Breaking the Breach: The Best Step in Hardware Trojan Detection

JoanMarie Leone

Department of Computer Science, Binghamton University

CS301 Ethical, Social, and Global Issues in Computing

Dr. George Weinschenk

Mar. 25, 2022

Abstract

With the constant threat of hardware Trojans releasing sensitive and confidential information into the wrong hands, its important to ensure that all information is protected with the most reliable and efficient hardware Trojan detection algorithm. The Run-Time Monitoring and Validation Using Reverse Function for Hardware Trojans detection algorithm successfully improves on some of the most widely used runtime monitoring HT detection algorithms (TMR and ATMR) by cutting down on power consumption and resource usage. Another HT detection algorithm, the Multi-Parameter Approach for FPGA Hardware Trojan detection, splits up its algorithm into three methods: run-time, logic testings, and side channel analysis. The cooperation between these three smaller methods creates a more reliable HT detection method as the multiple stages all work together to ensure the least amount of HTs go undetected. Additionally, the algorithm gathers the details of the attempted HT which the algorithm then learns from to alter and improve on its HT detection in the future. Improvements including the limited run-time monitoring portion of the algorithm and the removal of golden measurements make the Multi-Paramater Approach for FPGA HT detection method much less power consuming than the RMVRF. When so much information is stored online, the potential dangers and consequences that can result from security breaches caused by hardware trojans makes it important that the superior HT detection algorithm, the multi-parameter approach for FPGA's, is implemented.

Breaking the Breach: The Best Step in Hardware Trojan Detection

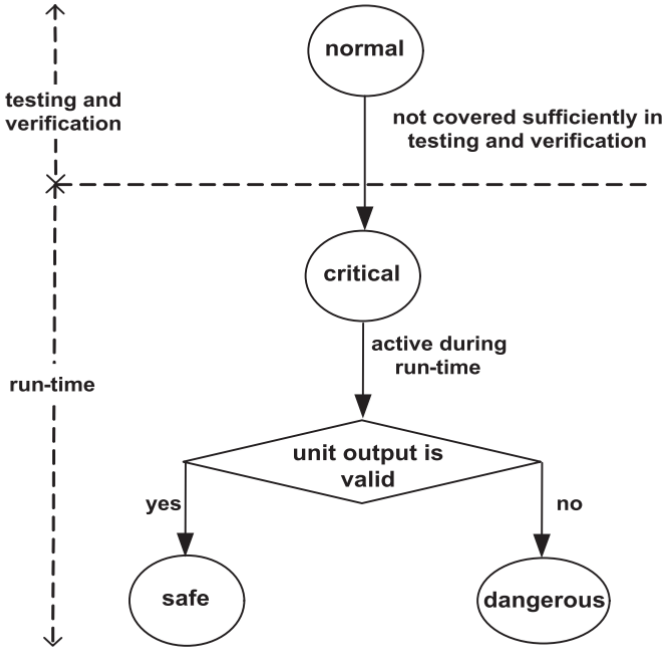
Implementing the best hardware trojan detection algorithm becomes crucial when an undetected hardware trojan horse could spawn a major security breach, putting sensitive and even life-threatening data into unwanted hands. Hardware Trojans (HTs) can have detrimental effects on the behavior and functionality of hardware as they maliciously modify integrated circuits including FPGAs (field-programmable gate arrays). An HT, if successful, can result in serious information leaks and performance degradation. As HTs prioritize avoiding detection, many algorithms exist for creating reliable and efficient detection methods that require the least amount of energy and resources possible. Although the Runtime Monitoring and Validation Using Reverse Function (RMVRF) improves upon previous runtime monitoring algorithms, the multi-parameter approach for FPGA hardware Trojan detection method provides all-inclusive coverage and refinement from HTs while also conserving more power which overall betters the protection against breaches of information and malfunctions of security-sensitive application domains. The RMVRF algorithm primarily seeks to decrease the power and resource usage of typical runtime monitoring approaches. While successful, its advancements fail at making RMVRF a superior algorithm over the multi-parameter approach for FPGAs.

