From manual to automated testing the energy consumption of embedded systems: A journey and some roadblocks

Nowadays, organizations need to deliver reliable and high-quality embedded products while having to consider more stringent energy constraints. Consider, for instance, the automotive domain where a change to a sub-system or the substitution of a hardware component can affect the system’s energy utilization. In such cases, it is not enough to show the functional correctness of each sub-system’s behavior, but one also needs to provide evidence that the system meets its energy requirements. Therefore, testing the system’s behavior with respect to not exceeding its provided energy budget, the so-called feasibility tests, as well as testing for the worst-case energy consumption are very important for ensuring the quality of service of the embedded system, and estimating its performance.

If software testing is severely constrained, this implies that less time is devoted to ensuring a proper level of software quality. Since load testers have limited resources to manually create test suites for all performance scenarios, researchers have proposed methods for automatically generating test suites by selecting the nominal and potentially troublesome simulations of an embedded system. In contrast to manual testing, test generation is automatic in the sense that test creation satisfying a given test goal is performed automatically. However, over the past few decades, it has been a challenge to develop strong and applicable test generation techniques and tools that are relevant in practice.

As an example, the work I do together with my colleagues at Mälardalen University in Sweden on automatically generating test suites [1] by selecting test cases using random test generation and mutation testing is a solution for improving the efficiency and effectiveness of testing. Specifically, we generate and select test cases based on the concept of energy-aware mutants, small syntactic modifications in the system architecture, intended to mimic real energy faults. Test cases that can distinguish a certain behavior from its mutations are sensitive to changes, and hence considered to be good at detecting faults. We applied this method on a brake by wire system and our results suggest that an approach that selects test cases showing diverse energy consumption can increase the fault detection ability. This kind of results should motivate both academia and industry to investigate the use of automatic test generation for energy consumption and identify the empirical evidence for, or against, the use of it in practice when developing industrial control software.