

## **Diversity, Diversity, Diversity: Three Words that Summarize the Impacts of Adaptive Random Testing**

T.H. Tse  
Department of Computer Science  
The University of Hong Kong  
Pokfulam, Hong Kong  
Email: thtse@cs.hku.hk

# **Adaptive Random Testing**

**The ART of Diversity** in Test Cases  
**The ART of Diversity** in Techniques  
**The ART of Diversity** in Applications

When we consider real estate, we know that “There are three things that matter in property: location, location, location” — a statement attributed to Lord Harold Samuel in 1944.

The article titled “Adaptive random testing: The ART of test case diversity” by T.Y. Chen, F.-C. Kuo, R.G. Merkel, and T.H. Tse was published in *Journal of Systems and Software* in 2010. We would like to mimic Lord Samuel and summarize the impacts and influences of our paper in three words: “diversity, diversity, diversity”. Intuitively speaking, the most effective way to reveal software failures is through the design of diversified test cases, because programmers may make all kinds of mistakes.

On the one hand, the impacts of the ART 2010 paper are at the fundamental level. It has enhanced the foundations of software testing, and opened up new research directions. On the other hand, the 2010 paper has also strongly influenced researchers to work on new ART algorithms, and motivated practitioners to apply ART to real-life software testing. These impacts have become apparent in the decade following the publication of the ART 2010 paper, as illustrated in this blog entry.

### **1. Adaptive Random Testing: The ART of Diversity in Test Cases**

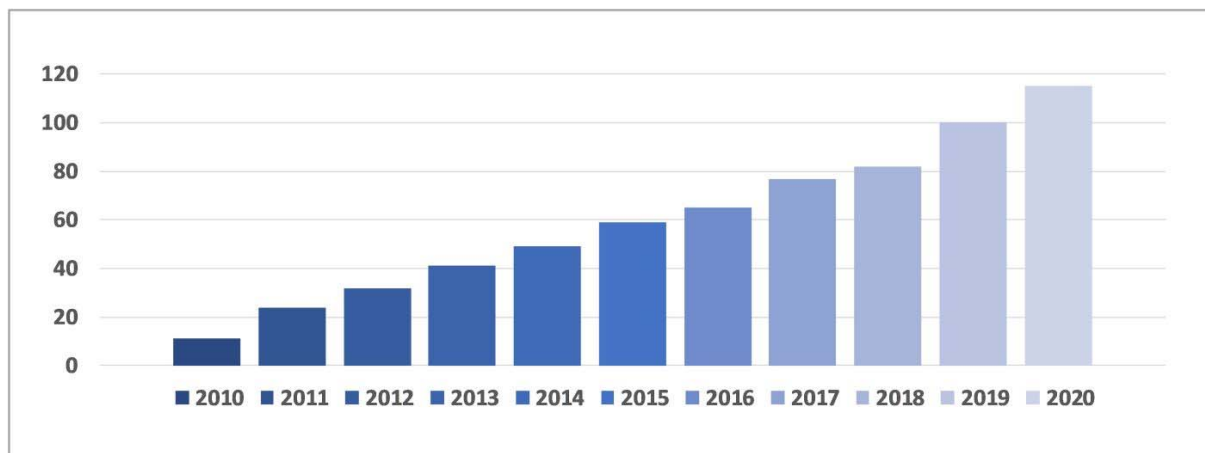
Random testing (RT) has been recognized as important in software quality assurance. It attempts to exercise a diversified choice of test cases through random selection. However, it has been criticized for insufficiently effective as it does not make use of any available information to guide the selection of test cases. Adaptive random testing (ART) was proposed by T.Y. Chen in the late 1990s to enhance random diversification of test cases by algorithmically spreading them more evenly.

The ART 2010 paper published in *Journal of Systems and Software* was a reflection on adaptive random testing after a decade of its inception. The article discussed the importance of test case diversity, justified by research results demonstrating that the simple notion of even spreading across the input domain can significantly enhance failure detection capability in software testing.

The concept of “even spreading across the input domain” is easily understood by any person, but is difficult to describe precisely. The difficulty lies in the fact that there are many perspectives to achieve such a phenomenon. Hence, numerous ART algorithms have been developed from various approaches or intuitions. For instance, the first and most popular ART algorithm, FSCS-ART, is based on the approach to choosing the farthest among a set of potential candidates as the next test case in order to achieve this phenomenon.

Thus, adaptive random testing has evolved into a large family of enhanced random testing methods built on the very simple notion of diversity and even spreading. This notion leverages on the phenomenon that neighboring inputs are more likely to behave similarly, because most programs mimic our daily lives, which are rarely stochastic.

