Adaptive Test Generation with Q-grams

Citation for the original published paper (version of record):
Biagiola, M., Feldt, R., Tonella, P. (2026). Adaptive Test Generation with Q-grams. Lecture Notes in Informatics, P-377: 129-130. http://dx.doi.org/10.18420/se2026-47

(article starts on next page)

Adaptive Test Generation with Q-grams

Matteo Biagiola [1], Robert Feldt [2], and Paolo Tonella [3]

Abstract: In this article, we summarize our paper *Adaptive Random Testing with Q-grams: The Illusion Comes True* [BFT25], which has been accepted for presentation at the 33rd ACM International Conference on the Foundations of Software Engineering (FSE 2025).

Keywords: Software testing, Adaptive random testing, Diversity-based testing

1 Summary

Adaptive Random Testing (ART) has been widely criticized for its computational inefficiency, most notably by Arcuri and Briand. Their analysis shows that ART incurs a quadratic increase in distance computations with the number of test executions, which severely limits its scalability in scenarios requiring extensive testing. Simulation studies reinforce this point, demonstrating that the cost of these distance calculations often outweighs the benefits ART provides. In particular, when a fault leads a failure only with very low probability, the extra computational overhead of ART is worthwhile only if test execution time is unrealistically larger than the cost of computing distances.

Although several ART variants attempt to reduce this overhead, they often do so at the expense of fault-detection effectiveness, lack complexity guarantees, or apply only to specific input domains such as numerical or discrete data.

In this paper, we present a new framework for adaptive random testing that eliminates pairwise distance computations by maintaining a compact summary of previous executions, for example, by counting Q-grams observed in earlier test runs. When a new test input is generated we only compare it once against this aggregation, instead of comparing it against all previously executed tests. Considering a set of candidate test cases to execute, our framework selects the most diverse (e.g., by maximizing the entropy of Q-gram counts) candidate, and incrementally updates the aggregate representation (e.g., by adding the counts of the selected test), effectively reducing computational complexity from quadratic to linear.

[1] Università della Svizzera Italiana (USI) and University of St. Gallen (HSG), Software Institute/Institute of Computer Science, Lugano/St. Gallen, Switzerland,
matteo.biagiola@{usi,unisg}.ch, https://orcid.org/0000-0002-7825-3409

[2] Chalmers University of Technology, Department of Computer Science and Engineering, Gothenburg, Sweden,
robert.feldt@chalmers.se, https://orcid.org/0000-0002-5179-4205

[3] Università della Svizzera italiana (USI), Software Institute, Lugano, Switzerland,
paolo.tonella@usi.ch, https://orcid.org/0000-0003-3088-0339

doi:10.18420/se2026-47

Across a benchmark of six web applications, our Q-gram-based ART achieves, on average, four times more unique target coverage than random testing and 3.5 times more than traditional distance-based ART methods.

2 Data Availability

The replication package of our work is available at https://github.com/testingautomated-usi/adaptive-tg-qgrams to facilitate independent reproduction of the experiments.

References

[BFT25] Biagiola, M.; Feldt, R.; Tonella, P.: Adaptive Random Testing with Q-grams: The Illusion Comes True. Proc. ACM Softw. Eng. 2 (FSE), pp. 533–553, 2025, https://doi.org/10.1145/3715740.