World Academy of Science, Engineering and Technology
International Journal of Cognitive and Language Sciences
Vol:17, No:07, 2023

A Machine Learning Approach for Detecting and Locating Hardware Trojans

Authors : Kaiwen Zheng, Wanting Zhou, Nan Tang, Lei Li, Yuanhang He

Abstract : The integrated circuit industry has become a cornerstone of the information society, finding widespread application in areas such as industry, communication, medicine, and aerospace. However, with the increasing complexity of integrated circuits, Hardware Trojans (HTs) implanted by attackers have become a significant threat to their security. In this paper, we proposed a hardware trojan detection method for large-scale circuits. As HTs introduce physical characteristic changes such as structure, area, and power consumption as additional redundant circuits, we proposed a machine-learning-based hardware trojan detection method based on the physical characteristics of gate-level netlists. This method transforms the hardware trojan detection problem into a machine-learning binary classification problem based on physical characteristics, greatly improving detection speed. To address the problem of imbalanced data, where the number of pure circuit samples is far less than that of HTs circuit samples, we used the SMOTETomek algorithm to expand the dataset and further improve the performance of the classifier. We used three machine learning algorithms, K-Nearest Neighbors, Random Forest, and Support Vector Machine, to train and validate benchmark circuits on Trust-Hub, and all achieved good results. In our case studies based on AES encryption circuits provided by trust-hub, the test results showed the effectiveness of the proposed method. To further validate the method's effectiveness for detecting variant HTs, we designed variant HTs using open-source HTs. The proposed method can guarantee robust detection accuracy in the millisecond level detection time for IC, and FPGA design flows and has good detection performance for library variant HTs.

Keywords : hardware trojans, physical properties, machine learning, hardware security
Conference Title : ICSLP 2023 : International Conference on Speech and Language Processing
Conference Location : Stockholm, Sweden
Conference Dates : July 06-07, 2023