## 2. Adaptive Random Testing: The ART of Diversity in Techniques



Recognizing the limitations in efficiency of the initial techniques, numerous papers have been published since 2010, proposing improved ART algorithms. Some notable examples are ant colony optimization [4], ART by exclusion through test profile [15], ART4SQLi [25], ARTsum [3], category selection-based ART [27], clustering [23], code coverage [7], divide and conquer [9], distance-aware forgetting strategy [17], dynamic partitioning [10], KD-ART [22], KDFC-ART [18], OMISS-ART [5], orthogonal recursive bisection [1], randomized quasi-random testing [14], and static mirror adaptive random testing [21]. For instance, ART by exclusion [15] is based on the concept of predefining a criterion and taking the first candidate that satisfies this criterion as the next test case. The ART algorithms by partitioning [10] are sparked by the intuition that the next test case should be selected from a partition without test cases.

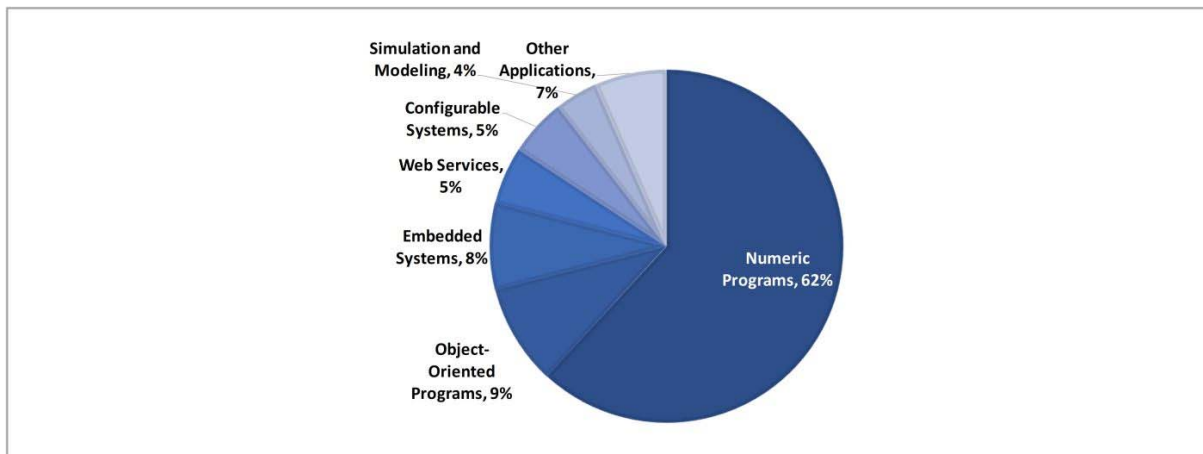
In particular, ART by bisection [1] and RBCVT-Fast using clustering by centroidal Voronoi tessellations [23] are linear-time algorithms for numeric programs. ARTsum [3] and category selection-based ART [26] are linear-time algorithms for nonnumeric programs. Two independent projects are currently working on linear-time algorithms for object-oriented programs.

Different ART methodologies have different favorable and unfavorable testing scenarios. Given the large family of techniques, the good news is that you can choose a cost-effective ART algorithm for the particular program under test. The bad news is that there is no free lunch — you need to understand ART techniques as well as the basic characteristics of the program under test before you can select the most appropriate one.

In addition to motivating the development of more effective and efficient algorithms, the ART 2010 paper has also suggested new insights, interpretations, and research directions. Most notably, the use of failure patterns, such as the distribution of failure-causing inputs, has given rise to the novel research direction of failure-based software testing [20]. Given such knowledge, we may successfully construct adversarial attacks of machine learning systems to test their robustness. Another example of failure patterns is the conceived geometry of the failure areas. This allows us to utilize any minimal information on hand to compare the effectiveness of software testing approaches, without the need to know the detailed algorithms or procedures.

Random ordering and random sequences have been commonly used as benchmarks for comparing test sequences generated by various algorithms in different problem contexts. Following ART, the new concept of adaptive random sequences [6][26] has evolved as a more effective alternative than random sequences.

### 3. Adaptive Random Testing: The ART of Diversity in Applications



ART has been adopted in numerous application domains. A plausible reason is that RT can be used alone or with other testing methods. It does not need the source code. As ART is an alternative to RT, it can be adopted in many application domains as well. The only limitation is the need for test oracles. With the increasing acceptance of metamorphic testing [8] as an effective means to alleviate the oracle problem, the applicability of RT and ART can be extended to the domains involving oracle problems.

