Predictive Analytics for Software Testing

Keynote Paper*

Federica Sarro
Department of Computer Science, University College London
London, United Kingdom
f.sarro@ucl.ac.uk

ABSTRACT

This keynote discusses the use of Predictive Analytics for Software Engineering, and in particular for Software Defect Prediction and Software Testing, by presenting the latest results achieved in these fields leveraging Artificial Intelligence, Search-based and Machine Learning methods, and by giving some directions for future work.

CCS CONCEPTS

• **Software and its engineering → Search-based software engineering**;

KEYWORDS

Predictive Analytics, Search-based Predictive Modelling

ACM Reference Format:
Federica Sarro. 2018. Predictive Analytics for Software Testing: Keynote Paper. In *SBST'18: IEEE/ACM 11th International Workshop on Search-Based Software Testing , May 28–29, 2018, Gothenburg, Sweden.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3194718.3194730

1 INTRODUCTION

The ever increasing volume of data produced by realising and using software calls for a new generation of analytical techniques that can help software engineers better understand their software processes, products and customers. Devising effective Predictive Analytics for Software Engineering holds the promise to maximise product quality, users' satisfaction and revenues.

This keynote talk illustrates how search-based heuristics can be used to build effective predictive models for software engineering [2, 7] and highlights some of the recent developments in the production of adaptive [6, 11, 12], human-competitive [13] and scalable solutions [3, 4, 10].

The keynote also introduces the mutually beneficial relationship between predictive modelling and software testing by illustrating new ways to strengthen current testing practise through the use of predictive models [14] and, viceversa, to construct better predictive models through the use of information available from software testing activities [1].

The talk concludes discussing how text mining, analytics and search-based techniques can be combined to obtain more efficient test automation, for example, exploiting artefacts written in natural language to infer behavioural models of software applications and to discover issues earlier in the process [5, 8, 9, 15].

*This one page paper provides an outline of the keynote talk given by Federica Sarro at SBST'18, with pointers to the literature for details of the results covered.

SBST'18, May 28–29, 2018, Gothenburg, Sweden

ACM ISBN 978-1-4503-5741-8/18/05.
https://doi.org/10.1145/3194718.3194730

ACKNOWLEDGMENTS

This work is supported by the EPSRC Programme Grant EP/J017515.

REFERENCES

[1] D. Bowes, T. Hall, M. Harman, Y. Jia, F. Sarro, and F. Wu. 2016. Mutation-aware Fault Prediction. In *Procs. of the 25th International Symposium on Software Testing and Analysis (ISSTA'16)*. ACM, 330–341. https://doi.org/10.1145/2931037.2931039
[2] F. Ferrucci, M. Harman, and F. Sarro. 2014. Search-Based Software Project Management. In *Software Project Management in a Changing World*. Springer, 373–399.
[3] F. Ferrucci, P. Salza, and F. Sarro. 2017. Using Hadoop MapReduce for Parallel Genetic Algorithms: A Comparison of the Global, Grid and Island Models. *Evolutionary Computation* (2017), 1–33. https://doi.org/10.1162/evco_a_00213
[4] L. Di Geronimo, F. Ferrucci, A. Murolo, and F. Sarro. 2012. A Parallel Genetic Algorithm Based on Hadoop MapReduce for the Automatic Generation of JUnit Test Suites. In *Procs. of the 5th International Conference on Software Testing, Verification and Validation (ICST'12)*. 785–793. https://doi.org/10.1109/ICST.2012.177
[5] M. Harman, A. Al-Subaihin, Y. Jia, W. Martin, F. Sarro, and Y. Zhang. 2016. Mobile App and App Store Analysis, Testing and Optimisation. In *Procs. of the International Conference on Mobile Software Engineering and Systems (MOBILESoft'16)*. ACM, 243–244. https://doi.org/10.1145/2897073.2897076
[6] M. Harman, S. Islam, Y. Jia, L. L. Minku, F. Sarro, and K. Srivisut. 2014. Less is More: Temporal Fault Predictive Performance over Multiple Hadoop Releases. In *Procs. of the International Symposium on Search-Based Software Engineering (SSBSE'14)*. Springer, 240–246. https://doi.org/10.1007/978-3-319-09940-8_19
[7] Ruchika M., Megha K., and Rajeev R. R. 2017. On the application of search-based techniques for software engineering predictive modeling: A systematic review and future directions. *Swarm and Evolutionary Computation* 32 (2017), 85 – 109.
[8] W. Martin, F. Sarro, and M. Harman. 2016. Causal Impact Analysis for App Releases in Google Play. In *Procs. of the 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering (FSE 2016)*. ACM, 435–446. https://doi.org/10.1145/2950290.2950320
[9] W. Martin, F. Sarro, Y. Jia, Y. Zhang, and M. Harman. 2017. A Survey of App Store Analysis for Software Engineering. *IEEE Trans. Software Engineering* 43, 9 (2017), 817–847. https://doi.org/10.1109/TSE.2016.2630689
[10] S. Di Martino, F. Ferrucci, V. Maggio, and F. Sarro. 2013. Towards Migrating Genetic Algorithms for Test Data Generation to the Cloud. (2013), 113–135. https://doi.org/10.4018/978-1-4666-2536-5.ch006
[11] F. Sarro, S. Di Martino, F. Ferrucci, and C. Gravino. 2012. A Further Analysis on the Use of Genetic Algorithm to Configure Support Vector Machines for Inter-release Fault Prediction. In *Procs. of the 27th Annual ACM Symposium on Applied Computing (SAC'12)*. ACM, 1215–1220. https://doi.org/10.1145/2245276.2231967
[12] F. Sarro, F. Ferrucci, M. Harman, A. Manna, and J. Ren. 2017. Adaptive Multi-Objective Evolutionary Algorithms for Overtime Planning in Software Projects. *IEEE Trans. Software Eng.* 43, 10 (2017), 898–917. https://doi.org/10.1109/TSE.2017.2650914
[13] F. Sarro, A. Petrozziello, and M. Harman. 2016. Multi-objective Software Effort Estimation. In *Procs.of the 38th International Conference on Software Engineering (ICSE'16)*. ACM, 619–630. https://doi.org/10.1145/2884781.2884830
[14] J. Zhang, Z. Wang, L. Zhang, D. Hao, L. Zang, S. Cheng, and L. Zhang. 2016. Predictive Mutation Testing. In *Procs. of the 25th International Symposium on Software Testing and Analysis (ISSTA'16)*. ACM, 342–353.
[15] Y. Zhang, M. Harman, Yue J., and F. Sarro. 2015. Inferring test models from Kate's bug reports using multi-objective search. In *Procs. of the International Symposium on Search Based Software Engineering (SSBSE'15)*. Springer, 301–307.