测试成本和缺陷检测能力的评价;在调研文献中,41%的研究只关注了单个类型的评价指标,但测试成本、缺陷检测能力和测试覆盖任意一种类型的评价都不够充分度量测试的质量,而在 Mix 混合主题中,75%的研究关注两个及两个以上类型的评价.

建议:

● 持续集成测试用例集优化的已有研究积累了较多的研究成果,在不同研究主题的研究取得了突破,但仍然存在不足,诸如时间成本太高、遗漏一些必要的测试,目前还少有研究在减少测试成本的同时,考虑如何保证缺陷的检测能力和测试的覆盖度.因此,综合考虑这三者的性能,研究具有较好成本效益的持续集成测试集优化方法,具有非常好的价值和应用前景.

● 已有研究提出的方法对大型项目,由于耗时、复杂等原因,限制了它的有效应用,但大型项目由于系统复杂、庞大,无选择地运行所有集成测试非常不经济,甚至不可能,所以更需要测试用例的优化选择.因此,大型项目下持续集成测试集优化仍然是目前该领域所面临的挑战.

● 动态方法追求精准测试选择,但由于耗时、第三方非预期中断等可能导致不安全或者不完整的测试;静态方法由于分析方法的局限性,有可能导致过多或者过少的测试,两者都有优缺点.但近年的研究 S25 表明,改进静态方法也可能获得全面的测试覆盖,并且成本效益更好,所以基于静态技术测试优化选择,可能是未来一个有价值的研究方向.另外,我们认为融合静态和动态的方法,譬如以静态技术搜索全局,以动态技术优化局部,也可能是一个值得研究的方向.

(2) 异构环境下的研究:研究文献指出,随着 Web、服务应用等程序的广泛使用,这类异构系统与传统软件存在两方面的差异:异质性和环境依赖.在调研文献中,多语言环境下的研究占调研文献的 15%,并且已有研究仍然存在一些问题.另外,已有研究中只有两项考虑了复杂异构环境下特定影响因子(例如数据库、配置文件、项目之间依赖、网络布局等).但研究指出,传统的以代码为中心的研究方法,忽视了复杂环境下特定因子,可能造成测试不完整和发布软件的质量不高的风险.

建议:随着 Web 服务应用等程序的广泛使用,复杂异构环境下特定的持续集成测试集优化,是面临的新的挑战.因此,除了传统的影响因子(包括测试历史、覆盖信息、代码变化和依赖关系等),还应当考虑复杂异构环境下特定的影响因子,提出一些新的研究方法,例如基于多语言环境下持续集成测试优化、非代码制品的文本检索、主题模型方法进行测试用例排序.

(3) 已有研究方法在工业界的验证:在已有研究中,只有 18%的研究在大型工业数据集上进行了验证,其他研究仅在中小型工业数据集或开源数据集上进行了验证,可能在大型集成项目上并不适用.此外,一些研究工作只是在单个数据集上验证了方法的有效性,这些研究的验证应该额外增加真实的场景或标准的实例.

建议:首先,应当鼓励研究者与工业界合作,把已有的研究应用到大型的工业项目中,是当务之急也是挑战.另外,应当鼓励研究者提供公开数据集,为其他相关研究提供比较数据.最后可能缺少一个公共技术框架,用来比较常见的测试用例集优化方法.

5 总 结

DevOps 的出现和广泛应用,使得持续集成越来越受到重视.在有限的资源约束和不降低测试质量的前提下,通过优化测试用例集来减少持续集成时间,已成为学术界和工业界研究的热点.

本文的研究从 4 个电子数据库中检索近十年的文献,共选取 39 篇重要文献,调研持续集成测试用例集优化的研究进展.从研究主题、影响因子、研究方法、研究对象和性能评价 5 个方面提取数据,定量分析回答设定的研究问题.总之,持续集成测试用例集优化的已有研究积累了较多的研究成果,但仍然存在一些问题,目前还少有研究在减少测试成本的同时,保证测试的缺陷检测能力和覆盖度.此外,大型复杂异构环境下的持续集成测试优化、已有研究在工业界的推广仍然是目前工业界和学术界面临的重大挑战.