Huawei has applied ART to an industrial case study involving 50 million test cases [27]. Originally, it took 50 days to test and analyze. It was necessary to reduce the number of test cases while revealing more failures. At the same time, control flow, data flow, or code instrumentation could not be applied for reasons of privacy and safety. Based on the linear time algorithm ARTsum [3], category selection-based ART (CSBART) was proposed, where partial categories were selected to compute test case distances. Experimentation showed that CSBART outperformed a more complex K-means clustering strategy.

Simula Research Laboratory has successfully applied ART to an industrial real-time embedded system [2][13][19]. Despite Simula's initial observation regarding the inefficiency of ART, the findings from the industrial system were positive. "Results on artificial problems and one industrial case study show that Adaptive Random Testing performs best." [13] "One main advantage of ART and SBT is that [they] can be tailored to whatever time and resources are available for testing: when resources are expended and time is up, we can simply stop their application without any side effect." [2] "Adaptive random output diversity algorithm is able to identify different failure types better than the search-based algorithms when test suites are above a certain size." [19]

Other notable practical applications include object-oriented programs [5], web services [24], configurable systems [3][7], mobile applications [16], database applications [25], regression testing [6], combinatorial testing [12], and mutation testing [22].

Furthermore, based on the new research direction proposed in the 2010 paper, a collaborative project between the industry and the academia is being conducted to study how failure patterns can help detect obstacle perception problems in automatic driving systems.

#### **4. Support, Support, Support: Three Further Things that Matter in ART**

There are three further things that matter in adaptive random testing: your support, your support, your support.

- ♦ If you are a researcher, we need your support in proposing better ART approaches and empirically evaluating the family of such techniques.
- ♦ If you are a practitioner, we need your support in applying various ART methodologies to large-scale industrial applications and sharing your experience with us.
- ♦ In either case, we need your immediate support in clicking the "applause" button at the end of the blog entry.

#### **References**

- [1] Ackah-Arthur, H., Chen, J., Towey, D., Omari, M., Xi, J., Huang, R., 2019. One-domain-one-input: Adaptive random testing by orthogonal recursive bisection with restriction. *IEEE Transactions on Reliability* 68 (4), 1404–1428.
- [2] Arcuri, A., Iqbal, M.Z., Briand, L., 2010. Black-box system testing of real-time embedded systems using random and search-based testing. In: *Testing Software and Systems: Proceedings of the 22nd IFIP WG 6.1 International Conference (ICTSS 2010)*, Lecture Notes in Computer Science, vol. 6435. Springer, Berlin, Germany, pp. 95–110.
- [3] Barus, A.C., Chen, T.Y., Kuo, F.-C., Liu, H., Merkel, R.G., Rothermel, G., 2016. A cost-effective random testing method for programs with non-numeric inputs. *IEEE Transactions on Computers* 65 (12), 3509–3523.

