Reliability and Security Assessment of Modern Embedded Devices

Annachiara Ruospo
Supervisor: Prof. Ernesto Sanchez

June 4, 2022

Summary

The complexity of modern embedded systems has increased rapidly over the past few decades. The integration of many various technologies as well as the adoption of more sophisticated algorithms pose significant challenges in ensuring the robustness and the reliability of these systems. Indeed, together with the shrinking of technology nodes, even more systems leverage artificial intelligence based algorithms to cope with their ever-growing computing requirements. In this light, worthy of note is the emerging trend toward the adoption of brain-inspired artificial models, i.e., Artificial Neural Networks (ANNs), in different fields, like automotive, robotic, avionic, etc. If from one side this results to be a very interesting solution, given their outstanding and near-human computational capabilities; from the other side this could be dangerous from the safety point of view. Actually, outsourcing important decisions to them (like deciding whether the car should stop to let a pedestrian pass), can be threatening for the following reason. As a matter of fact, being predictive model, they are not 100% accurate: even in fault-free scenarios, they can provide a wrong answer. Then, despite their very interesting capabilities which are appealing for both industry and academic, it is necessary to deeply investigate the reliability and the behaviours of those systems in faulty scenarios, especially because it has been demonstrated that with the shrinking of the technology, emerging devices are more prone to physical defects.

The primary focus of this Ph.D. thesis is on artificial intelligence-based systems and on the assessment and the improvement of their reliability. This topic has been addressed comprehensively at very different abstraction levels and from different perspectives. Firstly, a characterization of the existing fault models is provided together with the identification of the possible vulnerabilities in ANN-based systems. Then, a key contribution of these research activities is the proposals of different fault injection tools and methodologies to easy and support the reliability assessment process. Specifically, this is done at different levels: by addressing only the ANN model, or by considering the entire system entailing both the ANN software running on a target hardware system. The advantages and disadvantages of the different categories are detailed in the manuscript. At this level, the principal novelties are the identification of the critical bits in different data type representations, the establishment of critical neurons depending on the importance of a neuron as a class- and as a network-dependent

entity; trade-offs analysis on data type representations based on two aspects, i.e., the memory footprint of the application and their reliability. In the end, relying on these analysis and results, strategies to mitigate the effect of faults have been proposed. The first proposed technique aims at redistributing neuronal computations on an AI-based SoC by leveraging integer linear programming. The second mitigation technique is an on-line testing solution based on the adoption of Software Test Libraries coexisting with the requirements of ANN algorithms.

Together with the reliability assessment of neural networks, this Ph.D. thesis covers also a further important topic: the hardware security of modern embedded systems. The complexity of emerging hardware systems led the industry to pursue new development processes: to build a SoC, a recent trend is to rely on third-party IP blocks to keep the cost low and to meet the deadlines. This outsourcing poses increasing concerns regarding the security of modern embedded devices. As an example, these IP blocks might come with malicious circuitry intentionally added by adversaries, namely Hardware Trojans (HTs). They are hidden inside the design and become active under certain rare conditions (such as input sequences). Otherwise, they can be always active since the power on of their host device. Their malicious functionalities can vary from leaking secret information, degrading the circuit's performance, creating backdoors for attackers, and many others. In the literature, new advances and progress have been done in this field, especially in proposing specific detection methodologies to discover hidden Trojans inside a given design. However, the existing gap lies in the HT benchmarks used to validate state-of-the-art methodologies. Indeed, they are obsolete and very simple, unable to represent the complexity of current designs and architectures. For example, they are injected on small 8-bit 8051 microprocessors, or simple cryptographic circuits. In this context, one of the main contributions of this Ph.D. thesis is the release of a set of Hardware Trojans benchmarks at the Register Transfer Level for an open-source pipelined RISC core. They have been designed by following the guidelines for creating hard-to-detect Trojans. Furthermore, the second contribution in this field is a pre-silicon detection methodology for detecting RTL Hardware Trojans. This is build on a deep learning analysis of the dynamic and static properties extracted from the design RTL model. The proposed technique is highly accurate in highlighting suspicious code sections, as reported by the experimental results. It means that, by carefully inspecting them, it is possible to unfold all the described hardware trojans.