References:

[1] Hilton M, Nelson N, Tunnell T, Marinov D, Dig D. Trade-Offs in continuous integration: Assurance, security, and flexibility. In: Proc. of the Joint Meeting on Foundations of Software Engineering. 2017. 197–207.

[2] Vasilescu B, Yu Y, Wang H, Devanbu P, Filko V. Quality and productivity outcomes relating to continuous integration in GitHub. In: Proc. of the Joint Meeting on Foundations of Software Engineering. 2015. 805–816.

[3] Hilton M, Tunnell T, Huang K, Marinov D, Dig D. Usage, costs, and benefits of continuous integration in open-source projects. In: Proc. of the IEEE/ACM Int'l Conf. on Automated Software Engineering. 2016. 426–437.

[4] Martensson T, Stahl D, Bosch J. Continuous integration impediments in large-scale industry projects. In: Proc. of the IEEE Int'l Conf. on Software Architecture. IEEE, 2017. 169–178.

[5] Virmani M. Understanding DevOps & bridging the gap from continuous integration to continuous delivery. In: Proc. of the 5th Int'l Conf. on Innovative Computing Technology (INTECH). 2015. 78–82.

[6] Rathod N, Surve A. Test orchestration a framework for continuous integration and continuous deployment. In: Proc. of the Int'l Conf. on Pervasive Computing. 2015. 1–5.

[7] Kitchenham BA, Charters S. Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report, EBSE-2007-01, ICSE IEEE Computer Society, 2007. 1051–1052.

[8] Badampudi D, Wohlin C, Kai P. Experiences from using snowballing and database searches in systematic literature studies. In: Proc. of the Int'l Conf. on Evaluation and Assessment in Software Engineering. 2015. 1–10.

[9] Kitchenham BA, Dyba T, Jorgensen M. Evidence-Based software engineering. In: Proc. of the Int'l Conf. on Software Engineering. IEEE Computer Society, 2004. 273–281.

[10] Yoo S, Harman M. Regression testing minimization, selection and prioritization: A survey. Software Testing, Verification and Reliability, 2012,22(2):67–120.

[11] Kazmi R, Jawawi DNA, Mohamad R, Ghani I. Effective regression test case selection: A systematic literature review. ACM Computing Surveys, 2017,50(2):29.

[12] Nanda A, Mani S, Sinha S, Harrold MJ, Orso A. Regression testing in the presence of non-code changes. In: Proc. of the 4th IEEE Int'l Conf. on Software Testing, Verification and Validation. 2011. 21–30.

[13] Qiu D, Li B, Ji S, Leung H. Regression testing of Web service: A systematic mapping study. ACM Computing Surveys, 2015,47(2):21.

[14] Stratis P, Rajan A. Test case permutation to improve execution time. In: Proc. of the IEEE/ACM Int'l Conf. on Automated Software Engineering. 2016. 45–50.

[15] Huang CY, Chen CS, Lai CE. Evaluation and analysis of incorporating fuzzy expert system approach into test suite reduction. Information and Software Technology, 2016,79(C):79–105.

[16] Qu X, Cohen MB, Woolf KM. Combinatorial interaction regression testing: A study of test case generation and prioritization. In: Proc. of the IEEE Int'l Conf. on Software Maintenance. 2007. 255–264.

[17] Engström E, Runeson P, Skoglund M. A systematic review on regression test selection techniques. Information and Software Technology, 2010,52(1):14–30.

[18] Catal C, Mishra D. Test case prioritization: A systematic mapping study. Software Quality Journal, 2013,21(3):445–478.

[19] Soetens QD, Demeyer S, Zaidman A, Pérez J. Change-Based test selection: An empirical evaluation. Empirical Software Engineering, 2016,21(5):1990–2032.

[20] Blondeau V, Etien A, Anquetil N, Cresson S, Croisy P, Ducasse S. Test case selection in industry: An analysis of issues related to static approaches. Software Quality Journal, 2016, 1–35.

[21] Spieker H, Gotlieb A, Marijan D, Mossige M. Reinforcement learning for automatic test case prioritization and selection in continuous integration. In: Proc. of the 26th ACM SIGSOFT Int'l Symp. on Software Testing and Analysis. 2017. 12–22.

[22] Busjaeger B, Xie T. Learning for test prioritization: An industrial case study. In: Proc. of the 24th ACM SIGSOFT Int'l Symp. on Foundations of Software Engineering. 2016. 975–980.