- [4] Bidgoli, A.M., Haghighi, H., 2020. Augmenting ant colony optimization with adaptive random testing to cover prime paths. *Journal of Systems and Software* 161, 110495.1–110495.17.
- [5] Chen, J., Kuo, F.-C., Chen, T.Y., Towey, D., Su, C., Huang, R., 2017. A similarity metric for the inputs of OO programs and its application in adaptive random testing. *IEEE Transactions on Reliability* 66 (2), 373–402.
- [6] Chen, J., Zhu, L., Chen, T.Y., Towey, D., Kuo, F.-C., Huang, R., Guoa, Y., 2018. Test case prioritization for object-oriented software: An adaptive random sequence approach based on clustering. *Journal of Systems and Software* 135, 107–125.
- [7] Chen, T.Y., Kuo, F.-C., Liu, H., Wong, W.E., 2013. Code coverage of adaptive random testing. *IEEE Transactions on Reliability* 62 (1), 226–237.
- [8] Chen, T.Y., Kuo, F.-C., Liu, H., Poon, P.-L., Towey, D., Tse, T.H., Zhou, Z.Q., 2018. Metamorphic testing: A review of challenges and opportunities. *ACM Computing Surveys* 51 (1), 4:1–4:27.
- [9] Chow, C., Chen, T.Y., Tse, T.H., 2013. The ART of divide and conquer: An innovative approach to improving the efficiency of adaptive random testing. In: *Proceedings of the 13th International Conference on Quality Software (QSIC 2013)*, IEEE Computer Society, Los Alamitos, CA, 268–275.
- [10] Huang, R., Liu, H., Xie, X., Chen, J., 2015. Enhancing mirror adaptive random testing through dynamic partitioning. *Information and Software Technology* 67, 13–29.
- [11] Huang, R., Sun, W., Xu, Y., Chen, H., Towey, D., Xia, X., 2019. A survey on adaptive random testing. *IEEE Transactions on Software Engineering*. DOI: <https://doi.org/10.1109/TSE.2019.2942921>.
- [12] Huang, R., Xie, X., Chen, T.Y., Lu, Y., 2012. Adaptive random test case generation for combinatorial testing. In: *Proceedings of the IEEE 36th Annual International Computer Software and Applications Conference (COMPSAC 2012)*. IEEE Computer Society, Los Alamitos, CA, pp. 52–61.
- [13] Iqbal, M.Z.Z., Arcuri, A., Briand, L.C., 2011. Automated system testing of real-time embedded systems based on environment models. Technical Report 2011-19, Simula Research Laboratory, Lysaker, Norway.
- [14] Liu, H., Chen, T.Y., 2016. Randomized quasi-random testing. *IEEE Transactions on Computers* 65 (6), 1896–1909.
- [15] Liu, H., Xie, X., Yang, J., Lu, Y., Chen, T.Y., 2010. Adaptive random testing by exclusion through test profile. In: *Proceedings of the 10th International Conference on Quality Software (QSIC 2010)*. IEEE Computer Society, Los Alamitos, CA, pp. 92–101.
- [16] Liu, Z., Gao, X., Long, X., 2010. Adaptive random testing of mobile application. In: *Proceedings of the 2nd International Conference on Computer Engineering and Technology (ICCET 2010)*, vol. 2. IEEE, Piscataway, NJ, pp. 297–301.

- [17] Mao, C., Chen, T.Y., Kuo, F.-C., 2017. Out of sight, out of mind: A distance-aware forgetting strategy for adaptive random testing. *Science China Information Sciences* 60, 092106:1–092106:21.
- [18] Mao, C., Zhan, X., Tse, T.H., Chen, T.Y., 2019. KDFC-ART: a KD-tree approach to enhancing Fixed-size-Candidate-set Adaptive Random Testing. *IEEE Transactions on Reliability* 68 (4), 1444–1469.
- [19] Matinnejad, R., Nejati, S., Briand, L.C., Bruckmann, T., 2015. Effective test suites for mixed discrete-continuous Stateflow controllers. In: *Proceedings of the 10th Joint Meeting on Foundations of Software Engineering (ESEC/FSE-10)*. ACM, New York, NY, pp. 84–95.
- [20] Merkel, R.G., Kuo, F.-C., Chen, T.Y., 2011. An analysis of failure-based test profiles for random testing. In: *Proceedings of the 35th Annual International Computer Software and Applications Conference (COMPSAC 2011)*. IEEE Computer Society, Los Alamitos, CA, pp. 68–75.
- [21] Omari, M., Chen, J., Ackah-Arthur, H., Kudjo, P.K., 2019. Elimination by linear association: An effective and efficient static mirror adaptive random testing. *IEEE Access* 7, 71038–71060.
- [22] Patrick, M., Jia, Y., 2017. KD-ART: Should we intensify or diversify tests to kill mutants? *Information and Software Technology* 81, 36–51.
- [23] Shahbazi, A., Tappenden, A.F., Miller, J., 2013. Centroidal Voronoi tessellations: A new approach to random testing. *IEEE Transactions on Software Engineering* 39 (2), 163–183.
- [24] Tappenden, A.F., Miller, J., 2014. Automated cookie collection testing. *ACM Transactions on Software Engineering and Methodology* 23 (1), 3:1–3:40.
- [25] Zhang, L., Zhang, D., Wang, C., Zhao, J., Zhang, Z., 2019. ART4SQLi: The ART of SQL injection vulnerability discovery. *IEEE Transactions on Reliability* 68 (4), 1470–1489.
- [26] Zhang, X., Xie, X., Chen, T.Y., 2016. Test case prioritization using adaptive random sequence with category-partition-based distance. In: *Proceedings of the 2016 IEEE International Conference on Software Quality, Reliability and Security (QRS 2016)*. IEEE Computer Society, Los Alamitos, CA, pp. 374–385.
- [27] Zhang, Z., Wang, Y., Wang, Z., Qian, J., 2019. How to effectively reduce tens of millions of tests: An industrial case Study on adaptive random testing. *IEEE Transactions on Reliability* 68 (4), 1429–1443.