Alternate Technology

RMVFR, a runtime monitoring method, captures unexpected differentiations such as increases in power consumption caused by the activation of an HT. As an invasive method of HT detection, runtime monitoring requires a physical circuit modification to monitor the circuits, such as the addition of an on-chip sensor. Bassam J Mohd, member of the computer engineering department at Hashemite University, with fellow computer science and engineering department

members from other universities discussed how runtime monitoring algorithms such as the RMVRF algorithm, while invasive, guarantee full coverage to all types of HT attacks compared to other non-invasive methods, making it a valid competitor to many other HT detection algorithms (Mohd et al., 2021, p.2). This method also requires a large amount of resources from RCDs (resource-constrained devices), which already has limitations in memory size, battery capacity, and computing power. The first step in lowering power consumption involves reducing the cipher algorithm complexity for the encryption method; the use of lightweight ciphers consists of smaller block designs and lighter transformation functions that optimize power and energy. Hard to activate nodes likely conceal Harmful hardware trojans, so the algorithm creates a list of critical nodes, including low coverage nodes, that the functional tests will activate. The next step in the algorithm works by detecting critical node switches during runtime as an indicator of an activated HT as seen in Figure 1. Another design then validates the detected node to see whether the activation was the product of an HT. An HT can only avoid detection by being embedded in two different designs and triggered simultaneously, causing the validation to compare the two HT outputs to each other. The algorithm determines each activated critical node as either safe or dangerous; if safe, the algorithm adapts as not to perform validation on the node if activated again in the future.

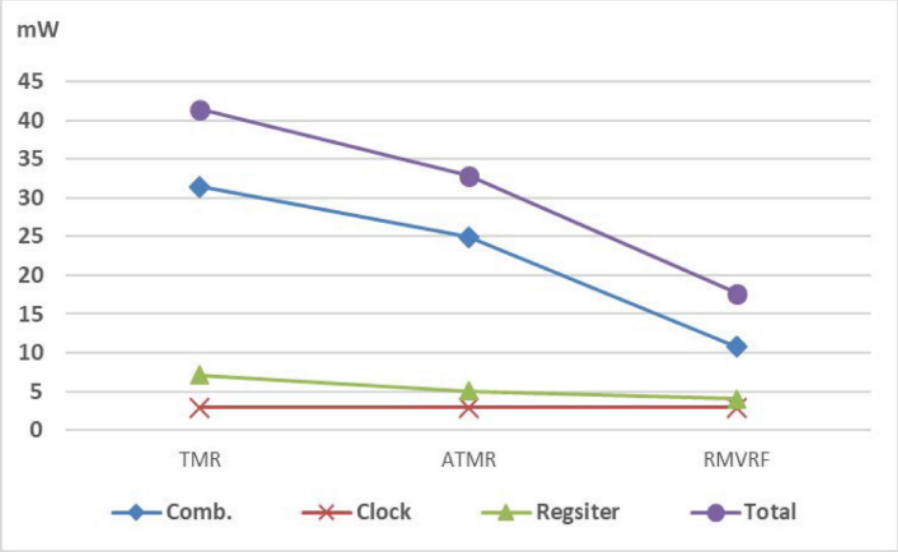
Figure 1. RMVFM Critical Node Classification



Mohd compared RMVRF to other popular runtime algorithms (TMR and ATMR), finding that RMVRF uses less power by at least a factor of 1.73 (Mohd et al., p.12) as seen in Figure 2. RMVFT greatly reduces the amount of combinational logic by having one main-unit active compared to two or three active units typical in other runtime monitoring algorithms, resulting in a significant decrease in power consumption. While RMVRF has the same resource utilization as other RT algorithms, Mohd et al found that it saves 25% more resources than the other algorithms by saving the reverse-function validating units in its blocks (Mohd et al., p.12). Although progressive in these aspects, the RMVRF algorithm has about the same time complexity and maximum frequency, showcasing essential aspects not improved upon. The RMVRF algorithm has improved the runtime monitoring method of HT detention, but its improvements do not prove substantial enough to make this runtime algorithm a more suitable

option over the multi-parameter approach for FPGAs, which provides better coverage and more refinement in its runtime monitoring portion.