[23] Elbaum S, Rothermel G, Penix J. Techniques for improving regression testing in continuous integration development environments. In: Proc. of the 22nd ACM SIGSOFT Int'l Symp. on Foundations of Software Engineering. 2014. 235–245.

[24] Parsai A, Soetens QD, Murgia A, Demeyer S. Considering polymorphism in change-based test suite reduction. In: Proc. of the Int'l Conf. on Agile Software Development. Cham: Springer-Verlag, 2014. 166–181.

[25] Shi A, Thummalapenta S, Lahiri SK, Bjorner N, Czerwonka J. Optimizing test placement for module-level regression testing. In: Proc. of the IEEE/ACM Int'l Conf. on Software Engineering. IEEE Computer Society, 2017. 689–699.

[26] Memon A, Gao Z, Nguyen B, Dhanda S, Nickell E, Siemborski R, Micco J. Taming Google-scale continuous testing. In: Proc. of the 39th Int'l Conf. on Software Engineering. 2017. 233–242.

[27] Marijan D, Liaaen M, Gotlieb A, Sen S, Ieva C. TITAN: Test suite optimization for highly configurable software. In: Proc. of the IEEE Int'l Conf. on Software Testing, Verification and Validation (ICST). 2017. 524–531.

[28] Ekelund ED, Engstrom E. Efficient regression testing based on test history: An industrial evaluation. In: Proc. of the IEEE Int'l Conf. on Software Maintenance and Evolution. 2015. 449–457.

[29] Noor TB, Hemmati H. A similarity-based approach for test case prioritization using historical failure data. In: Proc. of the 26th IEEE Int'l Symp. on Software Reliability Engineering (ISSRE). 2015. 58–68.

[30] Hemmati H, Fang Z, Mantyla MV. Prioritizing manual test cases in traditional and rapid release environments. In: Proc. of the 8th IEEE Int'l Conf. on Software Testing, Verification and Validation (ICST). 2015. 1–10.

[31] Herzig K, Greiler M, Czerwonka J, Murphy B. The art of testing less without sacrificing quality. In: Proc. of the IEEE/ACM Int'l Conf. on Software Engineering. 2015. 483–493.

[32] Bach T, Andrzejak A, Pannemans R. Coverage-Based reduction of test execution time: Lessons from a very large industrial project. In: Proc. of the IEEE Int'l Conf. on Software Testing, Verification and Validation Workshops. IEEE, 2017. 3–12.

[33] Beszédes Á, Gergely T, Schrettner L, Jász J, Langó L, Gyimóthy T. Code coverage-based regression test selection and prioritization in WebKit. In: Proc. of the 28th IEEE Int'l Conf. on Software Maintenance (ICSM). 2012. 46–55.

[34] Strandberg PE, Sundmark D, Afzal W, Ostrand TJ, Weyuker EJ. Experience report: Automated system level regression test prioritization using multiple factors. In: Proc. of the 27th IEEE Int'l Symp. on Software Reliability Engineering (ISSRE). 2016. 12–23.

[35] Engström E, Runeson P, Ljung A. Improving regression testing transparency and efficiency with history-based prioritization-an industrial case study. In: Proc. of the 4th IEEE Int'l Conf. on Software Testing, Verification and Validation (ICST). 2011. 367–376.

[36] Srikanth H, Cashman M, Cohen MB. Test case prioritization of build acceptance tests for an enterprise cloud application: An industrial case study. Journal of Systems and Software, 2016,119:122–135.

[37] Nguyen CD, Perini A, Tonella P, Kessler FB. Automated continuous testing of multiagent systems. In: Proc. of the 5th European Workshop on Multi-Agent Systems (EUMAS). 2007.

[38] Soetens QD, Demeyer S, Zaidman A. Change-Based test selection in the presence of developer tests. In: Proc. of the 17th IEEE European Conf. on Software Maintenance and Reengineering (CSMR). 2013. 101–110.

[39] Vöst S, Wagner S. Trace-Based test selection to support continuous integration in the automotive industry. In: Proc. of the Int'l Workshop on Continuous Software Evolution and Delivery. ACM, 2016. 34–40.

