

Report for 08-Towards Effective Adaptive Random Testing for Higher-Dimensional Input Domains

Xinyu Chen

March 31, 2017

1 Summary

This paper argues the reference method in the Distance-Based Adaptive Random Testing(D-ART) is not effective for higher dimensional input domains. They proposes a new method ID-D-ART(Increasing Domain) to improve the efficiency of Adaptive Random Testing. The authors argue that the efficiency of D-ART drops in higher dimensions is because the search domain of D-ART is too large. They propose the ID-D-ART which starts searching from the domain center, and increases the search domain during iterations. They tested the efficiency by a lot of simulations and concluded ID-D-ART was more efficient for the block failure pattern. The Random Testing looks at the test cases and failure cases in a geometric way is an interesting approach. The D-ART and ID-D-ART is analogous to clustering algorithms in the aspect that they do not scale well with high dimensional data. The method D-ART uses to choose test case is similar with how k-means++ chooses cluster medoids. But the metrics of f-measure is quite different.

2 Key Takeaway

The authors used some simulations to test their improved method. However, they assume the block failure patterns locate near the center of the domain. This assumption does not always hold. The take away is sometimes our

experiments and results may look good, but we should be careful to check our models. There may be some pitfalls in our assumptions.

3 Discussions

The paper is interesting. The following questions are not clear and relevant to their researches.

- *Failure Pattern.* The authors argue Adaptive Random Testing is good. This is reasonable when they know the shape of failure pattern a priori. However, when we know nothing about the domain, is the search strategy that assumes uniform random distribution more efficient?
- *Simulation.* The authors designed a simulation. They mutated some programs to generate *bug* programs. Although the number is large, we still cannot be assure the family of mutated programs can generalize the failure patterns. It is very possible they are only testing on a family of specific center-based failure patterns programs.
- *The metrics.* The authors used a f-measure to describe how their method can help to improve testing tasks. Some later papers argue that although the Adaptive Random Testing uses less iterations to find the first failure, but in each iteration they do much computations.
- *High dimensional input.* The Random Testing looks the test cases in a geometrics way. This is very interesting. They didn't explain why the width of hyper-cuboid increases in higher dimensional spaces. This is similar to clustering tasks in some extent . The method does not scale well even when dimensionality only goes to 6 or 8. That is much smaller than k-means algorithms can scale in high dimensional spaces.