Figure 2. Power Consumption Comparison When HT Detected



Support

Technical Details

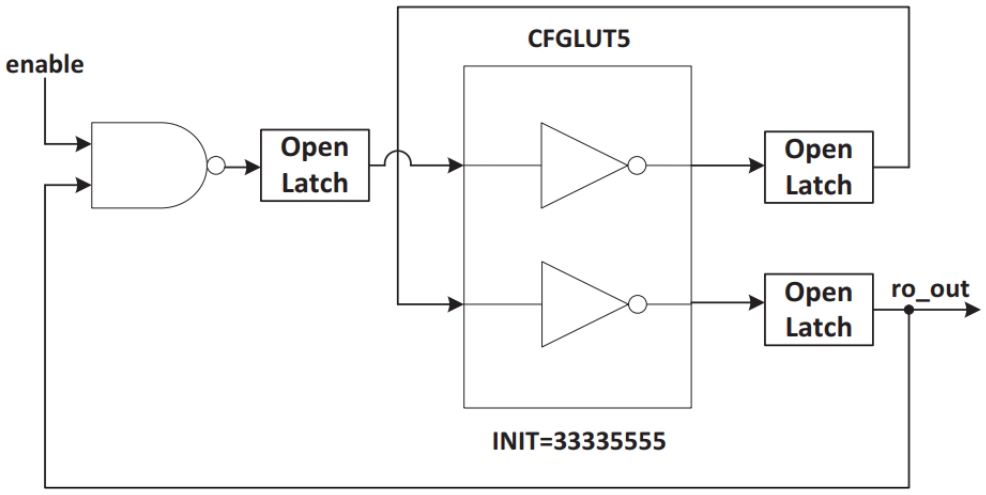
Creating the most foolproof HT detection method, the multi-parameter approach for FPGAs combines three different harmonious methods: runtime monitoring, logic testing, and side-channel analyzing. While RMVRF consists of one invasive runtime monitoring method, the multi-parameter approach only contains a runtime portion in favor of delegating more of the protection plan to other non-invasive methods. Computer engineering professors in Greece, including Apostolos Fournaris who hold a Ph.D in electrical and computer engineering, highlight an important feature added to the runtime portion of their multi-parameter algorithm: refining and comparing the result of the runtime method to determine how the HT is triggered to sharpen the detection methods by reducing the number of false positives (Fournaris et al., 2019, p.3).

While the RMVRF algorithm learns from previous critical node activations, the multi-parameter algorithm excels in refinement through its post trace collection analysis, implemented in the last stage of the algorithm. Beyond refining the runtime monitoring portion, the multi-parameter algorithm simultaneously improves on other logic testing and side-channel analysis methods. Logic testing examines illogical outputs or changes in power consumption possibly caused by an HT using special test vectors. Then when the circuit experiences a structural modification, the side-channel analysis measures the changes to the integrated circuit parameters, including electromagnetic emission (EM). Typical side-channel analysis HT detection methods require a golden chip or model, where a small set of chips in the integrated circuit becomes the trusted set of trojan-free chips. These chips get compared against possible trojan-infested chips to catch the HTs; however, creating the availability of trusted measurements and simulation models has an exceptionally high cost that is often unachievable. Fortunately the multi-parameter approach method completely removes the need for any golden chip or model. These three methods all come together in a three-step plan to form the complete and efficient multi-parameter approach algorithm.

The first step in the multi-parameter approach for FPGAs, the installment of an on-chip digital sensor grid based on Ring-Oscillators (RO), happens during the FPGAs design. This HT algorithm introduces a smaller and more efficient RO implementation involving only two look-up tables (LUTs) for maintaining data integrity and cutting down on power usage. The RO model also includes open latches, as seen in Figure 3, which increases the sensitivity of the RO by amplifying, controlling, and generating electrical signals (Fournaris et al., 2019, p.3). The frequency of ROs change when the circuit or the operations of its surrounding sensor are altered, which the insertion of an HT can cause. Paris Kristos, who would later co-author the

multi-parameter FPGA algorithm, researched the effect of RO length on HTs with fellow computer engineers in Greece and found that monitoring changes from the RO not only helps detect HTs but also helps support non-destructive testing methods for hardware designs and HT detection (Pirpilidis et al., 2017, p.1). The improvements on the in-chip sensor grids shape the main component of the runtime method portion of the multi-parameter detection algorithm and build the foundation for the next steps in the HT detection method.

