surrounds, since the active region of a resonant tunneling diode in operation is significantly higher than the temperature outside of the device. Additionally, the

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TETC.2019.2893040, IEEE Transactions on Emerging Topics in Computing

predictable nature of any variation the diodes may experience with external temperature can be leveraged to enhance security. One example case would be including a temperature sensor and examining variation due to temperature as an additional verification step, since a simulated component could not as easily or accurately adjust with temperature compared to the genuine semiconductor counterpart.

Cost: The ability to use parallel fabrication has transformed the semiconductor industry, making it possible to make a large quantity of the same device in a small number of processing steps. Whereas in other fabrication processes you might have to make one device at a time, here we can make a huge number of devices simultaneously. Using RTDs made from silicon as an example, it is common that 100 silicon wafers with a diameter of $11''$ would cost approximately $20. Therefore, around 2.5 million devices could be made for less than $10 in terms of cost of material. The main cost arises from the need for expertise and running the equipment and machinery necessary. However, as this many devices can be made in parallel, this cost is insignificant per device.

Material, Fabrication and Size: The RTDs examined in our study are formed from a double barrier of Indium Gallium Arsenide (InGaAs) and Aluminium Arsenide (AlAs) upon an Indium Phosphide (InP) substrate, with gold (Au) contacts [33]. The material was layered through molecular beam epitaxy, with top contacts defined by optical lithography and fabricated using thermal evaporation. Reactive Ion Etching (RIE) and wet etching was then employed to define the side walls of the RTD and air-bridge undercut, with the bottom contacts again thermally evaporated. RTDs can be of variable size and makeup, and while those studied here are formed of $4\,\mu\text{m}^2$ III-V semiconductor material, these devices can meet size and material requirements for easy integration into typical CMOS processes [34], as cost evaluated in Section 5.3.

6 Conclusion and Future Work

Our results show that RTDs can be used to construct a PUF, which is suitable for secure key generation (with limited number of CRPs) and low-cost authentication (with exponential number of CRPs). An RTD represents a physical system which is extremely hard to clone due to the device's nanoscale size and complexity. Additionally, they are simple to mass-manufacture and easy to operate. We have demonstrated that RTDs can generate bits which are unique and robust.

In this work we used data from 130 fabricated RTDs. In future we aim to investigate different ways of constructing PUFs using RTDs. We plan to produce and analyse a larger number of improved devices which integrate the measurement circuitry to enable peak finding.

Acknowledgment

R.J.Y. acknowledges support by the Royal Society through a University Research Fellowship (Grants UF110555 and UF160721). This material is based upon work supported by the Air Force Office of Scientific Research under Award FA9550-16-1-0276.