[40] Campos J, Arcuri A, Fraser G, Abreu R. Continuous test generation: Enhancing continuous integration with automated test generation. European Journal of Oral Sciences, 2010,109(4):241–248.

[41] Marijan D. Multi-Perspective regression test prioritization for time-constrained environments. In: Proc. of the IEEE Int'l Conf. on Software Quality, Reliability and Security (QRS). 2015. 157–162.

[42] Gligoric M, Eloussi L, Marinov D. Practical regression test selection with dynamic file dependencies. In: Proc. of the 2015 Int'l Symp. on Software Testing and Analysis. ACM, 2015. 211–222.

[43] Legunsen O, Hariri F, Shi A, Yu YF, Zhang LM, Marinov D. An extensive study of static regression test selection in modern software evolution. In: Proc. of the ACM SIGSOFT Int'l Symp. on Foundations of Software Engineering. 2016. 583–594.

[44] Mirarab S, Akhlaghi S, Tahvildari L. Size-Constrained regression test case selection using multicriteria optimization. IEEE Trans. on Software Engineering, 2012,38(4):936–956.

[45] Vasic M, Parvez Z, Milicevic A, Gligoric M. File-Level vs. module-level regression test selection for .NET. In: Proc. of the Joint Meeting on the Foundations of Software Engineering. ACM, 2017. 848–853.

[46] Jiang B, Zhang Z, Tse TH, Chen TY. How well do test case prioritization techniques support statistical fault localization. In: Proc. of the 33rd Annual IEEE Int'l Computer Software and Applications Conf. 2009,1:99–106.

[47] Do H, Mirarab S, Tahvildari L, Rothermel G. The effects of time constraints on test case prioritization: A series of controlled experiments. IEEE Trans. on Software Engineering, 2010,36(5):593–617.

[48] Knauss E, Staron M, Meding W, Söder O, Nilsson A, Castell M. Supporting continuous integration by code-churn based test selection. In: Proc. of the IEEE Int'l Workshop on Rapid Continuous Software Engineering. 2015. 19–25.

[49] Dösinger S, Mordinyi R, Biffl S. Communicating continuous integration servers for increasing effectiveness of automated testing. In: Proc. of the 27th IEEE/ACM Int'l Conf. on Automated Software Engineering. 2012. 374–377.

[50] Marijan D, Liaaen M. Effect of time window on the performance of continuous regression testing. In: Proc. of the IEEE Int'l Conf. on Software Maintenance and Evolution. 2016. 568–571.

[51] Cho Y, Kim J, Lee E. History-Based test case prioritization for failure information. In: Proc. of the 23rd IEEE Asia-Pacific Software Engineering Conf. (APSEC). 2016. 385–388.

[52] Marijan D, Liaaen M. Test prioritization with optimally balanced configuration coverage. In: Proc. of the IEEE Int'l Symp. on High Assurance Systems Engineering. IEEE, 2017. 100–103.

[53] Yoo S, Nilsson R, Harman M. Faster fault finding at Google using multi objective regression test optimization. In: Proc. of the 8th European Software Engineering Conf. and the ACM SIGSOFT Symp. on the Foundations of Software Engineering (ESEC/FSE 2011). Szeged, 2011.

[54] Marijan D, Gotlieb A, Sen S. Test case prioritization for continuous regression testing: An industrial case study. In: Proc. of the 29th IEEE Int'l Conf. on Software Maintenance (ICSM). 2013. 540–543.

[55] Biswas S, Blanton RD. Reducing test execution cost of integrated, heterogeneous systems using continuous test data. IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, 2011,30(1):148–158.

[56] Celik A, Vasic M, Milicevic A, Gligoric M. Regression test selection across JVM boundaries. In: Proc. of the 11th Joint Meeting on Foundations of Software Engineering. 2017. 809–820.

[57] Mondal D, Hemmati H, Durocher S. Exploring test suite diversification and code coverage in multi-objective test case selection. In: Proc. of the IEEE Int'l Conf. on Software Testing, Verification and Validation. 2015. 1–10.

李英玲(1984－),女,湖南衡阳人,博士生,主要研究领域为持续集成测试,生产力分析.

王青(1964－),女,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为软件过程方法与技术,软件知识工程,经验软件工程.