Figure 3. RO Implementation Consisting of Three Stages and Two LUTs

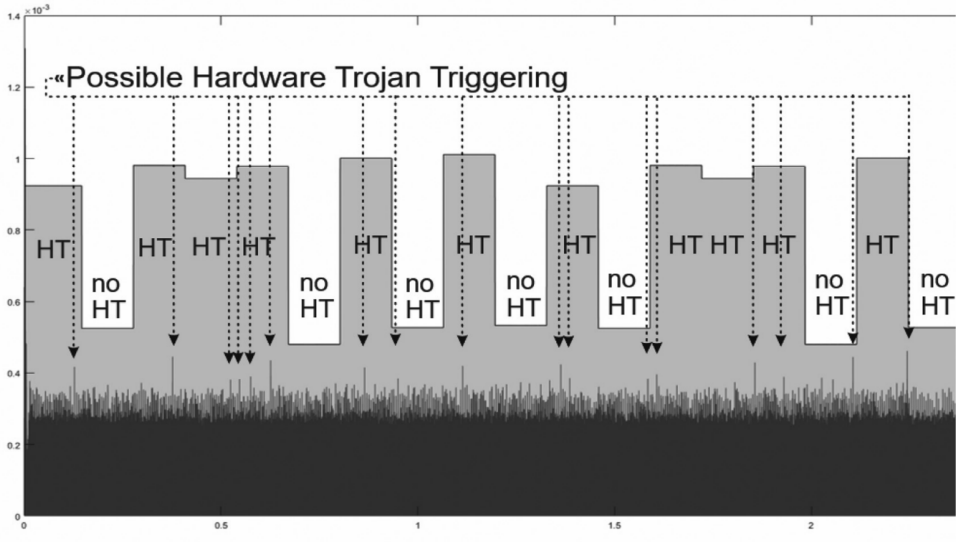


Step two uses a combination of the runtime monitoring and logic testing. The runtime method detects and collects the most plausible side-channel leakage signals from step 1. The logic testing method then analyzes the signals against a small set of reliable test vectors that can guarantee an HT's activation. This method is reliable against HTs that alter the output of an integrated circuit; more advanced HTs likely target other areas of the IC besides the output; these HTs aren't detected through this method. The RMVRF algorithm covers such advanced HTs, but in combating this, the multi-parameter algorithm implements other detection methods for these types of attacks. By detecting less advanced HTs with a simpler method than ones of greater

complexity, the multi-parameter detection algorithm helps preserve resource and power consumption. For the more advanced HTs, the collection of side-channel measurements occurs simultaneously in the background of step two to be used for the final stage of the HT detection algorithm.

The third and last step of the multi-parameter FPGA HT detection algorithm analyzes the side channel information collected in the previous step and combines all the methods above to create reliable HT analysis results as seen in Figure 4. Due to noise, the data from the side channel must undergo further processing. The already existing Fast Fourier Transformation (FFT) algorithm works to reveal significant high and low frequencies to denoise the side channel data. Then the side-channel analysis marks sample values that have a considerable difference from other values in the same trace, identifying them as interesting points. Multiple calculations are enacted on the interesting points, which the algorithm utilizes later in its automation and refinement process to understand HT patterns to better the detection algorithm for future HTs (Fournaris et al., p.3). This sophisticated polishing by the multi-parameter approach for FPGAs HT detection algorithm surpasses the minor refinements implemented in the RMVRF algorithm. The multi-parameter approach combines three innovative detection methods to create a dependable algorithm where the users can sleep easy knowing they won't be subject to a malicious hardware trojan attack. In proving the importance of having the best detection algorithm, **it's essential to understand what happens when an HT slips its way under the radar and infiltrates a computer system.**

Figure 4. Combination of logic testing, on-chip sensors, and side-channel analysis results to detect HT's and the triggered test vectors



Social Impact

The consequences of a successful hardware trojan are immeasurable. Hardware attacks such as hardware Trojans are not as common as other cyberattacks. Still, as discussed by Mark D. Hill, a computer science professor at the University of Wisconsin, these attacks can cause even correct and fully functioning software to leak information and also physically damage the hardware itself (Hill, 2020, p.1). Therefore, protecting against HTs holds just as much, if not more importance than other types of cyberattacks. The repercussions of an HT include severe damages to the government, companies, and people involved, which is why having the HT detection algorithm with the most coverage, aka the multi-parameter approach for FPGAs, holds so much importance. The costly consequences of HTs expand way beyond the physical damage it causes; business data breaches can be detrimental to the company; the combination of fines, penalties, legal fees, customer compensation, and customer loss has the power to run a business into the ground. The 2019 Cost of Data Breach Report from IMB calculated that the global average financial loss from data breaches was around 3.92 million dollars, with the US average estimating 8.19 million dollars (IBM Security, 2019). Businesses can be uprooted entirely if

protected information becomes accessed by outsiders. They now have to put a large amount of money and resources into finding the cause of the information breach and who did it in hopes of mitigating its damages. Lost work needs to be recovered, and business operations must change or even be brought to a complete halt as work needs to be done to fix the mess and disruption caused by the data breach. Even after the company settles the HT, the scar the HT left on the company's reputation won't fade away so quickly. Customers lose trust in the company; therefore, the company loses business. The IMB 2019 Cost of a Data Breach report found that companies that experienced user information leakage had almost a third of their customers discontinue their interaction with the breached company (IBM Security, 2019). As companies get hurt from these data breaches, so do their everyday employees. They might take pay cuts, lay off employees, or even close down the business, leaving everyone who worked for the company jobless. Companies and organizations aren't the only victims of data breaches; subjects of the data breach are also highly impacted.