References

[1] R. Pappu *et al.*, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002.
[2] J. W. Lee *et al.*, "A technique to build a secret key in integrated circuits for identification and authentication applications," in *Proc. of the Symposium on VLSI Circuits*. IEEE, 2004.
[3] J. Roberts *et al.*, "Using Quantum Confinement to Uniquely Identify Devices," *Scientific reports*, vol. 5, 2015.
[4] D. S. Boning *et al.*, "Statistical Metrology: Understanding Spatial Variation in Semiconductor Manufacturing," in *Proc. of SPIE 1996 Symposium on Microelectronic Manufacturing*, 1996.
[5] S. R. Nassif, "Modeling and forecasting of manufacturing variations," in *Proc. ASP-DAC'01*, 2001.
[6] K. A. Bowman *et al.*, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 2, pp. 183–190, 2002.
[7] K. Lofstrom *et al.*, "IC identification circuit using device mismatch," in *Proc. ISSCC'00*, 2000.
[8] P. S. Ravikanth, "Physical One-Way Functions," Ph.D. dissertation, MIT, 2001.
[9] B. Gassend *et al.*, "Silicon physical random functions," in *Proc. CCS'02*, 2002.
[10] G. E. Suh *et al.*, "Physical unclonable functions for device authentication and secret key generation," in *Proc. DAC'07*, 2007.
[11] J. Guajardo *et al.*, "FPGA intrinsic PUFs and their use for IP protection," in *Proc. CHES'07*, 2007.
[12] D. E. Holcomb *et al.*, "Initial SRAM state as a fingerprint and source of true random numbers for RFID tags," in *Proc. RFID-Sec'07*, 2007.
[13] A. Schaller *et al.*, "Intrinsic rowhammer PUFs: leveraging the Rowhammer effect for improved security," in *Proc. IEEE HOST'17*, 2017.
[14] P. Tuyls *et al.*, "Read-proof hardware from protective coatings," in *Proc. CHES'06*, 2006.
[15] L. Wei *et al.*, "BoardPUF: Physical unclonable functions for printed circuit board authentication," in *Proc. ICCAD'15*, 2015.
[16] U. Rührmair *et al.*, "On the Foundations of Physical Unclonable Functions," *Cryptology ePrint Archive, report 2009/277*, 2009.
[17] C. Herder *et al.*, "Physical unclonable functions and applications: A tutorial," *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1126–1141, 2014.
[18] B. L. Gassend, "Physical random functions," Ph.D. dissertation, Massachusetts Institute of Technology, 2003.
[19] C. Helfmeier *et al.*, "Cloning physically unclonable functions," in *Proc. IEEE HOST'13*, 2013.
[20] D. Merli *et al.*, "Side-channel analysis of PUFs and fuzzy extractors," in *Proc. TRUST'11*, 2011.
[21] S. A. Goorden *et al.*, "Quantum-secure authentication of a physical unclonable key," *Optica*, vol. 1, no. 6, pp. 421–424, 2014.
[22] U. Rührmair *et al.*, "Modeling attacks on physical unclonable functions," in *Proc. CCS'10*, 2010.
[23] ——, "Security applications of diodes with unique current-voltage characteristics," in *Proc. FC'10*, 2010.
[24] C. Jaeger *et al.*, "Random pn-junctions for physical cryptography," *Applied Physics Letters*, vol. 96, no. 17, p. 172103, 2010.
[25] L. Britnell *et al.*, "Resonant tunnelling and negative differential conductance in graphene transistors," *Nature communications*, vol. 4, p. 1794, 2013.
[26] J. N. Schulman *et al.*, "Physics-based RTD current-voltage equation," *IEEE Electron Device Letters*, vol. 17, no. 5, pp. 220–222, May 1996.
[27] P. Dasmahapatra *et al.*, "Thickness control of molecular beam epitaxy grown layers at the 0.01–0.1 monolayer level," *Semiconductor Science and Technology*, vol. 27, no. 8, p. 085007, 2012.
[28] S. Miyazaki *et al.*, "Resonant tunneling through amorphous silicon–silicon nitride double-barrier structures," *Physical review letters*, vol. 59, no. 1, p. 125, 1987.
[29] Q.-Y. Ye *et al.*, "Resonant tunneling via microcrystalline-silicon quantum confinement," *Physical Review B*, vol. 44, no. 4, p. 1806, 1991.
[30] K. Ismail *et al.*, "Electron resonant tunneling in Si/SiGe double barrier diodes," *Applied physics letters*, vol. 59, no. 8, pp. 973–975, 1991.
[31] J. Bergman *et al.*, "RTD/CMOS nanoelectronic circuits: Thin-film InP-based resonant tunneling diodes integrated with CMOS circuits," *Electron Device Letters, IEEE*, vol. 20, no. 3, pp. 119–122, 1999.
[32] U. DoD, "Mil-STD-810F: Department of defense test method standard for environmental engineering considerations and laboratory tests," *US Department of Defense (DoD)(01.01. 2000)*, 2000.
[33] M. A. M. Zawawi *et al.*, "Fabrication of Submicrometer InGaAs/AlAs Resonant Tunneling Diode Using a Trilayer Soft Reflow Technique With Excellent Scalability," *IEEE Transactions on Electron Devices*, vol. 61, no. 7, pp. 2338–2342, July 2014.
[34] R. Lavieville *et al.*, "Quantum dot made in metal oxide silicon-nanowire field effect transistor working at room temperature." *Nano letters*, vol. 15, no. 5, pp. 2958–2964, 2015.

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/.