Other people not directly involved, such as clients or citizens, also face harmful consequences from information breaches. Kieth Rozario from UCSI University in Malaysia even argues during his TED Talk on data breaches that the clients and customers are more a victim than the organization itself since it's *their* privacy that is exposed. Information privacy is about context and choice, and with a data breach such as ones caused by HTs, the affected users have no choice in the releasing of their information into unwanted hands and are given no context for such exposure (Rosario, 2018, 9:24). Any and all information that the government or business possesses has the possibility of being stolen and exposed. No matter how minuscule the breach may seem, one's consent and right to privacy are always broken. In extreme cases such as identity theft, there holds the possibility of completely ruining people's lives. Confidential

information such as from the government is typically kept secret for the sake of society. If this information gets out through data breaches, it can put millions of people in danger and maybe even the country as a whole. Protected and essential people's information could be released, prompting people to try and harm and even kill these people or their families. There's no telling what disasters could happen if certain confidential information gets released, which is all the more reason why it's so essential to implement the multi-parameter approach for the FPGAS HT detection algorithm, which would provide companies and people with better protection against data breaches. While there could be some good to secret information being released in certain contexts, the threat of harm and destruction that can be done with the release of confidential information greatly outweighs its benefits. Besides such incidents as these, there are also less dire reasons for implementing a more efficient HT detection algorithm.

While the difference in power consumption and cost between these two HT detection algorithms might not have as much weight as the successfulness of the algorithm, the extra expense and power consumption from the RMVRF algorithm still influences companies and government organizations. HT detection algorithms implemented on a large-scale can significantly increase power consumption and cost. The extra money and power saved by using the multi-parameter approach for the FPGAs algorithm could be delegated to other valuable projects, protections, or implementations. The extra money could be used to hire more people, produce better products, further enhance other security measures, and in various other beneficial ways. Not only does the multi-parameter approach provide better protection from breaches, but it helps at lessening the resources companies and organizations need to delegate to such protections.

The combination of all of the improvements created by the multi-paramater approach for FPGA hardware Trojan detection algorithm constitutes superiority over the RMVRF runtime monitoring detection algorithm. Assigning the duties of a HT algorithm to multiple methods including runtime monitoring, logic testing, and side channel analysis, as achieved by the multi-parameter approach, proves more full-proof and resourceful than delegating the HT detection to solely a runtime monitoring method as implemented by RMVRF. The substantial advancements achieved on each method in the multi-paramater approach for FPGA hardware trojan detection provides less invasitivy and power consumption than the RMVRF algorithm, while also yielding a more reliable and self improving detection method for HTs. Implementing the most efficient HT deception algorithm supplies the best security from important, confidential, and private information thats present in all aspects of computer science. Protection from hardware Trojans doesn't just help big corporations and the government, is an integral step in protecting everyone inaleable right to privacy.

References

Fournaris, A. P., Pyrgas, L., & Kitsos, P. (2019). An efficient multi-parameter approach for FPGA hardware trojan detection. *Microprocessors and Microsystems*, *71*. https://www.sciencedirect.com/science/article/pii/S0141933118305106

Hill, M. D. (2020) Technical Perspective: Why 'Correct' Computers Can Leak Your Information. *Communications of the ACM,* 63(7), 92. https://cacm.acm.org/magazines/2020/7/245683-technical-perspective-why-correct-computers-can-leak-your-information/fulltext

IBM: Cost of a data breach report 2019. (2019). *Computer Fraud & Security*, *2019*(8), 1–76. https://doi.org/10.1016/s1361-3723(19)30081-8

Mohd, B. J., Abed, S., Hayajneh, T., & Alshayej, M. H. (2021). Run-time monitoring and validation using reverse function (RMVRF) for hardware trojans detection. *IEEE Transactions on Dependable and Secure Computing, 18*(6). https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8941304&tag=1

Pirpilidis, F., Stefanidis, K. G., Voyiatzis, A. G., & Kitsos, P. (2017). On the effects of ring oscillator length and hardware trojan size on an FPGA-based implementation of AES. *Microprocessors and Microsystems, 54*. https://www.sciencedirect.com/science/article/pii/S0141933117300819

Rosario, K. [TED] (2018, October 29). Data Breaches And What To Do About It | Keith Rozario | TEDxUCSIUniversity [Video file]. YouTube. https://www.youtube.com/watch?v=sU2lvkMpv1A